



# Essais cliniques de recherche de dose en oncologie : d'un schéma d'essai permettant l'inclusion continue à l'utilisation des données longitudinales de toxicité

Adélaïde Doussau de Bazignan

## ► To cite this version:

Adélaïde Doussau de Bazignan. Essais cliniques de recherche de dose en oncologie : d'un schéma d'essai permettant l'inclusion continue à l'utilisation des données longitudinales de toxicité. Santé. Université Paris Sud - Paris XI, 2014. Français. <NNT : 2014PA11T013>. <tel-01138040>

**HAL Id: tel-01138040**

**<https://tel.archives-ouvertes.fr/tel-01138040>**

Submitted on 1 Apr 2015

**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.

THÈSE  
POUR LE DOCTORAT DE SANTÉ PUBLIQUE

SPÉCIALITÉ : RECHERCHE CLINIQUE, INNOVATION  
TECHNOLOGIQUE, SANTÉ PUBLIQUE

Soutenu le 31/03/2014

Par

**Adélaïde DOUSSAU de BAZIGNAN**

Essais cliniques de recherche de dose en  
oncologie : d'un schéma d'essai permettant  
l'inclusion continue à l'utilisation des données  
longitudinales de toxicité

**Directeur de thèse :** Monsieur Xavier PAOLETTI  
**Co-directeur de thèse :** Monsieur Rodolphe THIÉBAUT

**Composition du jury :**

*Présidente du jury :* Madame Pascale TUBERT-BITTER, Centre de Recherche en  
Epidémiologie et Santé des Populations, INSERM, Université Paris-Sud  
*Rapporteurs :* Monsieur Andrew KRAMAR, Centre Oscar Lambret, Lille  
Monsieur Raphaël PORCHER, Université Paris Descartes, Hôpital Saint-  
Louis  
*Examineurs :* Monsieur Nicolas ISAMBERT, Centre Georges-François Leclerc, Dijon  
Monsieur Xavier PAOLETTI, Institut Curie, Paris  
Monsieur Rodolphe THIÉBAUT, Université et CHU de Bordeaux

« Sapience n'entrez point en âme malivoie, et science sans conscience n'est que ruine de l'âme »

François Rabelais, Pantagruel, 1532.

Je dédie ce travail à tous les patients dont les données ont été analysées dans le cadre de ce travail, et aux médecins et professionnels de santé qui prennent soins des patients atteints de maladie cancéreuse. Je dédie aussi ce travail aux patients imaginaires qui ont consenti à être inclus dans des milliers d'essais simulés... Puisse leur « participation » être bénéfique pour des patients réels.

## Résumé

L'objectif des essais de phase I en oncologie est d'identifier la dose maximale tolérée (DMT). Le schéma « 3+3 » nécessite d'interrompre les inclusions en attendant l'évaluation d'une cohorte de trois patients pour définir la dose à attribuer aux patients suivants. Les investigateurs d'oncologie pédiatrique ont proposé l'adaptation *Rolling 6* pour éviter cette suspension temporaire des inclusions. Dans une étude de simulation, nous avons montré qu'un schéma adaptatif avec attribution des doses basées sur un modèle statistique permettait de pallier ce problème, et identifiait plus fréquemment la DMT.

Néanmoins ces trois schémas restent limités pour identifier la DMT, notamment du fait que le critère de jugement est un critère binaire, la survenue de toxicité dose-limitante sur un cycle de traitement. Nous avons proposé un nouveau schéma adaptatif utilisant les données ordinales répétées de toxicité sur l'ensemble des cycles de traitement. La dose à identifier est celle associée au taux de toxicité grave maximal par cycle que l'on juge tolérable. Le grade maximal de toxicité par cycle de traitement, en 3 catégories (grave / modéré / nul), a été modélisé par le modèle mixte à cotes proportionnelles. Le modèle est performant à la fois pour détecter un effet cumulé dans le temps et améliore l'identification de la dose cible, sans risque majoré de toxicité, et sans rallonger la durée des essais. Nous avons aussi étudié l'intérêt de ce modèle ordinal par rapport à un modèle logistique mixte plus parcimonieux.

Ces modèles pour données longitudinales devraient être plus souvent utilisés pour l'analyse des essais de phase I étant donné leur pertinence et la faisabilité de leur implémentation.

**Mots clés :** essais cliniques adaptatifs, essais de recherche de dose, phase I, modèle à cotes proportionnelles, modèles à effet aléatoire, oncologie.

## Abstract

### **Dose-finding clinical trials in oncology: from continuous enrolment, to the integration of repeated toxicity measurements**

Phase I dose-finding trials aim at identifying the maximum tolerated dose (MTD). The “3+3” design requires an interruption of enrolment while the evaluation of the previous three patients is pending. In pediatric oncology, investigators proposed the Rolling 6 design to allow for a more continuous enrollment. In a simulation study, we showed that an adaptive dose-finding design, with dose allocation guided by a statistical model not only minimizes accrual suspension as with the rolling 6, and but also led to identify more frequently the MTD.

However, the performance of these designs in terms of correct identification of the MTD is limited by the binomial variability of the main outcome: the occurrence of dose-limiting toxicity over the first cycle of treatment. We have then proposed a new adaptive design using repeated ordinal data of toxicities experienced during all the cycles of treatment. We aim at identifying the dose associated with a specified tolerable probability of severe toxicity per cycle. The outcome was expressed as the worst toxicity experienced, in three categories (severe / moderate / no toxicity), repeated at each treatment cycle. It was modeled through a proportional odds mixed model. This model enables to seek for cumulated toxicity with time, and to increase the ability to identify the targeted dose, with no increased risk of toxicity, and without delaying study completion. We also compared this ordinal model to a more parsimonious logistic mixed model.

Because of their applicability and efficiency, those models for longitudinal data should be more often used in phase I dose-finding trials.

**Keywords:** adaptive clinical trial, dose-finding trials, phase I, proportional odds model, mixed models, oncology.

## Travaux issus de la thèse

### Publications

#### Articles publiés

Doussau A, Asselain B, Le Deley MC, Geogerger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework*. **Contemp Clin Trials**. 2012 Jul;33(4):657-65.

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials*. **Stat Med**. 2013 Sep 10;32(30):5430-347.

#### Article soumis

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes*.

#### Article en cours de soumission

Doussau A, Thiébaud R, Geogerger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials*.

#### Chapitre de livre

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghiddey W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

## Autres travaux publiés en lien avec la thèse

Floquet A, Doussau A, Brouste V, Cany L, Dutin J, Mathoulin-Pélissier S. *Pegylated liposomal doxorubicin and cyclophosphamide in early recurrent ovarian carcinoma: Phase I dose-finding study*. **Cancer Chemother Pharmacol**. 2014 Jan;73(1):61-8.

Paoletti X, Doussau A, Georger B, Aerts I, Ledele MC, Vassal G, Doz F. *Multi-group Dose-finding Trials in Children: Learning from Adults Trials*. **Treatment Strategies – Oncology**, 2010;1:189-93.

Paoletti X, Postel-Vinay S, Servois V, Doussau A, Ollivier L, Le Tourneau C. *Méthodes pour la recherche de dose de traitements ciblés : nouvelles pistes*. **Bulletin du cancer** 2010;97(12):1485-95.

## Communications

### Communications dans des congrès internationaux

#### Orales

- Doussau A, Thiébaud R, Paoletti X. *Phase I dose finding methods using longitudinal data and proportional odds model in Oncology*. 33<sup>st</sup> annual conference of the International Society for Clinical Biostatistics, Bergen, Norvege 2012.
- Doussau A, Paoletti X, Thiébaud R. *Phase I dose finding methods using longitudinal data and proportional odds model in oncology*, International Biometric Society, International Biometric Conference, Kobe, Japon, Aout 2012 (présentation par Paoletti X).
- Doussau A, Le Deley MC, Paoletti X. *Dose finding methods in pediatric trials: methods comparison in a temporal framework*. 31<sup>st</sup> annual conference of the International Society for Clinical Biostatistics, Montpellier, 2010.

#### Affichée

- Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Paoletti X. *A new approach to integrate the grade of toxicity and later cycles in the analysis and reporting of phase I dose finding trials*. European Cancer Congress (ECCO) Amsterdam, sept 2013. Eur J Cancer. 2013;49(Suppl 2):S163

### Communications orales dans des congrès ou événements scientifiques nationaux

- Doussau A, Thiébaud R, Paoletti X. *A new approach to integrate the grade of toxicity and later cycles in the analysis and reporting of phase I dose finding trials*. Séminaire scientifique annuel du SIRIC « Bordeaux Recherche Intégrée Oncologie, BRIO », Bordeaux, novembre 2013.
- Doussau A. *Utilisation de modèles mixtes pour données ordinales répétées de toxicité dans les essais de phase I de recherche de dose*, Séminaires du service de Biostatistique et Epidémiologie de l'Institut Gustave Roussy, Villejuif, mai 2013.
- Doussau A. *Méthodologie des essais précoces*. Journées du Cancéropole Grand Sud-Ouest, Montpellier, 2012.

- Doussau A, Thiébaud R, Paoletti X. *Analyse de données de toxicité longitudinales dans les essais cliniques de phase I : identification d'une dose cible et détection d'un effet cumulé au cours du temps, à l'aide d'un modèle à odds proportionnels*. Sixième conférence francophone d'épidémiologie clinique (congrès thématique de l'ADLEF-RFUEC) et Journées des Statisticiens des Centres de Lutte contre le Cancer, Lyon 2012.
- Doussau A, Thiébaud R, Postel-Vinay S, Gomez-Rocca C, Kaye S, Soria JC, Paoletti X. *Méthodes de modélisation pour l'analyse de données de toxicité répétées dans les essais cliniques – application sur des données d'essais cliniques de phase I en cancérologie*. Journées des Statisticiens des Centres de Lutte contre le Cancer, Lille 2011.

## Affiliations et financement

### Equipes d'accueil

Institut Curie, service biostatistique (responsable Docteur Bernard Asselain) / INSERM U900 « Cancer et génome : bioinformatique, biostatistiques et épidémiologie d'un système complexe » (responsable Monsieur Emmanuel Barillot), Paris

Centre INSERM U897 Epidémiologie et biostatistique (directeur Professeur Roger Salamon), équipe Biostatistique (responsable Monsieur Daniel Commenges), Bordeaux.

### Financement

Institut National du Cancer (appel d'offre « recherche en santé publique et épidémiologie » 2010, pour le projet **Optidose : Design and statistical methods for phase I/II clinical trials of molecularly targeted agents in oncology**, investigateur coordonnateur : Monsieur Xavier Paoletti).

Division Biostatistique / Bioinformatique (Responsable Professeur Rodolphe Thiébaud de l'INSERM U897) du Labex Vaccine Research Institute.

Ecole doctorale Santé Publique ED 420 (Directeur Mr Jean Bouyer).

Centre Hospitalier-Universitaire de Bordeaux.

Université de Bordeaux.

## Remerciements

P1

**P2**

**P3**

**P4**

## Sommaire

Résumé .....	3
Abstract .....	4
Travaux issus de la thèse .....	5
Publications .....	5
Autres travaux publiés en lien avec la thèse.....	6
Communications.....	7
Affiliations et financement.....	9
Remerciements .....	10
Sommaire .....	14
Table des annexes.....	17
Abréviations .....	18
1 Introduction .....	19
1.1 Préambule.....	19
1.2 Essais de recherche de dose en oncologie : objectifs, spécificités .....	20
1.2.1 Objectif et contexte.....	20
1.2.2 Aspects éthiques et critères d'inclusion.....	22
1.2.3 Doses à explorer .....	23
1.2.4 Critères de jugement.....	24
1.3 Spécificités des essais en oncopédiatrie .....	26
1.4 Schémas d'essai.....	28
1.4.1 Schémas algorithmiques.....	29
1.4.2 Schémas guidés par un modèle statistique .....	32
1.4.3 Résumé des schémas .....	39
1.4.4 Extensions de la méthode de réévaluation continue.....	41
1.4.5 Performance des méthodes de recherche de dose.....	44
1.5 Schémas prenant en compte le caractère ordinal des toxicités .....	47
1.5.1 Incorporation des grades de toxicité dans les schémas algorithmiques.....	47
1.5.2 Modèle logistique à cotes proportionnelles .....	47
1.5.3 Modèle à cote proportionnelles et contrôle du surdosage .....	52
1.5.4 Modèle puissance .....	53
1.5.5 Approches utilisant des scores de toxicité.....	55
1.6 Schémas prenant en compte les toxicités au-delà du premier cycle.....	57
1.6.1 Notations .....	57
1.6.2 Schéma algorithmique avec escalade de dose intrapatient.....	57
1.6.3 Analyse longitudinale de données binaires .....	59

1.7	Objectif de la thèse .....	61
2	Evaluation comparative de deux schémas d'étude destinés à limiter l'interruption des inclusions en oncologie pédiatrique .....	62
2.1	L'essai erlotinib + radiothérapie dans les tumeurs cérébrales de l'enfant.....	63
2.2	Réanalyse algorithmique rétrospective à partir des données de l'essai.....	66
2.2.1	Méthodes .....	66
2.2.2	Résultats .....	67
2.2.3	Discussion .....	68
2.3	Résumé de l'étude de simulation.....	69
2.3.1	Introduction .....	69
2.3.2	Objectif.....	69
2.3.3	Méthodes .....	69
2.3.4	Principaux résultats .....	71
2.4	Article.....	74
2.5	Discussion .....	84
3	Proposition d'un nouveau schéma d'étude basé sur le modèle mixte de régression logistique à cotes proportionnelles.....	86
3.1	Introduction .....	86
3.2	Méthodes .....	87
3.2.1	Explicitation du modèle POMM .....	87
3.2.2	Programmation .....	96
3.3	Etude de simulation évaluant l'intérêt du modèle à cotes proportionnelles dans les essais de recherche de dose .....	97
3.3.1	Objectifs .....	97
3.3.2	Méthodes .....	97
3.3.3	Résultats .....	99
3.3.4	Article.....	101
3.3.5	Analyse ancillaire : impact du nombre de paliers de doses sur les performances.....	120
3.4	Etude appliquée sur les données de trois essais.....	121
3.4.1	Objectif.....	121
3.4.2	Méthodes .....	121
3.4.3	Principaux résultats .....	122
3.4.4	Valorisation .....	123
4	Modèle de régression logistique à cotes proportionnelles <i>versus</i> modèle logistique longitudinal	
4.1	Introduction .....	146
4.2	Méthodes .....	146

4.3	Résultats .....	147
4.4	Synthèse des estimations appliquées à l'essai erlotinib + radiothérapie .....	150
5	Discussion générale.....	152
5.1	Synthèse des principaux résultats.....	152
5.2	Précautions d'emploi.....	154
5.2.1	Impact des scénarios cliniques sur les performances des schémas évalués par simulation 154	
5.2.2	Modèle mixte à <i>intercept</i> aléatoire.....	155
5.2.3	Modèle logistique à cotes proportionnelles.....	156
5.2.4	Codage des variables explicatives.....	157
5.2.5	Définition de la cible de toxicité dans les modèles longitudinaux .....	159
5.2.6	Définition du critère de jugement principal.....	160
5.3	Mise en perspective .....	162
5.3.1	Impact à l'ère des thérapies moléculaires ciblées.....	162
5.3.2	Validité des données longitudinales .....	165
5.3.3	Amélioration des schémas et méthodes d'analyse .....	166
5.3.4	Amélioration de la qualité des publications .....	167
5.3.5	Essais oncologiques : des spécificités méthodologiques à remettre en cause ?.....	167
5.3.6	Rôle du biostatisticien .....	169
5.4	Recommandations pour la pratique .....	171
5.4.1	Indication des méthodes développées.....	171
5.4.2	Pistes d'amélioration à envisager .....	173
6	Conclusion.....	174
7	Références .....	175
	Table des illustrations, modèles et équations .....	181
	Table des notations.....	184
8	Annexes.....	185

## Table des annexes

### Annexes

Annexe 1. Déclaration d'Helsinki de l'Association Médicale Mondiale (Association Médicale Mondiale 2013) .....	185
Annexe 2. Annexes D et F à J de l'article (Doussau, Asselain et al. 2012).....	189
Annexe 3. Calcul de la pente pour le modèle logistique à un paramètre de la CRML .....	192
Annexe 4. Présentation des données d'un essai « essai 64 ». Syntaxe R, SAS et test de l'hypothèse de proportionnalité des cotes .....	193
Annexe 5. Communication affichée présentée au Congrès européen du cancer 2013 (Doussau, Thiebaut et al. 2013).....	198
Annexe 6. Chapitre de livre (Paoletti and Doussau 2014) .....	200
Annexe 7. Distribution des doses recommandées dans les nouveaux schémas adaptatifs utilisant les modèles mixtes logistique ou ordinal pour l'attribution des doses.....	227
Annexe 8. Article comparant le modèle mixte longitudinal au modèle mixte à cotes proportionnelles, et l'impact de la simplification du modèle (soumis).....	228

## Abréviations

### Abréviations utilisées en français

Abréviations en français		Abréviations en anglais	
DMT	Dose maximale tolérée	MTD	<i>Maximum tolerated dose</i>
DR	Dose recommandée	RD	<i>Recommended dose</i>
G	Grade (indice de sévérité)	G	<i>Grade</i>
IC	Intervalle de confiance	CI	<i>Confidence interval</i>
MRC	Méthode de réévaluation continue	CRM	<i>Continual reassessment method</i>
TDL	Toxicité dose limitante	DLT	<i>Dose limiting toxicity</i>

**Abréviations / glossaire anglophones, des expressions qui n'ont pas été traduites** (termes consacrés dans la littérature du domaine dont la traduction aurait alourdi la lecture).

Abréviations		(Référence)
CR-CRM	<i>Continuation ratio-CRM</i>	(Van Meter, Garrett-Mayer et al. 2012)
CRM dnd	<i>CRM with defined number of patients at a dose</i>	(Doussau, Asselain et al. 2012)
CRM fss	<i>CRM with fixed sample size</i>	(Doussau, Asselain et al. 2012)
CRMG	<i>CRM-Grade</i>	(Iasonos, Zohar et al. 2011)
CRML	<i>Likelihood CRM</i>	(O'Quigley and Shen 1996)
EWOC	<i>Escalation with overdose control</i>	(Babb, Rogatko et al. 1998)
GOM	<i>Generalized ordinal logistic model</i>	(Carriere and Bouyer 2006)
GOMM	<i>Mixed generalized ordinal logistic model</i>	(Carriere and Bouyer 2006)
Intercept	<i>Ordonnée à l'origine</i>	
logMM	<i>Logistic mixed model</i>	(Paoletti, Doussau, soumis)
NCI-CTC	<i>National Cancer Institute – Common Toxicity Criteria</i>	(National Cancer Institute and US National Institutes of Health 2010)
POM	<i>Proportional odds model</i>	(Van Meter, Garrett-Mayer et al. 2011)
PO-CRM	<i>Proportional odds model-CRM</i>	(Van Meter, Garrett-Mayer et al. 2011)
POM-EWOC	<i>Proportional odds model-Escalation with overdose control</i>	(Tighiouart, Cook-Weins et al. 2012)
POMM	<i>Mixed proportionnal odds model</i>	(Doussau, Thiebaut et al. 2013)
R6	<i>Rolling 6</i>	(Skolnik, Barrett et al. 2008)
RECIST	<i>Response evaluation criteria in solid tumours</i>	(Eisenhauer, Therasse et al. 2009)
RP2D	<i>Recommended phase II dose</i>	
TITE-CRM	<i>Time to event CRM</i>	(Cheung and Chappell 2000)
TITE-EWOC	<i>Time to event - Escalation with overdose control</i>	(Mauguen, Le Deley et al. 2011)

# 1 Introduction

## 1.1 Préambule

L'objet de cette thèse a été d'étudier deux types de problématiques temporelles des essais de phase I de recherche de dose.

La première est celle de la suspension temporaire des inclusions dans les essais adaptatifs lorsque des patients sont en cours d'évaluation. Dans le contexte des essais d'oncopédiatrie, pour lesquels l'entrée dans un essai de phase I est envisagée comme un moyen d'accès au meilleur traitement disponible, des adaptations pour permettre de limiter cette situation ont été envisagées.

Un autre aspect temporel est celui de l'utilisation des données de tolérance collectées tout au long du suivi des patients sous traitement. En effet les schémas de recherche de dose utilisent habituellement les données du premier cycle de traitement pour identifier la dose recherchée qui sera étudiée en phase II. De plus le critère pris en compte est le plus souvent binaire : la survenue d'au moins une toxicité grave sur cette période restreinte. Le grade de toxicité est une source d'information supplémentaire qui peut être prise en compte de surcroît, et pourrait permettre d'améliorer les schémas de recherche de dose.

Dans cette introduction, nous allons présenter l'ensemble de ces notions : le contexte général des essais de recherche de dose en oncologie, les spécificités de ces essais en oncopédiatrie, les schémas historiques, les développements méthodologiques de ces schémas et les techniques d'analyse rétrospective des données.

Pour les développements méthodologiques, nous détaillerons les méthodes les plus communément utilisées (Le Tourneau, Lee et al. 2009; Le Tourneau, Gan et al. 2012) ou celles qui ont servi de base à des approches plus riches intégrant des données ordinales ou des données longitudinales afin de présenter le contexte dans lequel se situe nos propres travaux de recherche.

## 1.2 Essais de recherche de dose en oncologie : objectifs, spécificités

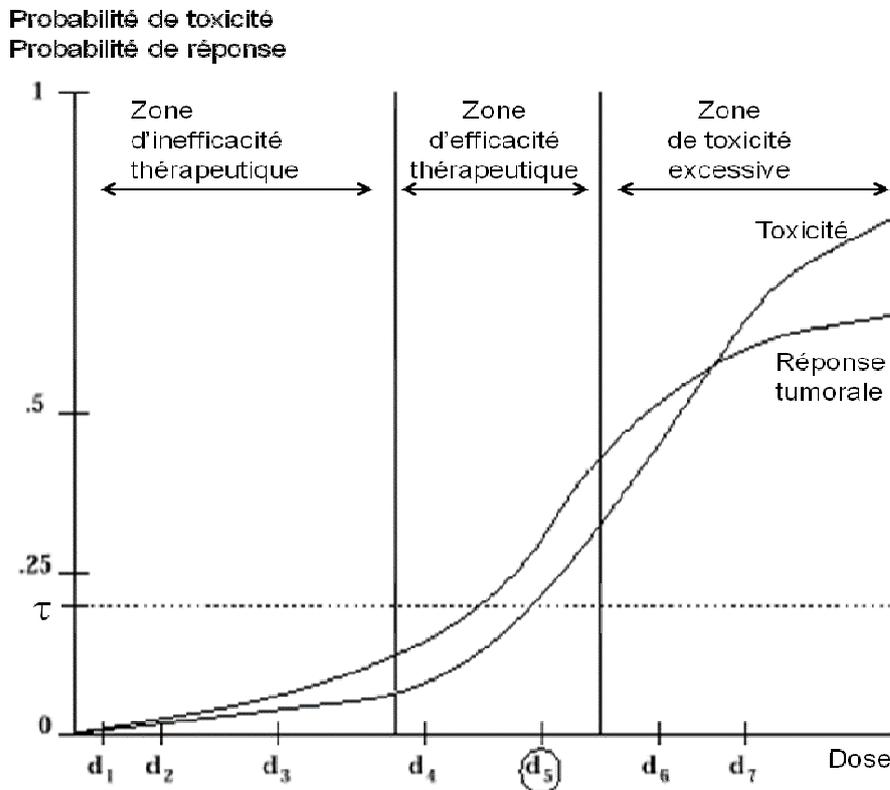
### 1.2.1 Objectif et contexte

L'objectif des essais de recherche de dose est d'identifier une ou plusieurs doses d'un nouveau médicament pour l'homme. Ce médicament entrera ensuite dans le processus de développement du médicament, jusqu'à ce que son efficacité puisse être confirmée, ou non, dans des essais randomisés, comparativement aux traitements de référence.

La recherche de doses a habituellement lieu en phase I ou II selon les phases classiques de développement des médicaments (European Medicines Agency 1998). La phase I de développement d'un nouveau médicament débute lors de la première administration chez l'homme, après les études de toxicologie animale. Les premiers essais sont alors spécialement dédiés à étudier la tolérance chez l'homme, ainsi que les propriétés pharmacologiques (pharmacocinétique, pharmacodynamique) du nouveau médicament, ou d'une nouvelle association de médicaments. Ces essais, dont l'objectif principal n'est pas d'obtenir un effet thérapeutique, peuvent être réalisés avec des volontaires sains. En cancérologie, ils sont habituellement réalisés avec des patients malades.

Dans le domaine de la cancérologie, les schémas d'essai de recherche de dose en phase I se sont développés pour évaluer la tolérance des traitements de chimiothérapie cytotoxiques. Leur mécanisme d'action, agissant à la fois sur les cellules cancéreuses mais aussi sur de nombreux tissus sains, a justifié que la première étape d'essais cliniques ait habituellement consisté à identifier la **dose maximale tolérée (DMT)**. C'est la dose la plus efficace que l'on puisse administrer au patients, avec néanmoins une tolérance jugée acceptable (Storer 1989; Le Tourneau, Lee et al. 2009). A ce titre, il a été communément admis, du fait de la gravité de la maladie, qu'une dose qui présenterait une fréquence de toxicité grave de 15 à 25 % en début de traitement pourrait être sélectionnée pour être étudiée dans les essais de phase II, études exploratoires d'efficacité (Korn, Midthune et al. 1994). Les chimiothérapies cytotoxiques ont en effet la caractéristique de présenter un index thérapeutique faible, c'est-à-dire que la dose efficace et la dose toxique sont proches. La relation dose-efficacité et la relation dose-toxicité suivent habituellement des relations monotones croissantes, permettant de définir deux seuils : un seuil en dessous duquel le traitement est peu toxique et inefficace, et un seuil au dessus duquel le traitement est trop toxique (Figure 1) ; les essais de phase précoce cherchent donc à identifier la ou les doses situées entre ces deux seuils. A l'ère des

chimiothérapies cytotoxiques, on résumait parfois cette quête selon l'adage : « le plus est le mieux » (« *the more, the better* »).



**Figure 1. Relations dose-toxicité et dose-efficacité typiques d'une chimiothérapie cytotoxique.**  $\tau$  correspond à un taux de toxicité de 20 % (ligne pointillée).

Dans ce contexte de première administration chez l'homme visant à expérimenter des traitements à haut risque de toxicité, les schémas d'étude ont donc cherché à introduire le traitement chez les premiers patients à des doses faibles, afin de minimiser les risques de toxicité excessive, et d'augmenter progressivement les doses pour les patients ultérieurs, une fois que la tolérance a été évaluée chez les patients inclus précédents. Les schémas classiques sont des essais séquentiels et ont donc intégré une interruption momentanée des inclusions en attendant l'évaluation de quelques patients inclus à un niveau de dose, avant de prendre une décision sur la dose à administrer pour le ou les patients suivants.

### 1.2.2 Aspects éthiques et critères d'inclusion

Ce type d'étude pose des questions éthiques spécifiques du fait que l'objectif principal est d'étudier la toxicité et que l'espoir d'obtenir une réponse thérapeutique est relativement faible : d'une part les patients inclus aux doses les plus basses auront une probabilité d'efficacité thérapeutique basse et d'autre part les patients qui seront traités aux doses plus élevées auront un risque plus élevé de présenter des toxicités graves. De ce fait, les schémas d'étude cherchent également à augmenter assez rapidement les doses, tout en garantissant la sécurité des patients, afin d'inclure néanmoins le plus possible de patients avec des doses que l'on espère être dans la zone thérapeutique. Une littérature abondante discute des aspects éthiques de ces essais, et plusieurs études ont cherché à évaluer notamment si l'information et le consentement éclairé étaient réellement possibles (Agrawal and Emanuel 2003; Anderson and Kimmelman 2012). Les difficultés de compréhension des participants sont en effets majeures, dans les essais chez l'adulte comme les essais pédiatriques (Daugherty, Ratain et al. 1995; Cousino, Zyzanski et al. 2012; Pentz, White et al. 2012), avec des problèmes de compréhension des objectifs de l'essai chez 2/3 des sujets environ. La déclaration d'Helsinki de l'Association Médicale Mondiale concernant les Principes éthiques applicables à la recherche médicale impliquant des êtres humains [Annexe 1] indique qu'une « recherche médicale impliquant des êtres humains ne peut être conduite que si l'importance de l'objectif dépasse les risques et inconvénients pour les personnes impliquées » (Association Médicale Mondiale 2013). La mise en œuvre des essais de recherche de dose met en tension la balance bénéfice-risque individuelle et l'intérêt collectif. En France, le Comité consultatif National d'éthique (CCNE) a été saisi à ce sujet (Comité consultatif national d'éthique pour les sciences de la vie et de la santé 2002). La première recommandation du CCNE était de considérer comme prioritaire le développement méthodologique des essais de phase I pour diminuer le risque de toxicité et rechercher conjointement toxicité et efficacité. Les autres recommandations étaient : d'avoir recours à des volontaires sains lorsque cela est possible ; rappeler aux participants que l'objectif majeur est d'étudier la tolérance, éviter le mot « traitement » et faire en sorte que l'espoir modeste de bénéfice ne cache pas l'incertitude sur le bénéfice ; informer la société de la réalité et de la nécessité des essais médicamenteux de ce type. La population à inclure était également l'objet de recommandations : patients en impasse thérapeutique, mais pas en fin de vie ; choix de patients dont la tumeur a une chance d'être sensible à la nouvelle molécule. En oncopédiatre, la difficulté du consentement et le rôle de la famille étaient soulignés. Enfin, il était recommandé d'offrir une prise en charge de tous les désagréments par des soins d'accompagnement optimaux, afin d'assurer une qualité

de vie la meilleure possible pour ces patients, et que l'approche palliative adaptée à l'état du patient reste toujours prioritaire sur les exigences de la pratique de l'essai clinique de phase I.

Afin de garantir que la participation des patients ne constitue pas une perte de chance dans leur prise en charge, ces essais sont donc réalisés chez des patients en échec thérapeutique pour tous les traitements actuellement disponibles dans leur indication.

D'après les données de revues de la littérature des essais de phase I, le risque de décès toxique serait faible, de l'ordre de 0,5 % et la probabilité de réponse tumorale est dans l'ensemble de 5 % (Agrawal and Emanuel 2003; Horstmann, McCabe et al. 2005; Piantadosi 2005; Italiano, Massard et al. 2008). La survie globale des patients inclus dans les essais de phase I est en médiane de 5 à 10 mois (Bachelot, Ray-Coquard et al. 2000; Roberts, Goulart et al. 2004; Italiano, Massard et al. 2008; Olmos, A'Hern R et al. 2012). Dans une étude pronostique sur environ 2 000 patients inclus entre 2005 et 2007, dans 14 centres, la médiane de survie était de 10 mois, avec un taux de mortalité à 30 jours de 3 % et de 16,5 % à 90 jours (Olmos, A'Hern R et al. 2012).

### 1.2.3 Doses à explorer

Pour les essais en première administration chez l'homme, dans les cas où l'index thérapeutique est faible, la première dose explorée doit être basse (European Medicines Agency 1994). La dose de départ est habituellement un dixième de la dose létale ayant entraîné la mort de 10 % d'une population de murins, rapportée au poids en mg/m<sup>2</sup>. Des méthodes plus sophistiquées prenant en compte les données de pharmacocinétique et pharmacodynamique sont disponibles (Committee for medicinal products for human use 2007).

Les doses à explorer consistent habituellement en une suite de 8 à 10 paliers de dose, souvent établis de manière empirique, avec des incréments additifs (n, n+10, n+20, n+30...), ou multiplicatifs (n, 2n, 4n, 8n...). Historiquement, la suite de Fibonacci modifiée a été utilisée, permettant des incréments élevés au début puis plus prudents par la suite (n, 2n, 3,3n, 5n, 7n, 9n, 12n, 16n..., ce qui correspondait à des incréments de 100 %, 65 %, 52 %, 40 %, 29 %, 33 %, 33 %). Les recommandations de l'Agence européenne du médicament sont de se baser sur les données précliniques pour définir les doses à explorer (Committee for medicinal products for human use 2007).

Le traitement est habituellement administré tant qu'il est correctement toléré et jusqu'à progression de la maladie. Dans la majorité des schémas, l'escalade de dose n'est pas autorisée chez un patient : le patient reçoit la dose qui lui est attribuée au départ (European Medicines Agency 1994) ; les toxicités peuvent nécessiter une diminution temporaire de la dose, jusqu'à l'interruption du traitement.

#### 1.2.4 Critères de jugement

- **Critère de jugement principal**

Le critère de jugement principal classique dans ces essais est la survenue d'une **toxicité dose-limitante (TDL)**, sur une fenêtre d'un à deux cycles de traitement. La sévérité des événements indésirables est habituellement cotée selon la classification National Cancer Institute Common Toxicity Criteria (NCI-CTC) (Trotti, Colevas et al. 2003; National Cancer Institute and US National Institutes of Health 2010) : le grade 1 correspond à des événements indésirables mineurs, le grade 2 à des événements modérés, le grade 3 à des événements graves (par exemple, nécessitant une hospitalisation), le grade 4 à la mise en jeu du pronostic vital, et le grade 5 au décès.

La définition du critère de jugement principal est habituellement explicitée dans le protocole et regroupe des événements indésirables graves, de grade 3, 4 ou 5. Les toxicités rapidement résolutive avec des traitements symptomatiques (nausées, vomissements, troubles hydro-électrolytiques), ou jugées cliniquement tolérables (alopécie) sont habituellement exclues. L'imputabilité au traitement de la recherche est prise en compte : seuls les effets indésirables (toxicités pour lesquelles le lien au traitement étudié est possible, probable ou certain) sont considérés. La durée d'évolution de l'effet indésirable, sa réversibilité, et les interruptions de traitement sont parfois également intégrées (Le Tourneau, Razak et al. 2011). La définition initiale intègre aussi parfois des toxicités modérées (grade 2) non-hématologiques (dans 20 % des essais, d'après une revue de 155 essais de thérapies moléculaires ciblées sur la période 2000-2010).

Ces critères sont très variables dans la littérature et sont parfois modifiés en cours d'essai (24 % des 155 essais de thérapies moléculaires ciblées en monothérapie de la revue de la littérature de Le Tourneau et al.). Ainsi, des toxicités modérées avaient été finalement classées comme TDL car considérées intolérables dans 8 % des essais, des toxicités survenues au delà de la fenêtre prédéfinie d'évaluation des TDL ont été prises en compte dans 8% des essais, et,

*a contrario*, des toxicités qui rentraient dans la définition des TDL n'ont finalement pas été classées comme TDL dans 8 % des essais.

- **Critères de jugement secondaires**

Dans ces essais, les objectifs secondaires regroupent des paramètres pharmacocinétique et pharmacodynamiques, et l'efficacité thérapeutique exploratoire selon les critères « *Response evaluation criteria in solid tumours* » (RECIST) évaluant la réponse tumorale pour les tumeurs solides (Eisenhauer, Therasse et al. 2009). La tolérance est surveillée pendant les cycles de traitement administrés au delà de la fenêtre d'évaluation de la TDL, mais ces données sont irrégulièrement présentées dans la littérature, et ne sont habituellement pas prises en compte dans la détermination de la **dose recommandée pour les essais de phase II (RP2D)**, sauf dans de très rares cas (Le Tourneau, Razak et al. 2011; Moroney, Fu et al. 2012).

### 1.3 Spécificités des essais en oncopédiatrie

Les essais de recherche de dose en oncologie pédiatrique présentent plusieurs spécificités par rapport aux essais menés chez l'adulte. Tout d'abord il s'agit bien évidemment d'une population protégée en termes de réglementation de la recherche clinique. Les essais pédiatriques sont néanmoins jugés nécessaires et actuellement recommandés de manière générale pour les médicaments pédiatriques, avec une réglementation européenne mise en place depuis 2006 (Agence Nationale de Sécurité du Médicament et des Produits de Santé 2006; The European Parliament and The Council of the European Union 2006).

Il s'agit également d'une maladie relativement rare ; en France, l'incidence annuelle des cancers de l'enfant est estimée à 153 cas par million, soit environ 1 700 nouveaux cas par an, tous types de cancers confondus (Institut National du Cancer 2010). Ces essais sont particuliers également du fait que la dose maximale tolérée est déjà connue chez l'adulte avant de mettre en place un essai pédiatrique (Smith, Bernstein et al. 1998), que l'on peut supposer que la DMT chez l'enfant soit liée à la DMT chez l'adulte, comme cela a été montré dans la littérature (Lee, Skolnik et al. 2005; Paoletti, Georger et al. 2013). De ce fait, il est classiquement recommandé de débiter les essais de recherche de dose à 80 % de la DMT de l'adulte, et l'exploration de 2 à 4 niveaux de dose est le plus souvent suffisante pour identifier la dose maximale tolérée. On considère également que le risque de toxicité grave inattendue est moindre, puisque le profil de toxicité est déjà connu chez l'adulte.

L'arsenal thérapeutique en oncologie pédiatrique est beaucoup plus réduit qu'en oncologie adulte. Dans une population d'enfants pour lesquels plus aucune autre option thérapeutique n'est disponible, le respect des principes éthiques de la recherche, est particulièrement délicat (Davous, Doz et al. 2007). Lorsque des essais cliniques de phase I sont ouverts dans un centre d'oncologie pédiatrique, étudiant l'effet de molécules qui sont déjà administrées chez l'adulte, la participation à un essai est souvent considérée par les oncologues pédiatriques comme une chance d'accéder au meilleur traitement accessible pour l'enfant. Par ailleurs, du fait de la rareté de la maladie, les essais de recherche de dose sont habituellement multicentriques. Du fait des délais pour la mise à jour de l'information entre les centres, il peut donc arriver que deux familles soient informées d'un essai en même temps, et il est alors éthiquement délicat de ne pas pouvoir accepter l'inclusion des deux enfants. Les investigateurs rapportent qu'il est difficile, auprès de la famille et de l'enfant, de refuser l'inclusion dans un essai ouvert, parce

que les inclusions sont suspendues dans l'attente de l'évaluation des patients précédents inclus [F Doz, Communication personnelle].

Globalement, d'après les données de la littérature, sur deux mille enfants inclus dans une soixantaine d'essais entre 2000 et 2004 (Lee, Skolnik et al. 2005), le taux de réponse tumorale (environ 8 à 10 %) était légèrement supérieur à celui observé chez l'adulte, avec une fréquence des décès toxiques équivalente à celle observée chez l'adulte (0,5 %).

## 1.4 Schémas d'essai

De nombreux schémas d'étude ont été proposés pour les essais de recherche de dose en oncologie.

Leurs caractéristiques communes sont de débiter en incluant les premiers patients à une dose raisonnablement basse, de définir une règle d'escalade de dose pour les patients suivants, en fonction des toxicités observées parmi les patients précédents.

On distingue classiquement deux grandes familles de schémas d'étude :

- les schémas d'escalade de dose algorithmiques,
- les schémas d'escalade de dose guidés par un modèle statistique de la relation dose-toxicité.

Nous présenterons dans cette partie les schémas historiques les plus utilisés de ces deux familles, pour lesquels le critère de jugement principal est la survenue d'une TDL au cours d'une fenêtre d'observation donnée (le premier ou les deux premiers cycles de traitement dans la majorité des cas). La majorité des schémas a été développée en oncologie adulte, et nous reverrons aussi une proposition spécifique faite pour l'oncologie pédiatrique. Notre propos se concentrera sur les méthodes destinées à rechercher la dose d'une monothérapie, ou rapportées à une seule dimension, c'est-à-dire que les paliers de dose peuvent être ordonnés selon une toxicité croissante.

De nombreuses extensions ont été mises en œuvre, et nous avons sélectionné les travaux des auteurs qui se sont intéressés à prendre en compte les toxicités tardives, ou ordinales, qui sont le socle des méthodes que nous avons développées.

## 1.4.1 Schémas algorithmiques

### 1.4.1.1 Le Schéma « 3+3 »

Le schéma « 3+3 » est le schéma le plus connu, et le plus utilisé dans la littérature (par exemple, utilisé dans 62 % des essais de phase I étudiant des thérapies moléculaires ciblées administrées en monothérapie, publiés entre 2000 et 2010 (Le Tourneau, Razak et al. 2011)). Il a été formalisé dans la littérature dans les années 90 (Storer 1989), bien que des schémas similaires escalade/désescalade (*up and down*) aient été proposés dans la littérature dès 1948 par Dixon et Mood. Ce schéma « 3+3 » a parfois été dit « de Fibonacci », par confusion entre la règle définissant les incréments de dose (dérivée de la suite de Fibonacci), et la règle d'escalade de dose, mais cette dernière n'a pas de rapport avec le mathématicien italien du XIII<sup>ème</sup> siècle.

Dans ce schéma d'escalade de dose, trois patients sont traités à la première dose. Si aucun patient ne présente de TDL, le palier de dose supérieur sera indiqué pour les trois patients suivants. Si deux patients présentent une TDL, les trois patients suivants recevront la dose inférieure. Si un patient sur trois présente une TDL, trois patients supplémentaires seront inclus au même palier de dose, et les patients suivants recevront la dose supérieure si un patient ou moins sur les six présente une TDL. Dans ce schéma, la dose recommandée à l'issue de l'essai est habituellement le palier de dose inférieur à celui pour lequel on aura observé un taux de TDL supérieur ou égal à 33 % (deux patients sur trois ou sur six), et pour lequel on aura observé au plus une TDL sur six patients ( $\leq 16,7\%$ ). En pratique, l'écriture probabiliste de ce schéma a permis de montrer que ce schéma permet d'identifier des doses dont le taux de toxicité est situé entre 10 % et 25 % (Korn, Midthune et al. 1994). Ce type de schéma est relativement conservateur car il est peu probable qu'il recommande une dose qui ait une toxicité supérieure à 30 % (pas plus de 10 à 15 % de recommandations par simulations d'essais), mais la dose cible recherchée est identifiée dans moins de 40 % des cas dans la majorité des scénarios (Simon, Freidlin et al. 1997; Reiner, Paoletti et al. 1999; Piantadosi 2005). Ce schéma est également apprécié des investigateurs car simple à mettre en place et à mener.

Une des critiques majeures est que la probabilité cible de toxicité (une TDL sur six patients, soit 16,7 %) n'est pas explicitement spécifiée pour s'adapter aux hypothèses de l'essai ; le nombre de patients traités à des doses basses, la lenteur de l'escalade de dose aux paliers inférieurs, ainsi que l'interruption des inclusions en attendant que les trois patients d'un palier

soient évalués pour la TDL sont également reprochés à ce schéma (Simon, Freidlin et al. 1997; Skolnik, Barrett et al. 2008; Le Tourneau, Lee et al. 2009).

#### *1.4.1.2 Escalade accélérée (Accelerated titration design)*

En 1997, plusieurs modifications ont été proposées pour pallier à la lenteur d'exploration des paliers inférieurs dans le schéma « 3+3 ». Simon et al ont proposé des schémas numérotés 1, 2, 3, 4. Le n°1 correspond au « 3+3 ». Les schémas 2, 3 et 4, sont des schémas d'escalade accéléré (*Accelerated titration design*) et sont déclinés sans possibilité d'escalade de dose intrapatient (2A, 3A et 4A), et avec escalade de dose intrapatient (2B,3B et 4B).

Parmi ceux-ci, les schémas 2A et 3A consistaient à permettre une escalade de dose plus rapide au début de l'essai, avec pour chaque nouveau patient inclus un incrément de dose de 40 % pour le schéma n°2A, et de 100 % pour le schéma n°3A, jusqu'à la survenue de la première TDL ou deux toxicités modérées sur le premier cycle. Une fois que la première TDL ou deux toxicités modérées au premier cycle ont été observées, l'escalade de dose se poursuit selon le schéma « 3+3 ». Les autres schémas proposés dans cette publication seront décrits ultérieurement.

Selon les études de simulations, ces schémas 2A et 3A n'amélioreraient pas la capacité à identifier la dose ciblée par rapport au schéma classique « 3+3 », mais permettait d'inclure moins de patients à des doses basses, de réduire le nombre de patients inclus, sans risques excessifs en termes de sécurité. D'après plusieurs revues de la littérature, la capacité à réduire le nombre de patients traités en dessous de la dose recommandée a été confirmée, néanmoins la durée des essais n'était pas significativement plus courte (Penel, Isambert et al. 2009; Le Tourneau, Gan et al. 2012).

Ces schémas ont souvent été implémentés en pratique ; ils ont été utilisés dans 30 % des essais de phase I étudiant des thérapies moléculaires ciblées publiés entre 2000 et 2010 (Le Tourneau, Razak et al. 2011).

### 1.4.1.3 Roulement par six (Rolling 6)

Dans le contexte des essais de pédiatrie oncologique, un schéma destiné à limiter la fermeture temporaire des inclusions dans les essais pendant l'évaluation de trois enfants du « 3+3 » a été proposé (Skolnik, Barrett et al. 2008). Dans ce schéma, il est possible d'inclure jusqu'à six patients au palier de dose en cours d'évaluation. La dose indiquée pour un nouveau patient est déterminée en fonction du nombre de patients en attente d'évaluation et du nombre de TDL observées sur les patients déjà évalués à ce palier. Si deux TDL ont été observées à ce palier, la dose inférieure est indiquée ; la dose supérieure est indiquée si 3/3, 4/4, 5/5, 5/6 ou 6/6 patients sont évalués sans TDL au palier en cours ; dans tous les autres cas, le patient est inclus à la dose en cours, jusqu'à un effectif maximal de six patients. Les règles d'escalade de dose sont présentées dans le Tableau 1, en fonction du nombre de patients déjà inclus au palier en cours d'exploration.

**Tableau 1. Règle d'escalade de dose avec le schéma *Rolling 6*, selon (Zhao, Lee et al. 2011).**

Nombre de patients inclus	Nombre de patients sans TDL	Nombre de patients en attente d'évaluation	Décision
1	-	-	Même dose
2	2	-	Désescalade*
	autre	-	Même dose
3	$\geq 2$	-	Désescalade*
	0	0	Escalade <sup>#</sup>
	autre	-	Même dose
4	$\geq 2$	-	Désescalade*
	0	0	Escalade <sup>#</sup>
	autre	-	Même dose
5	$\geq 2$	-	Désescalade*
	0	0	Escalade <sup>#</sup>
	autre	-	Même dose
6	$\geq 2$	-	Désescalade*
	$\leq 1$	1	Escalade <sup>#</sup>
	0	-	Escalade <sup>#</sup>
	autre	-	Suspension

\* Si la désescalade intervient lors de l'exploration du premier niveau de dose, l'étude est interrompue ; la dose maximale tolérée a été atteinte si six patients ont déjà été inclus à la dose inférieure. <sup>#</sup> Si le dernier palier à explorer a été atteint, la dose maximale tolérée est atteinte.

D'après l'étude de simulation qui a comparé le schéma *Rolling 6* au « 3+3 », cette variation permettait non seulement de limiter le nombre d'enfants refusés à l'inclusion, mais également d'accélérer la durée des essais, sans risque majoré de toxicité.

## 1.4.2 Schémas guidés par un modèle statistique

### 1.4.2.1 Méthode de réévaluation continue

La méthode de réévaluation continue (MRC, ou *Continual reassessment method, CRM*) a été proposée dans les années 1990 (O'Quigley, Pepe et al. 1990) pour tenter de pallier les principales limites des schémas d'escalade de dose algorithmiques, et constitue un tournant majeur dans les schéma de recherche de dose en oncologie.

C'est un schéma d'étude adaptatif pour lequel le choix de la dose à administrer à un nouveau patient inclus dans l'essai est basé sur l'estimation de la relation dose-toxicité par un modèle statistique, et en fonction de la cible de toxicité choisie pour l'essai. Contrairement aux schémas algorithmiques, la probabilité de toxicité ciblée,  $\tau$ , est explicitement fixée. Le plus souvent la valeur est fixée à 20 ou 25 %, mais d'autres valeurs ont été étudiées (par exemple 30 % ou 35 %).

Cette méthode a été développée dans un contexte d'inférence Bayésienne, puis ultérieurement adaptée à l'inférence par le maximum de vraisemblance (*CRML* pour *likelihood CRM*) (O'Quigley and Shen 1996). Elle s'implémente avec les mêmes contraintes : permettre de prendre en compte l'inefficacité du traitement administré aux doses basses, le risque de toxicité aux doses élevées, la méconnaissance de la relation dose-toxicité en début d'essai et, enfin, en l'absence de preuve d'efficacité du traitement, la contrainte de mettre en place des schémas d'étude incluant peu de patients (Chevret and Zohar 2006).

La caractéristique commune des modèles utilisés pour modéliser la relation dose-toxicité est leur propriété monotone croissante en fonction de la dose, pour respecter l'hypothèse sous-jacente que la dose tolérable la plus élevée sera la plus efficace que l'on puisse administrer.

Soit  $d_\ell$ , une des doses parmi  $L$  doses croissantes à explorer  $D_L = (d_1, \dots, d_\ell, \dots, d_L)$ . Les probabilités de toxicité de chacune des  $L$  doses croissantes ( $\ell = 1, \dots, L$ ) explorées,  $p_\ell$ , répondront à la contrainte  $p_1 \leq \dots \leq p_\ell \leq \dots \leq p_L$ .

Soit  $Y_i$  une variable binaire, la survenue d'une toxicité dose-limitante au premier cycle de traitement pour le patient  $i$  ( $Y_i = 1$ ), *versus*  $Y_i = 0$  (absence de TDL),  $x_i$  la dose  $d_\ell$  reçue par ce patient parmi les  $L$  doses à explorer, et  $\theta$  le vecteur des paramètres du modèle.

Soit  $\Psi(x_i, \theta)$ , une fonction de la dose  $x_i$ , qui permet de modéliser la probabilité de toxicité à l'aide des paramètres  $\theta$ .

$$\psi(x_i, \theta) = P(Y_i = 1 | x_i = d_l) = E(Y_i | x_i)$$

- **Inférence et estimation dans le cadre Bayésien**

Dans la mise en œuvre initiale de la MRC, implémentée dans le cadre conceptuel Bayésien, le premier patient était inclus à la dose à explorer supposée la plus proche de la DMT par les investigateurs. Cette dose pouvait être considérée comme une hypothèse *a priori* de la DMT, même si elle n'était parfois pas la dose la plus basse à explorer. Une fois la réponse observée chez un patient, la dose à assigner au nouveau patient était déterminée en fonction de cet *a priori* et des données observées sur les patients précédemment inclus.

Soit  $\Omega_n = \{(x_1, y_1) \dots (x_n, y_n)\}$  regroupant les données accumulées après  $n$  inclusions,  $\pi(\theta)$  la densité *a priori* de  $\theta$  qui reflète la connaissance de la relation entre la dose et la toxicité avant que l'essai ne commence, et  $f(\Omega_n | \theta) = L_n(\theta)$  la vraisemblance des  $n$  observations.

$$f(\Omega_n | \theta) = L_n(\theta) = \prod_{i=1}^n [\psi(x_i; \theta)]^{y_i} [1 - \psi(x_i; \theta)]^{1-y_i} \quad (\text{Équation A})$$

La distribution de  $\theta$ , vecteur des paramètres du modèle, est modifiée, selon le théorème de Bayes, de la manière suivante :

$$\pi(\theta | \Omega_n) = \frac{f(\Omega_n | \theta) \pi(\theta)}{\int_A f(\Omega_n | \theta) \pi(\theta) d\theta}, \text{ avec } A, \text{ domaine de définition de } \theta.$$

La moyenne *a posteriori* de  $\theta$  peut ainsi être calculée :

$$\tilde{\theta}_n = E(\theta | \Omega_n) = \int_A \theta \pi(\theta | \Omega_n) d\theta$$

La distribution de probabilité de toxicité  $p_{ln}$  de chaque dose  $d_l$  après les  $n$  premiers patients inclus,  $\Psi(d_l, \theta_n)$ , peut alors être estimée par sa moyenne :

$$p_{ln} = \int_A \psi(x_i; \theta) \pi(\theta | \Omega_n) d\theta \approx \psi(d_l; \tilde{\theta}_n)$$

Ce processus est répété de manière itérative pour tout nouveau patient inclus. La dose  $x_{n+1}$  attribuée pour un nouveau patient inclus, le patient  $n+1$ , est la dose  $d_l$  dont la probabilité de

toxicité  $p_{tn}$  est la plus proche de  $\tau$ , cible de probabilité de toxicité maximale tolérable. Cela revient à chercher la dose  $d_l$  qui minimise la quantité  $|p_{tn} - \tau|$ . On déterminera de la même manière la dose finale recommandée à l'issue de l'essai.

- **Inférence et estimation par le maximum de vraisemblance**

Le schéma MRC, initialement proposé dans un cadre Bayésien, a été critiqué notamment à cause de l'impact potentiel du choix de l'*a priori* sur les résultats d'études incluant au total un petit échantillon, et du choix d'une dose de départ différente de la première dose (O'Quigley and Shen 1996). La méthode a donc été proposée avec estimation par le maximum de vraisemblance, avec des propriétés opératoires équivalentes.

La vraisemblance  $f(\Omega_n|\theta)$  des données observées sur les  $n$  premiers patients évalués s'écrit de la même manière que précédemment (Equation A), et  $\theta_n$  est estimé par la méthode du maximum de vraisemblance.

Par extension, un intervalle de confiance peut être calculé par approximation normale, pour chaque niveau de dose :  $IC_{(1-\alpha)\%}(d_l) = \psi \left( d_l; \widehat{\theta}_n \mp z_{1-\alpha} \sqrt{v(\widehat{\theta}_n)} \right)$ ,  $z_\alpha$  étant le percentile de rang  $\alpha$  de la distribution Normale. Une correction de Cornish-Fisher peut être utilisée (O'Quigley 1992), adaptée aux petits échantillons.

Pour la suite de notre propos nous nous concentrerons principalement sur l'approche utilisant des estimations par le maximum de vraisemblance de la MRC.

- **Choix du modèle de la relation dose-toxicité**

Le modèle initialement utilisé était un modèle, courbe convexe-concave, explicité par une fonction tangente hyperbolique de la dose, équivalent à un modèle puissance avec reparamétrisation de la dose  $\mu_i = [\tanh(x_i) + 1] / 2$  (O'Quigley 1992; Shen and O'Quigley 1996).

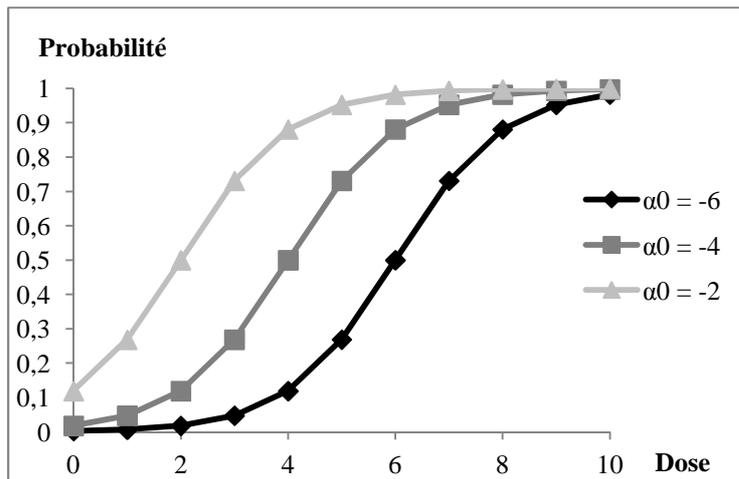
$$\psi(x_i, \theta) = \mu_i^\theta \quad \text{(Modèle 1)}$$

La fonction logistique a été ensuite proposée :

$$\psi(x_i, \theta) = \psi(x_i, \alpha_0, \beta_1) = \frac{e^{(\alpha_0 + \beta_1 x_i)}}{1 + e^{(\alpha_0 + \beta_1 x_i)}} \quad (\text{Modèle 2})$$

Le modèle 2 nécessite l'estimation de 2 paramètres : l'*intercept*  $\alpha_0$  (ordonnée à l'origine) et la pente  $\beta_1$ . D'emblée, il a été proposé d'utiliser un modèle à un paramètre en fixant soit l'*intercept*, soit la pente. Des études de simulations ont montré qu'un modèle à deux paramètres n'était pas plus performant qu'un modèle à un paramètre (Paoletti and Kramar 2009). Le principal argument sous-jacent est que du fait de l'échantillonnage à la meilleure estimation de la DMT, l'information observée est concentrée sur un petit nombre de paliers de dose (asymptotiquement la fréquence d'attribution de la dose cible tend vers un) et un modèle à un paramètre devient saturé. Le plus important n'est pas d'estimer la courbe dose-toxicité dans son ensemble, mais uniquement d'estimer correctement la probabilité de toxicité aux alentours de la dose maximale tolérée (O'Quigley, Pepe et al. 1990). Pour assurer la convergence asymptotique dans le modèle à un paramètre, le modèle doit vérifier les conditions suivantes : croissance en la dose, monotonie en  $\beta_1$ , le paramètre d'effet de la dose, et une troisième condition d'adéquation locale. Shen et O'Quigley, Cheung et Chappell, ont montré les conditions sous lesquelles on obtiendrait une convergence asymptotique, même si le modèle est mal spécifié (Shen and O'Quigley 1996; Cheung and Chappell 2002). Il a été montré dans la littérature que l'estimation de la probabilité de toxicité converge asymptotiquement vers la vraie valeur de la probabilité de toxicité de la DMT avec plusieurs modèles (Shen and O'Quigley 1996).

La relation entre la dose et la probabilité d'événement d'intérêt est représentée sur la Figure 2 pour des doses allant de 0 à 10, une pente  $\beta_1 = 1$  et des *intercept*  $\alpha_0$  prenant les valeurs -2, -4 et -6.



**Figure 2. Probabilité d'évènement en fonction de la dose, dans un modèle logistique de pente 1 et d'intercept -6, -4 ou -2.**

- **Mise en œuvre pratique**

Le principe de la MRC consiste donc à utiliser toutes les données collectées sur les patients précédents pour sélectionner la dose  $d_i$  à administrer au nouveau patient à inclure dans l'essai.

La dose sélectionnée sera celle dont la probabilité de toxicité estimée est la plus proche de la cible  $\tau$ . Le modèle de la relation dose-toxicité et le type d'inférence ne suffisent pas à définir le schéma d'étude. Comme cela a été introduit pour les méthodes Bayésiennes, il faut définir une méthode de démarrage pour le cadre d'estimation par le maximum de vraisemblance. Il faut également définir une règle d'arrêt de l'essai.

**a) Démarrage**

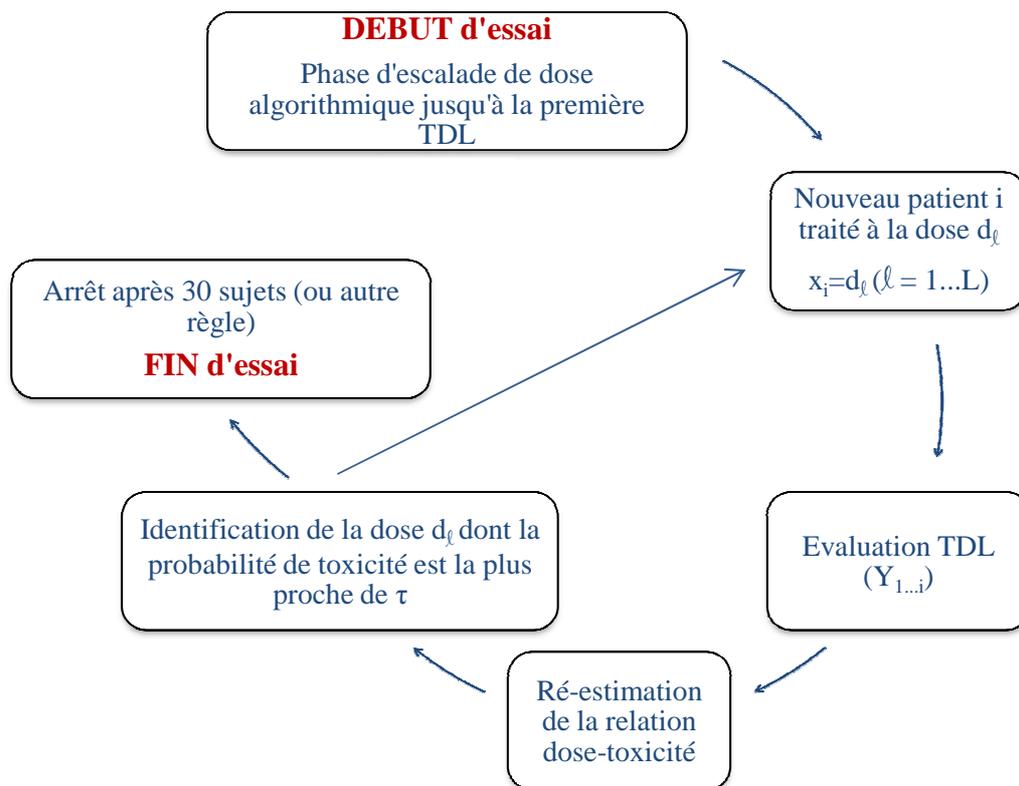
Dans le contexte d'inférence Bayésienne, le premier patient est traité selon la probabilité de toxicité *a priori*. Dans le contexte des schémas utilisant l'estimation par le maximum de vraisemblance, il est nécessaire, pour un modèle à un paramètre, d'observer à la fois une première TDL ainsi qu'un patient sans TDL pour pouvoir estimer le paramètre du modèle, c'est-à-dire une hétérogénéité des réponses observées. Dans le cadre d'un modèle à deux paramètres, la condition d'hétérogénéité est encore plus contraignante (au moins une toxicité et une non toxicité à deux doses explorées). Un schéma algorithmique classique ou une inférence Bayésienne peuvent alors être utilisés en début d'essai jusqu'à cette condition d'hétérogénéité. Une escalade rapide peut être mise en place en début d'essai, par exemple escalader la dose tant qu'aucune TDL n'a été observée au palier inférieur.

## b) Arrêt

Plusieurs règles d'arrêt ont été proposées. L'essai peut s'arrêter après un nombre total fixe de patients inclus (par exemple, 30 patients). Certaines règles sont basées selon l'arbre décisionnel anticipé en fonction des réponses futures possibles. Un arbre décisionnel peut être construit en calculant la dose recommandée aux patients à venir selon toutes les combinaisons de réponses possibles. Dans ce cas, l'essai peut s'arrêter si la probabilité que tous les patients futurs reçoivent le même palier de dose est élevée (O'Quigley and Reiner 1998).

D'autres règles consistent à s'arrêter lorsqu'un nombre suffisant de patients ont été traités consécutivement à une même dose, par exemple 10 patients traités à la même dose. D'autres règles sont conçues selon la précision de l'estimation de la probabilité de toxicité à la dose maximale tolérée (Zohar and Chevret 2001; O'Quigley 2002).

Le schéma de la méthode continue de réévaluation peut être résumé par la Figure 3 3.



**Figure 3. Représentation schématique d'un essai mené avec une méthode de réévaluation continue.**  $\tau$  = probabilité de toxicité cible,  $Y$  = critère de jugement principal, (TDL = toxicité dose-limitante).

- **Avantages et inconvénients de la méthode de réévaluation continue**

Les propriétés de ce schéma ont été étudiées par simulations, par rapport aux méthodes algorithmiques (Korn, Midthune et al. 1994; Piantadosi 2005; Iasonos, Wilton et al. 2008; Le Tourneau, Lee et al. 2009). Par opposition aux schémas algorithmiques qui sont dits sans mémoire puisque la dose attribuée pour un nouveau patient ne dépend que des observations au palier en cours d'exploration, le schéma MRC permet d'utiliser, pour recommander la dose de chaque nouveau patient inclus, les données collectées dans la fenêtre d'évaluation des TDL sur l'ensemble des patients inclus.

Ce schéma permet dans l'ensemble d'atteindre la dose maximale tolérée plus rapidement que les schémas algorithmiques, c'est-à-dire de traiter moins de patients aux doses infrathérapeutiques, en concentrant les patients autour de la dose maximale tolérée. Asymptotiquement, il est démontré que l'estimation de la probabilité de toxicité converge vers l'identification de la DMT. La MRC permet également de ne pas spécifier à l'avance les paliers de doses à explorer, bien que le plus souvent ces paliers soient néanmoins déterminés à l'avance.

Actuellement, il y a peu d'arguments pour préférer une méthode algorithmique à une méthode de MRC (Iasonos and O'Quigley 2011). Les inconvénients de ce schéma MRC sont néanmoins sa plus grande sensibilité aux TDL parfois enregistrées précocement, soit du fait du hasard, soit du fait de la difficulté à distinguer une toxicité d'un effet de la progression de la maladie lors des toutes premières expérimentations chez l'homme.

Les schémas MRC nécessitent par ailleurs le concours du statisticien tout au long de la conduite de l'essai, l'actualisation des données informatisées, mais également un plus longue période de maturation méthodologique, si l'on souhaite réaliser des études de simulation pour vérifier au préalable les caractéristiques opératoires d'un ou plusieurs schémas dans un scénario donné, comme cela est le plus souvent conseillé (Thall, Lee et al. 1999; Paoletti, O'Quigley et al. 2004).

#### ***1.4.2.2 Escalade de dose avec contrôle du surdosage***

Le schéma d'escalade de dose avec contrôle du surdosage (*escalation with overdose control, EWOC*), une adaptation de la MRC Bayésienne, a été construit pour limiter le nombre de patients traités au dessus de la DMT dans les schémas initiaux de MRC. Dans le schéma

EWOC, la dose choisie pour le nouveau patient à inclure est celle pour laquelle la probabilité que cette dose excède la dose maximale tolérée est inférieure à un seuil donné (Babb, Rogatko et al. 1998),  $\delta$ . Le concept de base de cette méthode repose sur la fonction de distribution cumulative *a posteriori* de la dose maximale tolérée,  $\pi$ .

Soit  $\delta$  la probabilité maximale acceptable de traiter un nouveau patient à une dose supérieure à la dose maximale tolérée, et  $\pi_n(d_l)$  la probabilité que la dose  $d_l$  soit une dose qui excède la DMT recherchée après avoir inclus  $n$  patients. Pour l'attribution de dose du patient  $n+1$ , le schéma EWOC cherche à garantir la contrainte suivante :

$$\pi_n(x_{n+1} = d_l) = \delta$$

Après avoir inclus  $n$  patients,  $\pi_n(d_l)$  peut s'écrire :

$$\pi_n(d_l) = P(d_l > DMT | \Omega_n)$$

Autrement dit :

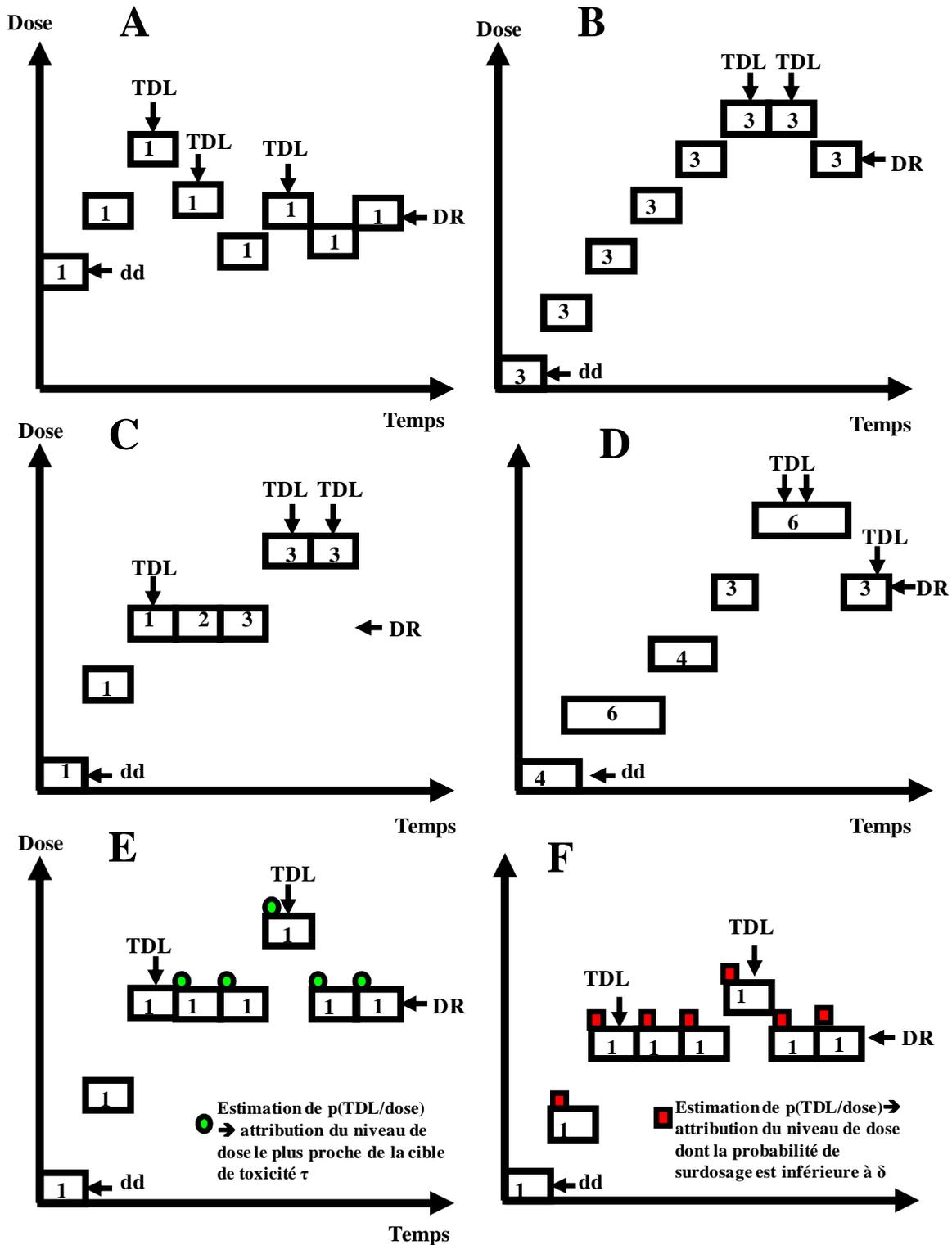
$$P(x_{n+1} > DMT | \Omega_n) = \delta$$

Un modèle logistique à deux paramètres est utilisé. Le schéma est conduit dans une approche Bayésienne. Une information *a priori* est spécifiée pour les paramètres  $\alpha_0$  et  $\beta_1$ , par l'intermédiaire d'une fonction de densité de probabilité. La dose attribuée au nouveau patient résulte de la fonction de densité de probabilité *a posteriori*, combinant les premières observations  $\Omega_n$  et l'*a priori*.

Une illustration d'un schéma de ce type est intégrée à la Figure 4 4.

### 1.4.3 Résumé des schémas

Des exemples d'essais menés selon les schémas classiques algorithmiques et de MRC que nous venons d'introduire sont présentés sur la Figure 4 4. Nous avons également présenté une représentation du schéma *up and down* qui consiste à diminuer la dose si une TDL est survenue sur le dernier patient inclus, et augmenter la dose si le patient précédent a bien toléré le traitement.



**Figure 4. Représentation graphique des principaux schémas de recherche de dose.**

Chaque rectangle représente une cohorte de patients dont l'effectif est inscrit dans le rectangle. A) schéma « up and down » ; B) Schéma « 3+3 » ; C) Schéma accéléré de Simon, sans escalade de dose intrapatient ; D) Rolling 6 ; E) méthode de réévaluation continue (CRML) ; F) « Escalation with overdose control » (EWOC). TDL = toxicité dose-limitante, dd = dose de départ, DR = dose recommandée à l'issue de l'essai,  $\tau$  = percentile cible de toxicité,  $\delta$  = percentile cible de surdosage. Schéma adapté d'après (Le Tourneau, Lee et al. 2009).

#### 1.4.4 Extensions de la méthode de réévaluation continue

Nous nous concentrons dans ce paragraphe sur les extensions de la méthode MRC, qui ont été développées afin de prendre en compte l'aspect temporel du suivi, en conservant un critère de jugement binaire : la survenue de toxicité dose-limitante pendant une fenêtre définie de temps.

##### 1.4.4.1 Cas des toxicités tardives

Dans le contexte des essais pour lesquels la survenue de toxicité est tardive, comme dans les essais de radiothérapie, attendre l'évaluation complète des patients précédemment inclus pour inclure de nouveaux patients aurait pour conséquence des essais d'une durée incompatible avec la réalisation pratique. Plusieurs méthodes ont été proposées pour s'adapter à cette spécificité.

Une modification des règles d'inclusion de la MRC a été proposée (Thall, Lee et al. 1999), permettant de poursuivre les inclusions tant que des patients sont en attente d'évaluation, selon plusieurs stratégies. Lorsque le nouveau patient éligible se présente et que des patients sont encore en attente d'évaluation dans l'essai, la première stratégie consiste à examiner le délai maximal raisonnable pour débiter son traitement. Soit le traitement ne peut être différé et le patient reçoit un traitement usuel en dehors du protocole, soit on peut attendre un peu et le patient attend pour être inclus dans l'essai dès que l'information nécessaire sera disponible pour l'attribution de sa dose. Dans la seconde stratégie, les nouveaux patients continuent à être inclus à la dose en cours d'exploration, tant que l'estimation de la dose cible n'est pas modifiée selon les données disponibles collectées ; cette approche se base sur le principe qu'il s'agit toujours alors de l'estimation la plus à jour disponible qui doit donc continuer à guider l'attribution de dose. La seconde stratégie permet d'atteindre les durées d'essais les plus courtes, sans aucun délai d'attente. Mais le risque de toxicité est important avec cette stratégie si la dose cible est proche de la dose de départ et que le rythme d'inclusion est rapide : en effet, de nombreux patients risquent d'être inclus au dessus de la dose maximale tolérée.

La méthode *Time to event CRM (TITE-CRM)* a ensuite été développée en modifiant le modèle statistique à utiliser (Cheung and Chappell 2000). Il ne s'agit pas d'une analyse de survie classique comme son nom pourrait le suggérer, mais d'une adaptation de la MRC pour

laquelle les observations sont pondérées en fonction de la durée du suivi. Dans cette méthode, une observation pour laquelle le critère de jugement n'a pas été observé peut être considérée comme censurée au moment où l'on souhaite analyser les données pour choisir la dose d'un nouveau patient. On suppose que la durée maximum de suivi est T, correspondant à la fenêtre d'évaluation des TDL. Une observation « censurée » est traitée comme une observation partielle et elle est pondérée avec un poids  $\omega$ , compris entre 0 et 1, en fonction du délai avant évènement.

Le modèle de la relation dose-toxicité  $\psi_t$ , monotone et croissant est :

$$\psi_t(x; \omega; \theta) = \omega \psi(x; \theta)$$

Il répond aux contraintes suivantes :  $\psi_t(x; 0; \theta) = 0$  et  $\psi_t(x; 1; \theta) = 1$

Dans cette méthode, la vraisemblance est modifiée :

$$L_n(\theta) = \prod_{i=1}^n [\psi_t(x_i; \omega_{i,n}; \theta)]^{y_{i,n}} [1 - \psi_t(x_i; \omega_{i,n}; \theta)]^{1-y_{i,n}}$$

où  $y_{i,n}$  et  $\omega_{i,n}$  sont respectivement la réponse et le poids assignés aux patients  $i = 1$  à  $n$ , juste avant que le patient  $n+1$  ne soit inclus.

Soit  $V_i$  le temps jusqu'à survenue d'une toxicité pour le patient  $i$ . Pour un patient suivi pour une durée  $t \leq T$ , on peut montrer (Cheung and Chappell 2000) que :

$$\text{pour } t \leq T, P(V_i \leq t) \equiv \omega(t; T)\psi(x_i; \theta),$$

où  $\psi$  est le modèle de la relation dose-toxicité pour données non censurées introduit précédemment (modèle 1).

Dans la fonction de pondération la plus simple proposée, l'observation prend un poids de 1 si une TDL est survenue ou si le suivi du patient est terminé, et un poids proportionnel à la durée de son temps d'observation par rapport à la période d'évaluation de la TDL s'il n'a pas présenté de TDL au moment de l'évaluation. C'est-à-dire :  $\omega(t; T) = t/T$ .

Une étude de simulation a comparé ce nouveau schéma TITE-CRM à la MRC classique, aussi bien par la méthode Bayésienne que par le maximum de vraisemblance (Cheung and Chappell 2000). Les modèles  $\psi$  et  $\psi_t$  de la relation dose-toxicité reposaient sur la fonction puissance.

Ces simulations montraient un gain sur la durée des essais, alors que la précision des estimations et la sécurité des patients semblaient garanties.

D'autres fonctions de pondération sont possibles notamment pour différencier les cas où les toxicités sont susceptibles d'arriver rapidement ou tardivement (Cheung and Chappell 2000), et pour prendre en compte le fait que le temps de survenue d'une toxicité raccourcit lorsque les doses augmentent (Braun 2006) ou pour modifier dynamiquement la fonction de pondération selon le délai avant toxicité. Une extension a été proposée (Braun, Levine et al. 2003), permettant de modifier la dose de chaque patient en cours de traitement, à chaque fois que la relation dose-toxicité est réestimée.

Ultérieurement, il a été montré qu'en cas de rythme d'inclusion très rapide et de toxicités survenant très tardivement, l'escalade de dose résultante peut être très rapide avec ce schéma. Des adaptations ont été proposées pour prévenir ce risque (Polley 2011).

#### *1.4.4.2 Toxicités tardives et escalade de dose avec contrôle du surdosage*

L'approche de la TITE-CRM a été combinée avec le schéma d'escalade de dose avec contrôle du surdosage, présentée précédemment (§1.4.2.2). Cette approche a été développée pour limiter le risque d'escalade de dose trop rapide. Dans le schéma combiné *time-to-event dose-escalation method with overdose control* (TITE-EWOC), l'attribution des doses se fait selon la démarche Bayésienne de contrôle du surdosage EWOC, mais les observations sont pondérées dans la vraisemblance comme dans l'approche de la TITE-CRM (Mauguen, Le Deley et al. 2011). Le modèle de la relation dose-toxicité est un modèle logistique à deux paramètres. Ce schéma permet d'accélérer considérablement la durée d'un essai mené selon la méthodologie EWOC, en utilisant toute l'information disponible, incluant celle des patients encore dans la fenêtre d'évaluation des TDL.

Par simulation, la capacité de cette méthode à identifier correctement la dose cible, la sécurité et le gain sur la durée des essais ont été évalués pour divers rythmes d'inclusion et diverses relations dose-toxicité. Cependant, la proportion de patients traités aux doses supérieures est légèrement augmentée par rapport à un schéma EWOC qui attendrait qu'un sujet soit complètement évalué pour inclure le suivant.

#### 1.4.5 Performance des méthodes de recherche de dose

Même si les schémas type MRC sont plus performants pour identifier la dose cible que les schémas algorithmiques, les performances de ces méthodes restent limitées, avec une probabilité toujours insuffisante d'identification de la dose cible. Les performances de l'ensemble des schémas sont influencées par le nombre de doses inférieures à la cible à explorer, et le nombre total de doses à explorer. En effet les performances seront d'autant meilleurs que la dose cible proche de la dose de départ, et lorsque la dose cible se situe au-delà, les données seront plus concentrées aux paliers adjacents à celui de la dose cible s'il n'y a pas ou peu de paliers de dose à explorer au dessus.

Des schémas optimaux ont cherché à évaluer comment utiliser au mieux l'information collectée pour améliorer l'identification de la dose cible. Plusieurs schémas « optimaux » ont ainsi été proposés dans la littérature. L'attribution des doses est déterminée pour optimiser l'information collectée, et guidée selon ces critères d'optimalité. Les critères d'optimalité peuvent être d'échantillonner aux doses qui minimiseront la matrice de variance-covariance des paramètres. Ces méthodes optimales dépendent donc du modèle utilisé pour estimer la relation dose-toxicité et supposent ce modèle correctement spécifié. Elles peuvent être utilisées avec des molécules dont la toxicité ne menace pas la santé du sujet et sur des populations qui n'attendent pas un bénéfice thérapeutique. En revanche, elles ne sont pas applicables aux essais en cancérologie, car l'attribution des doses selon des critères statistiques sans contrôle du risque de toxicité excessive mettrait en danger les patients.

Par exemple, une méthode optimale permet d'étudier la distribution théorique des doses recommandées dans des simulations où l'on considère que tous les patients pourraient être traités à toutes les doses (Paoletti, O'Quigley et al. 2004). La méthode part du principe qu'un sujet qui présente une toxicité à une dose aurait présenté aussi une toxicité aux doses supérieures. En revanche, le résultat d'un traitement aux doses inférieures reste inconnu. Cette situation produit une information dite incomplète. Lorsqu'on connaît la vraie relation entre la dose et la probabilité de toxicité, il est possible de déterminer la distribution d'une information complète sur des données générées par cette relation. En d'autres termes, il est possible de déterminer la distribution de la dose la plus faible à partir de laquelle un sujet souffrira d'une TDL. Cette information est ensuite utilisée pour estimer la relation dose-toxicité à partir d'un échantillon de taille donnée.

Soit  $N$  une variable aléatoire qui correspond au niveau de dose le plus bas auquel une TDL survient, et  $\eta_i$  la réalisation de cette variable pour le sujet  $i$ . Soit  $Y_{li}$  la réponse du patient  $i$  au niveau de dose  $l$  (TDL : oui / non). Conditionnellement à la valeur  $\eta_i$ , la probabilité pour le patient de présenter une TDL aux différentes doses étudiées est soit 0, soit 1. Soit  $\phi$  la fonction de distribution cumulative de  $\eta_i$ .

La probabilité que le patient présente une toxicité à la dose  $d_l$  est  $R_l$  :

$$R_l = P(Y_l = 1) = \sum_{m=1}^l P(N = N_m) = P(N \leq d_l) = \phi(d_l)$$

Pour générer l'information complète du patient  $i$ , on compare la valeur de  $\eta_i$  à chacun des niveaux de dose : si  $\eta_i \leq d_l$  le patient a une TDL à cette dose, et si  $\eta_i > d_l$  alors le patient n'a pas de TDL. A partir d'une relation dose-toxicité connue et d'une taille d'étude donnée, on peut générer les données par simulation à toutes les doses, et la méthode optimale fournira la probabilité de recommander chaque dose étudiée.

La proportion de recommandation correcte de la dose cible, donnée par cette méthode optimale, peut être considérée comme la meilleure performance possible d'un schéma d'escalade de dose pour une taille d'étude donnée, c'est-à-dire une limite au-delà de laquelle on ne pourra pas s'améliorer sans hypothèses supplémentaires. Compte tenu que ces méthodes ne sont pas applicables en pratique, et que l'on n'envisage le plus souvent pas d'augmenter la taille de l'échantillon pour des essais chez des sujets malades avec des traitements n'ayant pas encore fait preuve d'efficacité, il est nécessaire d'explorer d'autres pistes d'amélioration.

Un des arguments avancés pour expliquer les limites des performances des schémas classiques pour identifier la dose cible est la variabilité inhérente à l'observation d'un critère de jugement binaire sur des petits échantillons. Dans les deux parties suivantes, nous allons aborder les innovations méthodologiques récentes proposées pour prendre en compte tout d'abord le caractère ordinal des toxicités en gardant la même fenêtre d'observation que dans les méthodes initiales : le premier cycle de traitement. Puis nous aborderons les méthodes qui utilisent les données des toxicités des cycles de répétés de traitement.

Ces méthodes ont été envisagées pour plusieurs raisons : d'une part, une quantité importante d'information est collectée et disponible, et fait partie de l'évaluation globale de la tolérance dans ces essais. Dans une analyse rétrospective de 445 patients inclus dans des essais de thérapies moléculaires ciblées, environ un tiers des patients ont expérimenté une toxicité de grade 2 au cours du premier cycle de traitement ; par ailleurs, plus de la moitié des patients ont subi des toxicités graves au-delà du premier cycle de traitement (Postel-Vinay, Gomez-Roca et al. 2011). Comme on peut l'observer dans certains essais de thérapies ciblées, les toxicités modérées sont d'ailleurs parfois intégrées à la définition des TDL (Le Tourneau, Razak et al. 2011). D'autre part, la survenue de toxicités cumulées ou tardives est crainte, bien que cela n'ait été que rarement recherché dans les essais de phase I. La seule étude qui l'ait étudiée a montré que 5 % des essais du National Cancer Institute conduits avant 1997 présentaient une toxicité cumulée statistiquement significative (Simon, Freidlin et al. 1997).

## 1.5 Schémas prenant en compte le caractère ordinal des toxicités

Nous allons à présent examiner différentes méthodes qui ont cherché à prendre en compte l'intensité des toxicités tout en se concentrant sur la période d'évaluation classique de la TDL (le premier cycle de traitement).

Peu de schémas ont considéré la toxicité autrement que de manière binaire. Certaines approches se sont intéressées à un critère de jugement ordinal. Nous reverrons également succinctement les approches qui ont considéré une variable réponse quantitative par la construction de scores de toxicité et une dernière catégorie de méthodes qui ont incorporé les toxicités modérées pour adapter un schéma algorithmique.

### 1.5.1 Incorporation des grades de toxicité dans les schémas algorithmiques

Les schémas accélérés de Simon 2A et 3A ont été introduits ci-dessus (§1.4.1.2) ; dans ces schémas, la bascule de la phase initiale d'escalade accélérée vers un schéma d'escalade « 3+3 » se fait après la survenue de deux toxicités modérées ou une toxicité grave (Simon, Freidlin et al. 1997).

### 1.5.2 Modèle logistique à cotes proportionnelles

En 2009 puis en 2011, l'utilisation d'un **modèle à cotes proportionnelles** (*proportional odds model, POM*) a été implémentée dans des schémas MRC (Liu, Rosenberger et al. 2009; Van Meter, Garrett-Mayer et al. 2011). Ce modèle, également appelé modèle logistique cumulatif, est une extension du modèle logistique pour les variables à expliquer ordinales.

- **Spécification du modèle à cotes proportionnelles**

Soit  $Y_i$ , une variable réponse ordonnée qui prend la valeur  $k$  parmi  $K$  modalités ordonnées (1, 2, ...,  $K$ ) et qui représente le grade maximal de toxicité sur la période d'évaluation des TDL.

$$\psi^{-1}(x_i, \theta) = \log \left( \frac{P(Y_i \leq k | x_i)}{1 - P(Y_i \leq k | x_i)} \right) = \alpha_k - \beta_1 x_i, \text{ avec } \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{K-1} \quad \text{(Modèle 3)}$$

Dans le modèle, le caractère ordonné de la réponse est pris en compte par les ordonnées à l'origine  $\alpha_k$ , l'effet sur les covariable étant supposé le même quelle que soit la modalité de réponse (Carriere and Bouyer 2006). Le nombre de paramètres à estimer correspond à la somme des effets fixes et de  $K-1$  *intercept*. Ce modèle est parfois présenté dans son écriture ascendante, et seul le signe des paramètres est modifié :

$$\log\left(\frac{P(Y_i \geq k|x_i)}{1-P(Y_i \geq k|x_i)}\right) = -\log\left(\frac{P(Y_i < k|x_i)}{1-P(Y_i < k|x_i)}\right) = -\alpha_k + \beta_1 x_i$$

Ce modèle permet d'estimer la probabilité d'observer les différentes modalités de réponse. Pour une variable  $Y_i$  qui pourrait prendre 3 modalités (1 = pas de toxicité, 2 = toxicité modérée ou 3 = toxicité grave), la probabilité de toxicité grave ( $p_3$ ) et la probabilité de toxicité modérée ou grave ( $p_{2+}$ ) sont calculées :

$$\left\{ \begin{array}{l} p_3(d_i|\hat{\theta}) = P(Y_i = 3|\hat{\alpha}_2, \hat{\beta}_1) = 1 - P(Y_i \leq 2|\hat{\alpha}_2, \hat{\beta}_1) = 1 - [e^{\hat{\alpha}_2 - \hat{\beta}_1 d_i} / (1 + e^{\hat{\alpha}_2 - \hat{\beta}_1 d_i})] \\ p_{2+}(d_i|\hat{\theta}) = P(Y_i \geq 2|\hat{\alpha}_1, \hat{\beta}_1) = 1 - P(Y_i \leq 1|\hat{\alpha}_1, \hat{\beta}_1) = 1 - [e^{\hat{\alpha}_1 - \hat{\beta}_1 d_i} / (1 + e^{\hat{\alpha}_1 - \hat{\beta}_1 d_i})] \end{array} \right. \quad \text{(Équation B)}$$

Pour chaque nouveau patient, on sélectionnera la dose  $d_\ell$  dont la probabilité de toxicité grave  $p_3(d_\ell)$  est la plus proche de  $\tau$ .

- **Hypothèse de proportionnalité des cotes**

Ce modèle repose sur l'hypothèse des cotes proportionnelles, qui impose que le ou les coefficients  $\beta$  soient indépendants du niveau de réponse  $k$ . Cela veut dire, en pratique, que l'effet de la dose sur la survenue d'une toxicité grave, par rapport à une toxicité nulle ou modérée, est estimé par le même coefficient  $\beta$  que l'effet de la dose sur la survenue d'une toxicité modérée ou grave par rapport à l'absence de toxicité. L'effet d'une covariable est supposé le même pour le passage d'un niveau de réponse au niveau suivant, quels que soient ces niveaux.

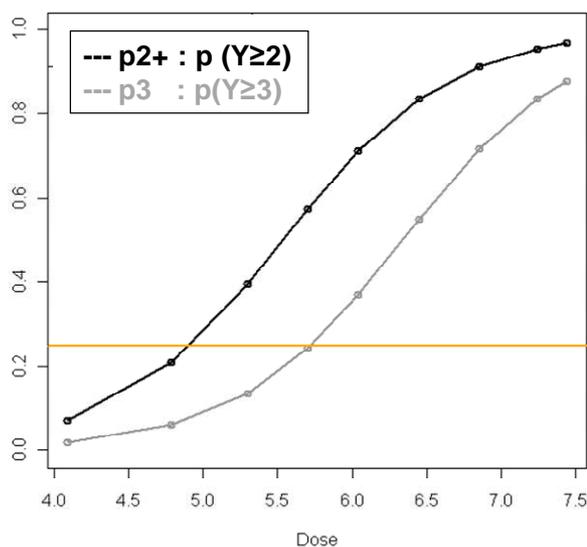
Cette hypothèse peut être jugée trop stricte. Afin de tester formellement cette hypothèse, un modèle logistique ordonné généralisé peut être utilisé (modèle 4) (Carriere and Bouyer 2006).

$$\log\left(\frac{P(Y_i \leq k|x_i)}{1 - P(Y_i \leq k|x_i)}\right) = \alpha_k - \beta_k x_i \quad (\text{Modèle 4})$$

Le modèle 3 est emboîté dans le modèle 4, ce qui permet de tester l'hypothèse de proportionnalité des cotes par le test du rapport de vraisemblance. Si l'on décompose le coefficient  $\beta_k$  en  $\beta_1 + \beta'_k$  (avec  $\beta'_k = 0$  pour  $k = 1$ ), cela revient à tester l'hypothèse  $H_0 : \beta'_k = 0$  pour  $k = 2 \dots K$ , avec  $K-2$  degrés de liberté.

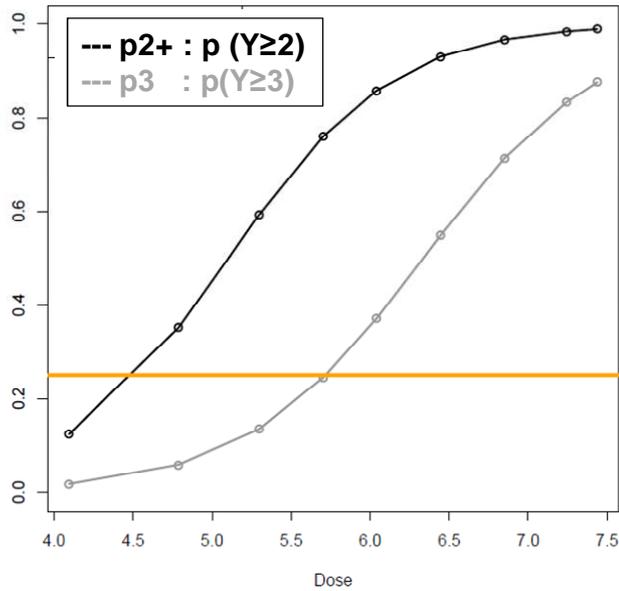
Un exemple de relation entre la dose et la probabilité de toxicité grave, et modérée ou grave, selon un modèle à cotes proportionnelles, est représenté sur la Figure 5. La Figure 6 6 représente une relation dose-toxicité qui ne respecterait pas l'hypothèse des cotes proportionnelles.

#### Probabilité de toxicité



**Figure 5. Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle à cotes proportionnelles, de paramètre  $\theta = (\alpha_1 = 9.85, \alpha_2 = 11.29, \beta_1 = 1.78)$ . La ligne orange correspond à une cible de probabilité de toxicité  $\tau = 25\%$ .**

## Probabilité de toxicité



**Figure 6.** Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle logistique ordonné généralisé, de paramètre  $\theta = (\alpha_1 = 9.85, \alpha_2 = 11.29, \beta_{1.1} = 1.63, \beta_{1.2} = 1.78)$ . La ligne orange correspond à une cible de probabilité de toxicité  $\tau = 25\%$ .

- **Vraisemblance du modèle à cotes proportionnelles**

A partir des données  $\Omega_n$  observées sur n sujets,  $\Omega_n = \{(x_1, y_1) \dots (x_n, y_n)\}$ , la vraisemblance peut s'écrire :

$$L(\theta | \Omega_n) = \prod_{i=1}^n p_1^{I^{[y_i=1]}} \times p_2^{I^{[y_i=2]}} \times p_3^{I^{[y_i=3]}} \quad (\text{Équation C})$$

où  $I^{[y_i=k]}$  est une fonction qui prend la valeur 1 si  $Y_i=k$ , et la valeur 0 dans les autres cas.

- **Evaluation du modèle logistique à cotes proportionnelles dans le contexte des essais de phase I**

Ce type d'approche a été évalué par simulations (Van Meter, Garrett-Mayer et al. 2011), dans le nouveau schéma adaptatif *proportional odds*-CRM (PO-CRM). Dans cette étude, la relation entre la dose et le risque de toxicité était modélisée par un modèle logistique à cotes proportionnelles. La toxicité sur la période d'évaluation des TDL était une variable en 5

modalités ( $k = 0..4$ ), correspondant aux grades équivalents de la classification des événements indésirables NCI-CTC (Trotti, Colevas et al. 2003).

Les paramètres étaient estimés par le maximum de vraisemblance, une fois l'hétérogénéité obtenue ou par estimation d'une distribution *a posteriori* dans un cadre Bayésien.

Pour l'approche Bayésienne développée par les auteurs, il est nécessaire de préciser la probabilité de toxicité des différents grades à deux doses, par exemple : au niveau de dose 600 mg, 10 % de TDL, correspondant à la répartition des grade de 60 % d'absence de toxicité, 20 % de grade 1, 10 % de grade 2, 6 % de grade 3 et 4 % de grade 4 ; et au niveau de dose 3000 mg : 2 % d'absence de toxicité, 3 % de grade 1, 5 % de grade 2, 45 % de grade 3 et 45 % de grade 4. Pour construire une distribution *a priori* de la DMT dans cette approche bayésienne, il faut également préciser la dose qui a un taux de TDL de 50 %.

Dans ce schéma la dose reçue était une variable continue, et non discrète comme dans les approches les plus fréquemment utilisées en pratique clinique. La dose  $\hat{d}$  à allouer à un nouveau patient à inclure était donc la dose dont la probabilité de toxicité est égale à la cible de toxicité  $\tau$ . Elle était calculée à partir des paramètres estimés selon la formule :

$$\hat{d} = \frac{\hat{\alpha}_3 - \log\left(\frac{1-\tau}{\tau}\right)}{\hat{\beta}_1}.$$

Le gain apporté par un schéma adaptatif basé sur ce modèle était modeste par rapport à un schéma reposant sur l'utilisation d'un modèle logistique classique (modèle logistique à 2 paramètres pour variable à expliquer binaire). La principale limite de cette méthode est de disposer de données pour définir les hypothèses à préciser pour l'initiation de l'essai.

Liu et al. se sont également intéressés à ce modèle dans un schéma adaptatif de recherche de dose, et le problème d'identification du modèle à cotes proportionnelles par la méthode du maximum de vraisemblance sur les petits échantillons du début d'essai a également été étudié (Liu, Rosenberger et al. 2009) : en considérant une variable réponse en 5 modalités, il était nécessaire d'avoir en moyenne au moins 12 sujets inclus pour pouvoir procéder à la première estimation par la méthode du maximum de vraisemblance. Le schéma de Liu et al. n'était en revanche pas adapté pour une mise en œuvre réelle car l'attribution des doses se base sur l'optimisation de l'information, et ce qui rend son implémentation improbable en pratique clinique du fait du risque de toxicité.

### 1.5.3 Modèle à cote proportionnelles et contrôle du surdosage

Le modèle ordinal a également été implémenté dans l'approche EWOC : POM-EWOC (Tighiouart, Cook-Weins et al. 2012 ). Les auteurs combinent l'approche par modèle à cotes proportionnelles PO-CRM (Van Meter, Garrett-Mayer et al. 2011) avec la règle de sélection des doses utilisée dans l'escalade avec contrôle du surdosage EWOC (Babb, Rogatko et al. 1998).

Pour une variable ordinaire discrète en 3 classes (toxicité grave, modérée ou négligeable, prenant les valeurs 3, 2 ou 1) qui représente la toxicité maximale expérimentée sur la fenêtre d'observation des TDL, la relation dose toxicité est supposée suivre la relation logistique à cotes proportionnelles du modèle 4 :  $\text{logit}[P(Y_i \leq k)] = \alpha_k - \beta_1 x_i$ , en supposant que les *intercept* sont ordonnés, et que le paramètre de dose est positif. La DMT  $\rho$  est définie comme la dose dont le risque de toxicité grave attendue est une probabilité  $\tau$ . La vraisemblance est la même que celle utilisée dans Van Meter et al. (Equation C).

Comme dans l'approche EWOC, le modèle est reparamétré en fonction des probabilités  $p_3$  et  $p_{2+}$  (telles qu'introduites précédemment dans l'équation B (§1.5.2)) : soit  $p_3$  la probabilité qu'un patient traité à la dose  $d_\ell$  présente une toxicité grave et  $p_{2+}$  la probabilité qu'un patient traité à la dose  $d_\ell$  présente une toxicité modérée ou grave.

Pour  $\rho$ , la dose maximale tolérée réelle que l'on cherche à identifier, on peut écrire :

$$\left\{ \begin{array}{l} \alpha_1 = \Psi^{-1}(p_{2+}) \\ \alpha_2 = \Psi^{-1}(p_3) \\ \beta_1 = [\Psi^{-1}(\tau) - \Psi^{-1}(p_3)] \times 1/\rho \end{array} \right.$$

Définissons à présent  $\Psi_1$  et  $\Psi_2$  tels que :

$$\Psi_1(p_3, p_{2+}, \rho ; x_i) = \Psi [\Psi^{-1}(p_{2+}) + (\Psi^{-1}(\tau) - \Psi^{-1}(p_3)) \times x_i/\rho]$$

$$\Psi_2(p_3, p_{2+}, \rho ; x_i) = \Psi [\Psi^{-1}(p_3) + (\Psi^{-1}(\tau) - \Psi^{-1}(p_3)) \times x_i/\rho]$$

La vraisemblance du modèle reparamétré après recueil des données sur  $n$  patients,  $\Omega_n$ , est :

$$\begin{aligned} L(p_3, p_{2+}, \rho \mid \Omega_n) &= \prod_{i=1}^n [1 - \Psi_1(p_3, p_{2+}, \rho ; x_i)]^{I^{[y_i=1]}} \\ &\times [\Psi_1(p_3, p_{2+}, \rho ; x_i) - \Psi_2(p_3, p_{2+}, \rho ; x_i)]^{I^{[y_i=2]}} \\ &\times [\Psi_2(p_3, p_{2+}, \rho ; x_i)]^{I^{[y_i=3]}} \end{aligned}$$

Une approche Bayésienne est utilisée pour estimer les risques de toxicité à chacune des doses. L'algorithme d'attribution de la dose au prochain patient suit le principe du schéma EWOC : le premier patient reçoit la dose  $x_1$  supposée tolérable par l'investigateur.

Soit  $\pi_n(\rho) = \pi(\rho \mid \Omega_n)$ , la distribution cumulative *a posteriori* de la DMT. Le patient  $n+1$  reçoit la dose  $x_{n+1} = \pi_n^{-1}(\delta)$ , de manière à ce que la probabilité que la dose attribuée excède la DMT soit égale à  $\delta$ , comme nous l'avons introduit au §1.4.2.2.

Les auteurs montrent que les performances du POM-EWOC ne sont pas meilleures comparées à EWOC, voire légèrement dégradées en termes de probabilité d'identifier la bonne dose. La proportion de patients exposés à des doses élevées, et le nombre de patients présentant une TDL est un peu moindre. Une certaine robustesse à l'écart à l'hypothèse POM est décrite, lorsque les données étaient générées selon un modèle ordonné généralisé de la relation dose-toxicité. Comme avec l'approche POM-CRM, cette méthode ne montre pas de bénéfice majeur, au prix d'une hypothèse supplémentaire sur le rapport des cotes.

#### 1.5.4 Modèle puissance

Un autre schéma a été proposé dans lequel les toxicités graves sont distinguées des toxicités modérées, en supposant un modèle puissance pour la relation dose-toxicité (Iasonos, Zohar et al. 2011).

Le nouveau schéma consistait à utiliser, un modèle prenant en compte le grade de toxicité, pour l'attribution des doses par MRC. Un modèle puissance à un ou deux paramètres était utilisé, pour un critère de jugement représentant la toxicité maximale sur la fenêtre d'évaluation des TDL, en trois modalités (toxicité absente, modérée ou grave).

Soit  $Y_i$  une variable réponse ordonnée qui prend les modalités 1, 2 ou 3 pour le patient  $i$ , et  $x_i$  la dose reçue par le patient  $i$ . La probabilité de survenue des différentes modalités de réponse est :

$$\left\{ \begin{array}{l} p_1 = P(Y_i = 1) = 1 - (x_i^a)^b \\ p_2 = P(Y_i = 2) = (x_i^a)^b - x_i^a \\ p_3 = P(Y_i = 3) = x_i^a \end{array} \right.$$

Deux modèles ont été proposés pour ce schéma de MRC prenant en compte le grade (*CRMG*). Dans le schéma *CRMG* (1,1), seul un paramètre est estimé,  $b$  étant considéré connu. Dans le schéma *CRMG* (1,2),  $a$  et  $b$  sont estimés.

Ce schéma a été comparé à deux schémas de MRC en deux étapes. Pour l'initiation de l'escalade de dose, le premier schéma commençait avec une escalade en « 3+3 » et, le second schéma commençait par une escalade accélérée qui prenait en compte les toxicités modérée (approche similaire à la phase initiale des schémas accélérés de Simon). Puis, dès que la première TDL avait été observée, chacun de ces schémas se poursuivait par une MRC classique. L'étude de simulation a évalué le gain de l'approche *CRMG* comme mineure par rapport aux autres schémas MRC (Iasonos, Zohar et al. 2011).

### 1.5.5 Approches utilisant des scores de toxicité

Chaque patient peut présenter plusieurs toxicités au cours de la période d'évaluation des TDL. Pour mieux prendre en compte les multiples toxicités présentées par les patients, la construction de scores de toxicité a été proposée. Ces scores ont été en particulier proposés dans le contexte des thérapies moléculaires ciblées, qui sont administrées au long cours, et pour lesquelles de nombreuses toxicités modérées entraînent des difficultés d'adhérence au traitement (Ezzalfani, Zohar et al. 2013).

Toutes ces méthodes s'intéressent à la fenêtre habituelle d'évaluation des TDL.

En 2004, Bekele et Thall ont exploré une mesure du fardeau global de toxicité (*total toxicity burden*) pour prendre en compte les toxicités multiples (Bekele and Thall 2006). La méthode consistait à attribuer un poids à chaque grade de chaque type de toxicité, et un score par patient résultait de la somme de tous les poids expérimentés. Les poids sont déterminés avec les investigateurs, préalablement au commencement de l'essai, et reflètent l'importance clinique des différentes toxicités. La mise en place de cette méthode est complexe du fait de la difficulté à élaborer les poids, ce qui explique son utilisation rare.

En 2007, Yuan et al. (Yuan, Chappell et al. 2007) ont proposé un schéma « quasi-CRM » pour prendre en compte le grade dans la décision d'escalade ; le score prend la valeur de 1 si la toxicité est de grade 3 (équivalent d'une TDL), une valeur de 0,5 si la toxicité est de grade 2, et une valeur de 1,5 si la toxicité est de grade 4. Le score est normalisé par rapport au score maximal obtenu parmi tous les patients, ce qui transforme sa valeur en une fraction d'évènement grave, comprise entre 0 et 1. Un modèle de travail à un paramètre utilisant la vraisemblance de quasi-Bernouilli est utilisé pour identifier la dose la plus proche du score cible, par une méthode Bayésienne. Cette méthode a montré des performances équivalentes à la précédente ; en revanche, cette méthode ne prend pas en compte les toxicités multiples survenues chez un même patient.

Un autre score a été proposé (Chen, Krailo et al. 2010), utilisé avec une régression isotonique. Le score consiste à attribuer un poids égal au grade pour tout effet indésirable et à rajouter 2 points si l'effet indésirable correspond à une TDL. Lorsque les patients ont présenté plusieurs toxicités, la multiplicité des toxicités observées chez un même patient ne contribue qu'à la partie décimale du score. Le score est normalisé, compris entre 0 et 1, en divisant les scores individuels par le score maximal observé sur tous les patients. Ce score ne prend pas en compte l'importance clinique relative entre différents types de toxicité de même grade.

Une difficulté particulière de ce type de méthode est de déterminer la cible de toxicité acceptable, exprimée sous cette forme de score, et nécessite habituellement des méthodes de consensus, de même que pour toutes les méthodes qui nécessitent une pondération spécifique à chaque essai pour les effets indésirables attendus. Les poids sont à définir pour chaque nouvelle étude, et nécessitent d'avoir un *a priori* sur les toxicités attendues. Ces méthodes sont donc beaucoup plus délicates à implémenter que les méthodes basées directement sur le grade de toxicité maximale observé au cours de la période d'observation des TDL.

## 1.6 Schémas prenant en compte les toxicités au-delà du premier cycle

Nous allons à présent nous intéresser aux méthodes prenant en compte les toxicités survenant au delà de la période classique d'évaluation des TDL, pour lesquelles les toxicités survenues à chaque cycle de traitement sont prises en considération.

### 1.6.1 Notations

Nous reprenons et étendons les notations introduites précédemment. Le patient  $i$  reçoit plusieurs cycles  $j$  ( $1, \dots, J$ ). Soit  $Y_{ij}$ , variable ordinale reflétant le grade maximal de toxicité observé par le patient  $i$  au cycle  $j$ ,  $x_{ij}$  la dose reçue au cycle  $j$ , et  $t_{ij}$ , le temps, que l'on assimile au cycle du fait que la durée des cycles de traitement est fixe.

### 1.6.2 Schéma algorithmique avec escalade de dose intrapatient

Le schéma accéléré n°4A de Simon et al. se distingue des schémas 2A et 3A qui ont été introduits précédemment (§1.4.1.2), par le fait qu'il prend en compte la survenue de toxicités modérées ou graves, quel que soit le cycle de traitement, pour basculer de la phase accélérée d'escalade à une méthode d'escalade plus prudente type « 3+3 » (Simon, Freidlin et al. 1997).

Parmi les propositions de Simon et al., les différents schémas accélérés ont été proposés avec la possibilité d'une escalade de dose intrapatient si le premier cycle est correctement toléré : schémas 2B, 3B et 4B, afin de permettre qu'un maximum de patients soient traités à la DMT. L'approche 4B combine donc, dans un schéma algorithmique, l'utilisation des données de toxicité des cycles ultérieurs à la période d'évaluation des TDL, et la possibilité d'escalade de dose intrapatient. Les règles d'escalade de dose intrapatient sont équivalentes aux règles d'escalade de dose entre patients.

Ces schémas de Simon et al. qui utilisent les données au-delà de la fenêtre habituelle des TDL pour l'escalade de dose restent des schémas algorithmiques avec une recommandation de la MTD basée sur la fenêtre classique des TDL, puisqu'ils basculent tous sur un schéma « 3+3 ».

Dans cette publication, Simon et al. proposaient également un modèle statistique qui permettait d'analyser les données en fin d'essai, en utilisant les données de toxicité de chaque cycle de traitement. Ils soulignaient en effet la nécessité de disposer de modèles spécifiques pour analyser les données en cas d'escalade de dose intrapatient. Un modèle dérivé du modèle

« Kmax » a été proposé, prenant en compte la variabilité interpatient et intrapatient dans la survenue de toxicité et de toxicité cumulée (Simon, Freidlin et al. 1997; Dancey, Freidlin et al. 2006). Dans cette analyse, les données étaient résumées à chaque cycle par la sévérité de la toxicité la plus grave expérimentée au cours du cycle de traitement.

Soit  $D_{ij}$ , la dose cumulée aux cycles antérieurs. La variable aléatoire  $u_i$  (distribuée selon la loi Normale, de moyenne nulle et de variance  $\sigma_0^2$ ) représente la variabilité interpatient vis-à-vis du risque de survenue de toxicité ;  $\varepsilon_{ij}$  (distribuée selon la loi Normale, de moyenne nulle et de variance  $\sigma_\varepsilon^2$ ) correspond à la variabilité intrapatient, pour un patient donné recevant une dose donnée. Ces paramètres déterminent l'amplitude de la variable latente  $s_{ij}$ , une variable continue de toxicité, pour le patient  $i$  au cycle  $j$ , selon le modèle :

$$s_{ij} = \log(x_{ij} + \gamma D_{ij}) + u_i + \varepsilon_{ij} \quad \text{(Modèle 5)}$$

Si l'on s'intéresse à une variable réponse en trois catégories (toxicité nulle, modéré ou grave), deux paramètres supplémentaires  $S_1$  et  $S_2$  sont à estimer de telle sorte que les valeurs de  $s_{ij} < S_1$  correspondent à l'absence de toxicité, les valeurs de  $s_{ij}$  entre  $S_1$  et  $S_2$  correspondent aux toxicités modérées, et les valeurs de  $s_{ij} \geq S_2$  correspondent aux toxicités graves (Dancey, Freidlin et al. 2006).  $\gamma$  est une mesure de l'effet de l'accumulation de la dose (*carry over*) ;  $\gamma = 0$  correspond à l'absence de toxicité cumulée. En absence d'escalade de dose intrapatient, l'effet de la dose cumulée et l'effet du temps sont confondus.  $\sigma_\varepsilon^2$  représente la quantité de variabilité liée au patient (intrapatient) inexpliquée par la dose courante et les doses précédemment reçues. Une variance  $\sigma_\varepsilon^2 = 0$  signifie que la toxicité survenue chez un patient est déterminée intégralement par la dose, et par des caractéristiques qui ne changent pas d'un jour à l'autre. Selon Simon, les valeurs importantes de  $\sigma_\varepsilon^2$  peuvent être interprétées comme reflétant la variabilité biologique réelle observable chez un même individu, ou une variabilité reflétant la difficulté de jugement concernant l'imputabilité d'un effet au traitement chez des sujets très malades.  $\sigma_0^2$  représente la quantité de variabilité interpatient. Une variance  $\sigma_0^2 = 0$  signifierait que les patients inclus dans l'étude ne diffèrent pas dans leur susceptibilité à tolérer le traitement.

Ainsi, 5 paramètres sont à estimer dans ce modèle :  $\theta = (\gamma, \sigma_0^2, \sigma_\varepsilon^2, S_1, S_2)$ .

Les auteurs ont analysé les données rétrospectives de 20 essais avec ce modèle, et un risque de toxicité cumulée a été détecté dans 3 essais (15 %). Cependant, il semble que ce modèle longitudinal n'ait pas été utilisé dans la mise en œuvre séquentielle d'essais de phase I.

Bien qu'adapté à l'analyse de variables ordinales, cette formulation revient à modéliser les réponses ordinales comme une variable latente continue sous-jacente dont la distribution est normale. Cependant, cette variable latente est un outil dont l'interprétation n'est pas toujours évidente (Carriere and Bouyer 2006) ; le modèle mixte à cotes proportionnelles lui est donc potentiellement préféré.

Le modèle logistique mixte à cotes proportionnelles a été utilisé dans les essais cliniques dans d'autres domaines depuis les années 1990, par exemple pour les essais croisés (Ezzet and Whitehead 1991), dans un essai randomisé vaccinal antipneumococcique (Pedrono, Thiebaut et al. 2009), ou dans un essai sur un antalgique (Molenberghs and Verbeke 2005a). Par exemple, dans ce dernier essai, les patients étaient évalués sur une échelle de satisfaction en 5 modalités à 3, 6, 9, 12 mois (Molenberghs and Verbeke 2005b), dont l'évolution a été analysée par ce modèle. En revanche le modèle logistique mixte à cotes proportionnelles n'a pas été implémenté dans des schémas adaptatifs de recherche de dose.

### 1.6.3 Analyse longitudinale de données binaires

Une simplification du modèle pour données répétées de Simon et al. (Simon, Freidlin et al. 1997) a été proposée (Legedza and Ibrahim 2000), mais n'a jamais été implémentée en pratique. Dans cette simplification, la variable à expliquer est binaire, un seul paramètre fixe  $\zeta$  est estimé concernant l'effet d'une variable combinant la dose courante, additionnée d'une fraction de la dose cumulée passée ; enfin un *intercept* aléatoire peut être intégré au modèle.

$$\text{Log} \left( \frac{P(Y_{ij}=1)}{1-P(Y_{ij}=1)} \right) = u_i + \zeta \log[d_{ij} + D_{ij} \exp(-\lambda)],$$

avec  $\zeta \sim G(c_0, d_0)$  ( $c_0, d_0$ ),  $\zeta > 0$ ,  $u_i \sim N(\mu_0, \sigma_0^2)$ ,  $\lambda = \text{constante}$  **(Modèle 6)**

Puisque  $\lambda$ , le taux de clairance du médicament, est supposé connu et constant, ce modèle suppose que tout effet du temps est lié à la clairance du médicament. Ce modèle est utilisé dans un contexte Bayésien ; la distribution des paramètres *a posteriori* est donnée à partir de la distribution *a priori* et d'un échantillon observé, selon le théorème de Bayes :

$$\pi(\zeta, u | \Omega_n) = f(\Omega_n | \zeta, u) \pi(\zeta, u) \propto L(\zeta, u) \pi(\zeta, u)$$

Son implémentation nécessite par ailleurs de déterminer une distribution *a priori* sur le coefficient  $\zeta$ , et sur la distribution de l'*intercept* aléatoire.

Par une étude de simulation, il a été montré que, pour conduire l'essai selon ce modèle, une simplification supplémentaire du modèle est nécessaire, et l'effet aléatoire est omis. Au final, un seul paramètre est estimé ( $\zeta$ ). Bien que ce modèle prenne en compte les données répétées de toxicité, il ne permet pas d'étudier le risque de toxicité au cours du temps, puisque le paramètre  $\lambda$  n'est pas estimé.

Cette approche est destinée à mettre chaque patient à la dose la plus haute tolérable, c'est-à-dire mettre chaque participant au plus proche de la cible  $\tau$  (titration individuelle). L'approche suppose aussi que le traitement a une durée illimitée. Cette approche diffère beaucoup de l'approche classique des essais de recherche de dose.

## 1.7 Objectif de la thèse

Au travers de cette introduction, deux problématiques temporelles ont été individualisées dans le contexte des essais de phase I de recherche de dose :

- d'une part la durée des essais pour laquelle le choix du schéma d'escalade de dose est déterminant, du fait de l'aspect séquentiel des inclusions.
- d'autre part la prise en compte des données de suivi des patients, alors que les approches classiques ne prennent en compte que les données du premier cycle de traitement, résumées par une variable binaire, pour déterminer la dose recommandée pour les essais de phase II.

Notre travail s'articule autour de ces deux problématiques.

La partie 2 présente une étude de simulation destinée à évaluer l'intérêt d'un schéma d'étude par réévaluation continue, comparé à deux schémas d'étude algorithmiques (le « 3+3 » et le *Rolling 6*), dans le but de limiter la fermeture temporaire des inclusions des essais en pédiatrie oncologique.

La partie 3 présente deux travaux évaluant l'intérêt de l'utilisation d'un modèle mixte à cotes proportionnelles pour les essais de phase I de recherche de dose, dans le but d'améliorer l'identification de la dose cible, en utilisant les données ordinales répétées. Le modèle sera présenté, puis évalué par une étude de simulation ; enfin, nous présenterons les résultats produits par ce modèle en analysant rétrospectivement des données réelles d'essais de phase I, afin d'illustrer très concrètement notre approche.

La partie 4 présentera un travail complémentaire à la partie 3, qui a cherché à étudier l'apport d'un modèle logistique longitudinal par rapport à un modèle à cotes proportionnelles longitudinal, et l'impact de la simplification du modèle en considérant la pente de la relation dose-toxicité connue.

L'ensemble de ces travaux sera discuté dans la partie 5.

## 2 Evaluation comparative de deux schémas d'étude destinés à limiter l'interruption des inclusions en oncologie pédiatrique

Nous avons mis en œuvre une étude de simulation pour confronter les performances de deux schémas d'étude construits pour limiter la suspension des inclusions dans les essais de recherche de dose, dans le contexte spécifique des essais de pédiatrie oncologique : une MRC modifiée et le *Rolling 6*.

Dans un premier temps, nous allons présenter l'essai erlotinib + radiothérapie pour les gliomes chez l'enfant, qui a été mené selon un schéma MRC modifié (Georger, Hargrave et al. 2011). Pour cet essai, les pédiatres souhaitaient que les interruptions d'inclusion soient limitées. En effet cet essai avait deux spécificités : il s'adressait à des patients nouvellement diagnostiqués, et naïfs de traitement, et le traitement médicamenteux étudié était associé avec la radiothérapie. Une adaptation de la MRC avait été proposée à cet effet. Nous avons évalué ce schéma en le comparant au schéma *Rolling 6*. Pour cela nous avons tenté de réanalyser les données de cet essai avec le schéma *Rolling 6* (Skolnik, Barrett et al. 2008), puis effectué l'évaluation comparative formelle de ces schémas par une étude de simulation, en les confrontant également au schéma « 3+3 ».

L'essai erlotinib + radiothérapie est tout d'abord présenté de manière détaillée, car il s'agit d'un « fil rouge » : en effet, il a été utilisé comme illustration des diverses approches (§2, §3, et §4) que nous développerons.

## 2.1 L'essai erlotinib + radiothérapie dans les tumeurs cérébrales de l'enfant

Le pronostic des tumeurs du système nerveux central en oncopédiatrie est sombre, notamment chez les enfants atteints de gliomes du tronc cérébral (médiane de survie de 9 mois). L'erlotinib a été étudié dans cette indication. Il s'agit d'est une thérapie moléculaire ciblée, inhibiteur sélectif de la tyrosine kinase du récepteur du facteur de croissance épidermique (EGFR) qui bloque la transduction du signal de croissance intracellulaire. Il a obtenu son autorisation de mise sur le marché en Europe en 2005. Il est administré par voie orale quotidiennement (100 à 150 mg/j), pour le traitement du cancer bronchique non à petites cellules localement avancé, métastatique ou stable chez les patients ayant une expression du récepteur de l'EGFR et le traitement du cancer du pancréas métastatique.

Un essai multicentrique a été mené entre juin 2005 et août 2007 ; il était composé de deux sous-essais : le premier groupe de patients correspondait à des enfant avec une tumeur cérébrale réfractaire ou en rechute, inclus pour recevoir de l'erlotinib seul avec un schéma d'escalade de dose « 3+3 » ; le second groupe correspondait à des enfants récemment diagnostiqués pour un gliome du tronc cérébral, inclus pour évaluer la tolérance de l'erlotinib en complément de la radiothérapie, avec un schéma d'escalade de dose MRC (Geoerger, Hargrave et al. 2011). Nous concentrerons notre propos sur le second sous-essai.

Le traitement était administré quotidiennement sur des cycles de 3 semaines, et quatre niveaux de dose devaient être explorés : 75, 100, 125 et 150 mg/m<sup>2</sup>. La dose de départ correspondait à 80 % de la dose recommandée chez l'adulte (Smith, Bernstein et al. 1998). Les patients recevaient l'erlotinib et une radiothérapie localisée sur le tronc cérébral (1,6 Gy/jour), pendant 6 semaines ; la période d'évaluation des TDL dans cet essai était de 6 semaines (2 cycles d'erlotinib). L'erlotinib était poursuivi après la radiothérapie jusqu'à progression ou toxicité nécessitant l'arrêt de traitement. Les patients étaient suivis toutes les trois semaines sous traitement, puis tous les trois mois jusqu'au décès.

Pour la mise en œuvre de la MRC, les règles suivantes ont été choisies :

- Démarrer l'attribution de dose basée sur les estimations d'un modèle logistique à pente fixe, dès que la première TDL serait survenue, avec une cible de toxicité  $\tau = 20\%$ .

- Autoriser l'inclusion de patients au même niveau de dose à concurrence de six tant que tous les patients ne sont pas évalués au palier en cours.
- Interrompre l'essai après le traitement de 8 patients à la même dose (O'Quigley and Reiner 1998).

La TDL était définie dans le protocole par la survenue, pendant les six premières semaines de traitement, d'un des évènements suivants :

- événement indésirable de grade 3 ou plus non hématologique, à l'exception des épisodes de fièvre de grade 3, des toxicités hépatiques transitoires, des nausées ou vomissements en l'absence d'un traitement préventif adéquat et des évènements indésirables liés à la progression de la maladie ;
- neutropénie de grade 4, thrombopénie de grade 4 de plus de 7 jours, thrombopénie de grade 3 ou 4 nécessitant des transfusions pendant plus de 7 jours.

Le grade des évènements indésirables a été établi selon la version 3.0 de la classification du NCI (Trotti, Colevas et al. 2003).

Vingt et un enfants ont été inclus dans l'essai évaluant l'association erlotinib + radiothérapie dans les gliomes du tronc cérébral (âgés de 2 à 16 ans, âge médian de 6 ans), et inclus dans les deux mois suivant le diagnostic. Six enfants ont été inclus à la première dose ; parmi eux, un enfant est décédé de convulsions compliquées d'une pneumopathie d'inhalation. Ce décès a été classé comme TDL, et a entraîné une longue interruption des inclusions pour réunir un comité indépendant de surveillance. Au second palier de dose, 6 enfants ont été inclus, sans TDL observée. Au troisième palier de dose, 8 enfants ont été inclus, et une TDL a été observée (folliculite et prurit de grade 3) ; un neuvième enfant avait été inclus à ce palier mais a été considéré non évaluable pour la TDL, car il n'a pas pu recevoir les deux cycles de traitement. L'essai fut ensuite interrompu, et la quatrième dose ne fut pas explorée, du fait que le premier sous-essai avait déjà atteint sa dose recommandée ( $125 \text{ mg/m}^2$ ) avec la survenue de deux TDL à la dose de  $150 \text{ mg/m}^2$  (hémorragies intratumorales) et du fait des décès déjà survenus.

La probabilité de TDL à la dose recommandée de  $125 \text{ mg/m}^2$  était de 16 % (IC 95 %, 4 % - 45 %). Sachant qu'un enfant de 28 kg a une surface corporelle de  $1 \text{ m}^2$  et qu'un adulte de 1,80 m et 80 kg a une surface corporelle de  $2 \text{ m}^2$ , la dose pédiatrique  $125 \text{ mg/m}^2$  correspond à environ 300 mg/j chez un adulte. La dose recommandée en fin d'essai était donc à la fois

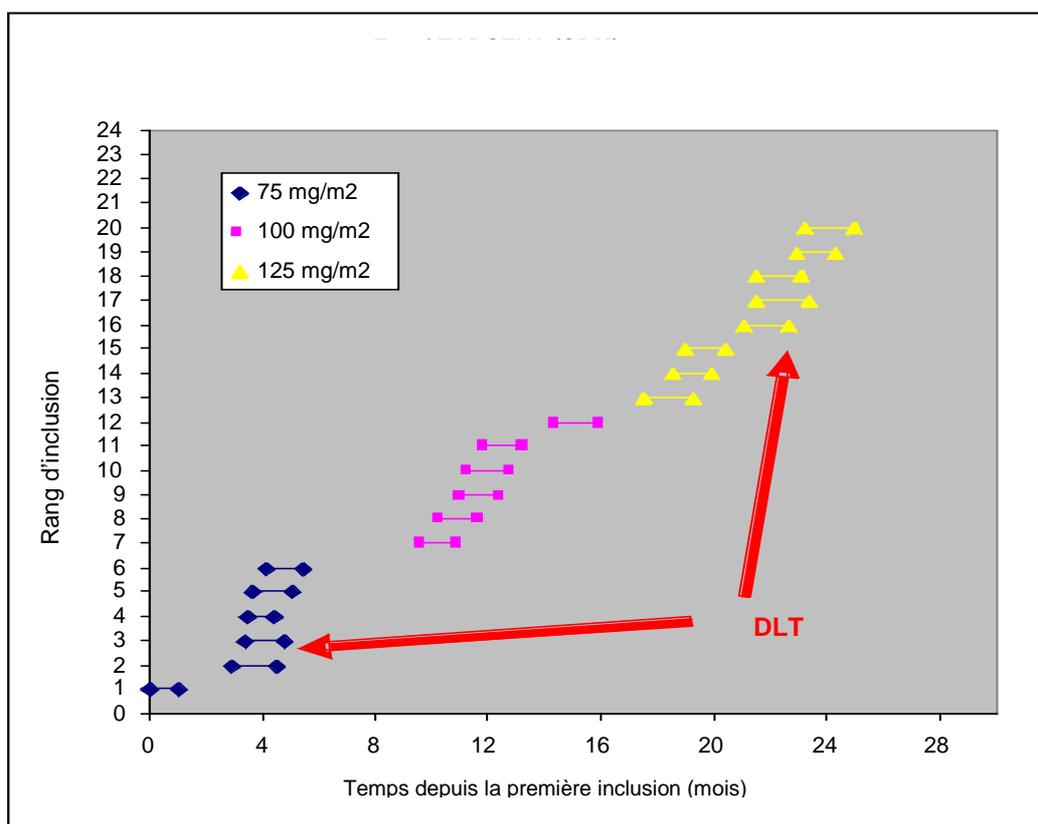
supérieure à celle habituellement utilisée chez l'adulte (100 à 150 mg/j) et à celle recommandée dans d'autres essais pédiatriques (85 mg/m<sup>2</sup>).

La durée médiane de traitement a été de 4 mois. Concernant l'efficacité, 18 enfants étaient évaluable. Parmi eux, 3 avaient une réponse partielle après 4 cycles, et 9 avaient une maladie stable. Dix enfants ont eu une stabilisation tumorale de plus de 6 mois. La survie médiane et son intervalle de confiance (IC) étaient de 12 mois (IC 95 %, 9,3 - 14,0).

## 2.2 Réanalyse algorithmique rétrospective à partir des données de l'essai

### 2.2.1 Méthodes

L'essai est présenté dans son aspect séquentiel sur la **Figure 7**. La participation de chaque enfant est représentée par sa date d'inclusion et la date d'évaluation des TDL. L'essai a duré 25 mois entre l'inclusion du premier enfant et l'évaluation du dernier enfant inclus.



**Figure 7. Dates d'inclusion et évaluation des TDL dans l'essai erlotinib + radiothérapie (Geoerger, Hargrave et al. 2011).** DLT: dose limiting toxicity. [« Appendix C » de l'article (Doussau, Asselain et al. 2012)]

En posant quelques hypothèses, nous avons tenté de réanalyser l'essai MRC erlotinib + radiothérapie, avec le schéma « 3+3 » et avec le schéma *Rolling 6*. C'est-à-dire que nous avons utilisé les données de patients réels inclus aux divers paliers de dose, et nous avons appliqué séquentiellement les règles d'escalade de dose algorithmiques du « 3+3 » et du *Rolling 6*. Lorsque nous ne disposions pas de données d'un patient réel pour cette réanalyse, nous avons utilisé des hypothèses présentées en Annexe 2.D.

## 2.2.2 Résultats

La réanalyse comparative rétrospective de l'essai selon les schémas d'étude « 3+3 » et *Rolling 6* tendrait à montrer que cet essai mené selon un schéma *Rolling 6* aurait probablement amené à la même conclusion et à un essai de durée similaire (Tableau 2). L'essai aurait alors identifié la même dose après inclusion de 19 enfants au lieu de 20, mais l'essai aurait duré 40 jours de plus que l'essai MRC et 3 enfants auraient été refusés à l'inclusion.

Pour le schéma « 3+3 », les conclusion dépendent de la présence ou non d'une TDL pour le 21<sup>ème</sup> enfant qui aurait été inclus à la dose de 125 mg/m<sup>2</sup>. En l'absence de TDL, la dose recommandée aurait été la même que dans l'essai réel (inclusion de 17 enfants, durée de 30 mois, 7 enfants refusés à l'inclusion). En cas de TDL pour cet enfant, il aurait été recommandé de diminuer la dose, et la dose recommandée aurait été la dose 100mg/m<sup>2</sup>. Vingt enfants auraient été inclus, huit enfants auraient été refusés à l'inclusion, et l'essai aurait duré 31 mois.

Tout comme avec la MRC modifiée, le *Rolling 6* aurait donc permis d'éviter la non-inclusion d'enfants liée au schéma d'essai, par rapport à un schéma « 3+3 ». Sous l'hypothèse que la vraie dose cible est bien la dose 125 mg/m<sup>2</sup>, on peut considérer dans cet exemple que la MRC était légèrement meilleure en terme de durée d'étude que le *Rolling 6*, avec un schéma d'étude n'imposant pas l'arrêt des inclusions.

**Tableau 2 . Résumé de l'essai erlotinib + radiothérapie [données réelles (Georger, Hargrave et al. 2011)] et des schémas adaptatifs rétrospectifs selon le *Rolling 6* et le « 3+3 ». [« Appendix B » de l'article (Doussau, Asselain et al. 2012)].**

	Erlotinib+RT CRM trial	Rolling 6 design	3+3 design	
			If no DLT in patient n° 21	If DLT in patient n° 21 *
Study start date	30/06/2005	30/06/2005	30/06/2005	30/06/2005
Study end date	30/07/2007	08/09/2007	30/10/2007	11/02/2008
Study duration (months)	25	26	30	31
Total number of patients enrolled	20	19	17	20
Number of patients skipped	**	3	7	8
Total	20	22	24	28
Recommended dose (mg/m <sup>2</sup> )	125	125	125	100

RT: radiotherapy, DLT: dose limiting toxicity. \*supposing that  $\leq 1$  DLT in pts at 100mg/m<sup>2</sup> dose: if  $\geq 2$  DLT occurred at dose 100mg/m<sup>2</sup>, dose would have been de-escalated to 75 mg/m<sup>2</sup> and the study would have been prolonged. \*\*the number of patients skipped was not recorded during the trial.

### 2.2.3 Discussion

Cet essai MRC illustre les avantages et inconvénients d'un schéma permettant une inclusion plus régulière. En effet, dans cet exemple où le rythme d'inclusion a été plus rapide qu'attendu, et du fait de la survenue d'une toxicité au premier palier de dose, le nombre de patients inclus au premier palier a été relativement important. Par ailleurs la première TDL était probablement liée à l'évolution de la maladie, plus qu'au traitement à l'essai, mais les investigateurs ont souhaité la considérer comme TDL du fait de sa particulière gravité. Le schéma d'essai permettant les inclusions régulières, et cet événement considéré comme TDL ont entraîné l'inclusion d'un nombre important d'enfants aux deux premiers paliers de dose.

La réanalyse séquentielle de cet essai de phase I a montré que les deux méthodes adaptées au contexte des essais pédiatriques (MRC et *Rolling 6*) avaient des comportements similaires et que chacune surpassait le « 3+3 ». La réanalyse séquentielle rétrospective est néanmoins délicate, parce que les différents schémas n'utilisent pas les mêmes observations. Au delà de l'analyse/réanalyse d'un essai unique, les simulations sont l'outil de choix pour évaluer comparativement les performances de différents schémas d'étude, et décrire précisément les caractéristiques opératoires de chaque schéma.

Afin de comparer formellement ces trois schémas en s'affranchissant des hypothèses nécessaires pour imaginer l'issue de l'essai, et de prendre en compte les fluctuations liées à la variabilité binomiale du critère de jugement, ainsi que la variabilité liée à l'arrivée aléatoire des enfants éligibles dans la pratique clinique réelle, nous avons donc mis en place une étude de simulation.

## 2.3 Résumé de l'étude de simulation

### 2.3.1 Introduction

Plusieurs adaptations des schémas ont été spécifiquement proposées dans le contexte des essais d'oncologie pédiatrique, notamment pour éviter d'interrompre les inclusions dans ces essais séquentiels. Le schéma de roulement par six « *Rolling 6* » a été proposé comme une adaptation du schéma « 3+3 » pour limiter la suspension des inclusions et accélérer l'identification de la dose cible (Skolnik, Barrett et al. 2008). Une adaptation de la méthode de réévaluation continue avait également été mise en œuvre dans l'essai d'oncologie pédiatrique de l'association erlotinib et radiothérapie dans les tumeurs cérébrales (Georger, Hargrave et al. 2011).

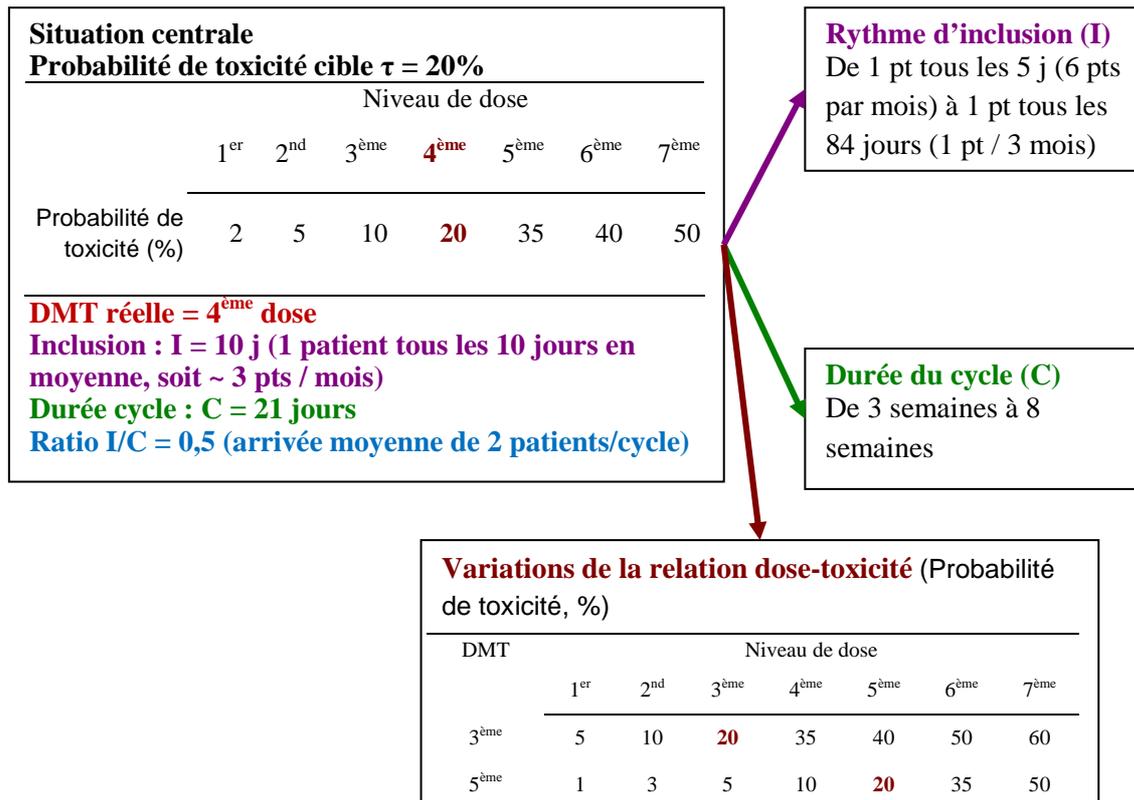
### 2.3.2 Objectif

Nous avons cherché à comparer quatre schémas (« 3+3 », *Rolling 6* et deux adaptations de la MRC qui diffèrent par leur règle d'arrêt), concernant leurs caractéristiques opératoires classiques (capacité à identifier la dose cible, risque global de toxicité), mais également leurs caractéristiques opérationnelles, c'est-à-dire le nombre d'enfants qui ne peuvent être inclus du fait de la suspension des inclusions, et la durée des essais.

### 2.3.3 Méthodes

Une étude de simulations a permis de reproduire des répétitions de 1 000 essais cliniques pour chaque schéma d'étude comparé, dans divers scénarios cliniques ou temporels. Nous avons comparé les schémas dans un scénario classique, dans lequel la dose à identifier ( $\tau = 20\%$ ) était la 4<sup>ème</sup> dose parmi 7 doses à étudier. La période d'évaluation des TDL correspondait au premier cycle, d'une durée de 21 jours. La durée moyenne entre l'inclusion de deux patients était de 10 jours. Les délais écoulés entre l'arrivée de deux patients consécutifs et les délais de survenue de TDL étaient générés par une loi exponentielle. Dans ce scénario qui a été défini comme scénario central, le rythme d'inclusion des patients était assez rapide, c'est-à-dire qu'en moyenne deux nouveaux patients se présentaient pendant la durée du premier cycle d'un patient.

Le plan d'expérience comportait des variations pour étudier l'impact des composantes du scénario sur les performances des schémas évalués (Figure 8). Dans tous les scénarios, l'essai débutait au premier palier de dose. Les variations ont porté sur le rythme d'inclusion, la durée du cycle (ou le ratio du rythme d'inclusion sur la durée du cycle), la probabilité de toxicité de chaque dose (dose cible au troisième ou au cinquième palier de dose).



**Figure 8. Scénarios étudiés dans l'étude de simulation comparant le 3+3, le *Rolling 6* et la MRC modifiée pour inclusion dans les essais pédiatriques : Situation centrale, et variations du rythme d'inclusion, de la durée du cycle, et de la relation dose-toxicité étudiées (Doussau, Asselain et al. 2012).** Ratio I/C = Délai moyen d'inclusion (I) / durée du cycle (C), DMT = dose maximale tolérée (dose ciblée),  $\tau$  = probabilité de toxicité ciblée, pt = patient.

Les deux schémas MRC se différençaient par leur règle d'arrêt : l'un avec un arrêt de l'essai après un nombre fixe de 20 enfants inclus, l'autre une fois que 10 enfants avaient été traités à une même dose, de manière consécutive ou non. Pour les deux schémas MRC, une escalade algorithmique permettait de débiter les inclusions, puis un modèle logistique à un paramètre était utilisé pour l'attribution des doses (pente  $\beta_1$  fixée, calculée selon la méthode présentée en Annexe 3).

Un schéma optimal (tel que présenté au §1.4.5) permettait également de confronter les performances de la distribution des doses recommandées que l'on obtiendrait si chaque patient inclus pouvait être traité à toutes les doses (Paoletti, O'Quigley et al. 2004). La taille d'échantillon pour le schéma optimal correspondait à la moyenne du nombre médian de patients inclus dans les quatre autres schémas d'étude.

### 2.3.4 Principaux résultats

Dans le scénario central où la dose cible réelle à identifier se trouvait au 4<sup>ème</sup> palier de dose sur 7 à explorer, les deux schémas algorithmiques permettaient de recommander la dose cible dans 38 % des simulations, contre 43 à 45 % pour les deux MRC (Tableau 3). Les méthodes MRC entraînaient néanmoins plus souvent la recommandation de la dose supérieure à la cible, dose dont la probabilité réelle de toxicité était de 35 % (environ 20 % de recommandation contre moins de 15 % pour les méthodes algorithmiques). Il n'y avait pas de majoration du risque de recommander les paliers de dose encore supérieurs (paliers de dose avec un risque de toxicité de 40 % ou plus).

Les schémas MRC et *Rolling 6* permettaient de diminuer le nombre de patients refusés aux inclusions par rapport au « 3+3 » (4 refus en médiane contre 11 refus, Tableau 4), sans impact majeur sur la durée (médiane de 9 à 11 mois pour tous les schémas). La distribution de doses attribuées au sein des essais différait néanmoins, le *Rolling 6* entraînant l'inclusion d'un nombre supérieur de patients aux doses basses, alors que la MRC tendait à inclure plus de patients à la dose cible réelle (Fig.2 de l'article). Cette différence résultait en un taux de toxicité global plus important avec le schéma MRC qu'avec le *Rolling 6* (16 % vs. 13 %), mais le taux de toxicité global restait inférieur à la cible (Tableau 4).

Les variations des scénarios montraient que, si le rythme d'inclusion s'avère finalement lent par rapport à la durée du cycle, tous les scénarios tendent à présenter les mêmes propriétés opératoires (nombre de patients inclus, nombre de patients refusés aux inclusions) que le « 3+3 », puisqu'il est beaucoup moins probable que le problème de refus des inclusions se pose. En revanche, la capacité à identifier la dose correcte est peu impactée par les paramètres temporels (rythme d'inclusion, durée du cycle, ratio).

Lorsque l'on fait varier la relation dose-toxicité en gardant les autres paramètres identiques au scénario central, les conclusions sont similaires concernant la capacité à identifier la dose

cible, entre les schémas algorithmiques et la MRC avec règle d'arrêt après 10 patients traités à la même dose, c'est-à-dire un gain de 5 % de recommandations de la dose cible avec la MRC par rapport aux deux méthodes algorithmiques (Tableau 3), avec une probabilité légèrement plus importante de recommander la dose supérieure à la dose cible, dont la probabilité de toxicité est de 35 %. La MRC qui entraîne l'arrêt de l'essai après une taille fixe de 20 patients est notablement limitée lorsque la dose cible se situe au 5<sup>ème</sup> palier de dose, car le schéma ne permet pas de correctement explorer les paliers supérieurs, par rapport aux autres schémas dont la taille d'étude augmente proportionnellement lorsque la dose à identifier se trouve à un palier plus élevé.

Selon le schéma optimal, on ne pourrait pas dépasser 53 % à 60 % de recommandations correctes de la dose-cible, même si tous les patients inclus pouvaient être traités et évalués à toutes les doses.

**Tableau 3.** [Tableau 3 de l'article (Doussau, Asselain et al. 2012)] *Distribution of the dose recommendations in proportion and median number of patients treated per dose level, over 1,000 trials simulated with 4 phase I dose escalation designs, according to true RP2D occurring at the 3<sup>rd</sup> (upper table) to 5<sup>th</sup> doses (lower table). Results from an optimal design are also presented. Mean interpatient arrival time = 10 days, cycle length = 21 days.*

Real RP2D=3		Dose recommendation (%)									Number of patients treated per dose (median)						
Tox. prob. of dose →	*	0.05	0.10	<b>0.20</b>	0.35	0.40	0.50	0.70	**	Total	0.05	0.10	<b>0.20</b>	0.35	0.40	0.50	0.70
3+3	2.8	8.7	30.0	<b>37.7</b>	14.9	5.0	0.9	0.0	0.0	100.0	3	3	<b>6</b>	3	0	0	0
R6	3.9	10.1	30.1	<b>39.1</b>	12.9	3.5	0.4	0.0	0.0	100.0	6	6	<b>6</b>	3	0	0	0
CRM-fss	0.4	3.3	23.8	<b>42.5</b>	23.8	4.8	1.3	0.0	0.1	100.0	3	4	<b>5</b>	2	0	0	0
CRM-dnd	0.4	8.3	21.3	<b>44.9</b>	20.8	4.0	0.2	0.0	0.1	100.0	3	3	<b>7</b>	2	0	0	0
optimal	-	1.8	19.4	<b>53.1</b>	19.4	5.0	1.3	0.0	-	100.0	18	18	<b>18</b>	18	18	18	18
Real RP2D=4		Dose recommendation (%)									Number of patients treated per dose (median)						
Tox. prob. of dose →	*	0.02	0.05	0.10	<b>0.20</b>	0.35	0.40	0.50	**	Total	0.02	0.05	0.10	<b>0.2</b>	0.35	0.40	0.50
3+3	0.3	2.2	10.0	28.5	<b>38.2</b>	14.3	5.1	0.0	1.4	100.0	3	3	3	<b>6</b>	3	0	0
R6	0.4	3.2	11.8	27.7	<b>38.4</b>	14.7	3.4	0.0	0.4	100.0	6	6	6	<b>6</b>	3	0	0
CRM-fss	0.0	0.2	3.2	25.0	<b>43.4</b>	19.9	7.0	0.9	0.4	100.0	3	3	3	<b>5</b>	2	0	0
CRM-dnd	0.0	0.7	4.5	23.7	<b>45.3</b>	22.1	3.2	0.1	0.4	100.0	3	3	3.5	<b>8</b>	2	0	0
optimal	-	0.0	1.3	19.0	<b>58.8</b>	17.3	2.9	0.6	-	100.0	23	23	23	<b>23</b>	23	23	23
Real RP2D=5		Dose recommendation (%)									Number of patients treated per dose (median)						
Tox. prob. of dose →	*	0.01	0.02	0.05	0.10	<b>0.20</b>	0.35	0.50	**	Total	0.01	0.02	0.05	0.10	<b>0.20</b>	0.35	0.50
3+3	0.0	0.4	2.7	9.2	27.1	<b>39.2</b>	17.9	0.0	3.5	100.0	3	3	3	3	<b>6</b>	3	0
R6	0.2	0.6	3.2	10.7	27.1	<b>41.5</b>	14.2	0.0	2.5	100.0	6	6	6	6	<b>6</b>	4	0
CRM-fss	0.0	0.0	0.5	4.2	25.7	<b>35.5</b>	28.2	4.3	1.6	100.0	3	3	3	3	<b>3</b>	1	0
CRM-dnd	0.0	0.1	0.5	4.2	23.3	<b>45.6</b>	23.4	1.3	1.6	100.0	3	3	3	3.5	<b>7</b>	2	0
optimal	-	0.0	0.1	1.4	21.2	<b>59.3</b>	17.6	0.5	-	100.0	24	24	24	24	<b>24</b>	24	24

RP2D: recommended phase 2 dose. tox. prob.: probability of toxicity. R6: Rolling 6. CRM-fss: CRM with fixed sample size stopping rule. CRM-dnd: CRM with defined number of patients at a dose stopping rule. \* Interruption because of 2 DLT at the first level. \*\* Interruption because the last level was reached and the design would require exploration of an upper level (3+3 and R6) or because no toxicity occurred up to the last level during initiation of CRM.

**Tableau 4.** [Tableau 4 de l'article (Doussau, Asselain et al. 2012)] *Median number of patients, duration and toxicity, over 1.000 trials simulated with 4 phase I dose escalation designs, according to true RP2D occurring at 3<sup>rd</sup> to 5<sup>th</sup> dose. Mean interpatient arrival time = 10 days, cycle length = 21 days.*

RP2D (dose)	N pts treated			N pts skipped			Duration (months)			Toxicity (%)			N DLT		
	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>
3+3	17	20	23	8	11	13	8.9	10.6	12.3	19.2	16.7	13.8	3	3	3
R6	20	25	30	3	4	5	7.8	10.0	12.0	16.7	13.3	11.1	3	3	3
CRM-fss	20	20	20	3	4	6	8.3	8.7	9.1	15.0	15.0	10.0	3	3	2
CRM-dnd	21	23	25	3	4	6	8.7	9.8	10.8	18.2	15.8	13.6	4	4	3

N pts: number of patients. N DLT: number of dose limiting toxicity. RP2D: recommended phase 2 dose. R6: Rolling 6. CRM-fss: CRM with fixed sample size stopping rule. CRM-dnd: CRM with defined number of patients at a dose stopping rule.

## 2.4 Article

Les annexes A, B, C et E de l'article (Doussau, Asselain et al. 2012) ont été présentées dans la partie §2.2. Les annexes D et F à J de l'article sont détaillées en Annexe 2.

## P1

Doussau A, Asselain B, Le Deley MC, Georger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework*. **Contemp Clin Trials**. 2012 Jul;33(4):657-65.

P2

Doussau A, Asselain B, Le Deley MC, Geoerger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework*. **Contemp Clin Trials**. 2012 Jul;33(4):657-65.

### P3

Doussau A, Asselain B, Le Deley MC, Geoerger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework*. **Contemp Clin Trials**. 2012 Jul;33(4):657-65.

P4

Doussau A, Asselain B, Le Deley MC, Geoerger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework*. **Contemp Clin Trials**. 2012 Jul;33(4):657-65.

P5

Doussau A, Asselain B, Le Deley MC, Geoerger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework*. **Contemp Clin Trials**. 2012 Jul;33(4):657-65.

P6

Doussau A, Asselain B, Le Deley MC, Geoerger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework*. **Contemp Clin Trials**. 2012 Jul;33(4):657-65.

P7

Doussau A, Asselain B, Le Deley MC, Geoerger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework*. **Contemp Clin Trials**. 2012 Jul;33(4):657-65.

P8

Doussau A, Asselain B, Le Deley MC, Geoerger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework.* **Contemp Clin Trials.** 2012 Jul;33(4):657-65.

P9

Doussau A, Asselain B, Le Deley MC, Geoerger B, Doz F, Vassal G, Paoletti X. *Dose-finding designs in pediatric phase I clinical trials: Comparison by simulations in a realistic timeline framework*. **Contemp Clin Trials**. 2012 Jul;33(4):657-65.

## 2.5 Discussion

Au final, lorsque le rythme d'inclusion attendu est rapide, il semble que le *Rolling 6*, tout comme la MRC, permet bien d'éviter la suspension des inclusions. Le *Rolling 6* entraîne l'inclusion d'un nombre important d'enfants supplémentaires aux doses les plus basses explorées ; dans le contexte des essais pédiatriques, ces premières doses sont proches de la zone d'efficacité du fait des connaissances déjà disponibles chez l'adulte, cela constitue donc un gain important par rapport au schéma « 3+3 ». Néanmoins la MRC s'avère supérieure au *Rolling 6*, avec une meilleure capacité à identifier la dose cible, et un nombre d'enfants inclus à la dose cible plus important. Malgré un gain de 5 % environ d'identification de la dose cible avec ce schéma MRC par rapport aux schémas algorithmiques, la dose cible reste identifiée dans moins de 50 % des essais.

Comme dans notre étude, Sposto et Groshen qui ont comparé le *Rolling 6* au schéma « 3+3 » ont relativisé le gain annoncé par le *Rolling 6* concernant la durée des études, ainsi que l'intérêt du schéma puisque les patients supplémentaires inclus sont inclus aux doses basses, ce qui n'améliore pas l'efficacité du schéma pour atteindre la dose cible, et augmente le coût des essais (Sposto and Groshen 2011).

Dans une autre étude de simulations, Onar-Thomas et al. trouvaient un gain en termes de durée similaire entre une méthode MRC et le *Rolling 6* comparativement au schéma « 3+3 », et soulignaient eux aussi l'intérêt de la MRC qui traite plus de patients à la dose cible que le *Rolling 6* (Onar-Thomas and Xiong 2010). Ils développaient également un schéma qui utilise la dose réelle attribuée en fonction de la surface corporelle (variations entre la dose choisie et la dose réellement administrée, dues aux dosages disponibles du médicament), problème fréquemment rencontré en pédiatrie du fait de la variabilité de la surface corporelle.

Une étude de simulations a aussi comparé le schéma *Rolling 6* à la TITE-CRM. Le modèle de la TITE-CRM était choisi pour utiliser toute l'information disponible à chaque nouvelle inclusion car la durée d'évaluation des TDL était courte (cycle de 21 jours, pour un délai moyen d'inclusion de 10 jours). Cette étude a montré la capacité de la TITE-CRM à éviter la suspension des inclusions, à améliorer l'identification de la cible, tout en traitant plus de patients à la cible ou à la dose qui lui est inférieure (Zhao, Lee et al. 2011).

Les estimations du schéma optimal (schéma introduit au §1.4.5) où chaque patient pourrait être traité à toutes les doses, que l'on peut interpréter comme l'efficacité maximale théorique que l'on pourrait atteindre avec une taille d'échantillon donnée, semblent montrer que l'on ne peut pas espérer dépasser 60 % de recommandations correctes de la dose-cible avec un critère de jugement binaire, même si tous les patients inclus étaient évalués à toutes les doses, sans hypothèse supplémentaire sur la relation dose-toxicité.

La piste d'amélioration que nous avons souhaité investiguer dans la suite de ce travail porte sur l'utilisation des données de tolérance des cycles de traitement administrés au-delà de la période d'évaluation classique des TDL, en utilisant également leur caractère ordinal.

### 3 Proposition d'un nouveau schéma d'étude basé sur le modèle mixte de régression logistique à cotes proportionnelles

#### 3.1 Introduction

Nous avons présenté les méthodes qui utilisent des données ordinales ou répétées. Peu de travaux ont été effectués concernant l'utilisation des données de toxicité répétées aux cours des cycles. Simon et al. ont utilisé ce type d'approche pour réanalyser les données de 20 essais avec une variable ordinaire mais leur approche n'a pas été utilisée par d'autres équipes. La version simplifiée de (Legedza and Ibrahim 2000) prenait en compte les données répétées, mais sous la forme d'une variable binaire.

Nous proposons dans cette seconde partie d'utiliser le **modèle mixte de régression logistique à cotes proportionnelles** (*mixed-effect proportional odds model, POMM*) pour conduire la recherche de dose mais également pour analyser rétrospectivement les données en fin d'essai si une autre méthode de recherche de dose a été utilisée.

L'intérêt de ce modèle est multiple : d'une part modéliser la variation du risque de toxicité dans le temps pour étudier de possibles effets cumulés ou retardés. D'autre part, augmenter la quantité d'information utilisée en prenant en compte le caractère ordinal des données de toxicité, c'est-à-dire l'information donnée par les patients présentant une toxicité grave mais aussi celle des patients présentant une toxicité modérée.

Notre démarche a été la suivante : Nous avons évalué les caractéristiques opératoires de ce modèle dans le contexte des essais de phase I dans une étude de simulations (Doussau, Thiebaut et al. 2013). Nous avons ensuite analysé rétrospectivement les données complètes de trois essais de phase I avec le modèle POMM. Nous avons également confronté ces résultats à ceux d'une analyse de la probabilité de toxicité grave cumulée sur plusieurs cycles avec le modèle proposé dans la TITE-CRM afin d'illustrer les informations complémentaires apportées par ces deux types de modélisation.

## 3.2 Méthodes

Le modèle logistique mixte à cotes proportionnelles est une extension du modèle logistique à cotes proportionnelles utilisé sur la période d'évaluation des TDL, introduit dans le paragraphe 1.5.2, tel qu'utilisé dans la PO-CRM (Van Meter, Garrett-Mayer et al. 2011).

### 3.2.1 Explicitation du modèle POMM

Soit  $Y_{ij}$  la variable ordinaire reflétant le grade maximal de toxicité observée par le patient  $i$  au cycle  $j$ ,  $x_{ij}$  la dose reçue au cycle  $j$  et  $t_{ij}$  le temps que l'on assimile au cycle du fait que la durée des cycles administrés est fixe, comme cela a été introduit au paragraphe §1.6.1.

L'extension consiste à intégrer un *intercept* aléatoire,  $u_i$ , au niveau de la combinaison linéaire des paramètres du modèle 3. Cela permet de prendre en compte la variabilité du risque de toxicité entre patients (Carriere and Bouyer 2006; Pedrono, Thiebaut et al. 2009), et par ce moyen prendre en compte la corrélation des mesures chez un même patient.

Ainsi on fait l'hypothèse que, conditionnellement à cette valeur individuelle, les observations d'un même patient aux différents cycles de traitement deviennent indépendantes. Par la suite, la dose étant supposée fixe pour un patient donné, telle qu'administrée au premier cycle, la dose sera indexée uniquement sur  $i$ .

La modèle POMM le plus complet s'écrit :

$$\psi^{-1}(x_i, t_{ij}, \theta) = \log \left( \frac{P(Y_{ij} \leq k | x_i, t_{ij}, u_i)}{1 - P(Y_{ij} \leq k | x_i, t_{ij}, u_i)} \right) = \alpha_k - \beta_1 x_i - \beta_2 t_{ij} - u_i, \quad (\text{Modèle 7})$$

pour  $k = (1, \dots, K - 1)$ , avec  $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{K-1}$ ,  $u_i \sim N(0, \sigma_0^2)$

La condition «  $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{K-1}$  », sera considérée implicite dans l'écriture des modèles pour données ordinales suivants.

En l'absence de variation log-linéaire du risque de toxicité avec le temps, testée par le non-rejet de l'hypothèse  $H_0 : \beta_2 = 0$ , le modèle 7 peut être simplifié ainsi :

$$\psi^{-1}(x_i, \theta) = \log \left( \frac{P(Y_{ij} \leq k | x_i, u_i)}{1 - P(Y_{ij} \leq k | x_i, u_i)} \right) = \alpha_k - \beta_1 x_i - u_i \quad (\text{Modèle 8})$$

Afin de vérifier l'hypothèse de proportionnalité des cotes, le test du rapport de vraisemblance peut être effectué grâce à l'estimation du modèle mixte ordonné généralisé (GOMM) suivant:

$$\psi^{-1}(x_i, \theta) = \log \left( \frac{P(Y_{ij} \leq k | x_i, u_i)}{1 - P(Y_{ij} \leq k | x_i, u_i)} \right) = \alpha_k - \beta_{1k} x_i - u_i \quad (\text{Modèle 9})$$

Si l'on décompose le coefficient  $\beta_{1k}$  en  $\beta_1 + \beta'_{1k}$  (pour  $k = 2, \dots, K-1$ ), avec  $\beta'_{1k} = 0$  pour  $k = 1$ , les modèle 8 et 9 deviennent emboîtés. Ce test peut être effectué par le test du rapport de vraisemblance, à  $k-2$  degrés de liberté, entre les modèles (9) et (8), emboîtés. Un test de Wald peut également être effectué (hypothèse  $H_0 : \beta'_{1k} = 0$ ).

On peut également tester la présence d'une variabilité interpatient, sous forme d'un *intercept* aléatoire, entre le modèle POMM (modèle 8) et le modèle POM équivalent (modèle 3), par le test du rapport vraisemblance. La statistique de test suit la distribution d'un mélange entre  $\chi^2(0$  degré de liberté) et  $\chi^2(1$  degré de liberté) sous l'hypothèse nulle.

- **Vraisemblance**

Compte tenu des observations  $\Omega_n = \{(x_i, y_{ij})\}$  sur les  $n$  premiers sujets inclus et évalués sur plusieurs cycles, et puisque l'effet aléatoire est inconnu, l'évaluation de la vraisemblance doit être intégrée sur la distribution de l'effet aléatoire,  $f$ , c'est-à-dire :

$$L(\theta | \Omega_n) = \prod_{i=1}^n \prod_j \int p_1(x_i | u_i)^{I^{[y_{ij}=1]}} \times p_2(x_i | u_i)^{I^{[y_{ij}=2]}} \times p_3(x_i | u_i)^{I^{[y_{ij}=3]}} \times f(u_i) du_i$$

**(Équation D)**

La quadrature adaptative de Gauss-Hermite peut être utilisée. La maximisation de la log-vraisemblance est obtenue par intégration numérique. L'algorithme de la quadrature adaptative de Gauss-Hermite est utilisé car particulièrement efficace pour évaluer des intégrales sur des espaces de probabilité où la densité associée est symétrique autour d'un mode. Elle est dite adaptative car le pas de l'incrément pour évaluer numériquement la log-vraisemblance est modifié en fonction de la densité de probabilité. L'algorithme est suffisamment efficace pour permettre de réaliser 1 000 simulations sur un ordinateur personnel en des temps raisonnables ( $\approx 5$  heures).

- **Mise en perspective des modèles introduits**

Dans le tableau suivant, nous donnons le nombre de paramètres à estimer dans les différents modèles présentés en introduction, et ceux qui sont illustrés dans la suite de ce travail, avec une variable réponse en  $k$  modalités pour les modèles ordinaux (Tableau 5).

**Tableau 5. Nombre de paramètres par modèle.** Les variables explicatives sont précisées entre parenthèses.

Modèle	Effets fixes		Effets aléatoires
	$\alpha$	$\beta$	
Simon (dose, dose cumulée, <i>intercept</i> aléatoire, erreur résiduelle)	$k-1^*$	1	2
Legedza (dose cumulée $\pm$ <i>intercept</i> aléatoire)	1	1	(1)
Modèle à cotes proportionnelles : POM (dose)	$k-1$	1	-
Modèle logistique ordonné généralisé : GOM (dose)	$k-1$	$k-1$	-
Modèle mixte à cotes proportionnelles : POMM (temps, dose, <i>intercept</i> aléatoire)	$k-1$	2	1
Modèle mixte logistique ordonné généralisé : GOMM (temps, dose, <i>intercept</i> aléatoire)	$k-1$	$2 + 2 \times (k-2)$	1

\*estimation des seuils de coupure de la variable continue latente de toxicité

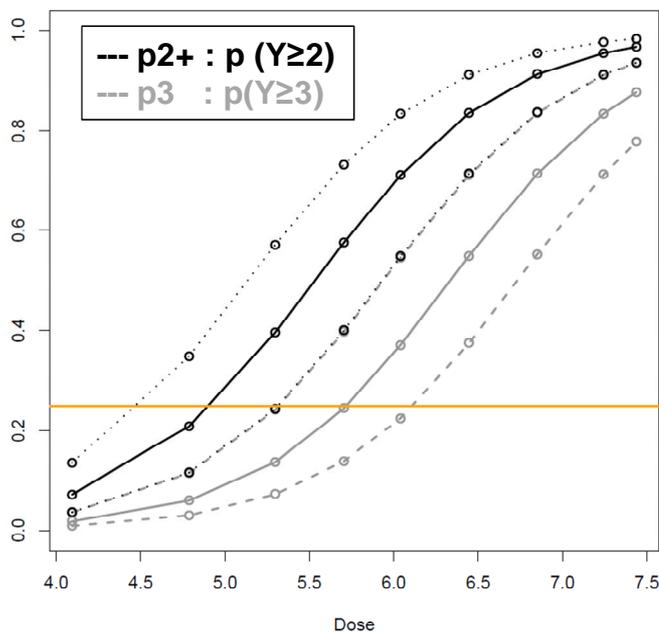
- **Probabilités estimées**

Plusieurs probabilités peuvent être calculées à partir des estimations du modèle. On peut calculer la probabilité de toxicité individuelle pour différentes valeurs de l'*intercept* aléatoire, c'est-à-dire une probabilité conditionnelle à une valeur donnée de  $u_i$  : par exemple pour un *intercept* aléatoire de valeur nulle, parfois assimilée à un patient « moyen », ou pour des valeurs seuil de la distribution de l'effet aléatoire, comme la valeur moyenne  $\pm 1$  écart-type, selon l'équation E (Figure 9). C'est ce que l'on ferait si on souhaitait prédire la relation dose-toxicité pour un patient dont on connaîtrait déjà la valeur de l'*intercept* aléatoire (prédiction individuelle) par un estimateur Bayésien empirique.

$$\left\{ \begin{array}{l} p_3(d_l | \hat{\theta}, u_i) = P(Y_{ij} = 3 | \hat{\alpha}_2, \hat{\beta}_1, u_i) = 1 - P(Y_{ij} \leq 2 | \hat{\alpha}_2, \hat{\beta}_1, u_i) = 1 - \left( \frac{e^{(\hat{\alpha}_2 - \hat{\beta}_1 d_l - u_i)}}{1 + e^{(\hat{\alpha}_2 - \hat{\beta}_1 d_l - u_i)}} \right) \\ p_{2+}(d_l | \hat{\theta}, u_i) = P(Y_{ij} \geq 2 | \hat{\alpha}_1, \hat{\beta}_1, u_i) = 1 - P(Y_{ij} \leq 1 | \hat{\alpha}_1, \hat{\beta}_1, u_i) = 1 - \left( \frac{e^{(\hat{\alpha}_1 - \hat{\beta}_1 d_l - u_i)}}{1 + e^{(\hat{\alpha}_1 - \hat{\beta}_1 d_l - u_i)}} \right) \end{array} \right.$$

(Équation E)

Probabilité de toxicité ( $u_i = 0 \pm \sigma_0^2$ )



**Figure 9.** Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle à cotes proportionnelles, de paramètre  $\theta = (\alpha_1 = 9,85, \alpha_2 = 11,29, \beta_1 = 1,78, \sigma^2=0,5)$ , pour les valeurs  $u_i = 0$  (lignes noire et grise pleines) et  $u_i = \pm \sigma_0$  (lignes pointillées). La ligne orange correspond à une cible de probabilité de toxicité  $\tau = 25\%$ . [Figure A1 (annexe) de l'article (Doussau, Thiebaut et al. 2013)]

On note que les courbes de différents individus sont parallèles, parce qu'elles diffèrent uniquement par l'*intercept* alors que le coefficient de la dose donc la pente, reste inchangée (Figure 9). Il en est de même pour les courbes dose-toxicité pour différents niveaux de réponse de la variable  $Y_{ij}$ , qui diffèrent également uniquement par l'*intercept*, du fait de l'hypothèse de proportionnalité des cotes. La courbe dose-toxicité de l'individu « moyen » correspond à ce qu'estime le modèle pour l'effet fixe de la dose (Carriere and Bouyer 2002).

La probabilité moyenne (obtenue par intégration de la distribution des effets aléatoires) est différente de la probabilité pour un sujet « moyen » ayant un *intercept* aléatoire de valeur  $u_i = 0$ . Cela est lié à la transformation logistique et à la non-linéarité de la variance qui en résulte ; la pente de la relation dose-toxicité est atténuée par l'intégration sur la distribution (Figure 10). Cette intégration peut être effectuée par des simulations de Monte Carlo avec tirage au sort de valeurs de l'effet aléatoire, selon une distribution Normale  $(0, \sigma_0^2)$ , et les probabilités prédites à chaque dose sont moyennées (Molenberghs and Verbeke 2005c).

Lorsque l'effet aléatoire est de faible amplitude, l'estimation des probabilités prédites de chaque dose par intégration est proche de la probabilité prédite pour  $u_i = 0$  (Figure 10). L'écart se creuse avec l'amplitude de l'effet aléatoire (Figure 11), sauf pour les probabilités extrêmes (proches de 0 ou de 1) ou pour la probabilité de 0,5, avec une inversion de courbure autour de la probabilité de 0,5, du fait de la symétrie de la fonction. Ainsi, en dessous d'une probabilité de 0,5, la probabilité prédite pour  $u_i = 0$  est systématiquement inférieure à la probabilité intégrée sur la distribution de l'*intercept* aléatoire.

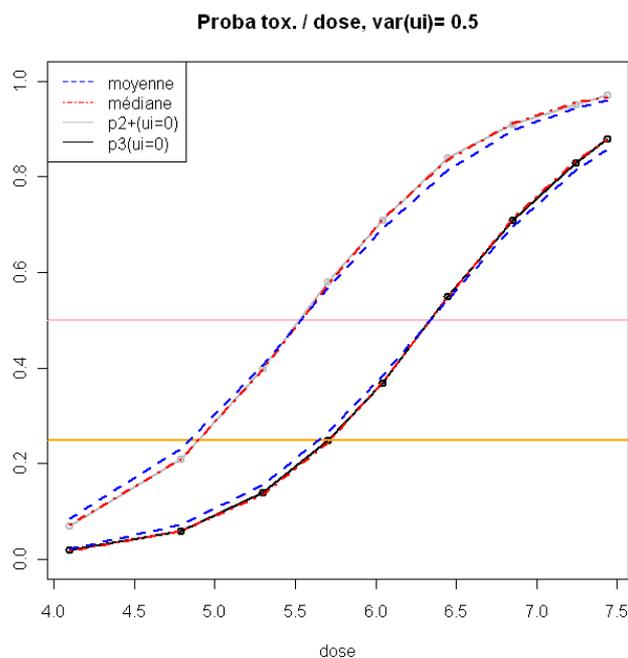
On peut appeler cette estimation via un modèle mixte découlant des modèles linéaires mixtes généralisés (GLMM) « moyenne marginale des prédictions » (MMP). Il ne faut pas confondre cette estimation avec l'estimation moyenne sur la population (EMP) qui aurait été estimée par un modèle marginal (estimation par équation généralisées, GEE) (Molenberghs and Verbeke 2005c) . Ces deux estimations MMP et EMP sont néanmoins le plus souvent proches.

Dans notre étude, nous avons utilisé un modèle à effet aléatoire de la famille GLMM, plutôt qu'un modèle marginal type GEE, pour deux raisons principales :

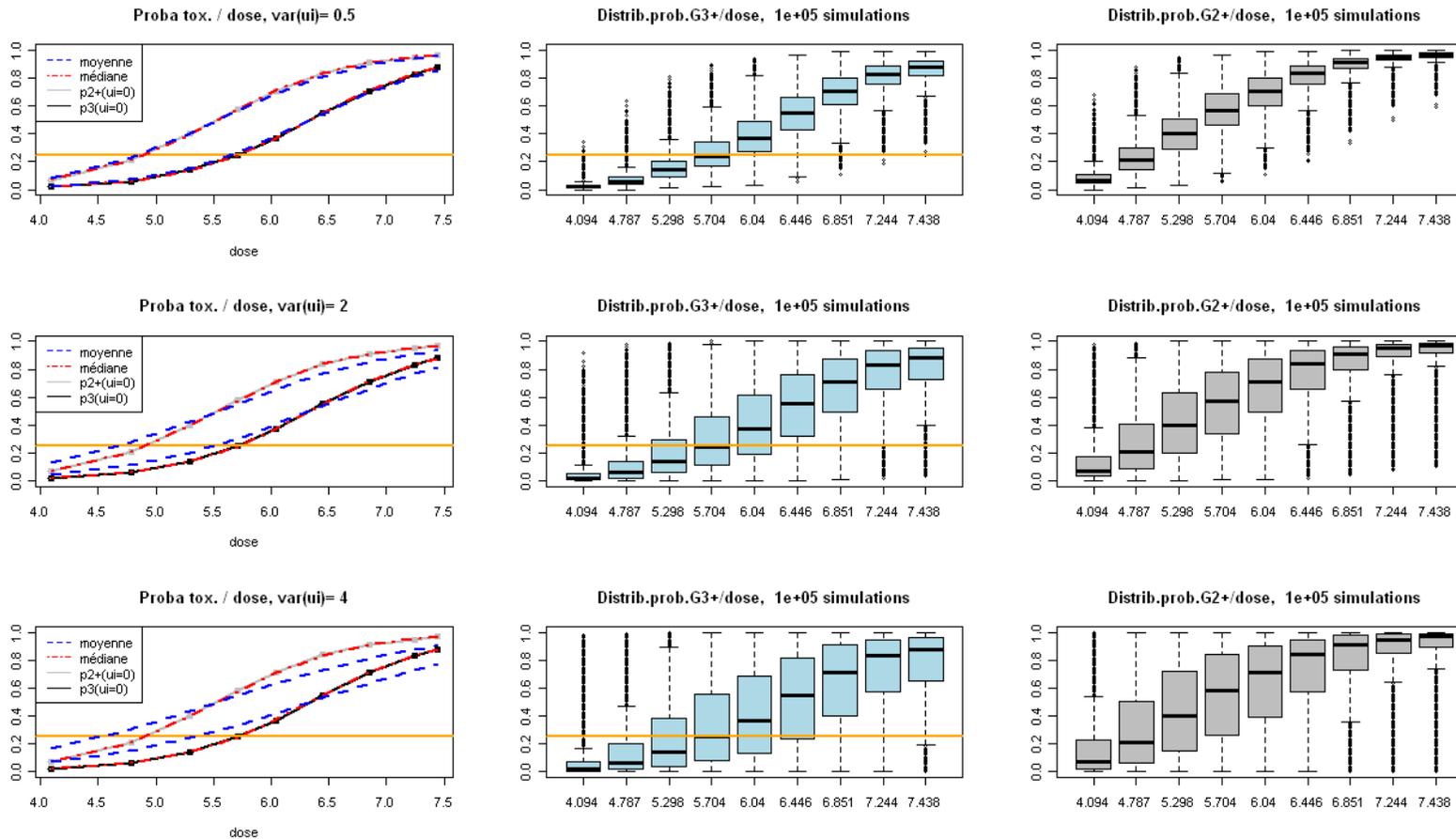
- d'une part les estimations par GLMM sont robustes aux données manquantes au hasard (Little and Rubin 1987; Carriere and Bouyer 2002; Molenberghs and Verbeke 2005d). En effet, les estimations sont robustes lorsque l'absence des données (toxicité ou covariables) dépend uniquement de données observées (Carriere and Bouyer 2002). Cette hypothèse nous semble tout à fait raisonnable dans le contexte clinique des essais de recherche de dose dans lesquels les patients interrompent le traitement pour toxicité ou progression, et cessent le suivi de toxicité à l'arrêt du traitement. En effet, si la sortie d'étude à un cycle  $j$  est liée à une toxicité grave observée, les données manquantes aux cycles  $j+1$  et suivants sont expliquées par une valeur observée. En revanche, ces données manquantes ne peuvent pas être considérées de type complètement au hasard (*missing completely at random*, MCAR), puisque la

probabilité de toxicité, donc la probabilité de censure, croissent avec la dose. De plus, nous effectuons l'hypothèse que, conditionnellement à la dose, la toxicité est indépendante de la progression de la maladie. En effet, pour les thérapies ciblées, les mécanismes d'action qui induisent la toxicité sont généralement différents de ceux qui permettent l'effet anti-tumoral. A ce jour, quelque soit la thérapie ciblée, on n'a pas mis en évidence de lien entre réaction toxique et effet anti-tumoral.

- d'autre part, l'effet aléatoire permet de faire des prédictions individuelles prenant en compte la variabilité interpatient, perspective qui avait été envisagée au départ comme une extension possible de la méthode.



**Figure 10. Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle à cotes proportionnelles, de paramètre  $\theta = (\alpha_1 = 9,85, \alpha_2 = 11,29, \beta_1 = 1,78, \sigma^2=0,5)$ . Les courbes noire et grise pleines correspondent aux prédictions conditionnelles à la valeur  $u_i = 0$ . Les courbes pointillées correspondent à la médiane (rouge) et à la moyenne (bleu) des probabilités prédites marginales. Probabilités marginales calculées par méthode de Monte Carlo avec  $10^5$  simulations de  $u_i$ . La ligne orange correspond à une cible de probabilité de toxicité  $\tau = 25\%$ . La ligne rose correspond au seuil de  $50\%$ .**



**Figure 11. Relations entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon l'amplitude de l'effet aléatoire.** Modèle à cotes proportionnelles, de paramètre  $\theta = [\alpha_1 = 9,85, \alpha_2 = 11,29, \beta = 1,78, \sigma_0^2 = (0,5 ; 2 ; 4, \text{ de haut en bas})]$ . Première colonne : Les courbes noire et grise pleines correspondent aux prédictions marginales ( $u_i = 0$ ). Les courbes pointillées correspondent à la médiane (rouge) et à la moyenne (bleu) des probabilités prédites marginales, après tirage au sort de  $10^5$  valeurs de  $u_i$ . La ligne orange correspond à une cible de probabilité de toxicité  $\tau = 25\%$ . La ligne rose correspond au seuil de  $50\%$ . Deuxième et troisième colonnes : distributions par dose des probabilités de G3+ et G2+, après tirage au sort de  $10^5$  valeurs de  $u_i$ .

- **Risque de toxicité au cours du temps**

Le modèle 7 permet d'estimer la probabilité de toxicité à chaque dose et à chaque cycle. Pour illustrer cela, nous avons représenté la variation du risque de toxicité de chaque dose au cours de cycles successifs selon que cette probabilité augmente un peu, beaucoup ou diminue au cours du temps (Figure 12).

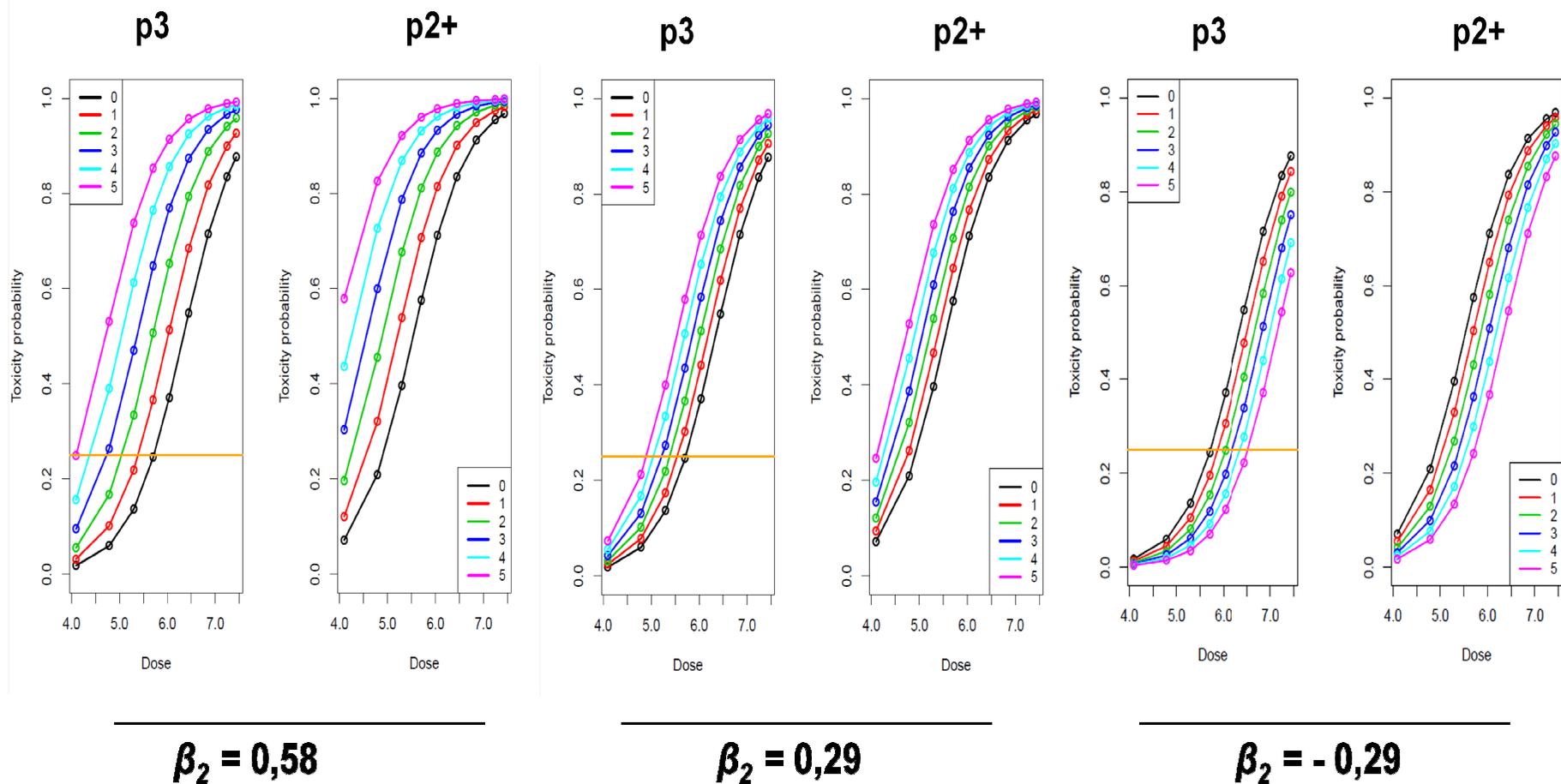


Figure 12. Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle à cotes proportionnelles, de paramètres  $\theta = [\alpha_1 = 9,85, \alpha_2 = 11,29, \beta_1 = 1,78, \beta_2 = (0,58 ; 0,29 ; -0,29, \text{ respectivement, de gauche à droite})]$ , pour  $u_i = 0$ . La ligne orange correspond à une cible de probabilité de toxicité  $\tau = 25\%$ . La couleur de chaque courbe correspond à un cycle différent (du temps 0 = 1er cycle en noir au temps 5 = 6<sup>ème</sup> cycle en rose). [Figure C.1 (annexe ) de l'article (Doussau, Thiebaut et al. 2013)]

### 3.2.2 Programmation

La mise en œuvre de ces modèles a été étudiée à l'aide de deux logiciels : SAS et R. Le *package* « ordinal » (Rune Haubo B Christensen, 2012.01-19) a été utilisé avec le logiciel R (v2.13.0). La procédure Nlmixed a été utilisée avec le logiciel SAS (v9.1). Le *package* « ordinal » sous R est moins flexible dans la spécification du modèle que la procédure Nlmixed qui permet d'estimer le modèle GOMM, utile pour tester l'hypothèse de proportionnalité des cotes. A titre d'illustration, nous avons présenté les syntaxes et estimations par les deux logiciels en parallèle sur le jeu de données d'un essai extrait des données publiées par (Postel-Vinay, Gomez-Roca et al. 2011). Les données portent sur 35 patients, qui ont reçu un total de 113 cycles à 5 doses différentes. Le test de l'hypothèse de proportionnalité des cotes est également présenté (Annexe 4).

### 3.3 Etude de simulation évaluant l'intérêt du modèle à cotes proportionnelles dans les essais de recherche de dose

#### 3.3.1 Objectifs

Nous avons souhaité évaluer la capacité du modèle POMM à améliorer l'identification de la dose cible. Nous avons envisagé son intérêt pour l'analyse rétrospective des données collectées d'une part, et d'autre part dans un nouveau schéma adaptatif pour lequel le modèle POMM serait utilisé pour l'attribution des doses, en utilisant comme cible de toxicité la probabilité de toxicité grave par cycle de traitement. Nous avons également souhaité évaluer la puissance du modèle POMM pour détecter une variation du risque de toxicité avec le temps lors de l'analyse finale des données d'essais menés par MRC.

#### 3.3.2 Méthodes

Le modèle POMM peut être appliqué sur les données rétrospectives collectées par un schéma CRML similaire à celui introduit précédemment (§2.3), en utilisant le modèle 7 (stratégie M1 : POMM sur données longitudinales collectées par MRC).

Par ailleurs, en cours d'essai, l'attribution des doses peut être effectuée à l'aide des estimations d'un modèle POMM incluant les données de tous les cycles administrés au moment de la nouvelle inclusion (schéma adaptatif POMM-CRML : stratégie M2). Dans le nouveau schéma adaptatif M2, l'utilisation du modèle POMM était autorisée à partir de l'inclusion du 16<sup>ème</sup> patient, à l'aide du modèle 8 (modèle fonction de la dose, avec un *intercept* aléatoire). Plusieurs simplifications du modèle étaient autorisées dans les cas où l'estimation n'était pas possible : (a) modèle POMM à pente fixée avec *intercept* aléatoire, (b) modèle POM à pente fixée sans *intercept* aléatoire.

La CRML classique, utilisant un modèle logistique pour l'attribution des doses (stratégie M0), était considérée comme stratégie de référence.

- **Schéma d'étude**

Dans une étude de simulation nous avons effectué 1 000 simulations pour plusieurs scénarios, avec une taille d'étude fixe de 30 patients et une cible de toxicité  $\tau = 25\%$  (probabilité de toxicité grave par cycle de traitement). La survenue de la première toxicité grave entraînait la censure. Le scénario principal étudié correspondait aux probabilités de toxicité de la Figure 9 :  $\theta = (\alpha_1 = 9,85, \alpha_2 = 11,29, \beta_1 = 1,78, \sigma_0^2 = 0,5)$ , sans variation du risque de toxicité dans le temps ( $\beta_2 = 0$ ). Dans ce scénario, les 9 paliers de dose explorés et la relation entre la dose et la probabilité de toxicité grave étaient issues d'une étude de simulation antérieure (Zhou and Whitehead 2003). Les variations ont consisté à étudier d'autres relations dose-toxicité (dose recommandée au 2<sup>ème</sup> ou 6<sup>ème</sup> palier sur 9 paliers à explorer) et des scénarios qui ne respectaient pas l'hypothèse de proportionnalité des cotes sur la dose (comme le scénario représenté sur la Figure 6). Nous avons également étudié dans une analyse ancillaire l'impact du nombre de paliers à explorer sur les performances des schémas, en étudiant des scénarios où 6 paliers de dose étaient à explorer au lieu de 9.

Les scénarios de la Figure 11 ont été étudiés pour la puissance de détection d'un effet du temps. Le scénario principal a été utilisé pour le risque de première espèce puisque les données étaient générées sans modification du risque de toxicité avec le temps. Pour cette partie de l'étude, le modèle POMM était estimé sur les données rétrospectives d'essais conduits par la CRML.

- **Critères d'évaluation**

Les performances des deux stratégies M1 et M2 ont été comparées sur leur capacité à identifier la dose cible, la distribution des doses recommandées, et la répartition des patients inclus par dose, par rapport à la CRML. Pour chaque analyse, nous avons également évalué la proportion d'essais pour lesquels les estimations ont été possibles, afin de vérifier la faisabilité de la stratégie. La puissance et le risque de première espèce pour détecter un effet du temps statistiquement significatif ont été évalués dans chaque scénario temporel.

- **Application**

Nous avons illustré les estimations du modèle POMM sur les données d'un essai de recherche de dose de l'aviscumine pour les tumeurs solides (Schoffski, Riggert et al. 2004). Cet essai mené par une méthode MRC a inclus 41 patients, à 14 niveaux de doses.

### 3.3.3 Résultats

Cette étude de simulation a montré que, lors de l'analyse rétrospective des données d'essais conduits sur 30 patients avec une CRML classique, le modèle POMM permettait de détecter une augmentation du risque de toxicité avec une puissance de 93 % à 45 % selon l'amplitude de l'effet temporel, avec un risque de première espèce de 5 % en l'absence de variation du risque de toxicité avec le temps.

La CRML conduisait à identifier la dose cible dans 46,7 % des essais dans le scénario central (stratégie M0). Avec la stratégie de réanalyse des données longitudinales obtenues par un schéma CRML (stratégie M1), 62,3 % des essais identifiaient correctement que la 4<sup>ème</sup> dose était celle la plus proche d'une toxicité de 25 % par cycle. Le nouveau schéma adaptatif (stratégie M2) ne permettait pas d'améliorer notablement la proportion d'essais recommandant la dose cible (62,7 %). Cependant, elle permettait d'allouer plus souvent les patients inclus à la dose cible (en moyenne 11 patients, contre 9 patients pour la stratégie M0). La stratégie M2, qui débutait l'attribution des doses selon le POMM à compter de la 16<sup>ème</sup> inclusion, a nécessité une simplification du modèle POMM utilisé dans 4,7 % des étapes (POMM à pente fixe) et 0,1% des étapes (POM à pente fixe).

Les résultats étaient robustes au travers des différents scénarios, même lorsque l'hypothèse des cotes proportionnelles n'était pas respectée (Tableau 6).

**Tableau 6. Distribution des doses recommandées et du nombre de patients moyen inclus par dose selon les différentes stratégies. Résultats sur données censurées après survenue de la première toxicité grave.** G3+: Severe toxicity, G2+: Moderate or severe toxicity. RP2D: true recommended phase II dose. DR: dose recommended at the end of the simulated trial, POMM: proportional odds mixed effect model, CRML: continual reassessment method. Les colonnes en gras correspondent aux résultats pour la dose cible, dont la toxicité est la plus proche de 25 %. ([Table II, III, V de l'article (Doussau, Thiebaut et al. 2013)]).

<b>Scenario : RP2D=4<sup>th</sup> dose (core scenario)</b>										
Dose level	1	2	3	<b>4</b>	5	6	7	8	9	<i>N cycles:</i>
Prob tox G3+	0.02	0.06	0.14	<b>0.24</b>	0.37	0.55	0.71	0.83	0.88	<i>mean</i>
Prob tox G2+	0.07	0.21	0.4	<b>0.58</b>	0.71	0.84	0.91	0.95	0.97	<i>(sd)</i>
<b>Classical CRML<sub>30</sub> (cycle 1)</b>										
Mean number of pts / dose	1.8	3.9	7.8	<b>9.3</b>	5.4	1.4	0.3	0.05	0	<i>30</i>
Distribution DR (%)	0.3	5.1	29.2	<b>46.7</b>	17.9	0.8				<i>(0)</i>
<b>POMM on longitudinal data collected from CRML<sub>30</sub></b>										
Mean number of pts / dose	1.8	3.9	7.8	<b>9.3</b>	5.4	1.4	0.3	0.05	0	<i>111.3</i>
Distribution DR (%)	0.2	1.9	18.5	<b>62.3</b>	15.4	1.0	0.0	0.1	0.5	<i>(19.3)</i>
<b>POMM-CRML adaptive design</b>										
Mean number of pts / dose	1.8	3.4	7.0	<b>11.1</b>	4.9	1.2	0.3	0.1	0.1	<i>110.1</i>
Distribution DR (%)	0.2	1.3	20.9	<b>62.7</b>	14.5	0.1	0.2	0.1	0	<i>(14.8)</i>
<b>Scenario : RP2D=4<sup>th</sup> dose - not PO</b>										
Dose level	1	2	3	<b>4</b>	5	6	7	8	9	<i>N cycles:</i>
Prob tox G3+	0.02	0.06	0.14	<b>0.24</b>	0.37	0.55	0.71	0.83	0.88	<i>mean</i>
Prob tox G2+	0.12	0.35	0.59	<b>0.76</b>	0.86	0.93	0.97	0.98	0.99	<i>(sd)</i>
<b>POMM on longitudinal data collected from CRML<sub>30</sub> design</b>										
Mean number of pts / dose	IDEM core Scenario									<i>111.3</i>
Distribution DR (%)	0.1	1.0	18.1	<b>63.8</b>	<b>16.0</b>	0.4	0.2	0.1	0.3	<i>(19.3)</i>
<b>POMM-CRML adaptive design</b>										
Mean number of pts / dose	1.8	3.3	6.9	<b>11.3</b>	5.0	1.2	0.3	0.1	0.1	<i>110,1</i>
Distribution DR (%)	0.1	1.3	18.7	<b>64.4</b>	15.2	0.3				<i>(14.8)</i>
<b>Scenario : RP2D=6<sup>th</sup> dose</b>										
Dose level	1	2	3	4	5	<b>6</b>	7	8	9	<i>N cycles:</i>
Prob tox G3+	0.01	0.02	0.04	0.09	0.15	<b>0.26</b>	0.42	0.60	0.68	<i>mean</i>
Prob tox G2+	0.02	0.07	0.16	0.28	0.41	<b>0.59</b>	0.75	0.86	0.89	<i>(sd)</i>
<b>Classical CRML<sub>30</sub> (cycle 1)</b>										
Mean number of pts / dose	1.2	1.5	2.2	3.7	7.7	<b>9.1</b>	3.7	0.7	0.2	<i>30</i>
Distribution DR (%)			2.0	6.6	33.4	<b>47.9</b>	9.8	0.3		<i>(0)</i>
<b>POMM on longitudinal data collected from CRML<sub>30</sub> design</b>										
Mean number of pts / dose	1.2	1.5	2.2	3.7	7.7	<b>9.1</b>	3.7	0.7	0.2	<i>115.9</i>
Distribution DR (%)			0.2	2.7	25.8	<b>62.1</b>	7.8	0.7	0.7	<i>(17.5)</i>
<b>POMM-CRML adaptive design</b>										
Mean number of pts / dose	1.2	1.4	1.9	3.0	7.3	<b>10.9</b>	3.3	0.7	0.4	<i>113.8</i>
Distribution DR (%)			0.2	2.6	26.6	<b>64.5</b>	6.0		0.1	<i>(12.5)</i>

### 3.3.4 Article

P1

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials*. **Stat Med**. 2013 Sep 10;32(30):5430-347.

P2

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials*. **Stat Med**. 2013 Sep 10;32(30):5430-347.

P3

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials*. **Stat Med**. 2013 Sep 10;32(30):5430-347.

P4

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

P5

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials*. **Stat Med**. 2013 Sep 10;32(30):5430-347.

P6

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

P7

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p8

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p9

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p10

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p11

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p12

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p13

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p14

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p15

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p16

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p17

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

p18

Doussau A, Thiébaud R, Paoletti X. *Dose-finding design using mixed effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials.* **Stat Med.** 2013 Sep 10;32(30):5430-347.

### 3.3.5 Analyse ancillaire : impact du nombre de paliers de doses sur les performances

L'impact du nombre de paliers de dose à explorer sur la MRC et l'analyse par le POMM des données longitudinales de toxicité des essais menés par MRC a été présenté dans le Tableau 7. Le schéma MRC est plus performant si la dose à identifier est proche de la dose de départ. Le nombre total de doses de 6 ou 9 n'a pas d'impact sur les performances de la MRC si la dose cible est le second ou le 4<sup>ème</sup> palier de doses. En revanche si la dose cible est le 6<sup>ème</sup> palier sur 6 à explorer, le schéma MRC est plus performant que s'il y a 9 paliers à explorer car les patients sont plus concentrés à la dose cible (13 patients en moyenne, contre 9). Le gain lié à l'utilisation du POMM dans l'analyse finale n'est pas impacté par le nombre de palier de doses (+ 15 % de recommandations correctes de la dose cible, par rapport aux résultats de la MRC utilisant les données du premier cycle, quel que soit le nombre total de doses à explorer, et le rang de la dose cible parmi celles-ci).

**Tableau 7. Impact du nombre total de doses à explorer sur le nombre de patients à inclus à la dose cible et la proportion de recommandations correctes de la dose cible.**

Essais menés selon une CRML avec arrêt après 30 patients. Les relations dose-toxicité sont celles décrites dans (Doussau, Thiebaut et al. 2013) selon que l'on puisse explorer 9 paliers ou seulement les 6 premiers paliers des relations dose-toxicité étudiées.

	RP2D =	RP2D =		RP2D =	
	2 <sup>ème</sup> dose	4 <sup>ème</sup> dose		6 <sup>ème</sup> dose	
	Sur 6 doses	Sur 9 doses	Sur 6 doses	Sur 9 doses	Sur 6 doses
Nombre moyen de patients inclus à la dose cible* par la MRC	15,2	9,3	9,8	9,1	13,1
Recommandations correctes* par la MRC (%)	64,7	46,7	48,4	47,9	53,9
Recommandations correctes# par analyse POMM sur données répétées (%)	80,2	62,3	60,7	62,1	70,7

\* dose dont la probabilité de toxicité est la plus proche de 25 % sur le premier cycle.

# dose dont la probabilité de toxicité est la plus proche de 25 % par cycle de traitement.

## 3.4 Etude appliquée sur les données de trois essais

### 3.4.1 Objectif

A l'issue de cette étude de simulation qui a permis d'évaluer les caractéristiques opératoires du modèle POMM dans des scénarios réalistes d'essais de phase I, nous avons souhaité illustrer notre approche par une analyse en parallèle de trois essais publiés (Schoffski, Riggert et al. 2004; Geoerger, Hargrave et al. 2011; Floquet, Doussau et al. 2014). L'essai erlotinib + radiothérapie sur les gliomes cérébraux en onco-pédiatrie conduit par la MRC a été déjà présenté dans la partie 2. Le second essai, l'essai aviscumine dans les tumeurs solides (Schoffski, Riggert et al. 2004), également conduit selon un design MRC, a fait l'objet de l'application de la publication de l'étude de simulation (§3.3.4). L'essai Cacy, est un essai de recherche de dose de la combinaison de la doxorubicine liposomale pégylée avec le cyclophosphamide, dans les rechutes précoces des cancers de l'ovaire, du péritoine et de la trompe (Floquet, Doussau et al. 2014), mené selon un schéma « 3+3 ».

Cette étude appliquée avait pour objectif d'illustrer de manière pragmatique les informations apportées par le POMM au travers de l'estimation de la probabilité de toxicité modérée ou grave et la probabilité de toxicité grave, par cycle de traitement, pour chaque dose. Nous avons également confronté les informations obtenues aux estimations apportées par une analyse par le modèle utilisé dans l'approche TITE-CRM, présentée au paragraphe §1.4.4.1. Ces essais représentaient deux situations cliniques distinctes : un essai avec peu de cycles répétés par patient, deux essais avec de nombreux cycles par patient.

### 3.4.2 Méthodes

L'analyse par le POMM a été explicitée au paragraphe §3.2.1. Les hypothèses de ce modèle sont : hypothèse de proportionnalité des cotes pour la dose et le temps, données manquantes au hasard (MAR). Puisque nous avons intégré le temps et la dose comme variables quantitatives dans le modèle, l'hypothèse de log-linéarité de l'effet de la dose et du temps s'applique également.

Dans les estimations par le modèle de l'approche TITE-CRM, les probabilités estimées par un modèle logistique pondéré selon la durée d'observation (jusqu'à six cycles) peuvent être considérées comme une estimation de la probabilité de présenter au moins une toxicité grave, par patient, sur la durée de traitement administré. Les hypothèses de ce modèle sont un risque

constant d'événement au cours du temps, et que la censure soit non informative. Puisque nous avons intégré la dose comme variable quantitative, l'hypothèse de log-linéarité de l'effet de la dose s'applique également. Nous ferons référence à l'utilisation du modèle de cette approche par « approche TITE ».

Dans ce travail nous avons calculé la probabilité marginale (par intégration de la distribution des effets aléatoires), et l'intervalle de confiance des probabilités prédites, afin de disposer d'une mesure de précision utile pour les conclusions de l'analyse rétrospective des essais. Pour fournir l'intervalle de confiance des probabilités afin d'affiner l'estimation du profil de toxicité de la dose recommandée dans les différents essais, par le modèle POMM ou par l'approche TITE, nous avons utilisé la Delta Méthode. Cette méthode permet de prendre en compte la covariance entre les paramètres dans l'estimation de la variabilité de la combinaison linéaire des paramètres,  $\psi^{-1}$ . Nous avons supposé la combinaison linéaire des paramètres estimés suivait une distribution Normale. Le gain de précision par rapport à la CRML peut être estimé par la différence de largeur de l'intervalle de confiance entre le POMM et la CRML, rapportée à la largeur de l'intervalle de confiance estimé par la CRML.

### 3.4.3 Principaux résultats

Les essais avaient inclus respectivement 20, 41 et 21 patients. Les 6 premiers cycles de traitement ont été pris en compte, c'est à dire respectivement 94, 96 et 72 cycles. Le résultat majeur de cette étude a été la découverte d'une augmentation du risque de toxicité au cours du temps dans l'essai de recherche de dose de la combinaison doxorubicine-cyclophosphamide. Ainsi la probabilité de toxicité grave à la dose qui a été recommandée, augmentait de 27,1 % (IC 95 % = 3,8 % - 56,3 %) au premier cycle à 59,2 % (IC 95 % = 25,6 % - 86,7 %) au sixième cycle. L'estimation de la probabilité de toxicité au cours des 6 cycles, par l'approche dérivée de la TITE, était estimée à 78,3 % (IC 95 % = 41,4 % - 94,8 %) pour la dose qui a été recommandée dans l'essai de combinaison. Bien entendu, cette approche est d'autant plus intéressante lorsque la majorité des patients reçoit un traitement prolongé, ce qui n'était pas le cas dans un des essais (aviscumine).

Dans ces trois essais, l'hypothèse de proportionnalité des cotes n'était pas rejetée, mais le test est néanmoins probablement peu puissant sur cette taille d'échantillon.

Les approches illustrées permettaient d'évaluer deux mesures complémentaires du poids global des toxicités au cours du traitement : un poids des toxicités pour chaque cycle et un poids des toxicités par patient.

#### **3.4.4 Valorisation**

Les résultats du modèle POMM ont été présentés au congrès de la société européenne d'oncologie médicale (ESMO/ECCO, Annexe 5).

Cette approche a été intégrée à un chapitre de livre (Annexe 6).

- **Article**

P1

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

P2

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

P3

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

P4

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

P5

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

P6

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

P7

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p8

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p9

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p10

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p11

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p12

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p13

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p14

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p15

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p16

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p17

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p18

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p19

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p20

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

p 21

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

Doussau A, Thiébaud R, Georger B, Le Deley MC, Schöffski P, Floquet A, Mathoulin-Pélissier S, Rizzo E, Fumoleau P, Le Tourneau C, Paoletti X. *A new approach to integrate toxicity grade and later treatment cycles in the analysis and reporting of phase I dose-finding trials.* (soumis)

## 4 Modèle de régression logistique à cotes proportionnelles *versus* modèle logistique longitudinal

### 4.1 Introduction

L'étude de simulation (Doussau, Thiebaut et al. 2013) et les analyses rétrospectives sur des données d'essais (§3) ont permis d'étayer l'intérêt du modèle à cotes proportionnelles pour données répétées. Cette approche augmente l'information classiquement utilisée dans deux dimensions : d'une part, les données de l'ensemble des cycles sont utilisées et, d'autre part, le caractère ordinal du critère de jugement permet d'utiliser l'information apportée par les toxicités modérées ainsi que d'estimer leur fréquence par cycle.

Il nous a semblé nécessaire de compléter ce travail par une étude évaluant **le modèle logistique mixte (logMM) longitudinal**, comparé au modèle mixte à cotes proportionnelles, afin de distinguer le gain lié à la prise en compte d'un critère ordinal, du gain lié à la prise en compte des données longitudinales. Cette étude avait également pour objectif d'étudier l'impact d'un modèle simplifié dans lequel on n'estimerait pas la pente de la relation dose-toxicité, par rapport aux modèles où ce paramètre est estimé (POMM et logMM), afin d'évaluer la balance bénéfice-risque d'un modèle plus flexible *versus* un modèle simplifié avec moins de paramètres donc plus facilement identifiable.

### 4.2 Méthodes

En gardant une syntaxe similaire à celles introduites précédemment, le modèle logistique mixte à *intercept* aléatoire peut s'écrire :

$$\psi^{-1}(x_i, t_{ij}, \theta) = \log \left( \frac{P(Y_{ij} \leq 2 | x_i, t_{ij}, u_i)}{1 - P(Y_{ij} \leq 2 | x_i, t_{ij}, u_i)} \right) = \alpha_2 - \beta_1 x_i - \beta_2 t_{ij} - u_i, \quad (\text{Modèle 10})$$

avec  $u_i \sim N(0, \sigma_0^2)$ , pour  $Y_{ij} = (1, 2, 3)$

Dans ce modèle, au plus 4 paramètres sont à estimer :  $\theta = (\alpha_2, \beta_1, \beta_2, \sigma_0^2)$ .  $\beta_1$  peut être soit estimé à partir des données, soit considéré comme connu et non estimé à partir des données, dans ce cas on remplace  $\beta_1$  par  $\beta_c$ ;  $\beta_c$  peut être calculé de la même manière qu'explicité pour les simulations de la partie §2.3 (Annexe 3). Nous nous référerons dans la suite à ces

approches comme POMM<sub>d</sub> et logMM<sub>d</sub>, par rapport aux approches où le paramètre de la dose est estimé (POMM et logMM).

En cas d'absence d'effet temps, le modèle peut être simplifié.

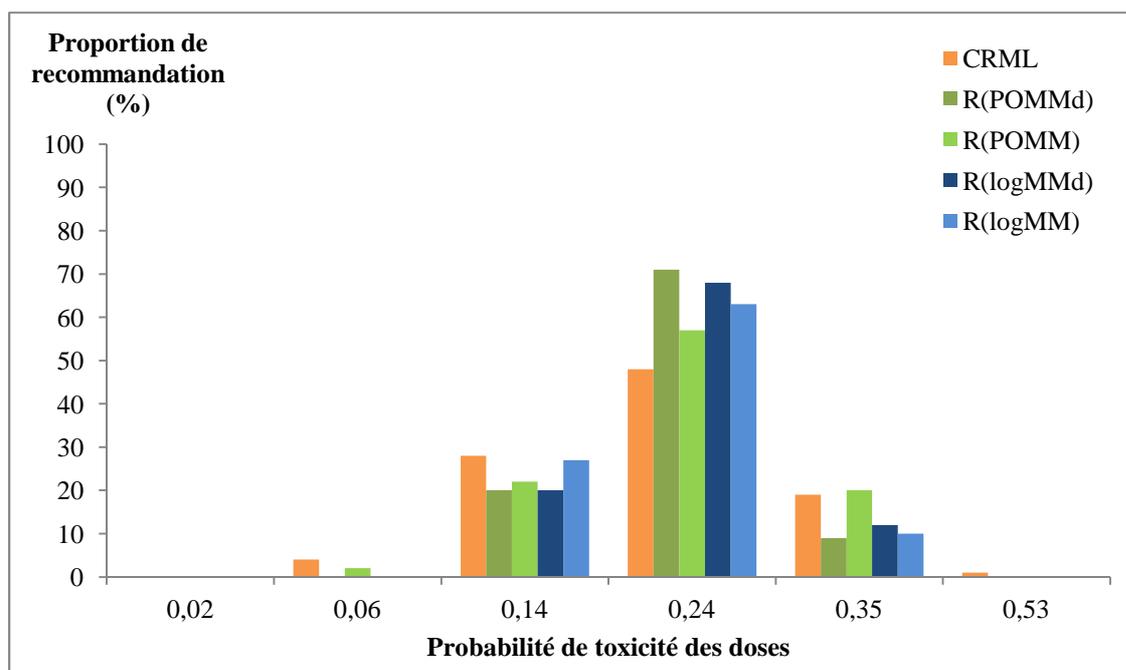
Les estimations de ces quatre différents modèles (POMM, logMM, POMM<sub>d</sub>, logMM<sub>d</sub>) permettent de sélectionner la dose dont la probabilité de toxicité grave par cycle est la plus proche de la probabilité cible  $\tau$ .

Nous avons envisagé les résultats selon que les modèles sont appliqués rétrospectivement sur des données d'essais conduits selon une MRC avec les données du premier cycle, ou selon que l'attribution des doses est menée par un modèle mixte pour données répétées de toxicité.

Le *package* « lme4 » a été utilisé avec le logiciel R.

### 4.3 Résultats

Nous avons mis en parallèle la distribution des doses recommandées selon les différentes méthodes (Figure 13). L'utilisation d'un modèle à cotes proportionnelles avec pente constante mal spécifiée (POMM<sub>d</sub>), est le meilleur modèle pour l'analyse rétrospective de données collectées par un schéma mené par une MRC classique : le modèle POMM<sub>d</sub> permet de recommander la dose cible dans 71 % des essais simulés, contre 57 % lorsque l'on estime l'effet de la dose. Le gain est majeur par rapport à une CRML classique (48 % de recommandations correctes de la dose cible). L'analyse par un modèle logistique permet 68 % d'identification correcte de la dose cible lorsque le paramètre de la dose n'est pas estimé (logMM<sub>d</sub>), contre 63 % lorsqu'on l'estime (logMM). Il faut néanmoins signaler que dans une fraction importante d'échantillons, les modèles dans lesquels la dose est une constante fixée, ne peuvent pas être estimés (18 % des cas avec le POMM<sub>d</sub>). La flexibilité du modèle génère une perte relative d'efficacité due à l'estimation d'un paramètre supplémentaire, mais est néanmoins estimable dans 100 % des cas sur 30 sujets.



**Figure 13. Distribution des doses recommandées sur 1 000 échantillons d’essais simulés selon un schéma CRML (n = 30 patients, recevant 1 à 6 cycles de traitement, jusqu’à la première toxicité grave), selon différents modèles d’analyse des données. La cible de probabilité de toxicité était  $\tau = 25\%$ . CRML : méthode de réévaluation continue, par le maximum de vraisemblance. POMM : proportionnal odds model. LogMM : modèle mixte logistique. Le d en indice signifie que le paramètre de la dose n’est pas estimé. Le R signifie que les données longitudinales sont utilisées pour analyse en fin d’essai (analyse rétrospective).**

Il est important de noter que cette amélioration de recommandation de la dose cible n’augmente pas la proportion d’essais recommandant une dose plus toxique que la cible. Au contraire l’amélioration se fait uniquement au bénéfice d’une diminution des recommandations de la dose inférieure à la dose cible.

Concernant les schémas prospectifs qui utilisent les modèles longitudinaux pour l’attribution adaptative de dose, les résultats sont similaires (Figure 17 en Annexe 7), avec à chaque fois des résultats légèrement améliorés lorsque l’on utilise les données longitudinales pour l’attribution des doses (3 à 5 % supplémentaires de recommandations de la dose cible par rapport au modèle équivalent appliqué sur des données rétrospectives de schémas conduits selon une CRML classique). Là encore, le modèle qui nécessite l’estimation du paramètre de la dose, modèle plus flexible, est plus souvent estimable que le modèle avec pente constante fixée.

En termes d'identification de la dose cible, on peut donc en déduire que l'utilisation des données longitudinales est la principale source d'amélioration des performances d'identification de la dose cible, alors que le caractère ordinal est une moindre source d'amélioration. Comme dans le travail précédent (Doussau, Thiebaut et al. 2013), on observe que le gain lié à l'utilisation des données longitudinales dans un schéma d'étude adaptatif n'est pas majeur comparé au gain de l'utilisation des données longitudinales dans l'analyse finale d'essais menés selon un schéma classique de MRC type CRML.

Par ailleurs cette étude révèle que le modèle POMM est substantiellement plus puissant que le modèle logistique longitudinal pour détecter une modification du risque de toxicité au cours des cycles de traitement (gain de 20 % de puissance du modèle POMM versus logMM lorsque le risque de toxicité augmente de manière linéaire avec le temps). Une erreur de spécification du modèle temps-toxicité conduit à une perte de puissance importante, non différentielle entre les deux modèles (perte de puissance de un tiers, Table 3, Annexe 8).

L'article complet est en Annexe 8.

#### 4.4 Synthèse des estimations appliquées à l'essai erlotinib + radiothérapie

Afin d'illustrer ces nouvelles méthodes d'analyse en parallèle, nous reprenons l'analyse rétrospective de notre « fil rouge », l'essai erlotinib + radiothérapie, par toutes les techniques d'analyse statistique envisagées dans ce travail (Tableau 8). Les résultats du POMM font partie des résultats de l'article présenté au §3.4.2, de même que les estimations de l'approche par modèle logistique pondéré de l'approche TITE. Dans cet essai, la probabilité de toxicité de chaque dose a été estimée avec une variable ordinale en 3 catégories pour le POMM, mais l'analyse d'une variable résumant le grade maximal de toxicité par cycle en 4 modalités (pas de toxicité, grade 1, grade 2, grade 3 ou plus) a également été effectuée. Dans ce tableau figurent également la probabilité cumulée de survenue de la première toxicité grave sur 6 cycles, stratifiée sur la dose (calculée par l'estimateur de Kaplan-Meier, avec censure à l'arrêt du traitement) et l'estimation de la probabilité de toxicité grave par cycle par le modèle logMM.

En l'absence d'effet statistiquement significatif du temps, les modèles longitudinaux ont été simplifiés, et l'*intercept* aléatoire et la dose ont été maintenus dans le modèle. La probabilité de toxicité grave estimée à la dose la plus élevée explorée varie peu selon les différents modèles longitudinaux ou le codage de la variable ordinale en 3 ou 4 catégories (9 à 12 % de probabilité de toxicité grave par cycle). Le gain de précision des modèles longitudinaux par rapport à la CRML, évalué par la variation de la largeur de l'intervalle de confiance, est de l'ordre de 50 % au moins. Selon le modèle logistique pondéré de l'approche TITE, la probabilité d'observer au moins une toxicité grave par patient serait de 37,2 % à cette dose.

**Tableau 8. Analyse de l'essai erlotinib + radiothérapie selon le modèle mixte à cotes proportionnelles, le modèle logistique mixte et le modèle logistique pondéré de l'approche TITE-CRM.** Les données en gras correspondent aux estimations de la dose qui a été recommandée. d = dose, p3 = probabilité de toxicité grave, p2+ = probabilité de toxicité modérée ou grave, p1+ = probabilité de toxicité mineure, modérée ou grave, CI = intervalle de confiance, KM = Kaplan Meier, POMM = proportionnal odds mixed model, logMM = logistic mixed model, TITE = time to event, DLT = toxicité dose limitante.

	Dose erlotinib (mg/m <sup>2</sup> )			Total
	75	100	125	
Number of patients	6	6	8	20
<b>DLT evaluation period</b>				
Number of DLT			<b>2</b>	
P(DLT/d=125)			<b>16</b>	
95%CI (pDLT/d=125)			<b>(4-45)</b>	
<b>Longitudinal data (cycle 1 to 6)</b>				
Number of cycles, worst grade:	26	34	<b>36</b>	96
mild tox (grade 1)	7	21	<b>18</b>	46
moderate tox (grade 2)	4	8	<b>7</b>	19
severe tox (≥ grade 3)	2	1	<b>4</b>	7
Observed mild, moderate or severe tox. (per cycle) %	50,0	88,2	<b>80,6</b>	75,0
Observed moderate or severe tox. (per cycle) %	23.1	26.5	<b>30.6</b>	27.1
Observed severe tox. (per cycle) %	7.7	2.9	<b>11.1</b>	7.3
<b>LogMM*</b>				
p3 (per cycle) %	5.3	7.1	<b>9.3</b>	
95%CI (p3/d=125) %			<b>(3.1-24.4)</b>	
<b>POMM, outcome in 3 categories**</b>				
p2+ (per cycle) %	22.7	26.6	<b>30.8</b>	
95%CI(p2+/d=125) %			<b>(18.7-46.3)</b>	
p3 (per cycle) %	5.8	7.1	<b>8.6</b>	
95%CI (p3/d=125) %			<b>(3.7-18.7)</b>	
<b>POMM, outcome in 4 categories***</b>				
p1+ (per cycle) %	61.0	73.4	<b>83.1</b>	
p2+ (per cycle) %	15.4	24.5	<b>36.5</b>	
p3 (per cycle) %	3.7	6.7	<b>11.6</b>	
95%CI (p3/d=125) %			<b>(4.6-25.6)</b>	
<b>Gain in precision of probability estimates of mixed models (as compared to CRML, %)</b>				
LogMM			<b>-48,0</b>	
POMM-3 modalities			<b>-63,4</b>	
POMM-4 modalities			<b>-48,8</b>	
<b>Data summarized over 6 cycles</b>				
N pts with severe tox (naïve incidence, %)	2 (33.3)	1 (16.7)	<b>3 (37.5)</b>	6 (30.0)
<b>KM incidence %</b>	44,4	16.7	<b>51.4</b>	36.5
95%CI %			<b>(0.0-80.5)</b>	(6.7-56.8)
<b>TITE model</b>				
Cumulated probability of severe toxicity %	30.8	33.9	<b>37.2</b>	
95%CI(p3/d=125) %			<b>(12.2-71.5)</b>	

\*  $\sigma_0=0.45$ . \*\*  $\sigma_0=négligeable$ . \*\*\*  $\sigma_0=0.72$ .

## 5 Discussion générale

### 5.1 Synthèse des principaux résultats

Notre travail a permis d'aborder plusieurs problématiques liées au cadre temporel des essais de phase I de recherche de dose en oncologie : l'aspect séquentiel des inclusions, et l'utilisation des données de suivi de tolérance.

La première partie a permis d'évaluer des **schémas permettant de poursuivre l'inclusion dans les essais de phase I en oncopédiatrie**. En effet, les traitements expérimentaux en phase I, qui ont déjà fait la preuve de leur activité chez l'adulte, apparaissent souvent comme une possibilité thérapeutique pour les enfants dont les cancers sont réfractaires aux traitements standards. Le schéma *Rolling 6* a été proposé à cet effet, afin d'éviter qu'un enfant ne puisse pas accéder à l'admission dans un essai du fait qu'une cohorte d'enfants soit en attente d'évaluation. Il permet un processus d'inclusion plus régulier qu'avec le 3+3 ; les données récentes tendent à modérer l'intérêt de ce schéma car de nombreux enfants sont inclus aux doses basses (Sposto and Groshen 2011; Doussau, Asselain et al. 2012). Les récentes données de la littérature montrent donc que ce nouveau schéma n'est pas plus intéressant qu'une méthode MRC adaptée, ou la Tite-CRM, qui permettent de limiter le nombre d'enfants refusés à l'inclusion et tendent à inclure plus d'enfants à la dose cible réelle, tout en augmentant la probabilité d'identifier la bonne dose, sans rallonger la durée des essais.

Dans la suite de notre travail nous nous sommes ensuite consacrés à étudier **l'intérêt de modèles longitudinaux** pour affiner l'analyse finale d'essais de recherche de dose conduits selon un schéma de classique (MRC ou autre). Les études de simulations nous permettent de prétendre à une meilleure identification de la dose cible par ces modèles longitudinaux, et la possibilité de détecter une augmentation du risque de toxicité dans le temps. Les analyses rétrospectives de données d'essais de recherche de dose ont à ce titre illustré l'intérêt du modèle POMM pour détecter un risque de toxicité cumulée dans un essai, et produire une plus fine estimation du risque de toxicité de la RP2D dans l'ensemble des essais, lorsque les patients reçoivent des cycles répétés.

Nous avons également étudié la faisabilité et les caractéristiques opératoires de **deux nouveaux schémas d'étude** dans lesquels l'attribution de dose se base sur les estimations de ces modèles longitudinaux (le modèle mixte ordinal à cotes proportionnelles (Doussau, Thiebaut et al. 2013) et le modèle logistique mixte). L'utilisation de ces modèles complexes est limitée par les difficultés d'estimation sur des petits échantillons. La taille d'échantillon

est d'autant plus réduite lorsque l'on souhaite utiliser ces modèles sophistiqués pour l'attribution de dose tôt dans le déroulement de l'essai.

L'étude du nouveau schéma adaptatif utilisant le modèle POMM a montré qu'une simplification du modèle était nécessaire dans 5 % des attributions de dose. Si on utilise ces schémas adaptatifs, une adaptation de la stratégie de modélisation sera envisageable, et cet aspect ne devrait pas limiter l'implémentation de la méthode ; une perspective d'adaptation des méthodes que nous avons proposées, est de transposer la démarche dans le cadre bayésien. Il faut néanmoins noter que le gain du nouveau schéma adaptatif avec attribution des doses basée sur le modèle POMM est modéré en terme de proportion d'essai recommandant la cible, par rapport à un essai classique CRML qui utiliserait simplement le POMM pour l'analyse des données en fin d'essai. En revanche, le nombre de patients inclus à la dose cible est légèrement augmenté lorsqu'on utilise les toxicités ordinales répétées pour l'attribution des doses. C'est l'intérêt majeur de cette approche.

L'étude sur le modèle logistique pour données binaires répétées complète ces résultats, en quantifiant le gain lié à des modèles plus ou moins sophistiqués et plus ou moins paramétrés. Cette étude a montré par les simulations que le modèle à cotes proportionnelles est supérieur au modèle logistique pour la détection d'un risque cumulé dans le temps. En revanche, pour l'identification de la dose cible, un modèle logistique longitudinal sur données binaires permet déjà d'obtenir un gain substantiel par rapport à l'approche d'identification de la DMT de la CRML (modèle logistique sur le premier cycle). Sur ce critère, le caractère ordinal du critère de jugement apporte finalement peu d'amélioration supplémentaire à celle apportée par les données longitudinales binaires.

Notre travail a montré une amélioration jusqu'à + 10 % ou + 20 % de recommandation de la bonne dose grâce à l'utilisation de l'ensemble des cycles de traitement. Ces résultats sont majeurs en regard des améliorations passées concernant les schémas d'escalade de dose. Néanmoins cela peut sembler encore insatisfaisant, étant donné qu'on ne dépasse pas trois quarts de recommandations correctes, dans les scénarios où la dose cible est la quatrième dose, un scénario très classique.

## 5.2 Précautions d'emploi

Il est difficile d'envisager tous les scénarios possibles, tant le nombre de paramètres définissant la réalité de ces essais est élevée. Il nous semble néanmoins que nos travaux couvrent un large panel de situations cliniquement réalistes pour évaluer de nouveaux schémas d'étude, y compris l'impact de la variation des paramètres ayant le plus d'impact potentiel, comme l'écart aux hypothèses des modèles (hypothèse de proportionnalité des cotes, type de données manquantes), ou différentes relations dose-toxicité.

### 5.2.1 Impact des scénarios cliniques sur les performances des schémas évalués par simulation

A l'aide de nos travaux, nous pouvons refaire un point sur l'impact de la relation dose-toxicité réelle sur les performances de différents schémas avec la MRC. L'impact du nombre de doses à explorer avant la dose maximale tolérée a été étudié dans divers simulations, selon que la DMT est la 3<sup>ème</sup>, 4<sup>ème</sup> ou 5<sup>ème</sup> dose sur 6 (Doussau, Asselain et al. 2012) ou la 2<sup>ème</sup>, 4<sup>ème</sup> ou 6<sup>ème</sup> dose sur 9 (Doussau, Thiebaut et al. 2013), ou la 2<sup>ème</sup>, 4<sup>ème</sup> ou 6<sup>ème</sup> dose sur 6 (§4). Ainsi, les résultats sont d'autant meilleurs que la DMT arrive tôt (2<sup>ème</sup> palier) : environ 65 % d'identification correcte de la dose cible pour la MRC avec une trentaine de patients. Le nombre de paliers au-delà de la dose cible a peu d'impact dans la majorité des cas (que la dose cible soit le 2<sup>ème</sup>, 3<sup>ème</sup> ou 4<sup>ème</sup> palier). En effet, avec ce schéma, la majorité des patients est habituellement incluse dans une fenêtre de +/- 2 paliers autour de la dose cible. D'après les études de simulation que nous avons envisagées, on explore rarement les paliers qui ont une probabilité de toxicité réelle supérieure à 50 % et quasiment jamais les paliers qui ont une probabilité de toxicité réelle supérieure à 70 % (Doussau, Asselain et al. 2012; Doussau, Thiebaut et al. 2013). Ainsi dans le scénario central des simulations évaluant le POMM (dose cible au 4<sup>ème</sup> palier), les paliers de dose 7, 8 et 9 ne sont quasiment jamais explorés.

Néanmoins le nombre total de doses à explorer a un impact sur les performances lorsque la dose cible est une des dernières doses à explorer car les inclusions sont plus concentrées à ces paliers. Par exemple si la dose 6 est la dose cible et que le 6<sup>ème</sup> palier est le dernier à explorer, alors les performances de la MRC sont bien meilleures que s'il y a 9 doses à explorer (inclusion de 13 patients en moyenne à la dose 6, avec 54 % d'identification correcte de la dose 6 lorsqu'il y a 6 doses à explorer, versus 9 patients en moyenne à la dose 6, avec 48 % de recommandations correctes de la dose cible lorsqu'il y a 9 paliers à explorer, Tableau 7).

L'impact du nombre de paliers de dose à explorer sur le schéma adaptatif POMM-MRC peut également être évalué entre les résultats des études de simulation (§3.3.4 et Annexe 8) ; les conclusions sont similaires à celles observées sur la CRML.

Il faut noter dans les cas où la dose cible est très éloignée de la dose de départ que la règle d'arrêt de la MRC a également un impact important. En effet, la MRC avec taille d'échantillon fixe à 20 patient ne permettait pas d'explorer correctement les niveaux de dose supérieurs (Doussau, Asselain et al. 2012) et il faut donc s'assurer d'avoir soit une taille d'étude fixe d'ampleur suffisante, soit une règle d'arrêt basée sur un nombre suffisant de sujets à une dose donnée (par exemple 10 patients à la même dose).

Bien entendu, le nombre de doses à explorer a également un impact sur la durée des essais lorsque la taille d'étude n'est pas fixée.

### 5.2.2 Modèle mixte à *intercept* aléatoire

Les modèles que nous avons implémentés sont des modèles mixtes à *intercept* aléatoire. Nous avons choisi ces modèles notamment du fait de leur propriété de robustesse pour produire des estimations non biaisées dans le cas de données manquantes de type « au hasard » (*missing at random*, MAR). En effet, dans ces essais, le processus de génération de données manquantes correspond à l'arrêt de traitement, qui a lieu soit pour toxicité majeure, soit pour progression ; dans ce cadre clinique particulier, il est extrêmement improbable que les patients soient perdus de vue, et il n'est donc par raisonnable de considérer ces données comme manquantes complètement au hasard (*missing completely at random*, MCAR).

Nous avons étudié les propriétés du modèle POMM (Doussau, Thiebaut et al. 2013) et du nouveau schéma adaptatif POMM-CRM dans un scénario de données manquantes au hasard, où les données étaient manquantes après la première toxicité grave du fait d'une sortie d'étude. Nous avons aussi étudié, le comportement du modèle en présence de données manquantes complètement au hasard lorsque les patients sortaient d'étude indépendamment de la dose. Les résultats sur données complètes, où les patients recevaient tous six cycles, ont également été évalués. Cela nous a permis de réaliser une analyse de sensibilité pour mesurer la robustesse des méthodes face à la présence de données manquantes.

Il est important de se rappeler l'interprétation des paramètres du modèle : l'*intercept* fixe mesure la probabilité moyenne s'il n'y a pas de variables explicatives. L'*intercept* aléatoire mesure la variabilité interindividuelle. Lorsque l'on prend en compte l'*intercept* fixe, et les

variables explicatives, l'intégration sur la distribution de l'*intercept* aléatoire donne la probabilité marginale de toxicité.

Lorsque la valeur de l'*intercept* aléatoire d'un patient donné est connue (par exemple en cours d'essai, estimée grâce à l'information des cycles précédents), cela pourrait permettre une adaptation individuelle de la dose pour atteindre la cible, selon une approche similaire à celle de (Legedza and Ibrahim 2000), où une adaptation des doses pour chaque patient en fonction des observations précédentes est réalisée.

Dans les simulations du nouveau schéma adaptatif qui utilise le modèle à cotes proportionnelles, nous avons utilisé une valeur nulle de l'effet aléatoire pour choisir la dose à attribuer à chaque nouveau patient à inclure. Lorsque la variabilité interpatient du risque de toxicité, non expliquée par la dose, est importante, la probabilité de toxicité selon l'approche d'un patient « moyen » ( $u_i = 0$ ) tendra à recommander une dose supérieure à la dose qui serait recommandée par les probabilités marginales : par exemple si la variance de l'effet aléatoire est de 4, la première approche indique que la dose cible est la 4<sup>ème</sup> dose, alors que la seconde indique que la dose cible est la 3<sup>ème</sup> dose (Figure 11).

Ces deux estimations sont différentes : la probabilité de toxicité pour un patient pour lequel la valeur d'*intercept* aléatoire est  $u_i = 0$ , est la probabilité de toxicité qu'il est le plus probable d'observer pour un patient tiré au hasard. C'est pour cela que nous l'avons choisie pour l'attribution des doses. Elle est équivalente à la valeur médiane des probabilités de toxicité d'une population. La probabilité marginale sera plus appropriée dans l'analyse finale de l'essai de phase I, pour évaluer le risque de toxicité auquel on devra faire face en phase II.

### 5.2.3 Modèle logistique à cotes proportionnelles

Pour l'utilisation du modèle à cotes proportionnelles, l'hypothèse sous-jacente doit également être gardée en mémoire. Cette hypothèse permet d'utiliser l'information apportée par les différents grades de toxicités, en absorbant le caractère ordonné uniquement par le biais des *intercept* supplémentaires à estimer. Dans les cas où cette hypothèse semble trop forte, un modèle logistique ordonné généralisé peut être utilisé. Nous n'avons pas étudié les propriétés opératoires du modèle ordonné généralisé sur des données de si petite taille, mais, considérant l'impact de l'estimation d'un paramètre supplémentaire sur les performances de ces méthodes échantillonnant principalement à une dose unique, on attend des performances dégradées. Par ailleurs, Van Meter et al. ont étudié le modèle POM et le modèle des ratios de continuations

(CR), plus complexe. Ce schéma (CR-CRM) nécessite plus d'information et n'a pas montré de gain supplémentaire. Les auteurs n'ont pas trouvé d'arguments pour préférer le modèle plus complexe (Van Meter, Garrett-Mayer et al. 2011; Van Meter, Garrett-Mayer et al. 2012).

Pour éviter cette hypothèse, un autre choix possible est de simplifier le modèle et utiliser le modèle logistique en dichotomisant le critère d'intérêt, puisque les performances restent relativement bonnes pour l'identification de la dose cible. Le modèle logistique est néanmoins moins puissant pour étudier la variation du risque de toxicité dans le temps.

#### 5.2.4 Codage des variables explicatives

Le codage des variables de la dose et du cycle doivent également être pris en compte.

##### A) La dose

Dans l'ensemble de ce travail une relation dose-toxicité croissante était supposée, et de plus la relation entre la dose et la probabilité de toxicité était log-linéaire. L'hypothèse de log-linéarité est difficile à tester sur ces petits jeux de données, mais les modèles simplifiés dans lesquels un seul paramètre est estimé montrent que les résultats de la CRML sont robustes au codage de la dose (Paoletti and Kramar 2009). Dans le contexte des essais de recherche de dose sur taille d'étude restreinte, il est raisonnable de choisir le modèle d'avantage pour ses bonnes caractéristiques opératoires, permettant d'identifier correctement la dose en minimisant le risque global de toxicité, plutôt que de choisir un modèle pour sa vraisemblance biologique. On peut donc supposer que les performances des nouvelles méthodes que nous avons développées sont robustes, tant que l'on reste sous l'hypothèse d'une relation dose-toxicité croissante.

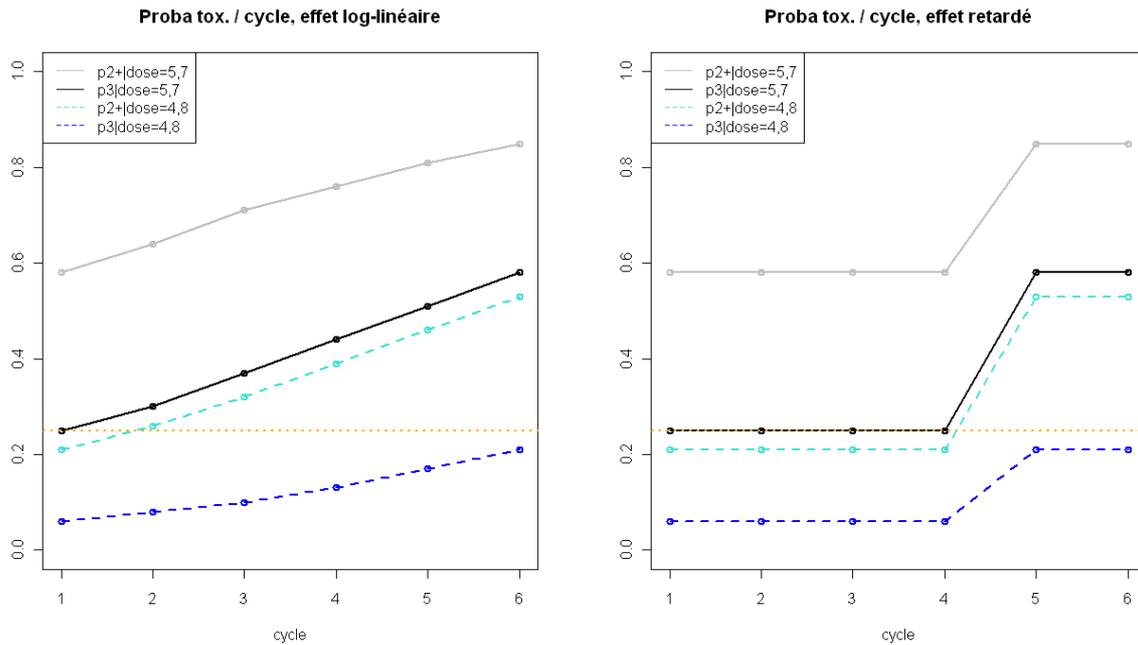
Nous avons également considéré que la dose était constante au cours du temps. Il s'agit de l'approche la plus fréquemment utilisée où l'on cherche, dans la mesure du possible, à ne pas modifier la dose pour pouvoir évaluer le profil de toxicité de chaque dose séparément (Le Tourneau, Lee et al. 2009). En situation réelle, les schémas dans lesquels une augmentation de la dose a été proposée, type *Accelerated titration design* (Simon, Freidlin et al. 1997), ont été rarement implémentés ; ces schémas posent de nombreux problèmes pour l'analyse et les agences réglementaires ne les recommandent pas : “*when all patients move through a series of rising doses (forced titration study), which is similar in concept and limitations to a randomized multiple cross-over dose-response study [...] the study design cannot distinguish*

*response to increased dose from response to increased time on drug therapy or a cumulative drug dosage effect*” (European Medicines Agency 1994).

En revanche, dans la pratique clinique de ces essais, il n’est pas rare que la dose du traitement soit diminuée en cas d’événement indésirable, et des interruptions temporaires de traitement ont également lieu, ayant pour conséquence une diminution de la dose-intensité reçue. Dans notre approche, le fait d’intégrer dans le modèle la dose initiale attribuée est, en quelque sorte, une démarche similaire à celle de l’analyse « en intention de traiter » utilisée dans les essais randomisés, c’est-à-dire que l’estimation du profil de toxicité est celle de la dose attribuée lors de l’inclusion du patient, quels que soient les changements ultérieurs. Cependant, dans ces essais précoces où la relation dose-effet est à évaluer, ce raisonnement est moins légitime que dans des essais de phase III. En pratique, le modèle longitudinal peut intégrer la dose courante qui pourrait varier d’un cycle à l’autre (augmentation ou diminution). Dans ce cas, si l’on utilisait le modèle développé dans notre approche (modèle 7), cela nécessiterait de supposer qu’une toxicité n’est liée qu’à la dose courante reçue, sans effet des doses précédentes ; le modèle de Simon et al. (Simon, Freidlin et al. 1997) permet en revanche de prendre en compte différemment la dose courante et les doses cumulées reçues et serait alors plus adapté.

## **B) Le temps**

La forme de la relation entre la probabilité de toxicité et le temps peut également être discutée. Dans l’étude de simulation évaluant le POMM, nous avons généré des données dans lesquelles le risque de toxicité est constant avec le temps et des données dans lesquelles le risque de toxicité augmente de manière log-linéaire avec le temps. C’est le scénario le plus probable si des toxicités cumulées survenaient : la fréquence des toxicités augmente progressivement avec le nombre de cycles reçus. En effet, dans notre modèle le plus complet (modèle 7, incluant un *intercept* aléatoire, un paramètre pour la dose, et un paramètre pour le temps), le temps depuis inclusion est proportionnel à la dose cumulée théorique reçue, une fois que la dose initiale attribuée a déjà été prise en compte dans le modèle. Pourtant si le mécanisme sous jacent est plutôt celui de toxicités retardées, alors le modèle recherchant un effet log-linéaire du temps serait moins puissant. Le modèle 7 que nous avons utilisé dans l’étude du POMM est donc plus puissant pour détecter les effets cumulés que les effets tardifs. La différence de forme de ces deux scénarios temporels est représentée pour deux doses (Figure 14).



**Figure 14. Représentation de deux types de variation du risque de toxicité dans le temps : effet cumulatif log-linéaire (gauche) et effet retardé (droite).** Les courbes de gauche sont générées avec les paramètres  $\theta = [\alpha_1 = 9,85, \alpha_2 = 11,29, \beta_1 = 1,78, \beta_2 = 0,29]$ , pour  $u_1 = 0$ .

La question de la détection d'un effet cumulé est cruciale car cela impactera de manière majeure le développement futur du traitement : la dose doit être diminuée ou le schéma d'administration doit être modifié pour que le traitement reste tolérable au long cours.

### 5.2.5 Définition de la cible de toxicité dans les modèles longitudinaux

Le choix d'une cible à 25 % ou 20 % est communément admis dans les essais de recherche de dose en oncologie. Dans la première étude, nous avons pris la cible de toxicité de l'étude erlotinib + radiothérapie (20 %). Dans la seconde étude, la cible de toxicité a été fixée à 25 %, car nous souhaitons utiliser la relation dose-toxicité d'une étude antérieure (Zhou and Whitehead 2003), et que la cible de 20 % de toxicité se trouvait à mi chemin entre 2 doses (3<sup>ème</sup> et 4<sup>ème</sup> doses), ce qui aurait complexifié l'interprétation du critère de jugement « proportion d'essais recommandant la bonne dose », puisque les deux doses auraient été aussi légitimes pour être la cible réelle de toxicité à identifier.

Dans l'approche longitudinale, nous avons considéré que le taux de toxicité cible observé sur le premier cycle correspondrait également à la cible de toxicité dans l'approche par cycle.

Cela pourrait néanmoins être discuté. Il est considéré tolérable que les patients présentent 20 à 25 % de toxicité grave au premier cycle dans les essais oncologiques, néanmoins une probabilité de 20 à 25 % de toxicité par cycle pourrait ne pas sembler acceptable pour autant. On pourrait par exemple imaginer que l'on souhaite finalement recommander une dose qui entraîne moins de 5, 10 ou 15 % de toxicité grave par cycle, ou encore une dose qui entraîne moins de 15, 25 ou 30 % de toxicité modérée ou grave par cycle.

Cette différence d'approche peut être illustrée par l'analyse rétrospective de l'essai erlotinib + radiothérapie (Georger, Hargrave et al. 2011). La probabilité de toxicité grave à la dose recommandée était de 16 % sur le premier cycle, de 9 % en moyenne par cycle de traitement, et le risque cumulé d'avoir au moins une toxicité grave après six cycles était estimé à 37 % (Tableau 8). Ceci montre qu'un taux de toxicité inférieur à la cible  $\tau$  au premier cycle ne garantit pas que le traitement soit réellement tolérable après plusieurs cycles.

Par ailleurs la dose à attribuer à chaque nouveau patient dans notre approche était celle dont la probabilité de toxicité était la plus proche de la cible de toxicité en valeur absolue. Actuellement le schéma EWOC pour lequel l'attribution de dose cherche à limiter la probabilité de recevoir une dose qui excède la cible est de plus en plus utilisé. Ce schéma qui a déjà été adapté aux données ordinales sur le premier cycle (POM-EWOC (Tighiouart, Cook-Weins et al. 2012)) pourrait être modifié pour prendre en compte les données répétées de toxicité dans un contexte Bayésien, dans un nouveau schéma : POMM-EWOC.

### **5.2.6 Définition du critère de jugement principal**

Dans les approches que nous avons développées, le critère de jugement modélisé reflète le grade maximal de toxicité à chaque cycle par patient. Cette approche est particulièrement adaptée pour la modélisation des toxicités réversibles : tolérance digestive, cutanée, hématologique...

Le critère de jugement étant le grade maximal de toxicité sur chaque cycle, notre approche ne prend néanmoins pas en compte l'ensemble des toxicités sur le cycle, ni le type clinique de toxicité. Des extensions seraient possibles en utilisant des variables restreintes à certains types de toxicité par organe. Cela pourrait permettre de mieux détecter des signaux spécifiques. Par exemple, dans l'essai erlotinib + radiothérapie (Georger, Hargrave et al. 2011), des effets précoces à type d'éruption sont survenus ; des effets plus tardifs, type panaris, sont également survenus. Dans ce cas, le risque de toxicité modélisé était stable dans le temps, alors qu'il se

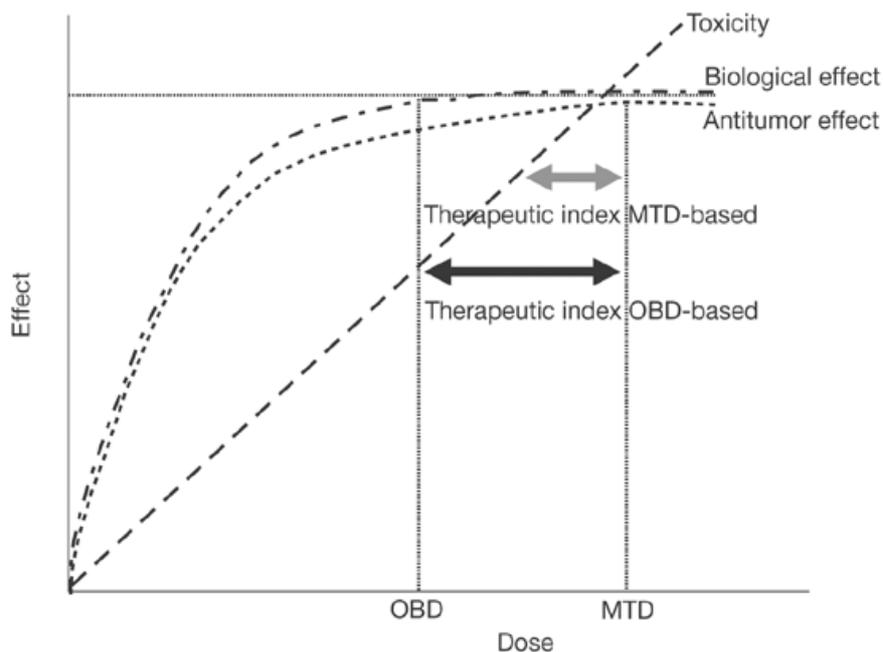
pourrait qu'en réalité deux processus cliniques distincts se superposent. Le fait de raffiner le critère de jugement à des types de toxicité pourrait permettre d'augmenter la spécificité du signal recherché ; néanmoins, on risque de manquer de puissance pour mettre en évidence des phénomènes rares, en particulier dans le contexte de ces essais où les informations sont pauvres au départ.

Un score longitudinal représentant la somme des grades de toxicités expérimentées à chaque cycle pourrait également être calculé. Ce score pourrait être analysé par un modèle longitudinal POMM ou un modèle linéaire généralisé, et il est possible qu'il soit plus puissant que le modèle POMM du grade maximal par cycle pour détecter l'évolution de la toxicité au cours du temps. En revanche, l'identification d'une dose cible par la modélisation de ce score, nécessiterait de définir une interprétation clinique et un seuil cible pour identifier la dose à recommander, ce qui est une limite majeure à la mise en œuvre de ce type d'approche. Il nous semble que le critère de jugement que nous avons utilisé (grade maximal à chaque cycle) est une grandeur clinique plus concrète à appréhender, et que les méthodes que nous avons proposées seront plus faciles à implémenter que celles basées sur des scores de toxicité (Yuan, Chappell et al. 2007; Lee, Hershman et al. 2012; Ezzalfani, Zohar et al. 2013).

## 5.3 Mise en perspective

### 5.3.1 Impact à l'ère des thérapies moléculaires ciblées

Dans le domaine de l'oncologie, l'ensemble du plan de développement des médicaments a été forgé sur mesure pour les traitements cytotoxiques. Malgré les différences majeures entre les thérapies moléculaires ciblées et les traitements cytotoxiques concernant le profil de toxicité, le mode d'administration et les durées de traitement, l'identification de la dose maximale tolérée reste actuellement l'objectif principal de la majorité des essais de phase I de thérapies moléculaires ciblées (Parulekar and Eisenhauer 2004), bien que 25 % de ces essais ne l'atteignent pas. On suspecte qu'avec ces traitements, la dose biologique optimale pourrait ne pas être la dose maximale tolérée (Figure 15). L'index thérapeutique pourrait être plus large qu'avec les traitements cytotoxiques et la dose identifiée comme dose maximale tolérée serait supérieure à la dose biologique minimale efficace.



**Figure 15. Schéma de l'effet de la dose sur la cible biologique, l'effet antitumoral, et de la relation dose-toxicité, supposé pour une thérapie ciblée associée à la radiothérapie.** OBD = optimal biologic dose, MTD = maximum tolerated dose. Selon (Deutsch, Soria et al. 2005).

Dans cette hypothèse, on peut concevoir plusieurs scénarios :

- Soit il reste nécessaire d'identifier en premier lieu la dose maximale tolérée. Dans ce cas, les schémas MRC restent les meilleurs connus à ce jour pour l'escalade de dose.
- Soit on considère qu'il faut d'emblée étudier conjointement l'efficacité et la tolérance afin de sélectionner une dose efficace lors de l'escalade de dose.

Dans la première hypothèse, si l'objectif reste d'étudier uniquement la tolérance en priorité, on peut néanmoins rediscuter certains aspects des essais de recherche de dose. En effet, puisque certains de ces traitements peuvent être peu toxiques (Booth, Calvert et al. 2008) et que l'objectif de l'essai de phase I est parfois uniquement d'étudier la tolérance, on pourrait envisager des essais sur volontaires sains, au lieu de patients atteints de cancer, comme cela est envisagé dans les essais pilotes dits de « phase 0 » (Marchetti and Schellens 2007). Cela permettrait de lever en partie la restriction liée à la taille d'étude, et résoudrait en partie le problème éthiques des essais sur sujet malade avec une probabilité de bénéfice faible. Dans le contexte d'essais sur des traitements moins toxiques que les chimiothérapies, on pourrait également plus facilement envisager des schémas d'étude qui combinent la sécurité des patients, et la recherche d'une efficacité basée sur des critères d'optimalité de l'information (Bartroff and Leung Lai 2011), ou des essais randomisés dès la phase I.

Actuellement, il est relativement admis pour les thérapies moléculaires ciblées, que les schémas qui sont basés sur l'hypothèse d'une relation dose-toxicité croissante semblent raisonnables pour vérifier, par escalade de dose, que la dose sélectionnée n'est pas trop toxique (Korn, Arbuck et al. 2001; Booth, Calvert et al. 2008). Il s'agit d'une étape préliminaire qui semble indispensable pour poursuivre le développement d'une nouvelle molécule.

La seconde hypothèse renvoie directement à la question de l'intégration de critères de jugement d'efficacité dans les essais de recherche de dose : biomarqueurs lorsqu'ils sont disponibles et validés pour mesurer un effet thérapeutique, évolution de la taille tumorale, paramètres pharmacocinétiques ou pharmacodynamiques (Parulekar and Eisenhauer 2004; Goodwin, Giaccone et al. 2012). L'intégration de critères conjoints de tolérance et d'efficacité est un champ de recherche à part entière (Thall and Cook 2006; Zhang, Sargent et al. 2006;

Mandrekar, Cui et al. 2007), qui sera probablement la piste majeure pour le développement futur de ces traitements.

La recherche de dose associant deux molécules est également un champ de développement, très important à l'ère des thérapies ciblées qui sont de plus en plus associées entre elles (Hamberg, Ratain et al. 2010). Il s'agit d'un champ méthodologique complexe (Wang and Ivanova 2005; Mandrekar, Cui et al. 2007; Yin and Yuan 2009; Wages, Conaway et al. 2011). Ces essais nécessitent des modèles capables de prendre en compte l'interaction potentielle entre les deux traitements étudiés, comme par exemple les modèles à copula. Il pourrait être intéressant d'intégrer le modèle POMM dans ces stratégies, comme nous l'avons fait de manière simplifiée en analysant rétrospectivement les données de l'essai Cacy.

Les essais de recherche de dose à l'ère des thérapies moléculaires ciblées sont déjà en pleine mutation. Actuellement, l'inclusion de cohortes d'expansion est déjà de plus en plus répandue : de 12 % en 2006 à 38 % en 2011 des essais sur médicaments en monothérapie (Manji, Brana et al. 2013). Les objectifs sont variés et parfois non clairement spécifiés. Lorsque l'objectif de la cohorte d'expansion était de poursuivre l'évaluation de la tolérance, la dose recommandée à la fin de l'essai était différente de celle identifiée à la fin de l'escalade de dose dans 13% des essais. Il semble donc essentiel d'utiliser les modèles les plus efficaces aux moments clef de prise de décision des essais pour utiliser toute l'information disponible aussi tôt que possible, et permettre de concentrer les patients autour des doses cibles de l'essai. Il nous semble que les approches que nous avons développées ont toute leur place ici.

Actuellement les patients traités par thérapies moléculaires ciblées peuvent être traités au long cours pendant plusieurs années, en cas de stabilisation de la maladie ou de rémission. La prise en charge de ces patients tend à se rapprocher de celle d'autres maladies chroniques. Dans de nombreuses maladies chroniques, des marqueurs longitudinaux sont modélisés pour suivre l'évolution de l'état des patients. On peut citer l'exemple du suivi des fonctions cognitives ou de l'autonomie dans la démence. Par exemple, un modèle POMM multivariable a été utilisé pour étudier les facteurs associés à l'évolution d'une échelle d'autonomie dans une étude de cohorte gériatrique EPIDOS (Carriere and Bouyer 2006). Un autre type de modèle ordinal longitudinal a été utilisé pour modéliser le déclin cognitif la cohorte gériatrique PAQUID (Jacqmin-Gadda, Proust-Lima et al. 2010). Notre approche prenant en compte les données

longitudinales est donc très intéressante dans le contexte des thérapies moléculaires ciblées, puisque ces traitements sont administrés parfois sur de très longues périodes. En dehors du cadre des essais de phase I, on pourrait donc utiliser le modèle POMM pour étudier la tolérance des thérapies ciblées au long cours, et la survenue d'effets éventuels cumulés retardés (néphrotoxicité, cardiotoxicité...).

### **5.3.2 Validité des données longitudinales**

Les méthodes qui utilisent les données collectées tout au long de l'essai doivent inclure des données dont la validité puisse être assurée. Pour les toxicités graves, le recueil est très attentif du fait des obligations réglementaires ; l'imputabilité aux traitements de la recherche est néanmoins parfois difficile. Par exemple dans l'analyse rétrospective de l'essai erlotinib, une chirurgie pour un kyste avait été notifiée comme toxicité grave possiblement liée au traitement à l'étude, de grade 3 ; pourtant lors de la réanalyse des données, il n'était pas si évident de considérer cet événement comme cliniquement pertinent.

La modélisation des données longitudinales nécessite donc que l'imputabilité de l'ensemble des événements indésirables soit effectuée, quel que soit le grade, de manière aussi attentive que cela est actuellement fait sur la période d'évaluation classique des TDL, et de la manière la plus reproductible possible. Ce point a une importance majeure dans les schémas proposés dans le cadre de ce travail, qui utilisent les données longitudinales pour l'attribution de la dose. Par exemple, une toxicité qui ne serait pas attribuée au traitement à tort, pourrait entraîner l'attribution d'une dose trop élevée pour un nouveau patient, du fait de l'erreur de classement. L'imputabilité pour les toxicités modérées pourrait être encore plus difficile que pour les toxicités graves, source de variabilité interinvestigateur et par conséquent source éventuelle de biais de mesure. Du fait de la fréquence importante des toxicités modérées, il pourrait être intéressant d'évaluer par simulation l'impact formel de ce type d'erreur sur l'attribution des doses et les recommandations finale du nouveau schéma POMM-CRM, comme cela a été fait pour la MRC par rapport au « 3+3 » (Iasonos, Gounder et al. 2012).

D'un point de vue pratique, la validité et la disponibilité de ces données pour une modélisation adaptative en temps réel impliquerait une forte réactivité de tous les intervenants sur le circuit des données : investigateurs, attachés de recherche clinique, promoteur, centre de méthodologie et de gestion des données.

### 5.3.3 Amélioration des schémas et méthodes d'analyse

Bien que de nombreux travaux méthodologiques aient étudié les propriétés des schémas basés sur un modèle et leur supériorité pour identifier une dose cible dans la grande majorité des scénarios plausibles, on observe une relative inertie à leur mise en œuvre (Rogatko, Schoeneck et al. 2007; Iasonos and O'Quigley 2011). La fréquence d'utilisation tend à augmenter progressivement : d'après une revue des essais publiés entre 1991 et 2006, 1,6 % des essais utilisaient un schéma basé sur un modèle, et les délais de mise en œuvre évalués entre l'année de publication du schéma d'étude et l'année de publication de l'essai étaient de l'ordre de 3 à 10 ans (Rogatko, Schoeneck et al. 2007) ; une revue de la littérature des essais sur les thérapies ciblées en monothérapie, publiés entre 2000 et 2010, avait trouvé 5 % d'essais menés selon un schéma guidé par un modèle (Le Tourneau, Razak et al. 2011).

Les schémas MRC sont pourtant tout à fait adaptés pour répondre aux recommandations de l'Agence européenne du médicament concernant les essais en première administration chez l'homme (Committee for medicinal products for human use 2007). Les agences réglementaires reconnaissant actuellement que les schémas basés sur un modèle statistique doivent être préférés (Porcher, Lecocq et al. 2011).

Le développement de nouveaux schémas sur mesure est difficile, du fait que la conception d'un nouveau schéma d'étude puisse nécessiter la mobilisation d'une équipe pendant plusieurs mois à années pour effectuer les études de simulations nécessaires à l'évaluation des propriétés opératoires dans les conditions d'essais séquentiels sur petits échantillons avant de débiter l'essai. Ces délais peuvent ne pas convenir aux investigateurs ou promoteurs pour la mise en place d'un essai précoce.

Pour favoriser l'aspect translationnel, de l'étude de simulation à la mise en œuvre de l'essai, il est essentiel :

- de constituer une communauté de méthodologistes experts qui puisse proposer les schémas d'étude et les méthodes d'analyse statistique les plus adaptées à chaque cas particulier d'essai,
- de communiquer avec les investigateurs afin qu'ils comprennent les enjeux scientifiques de la modernisation des schémas d'étude de recherche de dose, par rapport aux schémas historiques, qui ne devraient plus être utilisés (Jaki, Clive et al. 2013),

- que les instances régulatrices (Agence nationale de sécurité du médicament et des produits de santé, Comité de protection des personnes) favorisent l'implémentation de ces méthodes adaptatives, à présent bien établies, et refusent la mise en place d'essais selon des méthodes inadaptées. Cela nécessite de leur part une expertise méthodologique de pointe dans ce domaine.

Si l'on se base sur le principe que la communauté scientifique considère les essais d'escalade de dose chez des patients atteints de cancer comme une bonne pratique clinique pour la première étape d'investigation chez l'homme, alors, le fait d'utiliser les schémas les plus efficaces fait partie de la conduite éthique de ces essais. Ces essais méritent en effet les mêmes standards d'éthique et d'excellence scientifique que ceux exigés dans les essais de phase ultérieure, comme c'est communément admis dans les essais de phase III (Iasonos and O'Quigley 2011).

#### **5.3.4 Amélioration de la qualité des publications**

L'amélioration de ces essais repose également sur une amélioration de la qualité des publications. Une liste spécifique à l'évaluation de la qualité des publications d'essais de recherche de dose de traitements cytotoxiques anticancéreux a été proposée en 2008 (Zohar, Lian et al. 2008). Il serait intéressant de reprendre cette grille pour l'adapter, en y intégrant les recommandations du TARGET-DLT group (Postel-Vinay, Rizzo et al. 2013), qui nécessiterait donc au minimum d'intégrer des items comme les toxicités graves survenues au delà de la fenêtre classique d'évaluation des TDL, ou les toxicités modérées qui entraînent une diminution de la dose (dose administrée inférieure à 75 % de la dose prévue).

#### **5.3.5 Essais oncologiques : des spécificités méthodologiques à remettre en cause ?**

Les traitements développés en oncologie ont un fort taux d'attrition en phase III comparativement aux autres spécialités médicales : parmi les dix plus grandes compagnies pharmaceutiques mondiales sur la période 1991-2000, le taux de succès global (depuis la première administration à l'homme jusqu'à l'enregistrement) était estimé à 11 % toutes spécialités confondues, et 5 % en oncologie. Le taux de succès en phase III était de 40 % en oncologie contre 55 % environ pour les autres spécialités (Kola and Landis 2004).

Cela pourrait être expliqué en partie par différentes spécificités méthodologiques des essais de cancérologie par rapport aux essais d'autres spécialités, comme l'identification de la DMT en phase I et la réalisation d'essais de phase II non randomisés à la DMT en phase II. Dans les autres disciplines, la dose optimale est le plus souvent identifiée en phase II ou III dans des essais randomisés, parfois en incorporant l'étude de critères de jugement intermédiaires ou de biomarqueurs (Ratain, Humphrey et al. 2008). La présence de biomarqueurs présents sur le chemin causal de la maladie donne en effet un niveau de preuve plausible en faveur de l'efficacité détectable très précocement dans le développement d'un médicament (par exemple la charge virale du VIH). Dans le domaine du cancer, il existe encore peu de biomarqueurs ou de critères de substitution validés.

Différentes raisons de ces spécificités de la recherche clinique en oncologie sont évoquées : le fait que les centres anticancéreux sont souvent isolés des centres hospitaliers universitaires, la rareté des traitements et cibles disponibles par le passé, le dogme traditionnel en oncologie que le plus est le mieux (« *more is better* »), la nature dose dépendante et réversible des toxicités liées aux agents cytotoxiques, l'acceptabilité des toxicités aussi bien par les oncologues que par les patients atteints de cancer, la conviction qu'un taux de réponse tumorale est une information suffisante pour guider le développement futur d'une molécule, l'acceptation de design de phase II qui identifient des agents inefficaces sans garantir le succès de la phase III et enfin la réticence à l'utilisation de placebo même lorsque cela serait envisageable. Dans le domaine de la polyarthrite rhumatoïde, maladie inflammatoire chronique possédant certaines cibles communes avec certains cancers, la *Federal Drug Administration* (FDA) des Etats-Unis d'Amérique recommande la sélection de la dose en phase II, par des essais randomisés en groupe parallèle (Ratain, Humphrey et al. 2008). Cela illustre bien la différence d'approche liée à la discipline oncologique.

Hors cancérologie, des types de schémas différents sont plus facilement acceptés comme des essais randomisés de phase II/III avec plusieurs doses en phase II dont certaines seulement sont maintenues pour la phase III. Par exemple ce type d'essai a été utilisé pour évaluer un bronchodilatateur dans la broncho-pneumopathie obstructive (Donohue, Fogarty et al. 2010), et cet exemple a été discuté dans une table ronde réunissant des investigateurs, des méthodologistes, des représentants des agences réglementaires et des industriels (Porcher, Lecocq et al. 2011). Peu d'essais de ce type ont eu lieu dans le domaine du cancer, mais ce type d'approche se développe (Parmar, Barthel et al. 2008). Par exemple un essai sur l'hormonothérapie dans le cancer de la prostate (*Systemic Therapy in Advancing or Metastatic*

*Prostate Cancer: Evaluation of Drug Efficacy*, STAMPEDE (Sydes, Parmar et al. 2012)) est actuellement en cours, selon un schéma randomisé multibras-multiétape. Depuis son ouverture, des bras ont été arrêtés et d'autres ajoutés suite à des analyses intermédiaires, avec une bonne acceptabilité des agences réglementaires. Un autre essai de phase II/III est en cours étudiant plusieurs doses de radiothérapie dans le cancer du poumon à petites cellules (Korn, Freidlin et al. 2012). Il est probable que ce type d'essais évaluant plusieurs doses en phase II, menés dans des équipes de pointe du fait de la complexité méthodologique et des tailles d'étude nécessaires, contribueront à améliorer le plan de développement des médicaments en oncologie.

### **5.3.6 Rôle du biostatisticien**

Lorsque l'on aborde les aspects éthiques des essais de recherche de dose, il est intéressant de décomposer le point de vue des différents intervenants, analyser la balance bénéfico-risque, ainsi que les éventuels conflits d'intérêt. Ainsi (Anderson and Kimmelman 2012) ont proposé une analyse du point de vue de l'autorité compétente, du Comité éthique de l'hôpital, de l'investigateur, du médecin traitant, et du patient participant à l'essai. Le biostatisticien n'apparaît pas dans cette décomposition des intervenants.

Pourtant, son rôle est crucial dans les essais de recherche de dose à l'ère des schémas adaptatifs, puisque c'est le schéma d'étude proposé qui déterminera, avec les données collectées sur les premiers patients inclus, la dose qui sera reçue par les différents patients inclus. Dans les essais de recherche de dose guidés par un modèle, le statisticien prend donc un rôle très actif dans l'investigation, par rapport à ce qui se passe, par exemple dans un essai randomisé de phase III. Pour le biostatisticien, il me semble que les enjeux sont donc de plusieurs ordres :

- bien connaître les modèles proposés et les avantages et inconvénients, pour proposer le meilleur schéma d'essai, et le meilleur plan d'analyse statistique à chaque étape cruciale de l'essai,
- être en mesure d'effectuer des études de simulation lors de la planification afin d'envisager avec les investigateurs les performances espérées en fonction de divers scénarios,
- veiller à la prudence des schémas mis en œuvre pour garantir la sécurité des patients, tout en essayant d'optimiser au maximum la validité des résultats,

- participer à l'éducation des investigateurs concernant l'incertitude de résultats obtenus sur de petits échantillons, et la puissance pour tester une hypothèse donnée dans ces conditions.
- savoir remettre en question les schémas d'étude communément admis, en tentant d'enrichir la pratique avec les avancées méthodologiques d'autres disciplines cliniques, et en revoyant cette phase d'essais dans l'ensemble du processus de développement des nouveaux médicaments.

Pour que la participation des patients dans les essais garantisse les critères éthiques de la recherche biomédicale, il convient que les investigateurs puissent informer correctement les patients, et, du fait du schéma adaptatif, cela nécessite donc une excellente interaction biostatisticien-investigateur.

## 5.4 Recommandations pour la pratique

### 5.4.1 Indication des méthodes développées

- **Analyse finale (ou intermédiaire) des essais cliniques de recherche de dose en phase I**

L'utilisation de modèles utilisant l'ensemble des cycles de traitement administrés est indispensable et devrait être systématique. La stratégie recommandée est de chercher une variation du risque de toxicité avec les cycles, puis, d'estimer la probabilité de toxicité par cycle de chaque dose. On peut utiliser un modèle logistique longitudinal ou un modèle mixte à cotes proportionnelles, si l'on souhaite analyser également les toxicités modérées. Le modèle mixte à cotes proportionnelles est plus puissant pour détecter un effet temporel du risque de toxicité.

Ces modèles devraient être systématiquement envisagés en fin d'étude, mais aussi lors des étapes clef de l'essai : comité indépendant de surveillance, choix de la dose pour une cohorte d'expansion... Le rationnel est simplement d'utiliser l'ensemble des informations disponibles et nous avons montré son intérêt.

Comme cela a été illustré dans l'analyse rétrospective de 3 essais, lorsque la majorité des patients interrompent leur traitement très rapidement, après 1 à 2 cycles, les modèles pour données longitudinales sont de moindre intérêt, et le modèle logistique pondéré pour estimer une probabilité cumulée sur six cycles n'est pas applicable ; c'est le cas de l'essai aviscumine (Schoffski, Riggert et al. 2004).

- **Nouveaux schéma adaptatifs utilisant les modèles longitudinaux pour l'attribution de dose**

L'utilisation de ces nouveaux schémas utilisant des modèles longitudinaux est encouragée, si l'estimation du risque de toxicité par cycle est un critère de jugement principal scientifiquement pertinent. L'attribution des doses peut se baser sur les estimations d'un modèle logistique mixte (logMM-CRM) ou d'un modèle POMM (POMM-CRM), selon la cible de toxicité pertinente. Cette mise en œuvre a des implications pratiques plus lourdes qu'une MRC classique : validité des données collectées, et imputabilité de tous les cycles en temps réel du côté investigation, et maîtrise des techniques de modélisation dans l'équipe méthodologique. La modélisation peut nécessiter une simplification des modèles. La mise en

œuvre réelle de ces nouveaux schémas dans de futurs essais prospectifs apportera un complément d'information concernant leur applicabilité.

Nous avons développé l'intérêt éventuel d'utiliser un modèle logistique pondéré pour estimer la probabilité cumulée de toxicité grave sur 6 cycles (approche TITE). Cela peut soulever la question de l'intérêt d'un schéma adaptatif TITE-CRM (Cheung and Chappell 2000; Braun 2006) pour mener l'essai, par rapport aux approches logMM-CRM et POMM-CRM que nous avons développées. Il est probablement plus pertinent d'utiliser l'approche POMM-CRM ou logMM-CRM lorsque l'on attend des toxicités réversibles, et l'approche TITE-CRM lorsque l'on anticipe des toxicités irréversibles pour lesquelles la survenue d'un événement par patient au moins est vraiment cliniquement un problème majeur. Ces schémas d'étude (TITE-CRM *versus* POMM-CRM ou logMM-CRM) pourraient être comparées par simulations, mais il nous semble qu'il faut plus les considérer comme deux approches complémentaires adaptées à des contextes cliniques différents, que de les considérer comme des approches concurrentes.

#### 5.4.2 Pistes d'amélioration à envisager

Notre travail montre que l'amélioration des schémas d'étude d'escalade de dose par utilisation de modèles statistiques longitudinaux sophistiqués sur une taille d'étude d'une trentaine de patient ne permet pas d'améliorer l'identification de la dose cible au delà de 70 % des essais. Il est peu probable que l'on parvienne à améliorer ce taux par des techniques d'analyse de données spécifiques tant que le critère d'évaluation sera binaire. Cela veut dire que les améliorations proviendront de modifications plus profondes des schémas.

Trois pistes pragmatiques nous semblent devoir être principalement envisagées pour améliorer cela :

- la conduite de plusieurs essais de phase I en parallèle, qui permettrait de faire une évaluation méta-analytique,
- la mise en place d'essais de phase II avec randomisation des doses adjacentes à la MTD,
- la conception de nouveaux schémas d'étude de phase I/II. Par exemple :
  - o des schémas dans lesquels, après l'escalade de dose, les patients seraient randomisés entre plusieurs doses adjacentes (cohortes d'expansion randomisées étudiant l'efficacité préliminaire et la tolérance)
  - o des schémas d'escalade basés sur des critères conjoints de tolérance et efficacité.

Les performances de ces schémas innovants mériteraient d'être évaluées par simulations, et certains d'entre eux sont déjà en cours de développement. Les groupes d'experts comme le Task Force on Methodology for the development of Innovative Cancer Therapies (MDICT) (Booth, Calvert et al. 2008) ou le DLT-TARGET Group (Postel-Vinay, Rizzo et al. 2013) contribuent à la promotion des progrès méthodologiques aux seins des consortium pluridisciplinaires mettant en œuvre ces essais.

## 6 Conclusion

A partir d'une étude appliquée en oncologie pédiatrique, ce travail a permis de conforter le fait que l'adaptation des schémas algorithmiques pour permettre des inclusions continues reste inférieure aux schémas avec attribution de la dose par un modèle statistique. Les schémas par méthode de réévaluation continue basée sur un modèle doivent être préférés aux schémas algorithmiques.

L'ensemble de ces travaux tend également à montrer que les modèles mixtes pour données longitudinales permettent d'améliorer les performances des schémas adaptatifs de recherche de dose, en utilisant l'ensemble des données cliniques collectées. Le modèle mixte à cotes proportionnelles permet de surcroit d'améliorer la puissance pour détecter une variation du risque de toxicité au cours des cycles successifs, par rapport au modèle logistique mixte. L'implémentation de ces modèles dans les essais de recherche de dose en oncologie est indispensable. A l'ère des thérapies moléculaires ciblées, ces modèles ont toute leur place notamment pour évaluer la tolérance dans le temps, pour ces traitements administrés au long cours.

Toutefois, les capacités d'identification de la dose cible avec les méthodes longitudinales que nous avons proposées restent en deçà de 75 % des simulations, et des changements plus fondamentaux sur les schémas et le plan de développement global des traitements en oncologie doivent être envisagés.

## 7 Références

- Agence Nationale de Sécurité du Médicament et des Produits de Santé. (2006). "Médicaments en pédiatrie." Retrieved from <http://ansm.sante.fr/Activites/Medicaments-en-pediatrie/Reglementation-relative-aux-medicaments-pediatriques/>. Accessed March, 12, 2014.
- Agrawal, M. and E. J. Emanuel (2003). "Ethics of phase 1 oncology studies: reexamining the arguments and data." *JAMA* **290**(8): 1075-1082.
- Anderson, J. A. and J. Kimmelman (2012). "Are Phase 1 Trials Therapeutic? Risk, Ethics, and Division of Labor." *Bioethics*.
- Association Médicale Mondiale (2013). Déclaration d'Helsinki de L'AMM - Principes éthiques applicables à la recherche médicale impliquant des êtres humains. Association Médicale Mondiale: 9.
- Babb, J., A. Rogatko and S. Zacks (1998). "Cancer phase I clinical trials: efficient dose escalation with overdose control." *Stat Med* **17**(10): 1103-1120.
- Bachelot, T., I. Ray-Coquard, G. Catimel, C. Ardiét, J. P. Guastalla, A. Dumortier, F. Chauvin, J. P. Droz, T. Philip and M. Clavel (2000). "Multivariable analysis of prognostic factors for toxicity and survival for patients enrolled in phase I clinical trials." *Ann Oncol* **11**(2): 151-156.
- Bartroff, J. and T. Leung Lai (2011). "Incorporating individual and collective ethics into phase I cancer trial designs." *Biometrics* **67**(2): 596-603.
- Bekele, B. N. and P. F. Thall (2006). Dose-finding based on multiple ordinal toxicities in phase I oncology trials. *Statistical methods for dose-finding experiments*. S. Chevret. Chichester, West Sussex (England), Hoboken (NJ): Wiley Press: 243-258.
- Booth, C. M., A. H. Calvert, G. Giaccone, M. W. Lobbzoo, L. K. Seymour and E. A. Eisenhauer (2008). "Endpoints and other considerations in phase I studies of targeted anticancer therapy: recommendations from the task force on Methodology for the Development of Innovative Cancer Therapies (MDICT)." *Eur J Cancer* **44**(1): 19-24.
- Braun, T. M. (2006). "Generalizing the TITE-CRM to adapt for early- and late-onset toxicities." *Stat Med* **25**(12): 2071-2083.
- Braun, T. M., J. E. Levine and J. L. Ferrara (2003). "Determining a maximum tolerated cumulative dose: dose reassignment within the TITE-CRM." *Control Clin Trials* **24**(6): 669-681.
- Carriere, I. and J. Bouyer (2002). "Choosing marginal or random-effects models for longitudinal binary responses: application to self-reported disability among older persons." *BMC Med Res Methodol* **2**: 15.
- Carriere, I. and J. Bouyer (2006). "Random-effect models for ordinal responses: application to self-reported disability among older persons." *Rev Epidemiol Sante Publique* **54**(1): 61-72.
- Chen, Z., M. D. Krailo, S. P. Azen and M. Tighiouart (2010). "A novel toxicity scoring system treating toxicity response as a quasi-continuous variable in Phase I clinical trials." *Contemp Clin Trials* **31**(5): 473-482.
- Cheung, Y. K. and R. Chappell (2000). "Sequential designs for phase I clinical trials with late-onset toxicities." *Biometrics* **56**(4): 1177-1182.
- Cheung, Y. K. and R. Chappell (2002). "A simple technique to evaluate model sensitivity in the continual reassessment method." *Biometrics* **58**(3): 671-674.
- Chevret, S. and S. Zohar (2006). The continual reassessment method. *Statistical methods for dose-finding experiments*. S. Chevret. Chichester, West Sussex (England), Hoboken (NJ): Wiley Press: 133-148.
- Comité consultatif national d'éthique pour les sciences de la vie et de la santé (2002). Les essais de phase 1 en cancérologie. Avis n° 73. Paris, République Française.
- Committee for medicinal products for human use (2007). Guidelines on strategies to identify and mitigate risks for first-in-human clinical trials with investigational medicinal products. London, European Medicines Agency,. 1-12.

- Cousino, M. K., S. J. Zyzanski, A. D. Yamokoski, R. A. Hazen, J. N. Baker, R. B. Noll, S. R. Rheingold, J. R. Geyer, S. C. Alexander, D. Drotar and E. D. Kodish (2012). "Communicating and understanding the purpose of pediatric phase I cancer trials." *J Clin Oncol* **30**(35): 4367-4372.
- Dancey, J., B. Freidlin and L. Rubinstein (2006). Accelerated titration designs. *Statistical methods for dose-finding experiments*. S. Chevret. Chichester, West Sussex (England), Hoboken (NJ): Wiley Press: 91–114.
- Daugherty, C., M. J. Ratain, E. Grochowski, C. Stocking, E. Kodish, R. Mick and M. Siegler (1995). "Perceptions of cancer patients and their physicians involved in phase I trials." *J Clin Oncol* **13**(5): 1062-1072.
- Davous, D., F. Doz and M. Heard (2007). "Fin de vie de l'enfant et recherche clinique en cancerologie pediatrique. ." *Arch Pediatr* **14**(3): 274-278.
- Deutsch, E., J. C. Soria and J. P. Armand (2005). "New concepts for phase I trials: evaluating new drugs combined with radiation therapy." *Nat Clin Pract Oncol* **2**(9): 456-465.
- Donohue, J. F., C. Fogarty, J. Lotvall, D. A. Mahler, H. Worth, A. Yorgancioglu, A. Iqbal, J. Swales, R. Owen, M. Higgins and B. Kramer (2010). "Once-daily bronchodilators for chronic obstructive pulmonary disease: indacaterol versus tiotropium." *Am J Respir Crit Care Med* **182**(2): 155-162.
- Doussau, A., B. Asselain, M. C. Le Deley, B. Geoerger, F. Doz, G. Vassal and X. Paoletti (2012). "Dose-finding designs in pediatric phase I clinical trials: comparison by simulations in a realistic timeline framework." *Contemp Clin Trials* **33**(4): 657-665.
- Doussau, A., R. Thiebaut, B. Geoerger, M. C. Le Deley, P. Schöffski, A. Floquet, S. Mathoulin-Pélissier, E. Rizzo, P. Fumoleau and X. Paoletti (2013). "A new approach to integrate the grade of toxicity and later cycles in the analysis and reporting of phase I dose finding trials [Abstract]." *Eur J Cancer* **49**(Suppl 2): S163.
- Doussau, A., R. Thiebaut and X. Paoletti (2013). "Dose-finding design using mixed-effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials." *Stat Med* **32**(30): 5430-5347.
- Eisenhauer, E. A., P. Therasse, J. Bogaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe and J. Verweij (2009). "New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)." *Eur J Cancer* **45**(2): 228-247.
- European Medicines Agency (1994). ICH Topic E 4 - Dose Response Information to Support Drug Registration. Note for guidance on dose response information to support drug registration. London.
- European Medicines Agency (1998). ICH Topic E 8 - General Considerations for Clinical Trials. Note for guidance on general considerations for clinical trials (CPMP/ICH/291/95). London: 1-14.
- Ezzalfani, M., S. Zohar, R. Qin, S. J. Mandrekar and M. C. Deley (2013). "Dose-finding designs using a novel quasi-continuous endpoint for multiple toxicities." *Stat Med* **32**(16): 2728-2746.
- Ezzet, F. and J. Whitehead (1991). "A random effects model for ordinal responses from a crossover trial." *Stat Med* **10**(6): 901-906; discussion 906-907.
- Floquet, A., A. Doussau, V. Brouste, L. Cany, J. P. Dutin and S. Mathoulin-Pelissier (2014). "Pegylated liposomal doxorubicin and cyclophosphamide in early recurrent ovarian carcinoma: phase I dose-finding study." *Cancer Chemother Pharmacol* **73**(1): 61-68.
- Geoerger, B., D. Hargrave, F. Thomas, A. Ndiaye, D. Frappaz, F. Andreiuolo, P. Varlet, I. Aerts, R. Riccardi, T. Jaspan, E. Chatelut, M. C. Le Deley, X. Paoletti, C. Saint-Rose, P. Leblond, B. Morland, J. C. Gentet, V. Meresse and G. Vassal (2011). "Innovative Therapies for Children with Cancer pediatric phase I study of erlotinib in brainstem glioma and relapsing/refractory brain tumors." *Neuro Oncol* **13**(1): 109-118.
- Goodwin, R., G. Giaccone, H. Calvert, M. Lobbezoo and E. A. Eisenhauer (2012). "Targeted agents: how to select the winners in preclinical and early clinical studies?" *Eur J Cancer* **48**(2): 170-178.

- Hamberg, P., M. J. Ratain, E. Lesaffre and J. Verweij (2010). "Dose-escalation models for combination phase I trials in oncology." *Eur J Cancer* **46**(16): 2870-2878.
- Horstmann, E., M. S. McCabe, L. Grochow, S. Yamamoto, L. Rubinstein, T. Budd, D. Shoemaker, E. J. Emanuel and C. Grady (2005). "Risks and benefits of phase 1 oncology trials, 1991 through 2002." *N Engl J Med* **352**(9): 895-904.
- Iasonos, A., M. Gounder, D. R. Spriggs, J. F. Gerecitano, D. M. Hyman, S. Zohar and J. O'Quigley (2012). "The impact of non-drug-related toxicities on the estimation of the maximum tolerated dose in phase I trials." *Clin Cancer Res* **18**(19): 5179-5187.
- Iasonos, A. and J. O'Quigley (2011). "Continual reassessment and related designs in dose-finding studies." *Stat Med* **30**(17): 2057-2061.
- Iasonos, A., A. S. Wilton, E. R. Riedel, V. E. Seshan and D. R. Spriggs (2008). "A comprehensive comparison of the continual reassessment method to the standard 3 + 3 dose escalation scheme in Phase I dose-finding studies." *Clin Trials* **5**(5): 465-477.
- Iasonos, A., S. Zohar and J. O'Quigley (2011). "Incorporating lower grade toxicity information into dose finding designs." *Clin Trials* **8**(4): 370-379.
- Institut National du Cancer (2010). La situation du cancer en France en 2010. Boulogne-Billancourt, Institut National du Cancer.
- Italiano, A., C. Massard, R. Bahleda, A. L. Vataire, E. Deutsch, N. Magne, J. P. Pignon, G. Vassal, J. P. Armand and J. C. Soria (2008). "Treatment outcome and survival in participants of phase I oncology trials carried out from 2003 to 2006 at Institut Gustave Roussy." *Ann Oncol* **19**(4): 787-792.
- Jacqmin-Gadda, H., C. Proust-Lima and H. Amieva (2010). "Semi-parametric latent process model for longitudinal ordinal data: Application to cognitive decline." *Stat Med* **29**(26): 2723-2731.
- Jaki, T., S. Clive and C. J. Weir (2013). "Principles of dose finding studies in cancer: a comparison of trial designs." *Cancer Chemother Pharmacol* **71**(5): 1107-1114.
- Kola, I. and J. Landis (2004). "Can the pharmaceutical industry reduce attrition rates?" *Nat Rev Drug Discov* **3**(8): 711-715.
- Korn, E. L., S. G. Arbuck, J. M. Pluda, R. Simon, R. S. Kaplan and M. C. Christian (2001). "Clinical trial designs for cytostatic agents: are new approaches needed?" *J Clin Oncol* **19**(1): 265-272.
- Korn, E. L., B. Freidlin, J. S. Abrams and S. Halabi (2012). "Design issues in randomized phase II/III trials." *J Clin Oncol* **30**(6): 667-671.
- Korn, E. L., D. Midthune, T. T. Chen, L. V. Rubinstein, M. C. Christian and R. M. Simon (1994). "A comparison of two phase I trial designs." *Stat Med* **13**(18): 1799-1806.
- Le Tourneau, C., H. K. Gan, A. R. Razak and X. Paoletti (2012). "Efficiency of new dose escalation designs in dose-finding phase I trials of molecularly targeted agents." *PLoS One* **7**(12): e51039.
- Le Tourneau, C., J. J. Lee and L. L. Siu (2009). "Dose escalation methods in phase I cancer clinical trials." *J Natl Cancer Inst* **101**(10): 708-720.
- Le Tourneau, C., A. R. Razak, H. K. Gan, S. Pop, V. Dieras, P. Tresca and X. Paoletti (2011). "Heterogeneity in the definition of dose-limiting toxicity in phase I cancer clinical trials of molecularly targeted agents: a review of the literature." *Eur J Cancer* **47**(10): 1468-1475.
- Lee, D. P., J. M. Skolnik and P. C. Adamson (2005). "Pediatric phase I trials in oncology: an analysis of study conduct efficiency." *J Clin Oncol* **23**(33): 8431-8441.
- Lee, S. M., D. L. Hershman, P. Martin, J. P. Leonard and Y. K. Cheung (2012). "Toxicity burden score: a novel approach to summarize multiple toxic effects." *Ann Oncol* **23**(2): 537-541.
- Legedza, A. T. and J. G. Ibrahim (2000). "Longitudinal design for phase I clinical trials using the continual reassessment method." *Control Clin Trials* **21**(6): 574-588.
- Little, R. J. A. and D. B. Rubin (1987). *Statistical Analysis with Missing Data*. New-York, Wiley.
- Liu, G., W. F. Rosenberger and L. M. Haines (2009). "Sequential designs for ordinal phase I clinical trials." *Biom J* **51**(2): 335-347.
- Mandrekar, S. J., Y. Cui and D. J. Sargent (2007). "An adaptive phase I design for identifying a biologically optimal dose for dual agent drug combinations." *Stat Med* **26**(11): 2317-2330.

- Manji, A., I. Brana, E. Amir, G. Tomlinson, I. F. Tannock, P. L. Bedard, A. Oza, L. L. Siu and A. R. Razak (2013). "Evolution of clinical trial design in early drug development: systematic review of expansion cohort use in single-agent phase I cancer trials." *J Clin Oncol* **31**(33): 4260-4267.
- Marchetti, S. and J. H. Schellens (2007). "The impact of FDA and EMEA guidelines on drug development in relation to Phase 0 trials." *Br J Cancer* **97**(5): 577-581.
- Mauguen, A., M. C. Le Deley and S. Zohar (2011). "Dose-finding approach for dose escalation with overdose control considering incomplete observations." *Stat Med* **30**(13): 1584-1594.
- Molenberghs, G. and G. Verbeke (2005a). *The Analgesic Trial. Models for discrete longitudinal data.* Springer. New York, Springer: 309-323.
- Molenberghs, G. and G. Verbeke (2005b). *Ordinal Data. Models for discrete longitudinal data.* Springer. New York, Springer: 324-345.
- Molenberghs, G. and G. Verbeke (2005c). *Marginal versus Random-Effects Models. Models for discrete longitudinal data.* Springer. New York, Springer: 297-306.
- Molenberghs, G. and G. Verbeke (2005d). *Missing data concepts. Models for discrete longitudinal data.* Springer. New York, Springer: 481-488.
- Moroney, J., S. Fu, S. Moulder, G. Falchook, T. Helgason, C. Levenback, D. Hong, A. Naing, J. Wheler and R. Kurzrock (2012). "Phase I study of the antiangiogenic antibody bevacizumab and the mTOR/hypoxia-inducible factor inhibitor temsirolimus combined with liposomal doxorubicin: tolerance and biological activity." *Clin Cancer Res* **18**(20): 5796-5805.
- National Cancer Institute and US National Institutes of Health (2010). "Cancer Therapy Evaluation Program. Common Terminology Criteria for Adverse Events (CTCAE) v4.03." available at: [http://ctep.cancer.gov/protocolDevelopment/electronic\\_applications/ctc.htm](http://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm). Accessed March, 12, 2014.
- O'Quigley, J. (1992). "Estimating the probability of toxicity at the recommended dose following a phase I clinical trial in cancer." *Biometrics* **48**(3): 853-862.
- O'Quigley, J. (2002). "Continual reassessment designs with early termination." *Biostatistics* **3**(1): 87-99.
- O'Quigley, J., M. Pepe and L. Fisher (1990). "Continual reassessment method: a practical design for phase 1 clinical trials in cancer." *Biometrics* **46**(1): 33-48.
- O'Quigley, J. and E. Reiner (1998). "A stopping rule for the continual reassessment method." *Biometrika* **85**(3): 741-748.
- O'Quigley, J. and L. Z. Shen (1996). "Continual reassessment method: a likelihood approach." *Biometrics* **52**(2): 673-684.
- Olmos, D., P. A'Hern R, S. Marsoni, R. Morales, C. Gomez-Roca, J. Verweij, E. E. Voest, P. Schoffski, J. E. Ang, N. Penel, J. H. Schellens, G. Del Conte, A. T. Brunetto, T. R. Evans, R. Wilson, E. Gallerani, R. Plummer, J. Tabernero, J. C. Soria and S. B. Kaye (2012). "Patient selection for oncology phase I trials: a multi-institutional study of prognostic factors." *J Clin Oncol* **30**(9): 996-1004.
- Onar-Thomas, A. and Z. Xiong (2010). "A simulation-based comparison of the traditional method, Rolling-6 design and a frequentist version of the continual reassessment method with special attention to trial duration in pediatric Phase I oncology trials." *Contemp Clin Trials* **31**(3): 259-270.
- Paoletti, X. and A. Doussau (2014). *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose. Developments in Statistical Evaluation of Clinical Trials.* K. van Montfort, J. Oud and W. Ghiddey. New York, Springer: 333-358.
- Paoletti, X., B. Georger, F. Doz, A. Baruchel, F. Lokiec and C. Le Tourneau (2013). "A comparative analysis of paediatric dose-finding trials of molecularly targeted agent with adults' trials." *Eur J Cancer* **49**(10): 2392-2402.
- Paoletti, X. and A. Kramar (2009). "A comparison of model choices for the Continual Reassessment Method in phase I cancer trials." *Stat Med* **28**(24): 3012-3028.
- Paoletti, X., J. O'Quigley and J. Maccario (2004). "Design efficiency in dose finding studies." *Computational Statistics & Data Analysis* **45**(2): 197-214

- Parmar, M. K., F. M. Barthel, M. Sydes, R. Langley, R. Kaplan, E. Eisenhauer, M. Brady, N. James, M. A. Bookman, A. M. Swart, W. Qian and P. Royston (2008). "Speeding up the evaluation of new agents in cancer." *J Natl Cancer Inst* **100**(17): 1204-1214.
- Parulekar, W. R. and E. A. Eisenhauer (2004). "Phase I trial design for solid tumor studies of targeted, non-cytotoxic agents: theory and practice." *J Natl Cancer Inst* **96**(13): 990-997.
- Pedrono, G., R. Thiebaut, A. Alioum, P. Lesprit, B. Fritzell, Y. Levy and G. Chene (2009). "A new endpoint definition improved clinical relevance and statistical power in a vaccine trial." *J Clin Epidemiol* **62**(10): 1054-1061.
- Penel, N., N. Isambert, P. Leblond, C. Ferte, A. Duhamel and J. Bonneterre (2009). "'Classical 3 + 3 design' versus 'accelerated titration designs': analysis of 270 phase 1 trials investigating anti-cancer agents." *Invest New Drugs* **27**(6): 552-556.
- Pentz, R. D., M. White, R. D. Harvey, Z. L. Farmer, Y. Liu, C. Lewis, O. Dashevskaya, T. Owonikoko and F. R. Khuri (2012). "Therapeutic misconception, misestimation, and optimism in participants enrolled in phase 1 trials." *Cancer* **118**(18): 4571-4578.
- Piantadosi, S. (2005). Dose Finding Designs. *Clinical Trials: A methodologic perspective*. Hoboken, John Wiley: 223-250.
- Polley, M. Y. (2011). "Practical modifications to the time-to-event continual reassessment method for phase I cancer trials with fast patient accrual and late-onset toxicities." *Stat Med* **30**(17): 2130-2143.
- Porcher, R., B. Lecoq and M. Vray (2011). "Adaptive methods: when and how should they be used in clinical trials?" *Therapie* **66**(4): 319-326, 309-317.
- Postel-Vinay, S., C. Gomez-Roca, L. R. Molife, B. Anghan, A. Levy, I. Judson, J. De Bono, J. C. Soria, S. Kaye and X. Paoletti (2011). "Phase I trials of molecularly targeted agents: should we pay more attention to late toxicities?" *J Clin Oncol* **29**(13): 1728-1735.
- Postel-Vinay, S., E. Rizzo, C. Le Tourneau, D. Olmos, C. Massard, P. Ivy, L. Seymour, L. Siu, D. Lacombe, X. Paoletti, S. B. Kaye, J. Verweij, L. Collette and J. Soria (2013). "Towards new methods for the determination of dose limiting toxicities and recommended dose of molecularly targeted agents. On the behalf of DLT-TARGETT (DLT and Toxicity Assessment Recommendation Group for Early Trials of Targeted therapies) [Abstract]." *Eur J Cancer* **49**(Suppl 3): S5.
- Ratain, M. J., R. W. Humphrey, G. B. Gordon, G. Fyfe, P. C. Adamson, T. R. Fleming, W. M. Stadler, D. A. Berry and C. C. Peck (2008). "Recommended changes to oncology clinical trial design: revolution or evolution?" *Eur J Cancer* **44**(1): 8-11.
- Reiner, E., X. Paoletti and J. O'Quigley (1999). "Operating characteristics of the standard phase I clinical trial design." *Computational Statistics & Data Analysis* **30**: 303-315.
- Roberts, T. G., Jr., B. H. Goulart, L. Squitieri, S. C. Stallings, E. F. Halpern, B. A. Chabner, G. S. Gazelle, S. N. Finkelstein and J. W. Clark (2004). "Trends in the risks and benefits to patients with cancer participating in phase 1 clinical trials." *JAMA* **292**(17): 2130-2140.
- Rogatko, A., D. Schoeneck, W. Jonas, M. Tighiourt, F. R. Khuri and A. Porter (2007). "Translation of innovative designs into phase I trials." *J Clin Oncol* **25**(31): 4982-4986.
- Schoffski, P., S. Riggert, P. Fumoleau, M. Campone, O. Bolte, S. Marreaud, D. Lacombe, B. Baron, M. Herold, H. Zwierzina, K. Wilhelm-Ogunbiyi, H. Lentzen and C. Twelves (2004). "Phase I trial of intravenous aviscumine (rViscumin) in patients with solid tumors: a study of the European Organization for Research and Treatment of Cancer New Drug Development Group." *Ann Oncol* **15**(12): 1816-1824.
- Shen, L. Z. and J. O'Quigley (1996). "Consistency of continual reassessment method under model misspecification." *Biometrika* **83**: 395-405.
- Simon, R., B. Freidlin, L. Rubinstein, S. G. Arbuck, J. Collins and M. C. Christian (1997). "Accelerated titration designs for phase I clinical trials in oncology." *J Natl Cancer Inst* **89**(15): 1138-1147.
- Skolnik, J. M., J. S. Barrett, B. Jayaraman, D. Patel and P. C. Adamson (2008). "Shortening the timeline of pediatric phase I trials: the rolling six design." *J Clin Oncol* **26**(2): 190-195.

- Smith, M., M. Bernstein, W. A. Bleyer, J. D. Borsi, P. Ho, I. J. Lewis, A. Pearson, F. Pein, C. Pratt, G. Reaman, R. Riccardi, N. Seibel, R. Trueworthy, R. Ungerleider, G. Vassal and T. Vietti (1998). "Conduct of phase I trials in children with cancer." J Clin Oncol **16**(3): 966-978.
- Sposto, R. and S. Groshen (2011). "A wide-spectrum paired comparison of the properties of the Rolling 6 and 3+3 Phase I study designs." Contemp Clin Trials **32**(5): 694-703.
- Storer, B. E. (1989). "Design and analysis of phase I clinical trials." Biometrics **45**(3): 925-937.
- Sydes, M. R., M. K. Parmar, M. D. Mason, N. W. Clarke, C. Amos, J. Anderson, J. de Bono, D. P. Dearnaley, J. Dwyer, C. Green, G. Jovic, A. W. Ritchie, J. M. Russell, K. Sanders, G. Thalmann and N. D. James (2012). "Flexible trial design in practice - stopping arms for lack-of-benefit and adding research arms mid-trial in STAMPEDE: a multi-arm multi-stage randomized controlled trial." Trials **13**: 168.
- Thall, P. F. and J. D. Cook (2006). Using both efficacy and toxicity for dose-finding. Statistical methods for dose-finding experiments. S. Chevret. Chichester, West Sussex (England), Hoboken (NJ): Wiley Press: 275-285.
- Thall, P. F., J. J. Lee, C. H. Tseng and E. H. Estey (1999). "Accrual strategies for phase I trials with delayed patient outcome." Stat Med **18**(10): 1155-1169.
- The European Parliament and The Council of the European Union (2006). "REGULATION (EC) No 1901/2006 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, of 12 December 2006, on medicinal products for paediatric use and amending Regulation (EEC) No 1768/92, Directive 2001/20/EC, Directive 2001/83/EC and Regulation (EC) No 726/2004;." Official Journal of the European Union 27.12.2006.
- Tighiouart, M., G. Cook-Weins and A. Rogatko (2012 ). "Escalation with Overdose Control Using Ordinal Toxicity Grades for Cancer Phase I Clinical Trials." Journal of probability and statistics **2012**: 1-18.
- Trotti, A., A. D. Colevas, A. Setser, V. Rusch, D. Jaques, V. Budach, C. Langer, B. Murphy, R. Cumberlin, C. N. Coleman and P. Rubin (2003). "CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment." Semin Radiat Oncol **13**(3): 176-181.
- Van Meter, E. M., E. Garrett-Mayer and D. Bandyopadhyay (2011). "Proportional odds model for dose-finding clinical trial designs with ordinal toxicity grading." Stat Med **30**(17): 2070-2080.
- Van Meter, E. M., E. Garrett-Mayer and D. Bandyopadhyay (2012). "Dose-finding clinical trial design for ordinal toxicity grades using the continuation ratio model: an extension of the continual reassessment method." Clin Trials **9**(3): 303-313.
- Wages, N. A., M. R. Conaway and J. O'Quigley (2011). "Continual reassessment method for partial ordering." Biometrics **67**(4): 1555-1563.
- Wang, K. and A. Ivanova (2005). "Two-dimensional dose finding in discrete dose space." Biometrics **61**(1): 217-222.
- Yin, G. and Y. Yuan (2009). "Bayesian dose finding in oncology for drug combinations by copula regression." Journal of the Royal Statistical Society **58**: 211-224.
- Yuan, Z., R. Chappell and H. Bailey (2007). "The continual reassessment method for multiple toxicity grades: a Bayesian quasi-likelihood approach." Biometrics **63**(1): 173-179.
- Zhang, W., D. J. Sargent and S. Mandrekar (2006). "An adaptive dose-finding design incorporating both toxicity and efficacy." Stat Med **25**(14): 2365-2383.
- Zhao, L., J. Lee, R. Mody and T. M. Braun (2011). "The superiority of the time-to-event continual reassessment method to the rolling six design in pediatric oncology Phase I trials." Clin Trials **8**(4): 361-369.
- Zhou, Y. and J. Whitehead (2003). "Practical Implementation of Bayesian Dose-Escalation Procedures." Drug Information Journal **37**: 45-59.
- Zohar, S. and S. Chevret (2001). "The continual reassessment method: comparison of Bayesian stopping rules for dose-ranging studies." Stat Med **20**(19): 2827-2843.
- Zohar, S., Q. Lian, V. Levy, K. Cheung, A. Ivanova and S. Chevret (2008). "Quality assessment of phase I dose-finding cancer trials: proposal of a checklist." Clin Trials **5**(5): 478-485.

## Table des illustrations, modèles et équations

### Figures

<b>Figure 1.</b> Relations dose-toxicité et dose-efficacité typiques d'une chimiothérapie cytotoxique. ....	21
<b>Figure 2.</b> Probabilité d'évènement en fonction de la dose, dans un modèle logistique de pente 1 et d'intercept -6, -4 ou -2. ....	36
<b>Figure 3.</b> Représentation schématique d'un essai mené avec une méthode de réévaluation continue. ....	37
<b>Figure 4.</b> Représentation graphique des principaux schémas de recherche de dose. ....	40
<b>Figure 5.</b> Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle à cotes proportionnelles, de paramètre $\theta = (\alpha_1 = 9.85, \alpha_2 = 11.29, \beta_1 = 1.78)$ . ....	49
<b>Figure 6.</b> Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle logistique ordonné généralisé, de paramètre $\theta = (\alpha_1 = 9.85, \alpha_2 = 11.29, \beta_{1.1} = 1.63, \beta_{1.2} = 1.78)$ . ....	50
<b>Figure 7.</b> Dates d'inclusion et évaluation des TDL dans l'essai erlotinib + radiothérapie (Georger, Hargrave et al. 2011). ....	66
<b>Figure 8.</b> Scénarios étudiés dans l'étude de simulation comparant le 3+3, le <i>Rolling 6</i> et la MRC modifiée pour inclusion dans les essais pédiatriques : Situation centrale, et variations du rythme d'inclusion, de la durée du cycle, et de la relation dose-toxicité étudiées (Doussau, Asselain et al. 2012). ....	70
<b>Figure 9.</b> Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle à cotes proportionnelles, de paramètre $\theta = (\alpha_1 = 9,85, \alpha_2 = 11,29, \beta_1 = 1,78, \sigma^2=0,5)$ , pour les valeurs $u_i = 0$ (lignes noire et grise pleines) et $u_i = \pm \sigma_0$ (lignes pointillées). ....	90
<b>Figure 10.</b> Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle à cotes proportionnelles, de paramètre $\theta = (\alpha_1 = 9,85, \alpha_2 = 11,29, \beta_1 = 1,78, \sigma^2=0,5)$ . ....	92
<b>Figure 11.</b> Relations entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon l'amplitude de l'effet aléatoire. ....	93
<b>Figure 12.</b> Relation entre la dose et la probabilité de toxicité grave (p3), et modérée ou grave (p2+), selon un modèle à cotes proportionnelles, de paramètres $\theta = [\alpha_1 = 9,85, \alpha_2 = 11,29, \beta_1 = 1,78, \beta_2 = (0,58 ; 0,29 ; -0,29, \text{respectivement, de gauche à droite})]$ , pour $u_i = 0$ . ....	95
<b>Figure 13.</b> Distribution des doses recommandées sur 1 000 échantillons d'essais simulés selon un schéma CRML (n = 30 patients, recevant 1 à 6 cycles de traitement, jusqu'à la première toxicité grave), selon différents modèles d'analyse des données. ....	148
<b>Figure 14.</b> Représentation de deux types de variation du risque de toxicité dans le temps : effet cumulatif log-linéaire (gauche) et effet retardé (droite). ....	159
<b>Figure 15.</b> Schéma de l'effet de la dose sur la cible biologique, l'effet antitumoral, et de la relation dose-toxicité, supposé pour une thérapie ciblée associée à la radiothérapie. ....	162

<b>Figure 16.</b> Répartition de distribution de la toxicité maximale par cycle, selon la dose (gauche) et le cycle (droite). Données d'un essai issu de (Postel-Vinay, Gomez-Roca et al. 2011).....	193
<b>Figure 17.</b> Distribution des doses recommandées sur 1 000 échantillons d'essais simulés selon un schéma CRML, ou selon des schémas adaptatifs utilisant un modèle longitudinal pour l'attribution des doses (n = 30 patients, recevant 1 à 6 cycles de traitement, jusqu'à la première toxicité grave). .....	227

## Tableaux

<b>Tableau 1.</b> Règle d'escalade de dose avec le schéma <i>Rolling 6</i> , selon (Zhao, Lee et al. 2011). .....	31
<b>Tableau 2 .</b> Résumé de l'essai erlotinib + radiothérapie [données réelles (Georger, Hargrave et al. 2011)] et des schémas adaptatifs rétrospectifs selon le <i>Rolling 6</i> et le « 3+3 ».....	67
<b>Tableau 3.</b> <i>Distribution of the dose recommendations in proportion and median number of patients treated per dose level, over 1,000 trials simulated with 4 phase I dose escalation designs, according to true RP2D occurring at the 3<sup>rd</sup> (upper table) to 5<sup>th</sup> doses (lower table). Results from an optimal design are also presented.</i> .....	73
<b>Tableau 4.</b> <i>Median number of patients, duration and toxicity, over 1.000 trials simulated with 4 phase I dose escalation designs, according to true RP2D occurring at 3<sup>rd</sup> to 5<sup>th</sup> dose.</i> .....	74
<b>Tableau 5.</b> Nombre de paramètres par modèle.....	89
<b>Tableau 6.</b> Distribution des doses recommandées et du nombre de patients moyen inclus par dose selon les différentes stratégies. ....	100
<b>Tableau 7.</b> Impact du nombre total de doses à explorer sur le nombre de patients à inclus à la dose cible et la proportion de recommandations correctes de la dose cible.....	120
<b>Tableau 8.</b> Analyse de l'essai erlotinib + radiothérapie selon le modèle mixte à cotes proportionnelles, le modèle logistique mixte et le modèle logistique pondéré de l'approche TITE-CRM. ....	151

## Modèles et équations

(Modèle 1) .....	34
(Modèle 2) .....	35
(Modèle 3) .....	47
(Modèle 4) .....	49
(Modèle 5) .....	58
(Modèle 6) .....	59
(Modèle 7) .....	87
(Modèle 8) .....	87
(Modèle 9) .....	88
(Modèle 10) .....	146
(Modèle 11) .....	192
(Équation A) .....	33
(Équation B) .....	48
(Équation C) .....	50
(Équation D) .....	88
(Équation E) .....	90

## Table des notations

Les principales notations utilisées dans le manuscrit sont regroupées ici.

$I^{[y_i=k]}$	fonction qui prend la valeur 1 si $y_i=k$ , et la valeur 0 dans les autres cas
A	domaine de définition des paramètres
$D_L$	ensemble des L doses croissantes à explorer $D_L = (d_1, \dots, d_\ell, \dots, d_L)$
i	indice du sujet, selon le rang d'inclusion ; $i = (1 \dots I)$
j	indice du cycle, ou du temps ; $j = (1 \dots J)$
k	modalité de réponse de la variable ordinaire y ; $k = (1 \dots K)$
$\ell$	indice d'une dose, parmi les L doses à explorer ; $\ell = (1 \dots L)$
$p_{2+}$	probabilité de toxicité grave (Doussau, Thiebaut et al. 2013)
$p_3$	probabilité de toxicité modérée ou grave (Doussau, Thiebaut et al. 2013)
$R_\ell$	probabilité que le patient présente une toxicité à la dose $d_\ell$ (Paoletti, O'Quigley et al. 2004)
$s_{ij}$	variable latente, reflétant un score de toxicité (Simon, Freidlin et al. 1997)
t	temps (= cycle)
$u_i$	<i>intercept</i> aléatoire du sujet i
$V_i$	temps jusqu'à survenue d'une toxicité pour le patient i
$\omega$	pondération (Cheung and Chappell 2000)
$x_i$	dose reçue par le sujet i
y	variable réponse
$\alpha$	ordonnée à l'origine
$\beta$	effet fixe (dose, temps)
$\gamma$	paramètre estimant un effet cumulé de la dose (Simon, Freidlin et al. 1997)
$\delta$	fraction maximale de population exposée à un surdosage (Babb, Rogatko et al. 1998)
$\varepsilon$	variabilité intrapatient (Simon, Freidlin et al. 1997)
$\zeta$	paramètre estimant l'effet de la dose (Legedza and Ibrahim 2000)
$\eta_i$	niveau de dose le plus bas auquel une TDL survient pour le patient i (Paoletti, O'Quigley et al. 2004)
$\theta$	ensemble des paramètres à estimer
$\lambda$	clairance du produit (Legedza and Ibrahim 2000)
$\pi(\cdot)$	fonction de distribution (méthodes Bayésiennes, (Shen and O'Quigley 1996; Babb, Rogatko et al. 1998))
$\rho$	dose maximale tolérée (Tighiouart, Cook-Weins et al. 2012)
$\sigma_0^2$	variance de l'effet aléatoire
$\tau$	cible de probabilité de toxicité à identifier
$\varphi(\cdot)$	fonction de distribution cumulative de $\eta_i$ (Paoletti, O'Quigley et al. 2004)
$\psi(\cdot)$	fonction de la dose exprimant la probabilité de toxicité de chaque dose
$\psi_t(\cdot)$	fonction de la dose exprimant la probabilité de toxicité de chaque dose, selon la méthode TITE-CRM (Cheung and Chappell 2000), avec pondération selon la durée d'observation
$\Omega_n$	ensemble des doses et des réponses pour les n premiers patients inclus $\Omega_n = \{(x_1, y_1) \dots (x_n, y_n)\}$

## 8 Annexes

### Annexe 1. Déclaration d'Helsinki de l'Association Médicale Mondiale (Association Médicale Mondiale 2013)

#### Principes éthiques applicables à la recherche médicale impliquant des êtres humains

Adoptée par la 18e Assemblée générale de l'AMM, Helsinki, Finlande, Juin 1964 amendée [...], 64e Assemblée générale de l'AMM, Fortaleza, Brésil, Octobre 2013.

#### Préambule

1. L'Association Médicale Mondiale (AMM) a élaboré la Déclaration d'Helsinki comme un énoncé de principes éthiques applicables à la recherche médicale impliquant des êtres humains, y compris la recherche sur du matériel biologique humain et sur des données identifiables.

La Déclaration est conçue comme un tout indissociable. Chaque paragraphe doit être appliqué en tenant compte de tous les autres paragraphes pertinents.

2. Conformément au mandat de l'AMM, cette Déclaration s'adresse en priorité aux médecins. L'AMM invite cependant les autres personnes engagées dans la recherche médicale impliquant des êtres humains à adopter ces principes.

#### Principes généraux

3. La Déclaration de Genève de l'AMM engage les médecins en ces termes: «La santé de mon patient prévaudra sur toutes les autres considérations » et le Code International d'Ethique Médicale déclare qu'un «médecin doit agir dans le meilleur intérêt du patient lorsqu'il le soigne».

4. Le devoir du médecin est de promouvoir et de sauvegarder la santé, le bien être et les droits des patients, y compris ceux des personnes impliquées dans la recherche médicale. Le médecin consacre son savoir et sa conscience à l'accomplissement de ce devoir.

5. Le progrès médical est basé sur la recherche qui, en fin de compte, doit impliquer des êtres humains.

6. L'objectif premier de la recherche médicale impliquant des êtres humains est de comprendre les causes, le développement et les effets des maladies et d'améliorer les interventions préventives, diagnostiques et thérapeutiques (méthodes, procédures et traitements). Même les meilleures interventions éprouvées doivent être évaluées en permanence par des recherches portant sur leur sécurité, leur efficacité, leur pertinence, leur accessibilité et leur qualité.

7. La recherche médicale est soumise à des normes éthiques qui promeuvent et assurent le respect de tous les êtres humains et qui protègent leur santé et leurs droits.

8. Si l'objectif premier de la recherche médicale est de générer de nouvelles connaissances, cet objectif ne doit jamais prévaloir sur les droits et les intérêts des personnes impliquées dans la recherche.

9. Il est du devoir des médecins engagés dans la recherche médicale de protéger la vie, la santé, la dignité, l'intégrité, le droit à l'auto-détermination, la vie privée et la confidentialité des informations des personnes impliquées dans la recherche. La responsabilité de protéger les personnes impliquées dans la recherche doit toujours incomber à un médecin ou à un autre professionnel de santé et jamais aux personnes impliquées dans la recherche même si celles-ci ont donné leur consentement.

10. Dans la recherche médicale impliquant des êtres humains, les médecins doivent tenir compte des normes et standards éthiques, légaux et réglementaires applicables dans leur propre pays ainsi que des normes et standards internationaux. Les protections garanties par la présente Déclaration aux personnes impliquées dans la recherche ne peuvent être restreintes ou exclues par aucune disposition éthique, légale ou réglementaire, nationale ou internationale.

11. La recherche médicale devrait être conduite de sorte qu'elle réduise au minimum les nuisances éventuelles à l'environnement.

12. La recherche médicale impliquant des êtres humains doit être conduite uniquement par des personnes ayant acquis une éducation, une formation et des qualifications appropriées en éthique et en science. La recherche impliquant des patients ou des volontaires en bonne santé nécessite la supervision d'un médecin ou d'un autre professionnel de santé qualifié et compétent.

13. Des possibilités appropriées de participer à la recherche médicale devraient être offertes aux groupes qui y sont sous-représentés.

14. Les médecins qui associent la recherche médicale à des soins médicaux devraient impliquer leurs patients dans une recherche uniquement dans la mesure où elle se justifie par sa valeur potentielle en matière de

prévention, de diagnostic ou de traitement et si les médecins ont de bonnes raisons de penser que la participation à la recherche ne portera pas atteinte à la santé des patients concernés.

15. Une compensation et un traitement adéquats doivent être garantis pour les personnes qui auraient subi un préjudice en raison de leur participation à une recherche.

#### **Risques, contraintes et avantages**

16. Dans la pratique médicale et la recherche médicale, la plupart des interventions comprennent des risques et des inconvénients.

Une recherche médicale impliquant des êtres humains ne peut être conduite que si l'importance de l'objectif dépasse les risques et inconvénients pour les personnes impliquées.

17. Toute recherche médicale impliquant des êtres humains doit préalablement faire l'objet d'une évaluation soigneuse des risques et des inconvénients prévisibles pour les personnes et les groupes impliqués, par rapport aux bénéfices prévisibles pour eux et les autres personnes ou groupes affectés par la pathologie étudiée.

Toutes les mesures destinées à réduire les risques doivent être mises en œuvre. Les risques doivent être constamment surveillés, évalués et documentés par le chercheur.

18. Les médecins ne peuvent pas s'engager dans une recherche impliquant des êtres humains sans avoir la certitude que les risques ont été correctement évalués et pourront être gérés de manière satisfaisante.

Lorsque les risques s'avèrent dépasser les bénéfices potentiels ou dès l'instant où des conclusions définitives ont été démontrées, les médecins doivent évaluer s'ils continuent, modifient ou cessent immédiatement une recherche.

#### **Populations et personnes vulnérables**

19. Certains groupes ou personnes faisant l'objet de recherches sont particulièrement vulnérables et peuvent avoir une plus forte probabilité d'être abusés ou de subir un préjudice additionnel.

Tous les groupes et personnes vulnérables devraient bénéficier d'une protection adaptée.

20. La recherche médicale impliquant un groupe vulnérable se justifie uniquement si elle répond aux besoins ou aux priorités sanitaires de ce groupe et qu'elle ne peut être effectuée sur un groupe non vulnérable. En outre, ce groupe devrait bénéficier des connaissances, des pratiques ou interventions qui en résultent.

#### **Exigences scientifiques et protocoles de recherche**

21. La recherche médicale impliquant des êtres humains doit se conformer aux principes scientifiques généralement acceptés, se baser sur une connaissance approfondie de la littérature scientifique, sur d'autres sources pertinentes d'informations et sur des expériences appropriées en laboratoire et, le cas échéant, sur les animaux. Le bien être des animaux utilisés dans la recherche doit être respecté.

22. La conception et la conduite de toutes les recherches impliquant des êtres humains doivent être clairement décrites et justifiées dans un protocole de recherche.

Ce protocole devrait contenir une déclaration sur les enjeux éthiques en question et indiquer comment les principes de la présente Déclaration ont été pris en considération. Le protocole devrait inclure des informations concernant le financement, les promoteurs, les affiliations institutionnelles, les conflits d'intérêts potentiels, les incitations pour les personnes impliquées dans la recherche et des informations concernant les mesures prévues pour soigner et/ou dédommager celles ayant subi un préjudice en raison de leur participation à la recherche.

Dans les essais cliniques, le protocole doit également mentionner les dispositions appropriées prévues pour l'accès à l'intervention testée après l'essai clinique.

#### **Comités d'éthique de la recherche**

23. Le protocole de recherche doit être soumis au comité d'éthique de la recherche concerné pour évaluation, commentaires, conseils et approbation avant que la recherche ne commence. Ce comité doit être transparent dans son fonctionnement, doit être indépendant du chercheur, du promoteur et de toute autre influence indue et doit être dûment qualifié. Il doit prendre en considération les lois et réglementations du ou des pays où se déroule la recherche, ainsi que les normes et standards internationaux, mais ceux-ci ne doivent pas permettre de restreindre ou exclure l'une des protections garanties par la présente Déclaration aux personnes impliquées dans la recherche.

Le comité doit avoir un droit de suivi sur les recherches en cours. Le chercheur doit fournir au comité des informations sur le suivi, notamment concernant tout événement indésirable grave. Aucune modification ne peut être apportée au protocole sans évaluation et approbation par le comité. A la fin de la recherche, les chercheurs doivent soumettre au comité un rapport final contenant un résumé des découvertes et des conclusions de celle-ci.

#### **Vie privée et confidentialité**

24. Toutes les précautions doivent être prises pour protéger la vie privée et la confidentialité des informations personnelles concernant les personnes impliquées dans la recherche.

#### **Consentement éclairé**

25. La participation de personnes capables de donner un consentement éclairé à une recherche médicale doit être un acte volontaire. Bien qu'il puisse être opportun de consulter les membres de la famille ou les responsables de la communauté, aucune personne capable de donner un consentement éclairé ne peut être impliquée dans une recherche sans avoir donné son consentement libre et éclairé.

26. Dans la recherche médicale impliquant des personnes capables de donner un consentement éclairé, toute personne pouvant potentiellement être impliquée doit être correctement informée des objectifs, des méthodes, des sources de financement, de tout éventuel conflit d'intérêts, des affiliations institutionnelles du chercheur, des bénéfices escomptés et des risques potentiels de la recherche, des désagréments qu'elle peut engendrer, des mesures qui seront prises après à l'essai clinique et de tout autre aspect pertinent de la recherche. La personne pouvant potentiellement être impliquée dans la recherche doit être informée de son droit de refuser d'y participer ou de s'en retirer à tout moment sans mesure de rétorsion. Une attention particulière devrait être accordée aux besoins d'informations spécifiques de chaque personne pouvant potentiellement être impliquée dans la recherche ainsi qu'aux méthodes adoptées pour fournir les informations. Lorsque le médecin ou une autre personne qualifiée en la matière a la certitude que la personne concernée a compris les informations, il doit alors solliciter son consentement libre et éclairé, de préférence par écrit. Si le consentement ne peut pas être donné par écrit, le consentement non écrit doit être formellement documenté en présence d'un témoin.

Toutes les personnes impliquées dans des recherches médicales devraient avoir le choix d'être informées des conclusions générales et des résultats de celles-ci.

27. Lorsqu'il sollicite le consentement éclairé d'une personne pour sa participation à une recherche, le médecin doit être particulièrement attentif lorsque cette dernière est dans une relation de dépendance avec lui ou pourrait donner son consentement sous la contrainte. Dans ce cas, le consentement éclairé doit être sollicité par une personne qualifiée en la matière et complètement indépendante de cette relation.

28. Lorsque la recherche implique une personne incapable de donner un consentement éclairé, le médecin doit solliciter le consentement éclairé de son représentant légal. Les personnes incapables ne doivent pas être incluses dans une recherche qui n'a aucune chance de leur être bénéfique sauf si celle-ci vise à améliorer la santé du groupe qu'elles représentent, qu'elle ne peut pas être réalisée avec des personnes capables de donner un consentement éclairé et qu'elle ne comporte que des risques et des inconvénients minimes.

29. Lorsqu'une personne considérée comme incapable de donner un consentement éclairé est en mesure de donner son assentiment concernant sa participation à la recherche, le médecin doit solliciter cet assentiment en complément du consentement de son représentant légal. Le refus de la personne pouvant potentiellement être impliquée dans la recherche devrait être respecté.

30. La recherche impliquant des personnes physiquement ou mentalement incapables de donner leur consentement, par exemple des patients inconscients, peut être menée uniquement si l'état physique ou mental empêchant de donner un consentement éclairé est une caractéristique nécessaire du groupe sur lequel porte cette recherche.

Dans de telles circonstances, le médecin doit solliciter le consentement éclairé du représentant légal. En l'absence d'un représentant légal et si la recherche ne peut pas être retardée, celle-ci peut être lancée sans le consentement éclairé. Dans ce cas, le protocole de recherche doit mentionner les raisons spécifiques d'impliquer des personnes dont l'état les rend incapables de donner leur consentement éclairé et la recherche doit être approuvée par le comité d'éthique de la recherche concerné. Le consentement pour maintenir la personne concernée dans la recherche doit, dès que possible, être obtenu de la personne elle-même ou de son représentant légal.

31. Le médecin doit fournir des informations complètes au patient sur la nature des soins liés à la recherche. Le refus d'un patient de participer à une recherche ou sa décision de s'en retirer ne doit jamais nuire à la relation patient-médecin.

32. Pour la recherche médicale utilisant des tissus ou des données d'origine humaine, telles que les recherches sur tissus et données contenues dans les biobanques ou des dépôts similaires, les médecins doivent solliciter le consentement éclairé pour leur analyse, stockage et/ou réutilisation. Il peut se présenter des situations exceptionnelles où il est impraticable, voire impossible d'obtenir le consentement. Dans de telles situations, la recherche peut être entreprise uniquement après évaluation et approbation du comité d'éthique de la recherche concerné.

### **Utilisation de placebo**

33. Les bénéfices, les risques, les inconvénients, ainsi que l'efficacité d'une nouvelle intervention doivent être testés et comparés à ceux des meilleures interventions avérées, sauf dans les circonstances suivantes : lorsqu'il n'existe pas d'intervention avérée, l'utilisation de placebo, ou la non intervention, est acceptable ; ou lorsque pour des raisons de méthodologie incontournables et scientifiquement fondées l'utilisation de toute intervention moins efficace que la meilleure éprouvée, l'utilisation d'un placebo, ou la non intervention, est nécessaire afin de déterminer l'efficacité ou la sécurité d'une intervention, et lorsque les patients recevant une intervention moins efficace que la meilleure éprouvée, un placebo, ou une non intervention, ne courent pas de risques

supplémentaires de préjudices graves ou irréversibles du fait de n'avoir pas reçu la meilleure intervention éprouvée.

Le plus grand soin doit être apporté afin d'éviter tout abus de cette option.

#### **Conditions de l'accès à l'intervention testée après l'essai clinique**

34. En prévision d'un essai clinique, les promoteurs, les chercheurs et les gouvernements des pays d'accueil devraient prévoir des dispositions pour que tous les participants qui ont encore besoin d'une intervention identifiée comme bénéfique dans l'essai puissent y accéder après celui-ci. Cette information doit également être communiquée aux participants au cours du processus de consentement éclairé.

#### **Enregistrement des recherches, publication et dissémination des résultats**

35. Toute recherche impliquant des êtres humains doit être enregistrée dans une banque de données accessible au public avant que ne soit recrutée la première personne impliquée dans la recherche.

36. Les chercheurs, auteurs, promoteurs, rédacteurs et éditeurs ont tous des obligations éthiques concernant la publication et la dissémination des résultats de la recherche. Les chercheurs ont le devoir de mettre à la disposition du public les résultats de leurs recherches impliquant des êtres humains. Toutes les parties ont la responsabilité de fournir des rapports complets et précis. Ils devraient se conformer aux directives acceptées en matière d'éthique pour la rédaction de rapports. Les résultats aussi bien négatifs et non concluants que positifs doivent être publiés ou rendus publics par un autre moyen. La publication doit mentionner les sources de financement, les affiliations institutionnelles et les conflits d'intérêts. Les rapports de recherche non-conformes aux principes de la présente Déclaration ne devraient pas être acceptés pour publication.

#### **Interventions non avérées dans la pratique clinique**

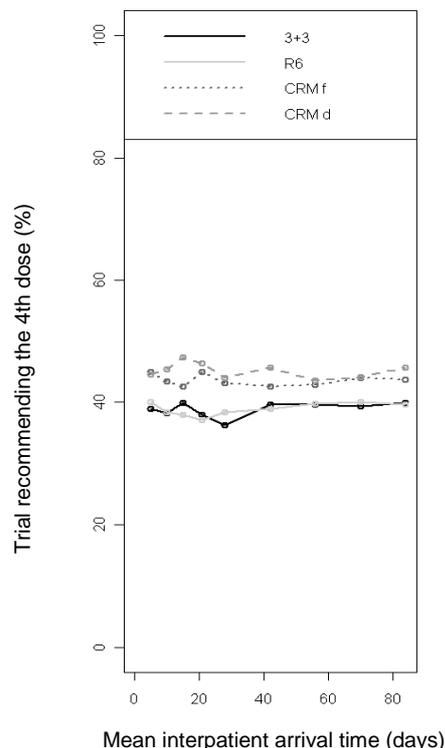
37. Dans le cadre du traitement d'un patient, faute d'interventions avérées ou faute d'efficacité de ces interventions, le médecin, après avoir sollicité les conseils d'experts et avec le consentement éclairé du patient ou de son représentant légal, peut recourir à une intervention non avérée si, selon son appréciation professionnelle, elle offre une chance de sauver la vie, rétablir la santé ou alléger les souffrances du patient. Cette intervention devrait par la suite faire l'objet d'une recherche pour en évaluer la sécurité et l'efficacité. Dans tous les cas, les nouvelles informations doivent être enregistrées et, le cas échéant, rendues publiques.

## Annexe 2. Annexes D et F à J de l'article (Doussau, Asselain et al. 2012)

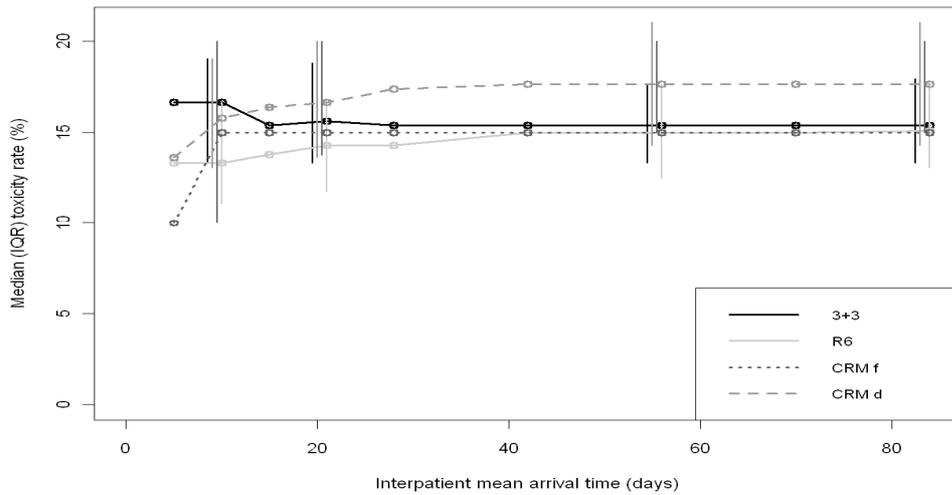
Les annexes A, B, C et E ne sont pas incluses, car les résultats ont été détaillés au §2.3.

### Appendix D. Rules for reanalysis of the erlotinib+RT trial

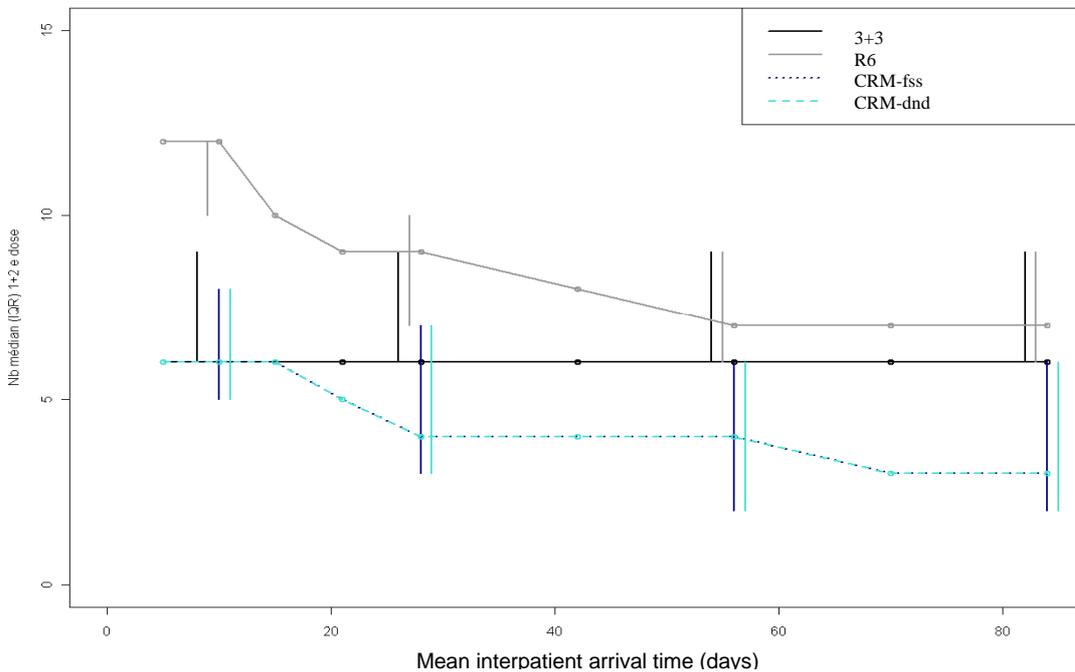
- If accrual was stopped according to the simulated designs at a timepoint at which a child had been enrolled in the erlotinib+RT clinical trial, then the child was not included and was counted as a skipped child.
- If a child did not develop DLT in the erlotinib+RT trial at a given dose, it was assumed that he/she would have tolerated treatment at all lower doses.
- If a child was supposed to receive a higher dose of treatment in the simulation than in the erlotinib+RT clinical trial, the same inclusion date was used but the toxicity outcome was based on the first patient treated at this higher dose. In this case, the toxicity outcomes of the remaining patients were all shifted.
- When all 20 children had been enrolled in the simulation, accrual of supplementary children was simulated using an arrival time of 26 days (corresponding to the observed mean arrival time in the erlotinib+RT trial). Different options were proposed according to toxicity or non-toxicity profile of the simulated children.



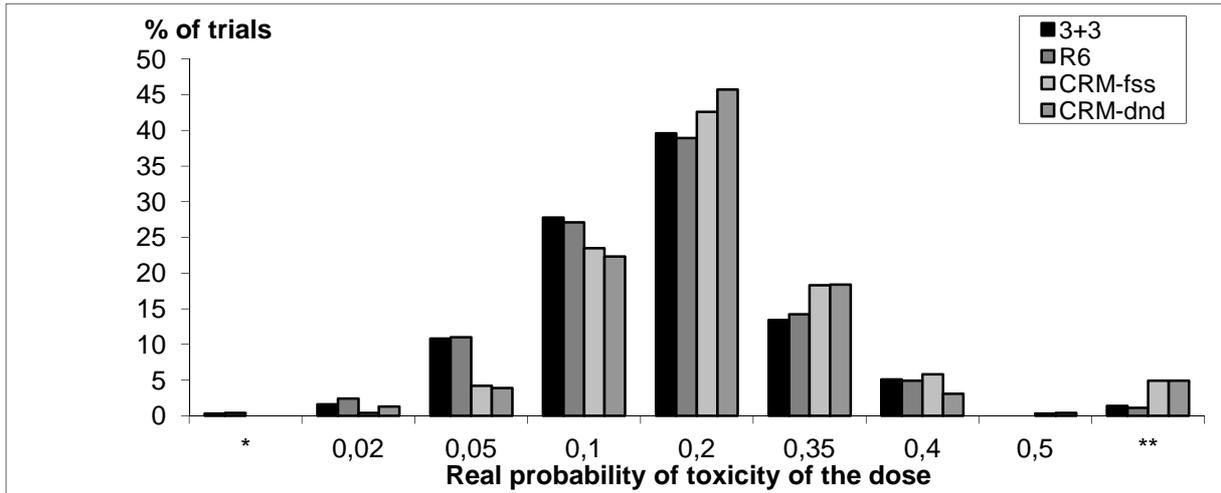
**Appendix F. Proportion of recommendations of true RP2D, according to various mean interpatient arrival times, in 1,000 simulated studies with 4 different designs for each situation.** Cycle length=21 days, true RP2D at the 4th dose. RP2D: recommended phase 2 dose, R6: Rolling 6, CRM-f: CRM with fixed sample size stopping rule, CRM-d: CRM with defined number of patients at a dose stopping rule.



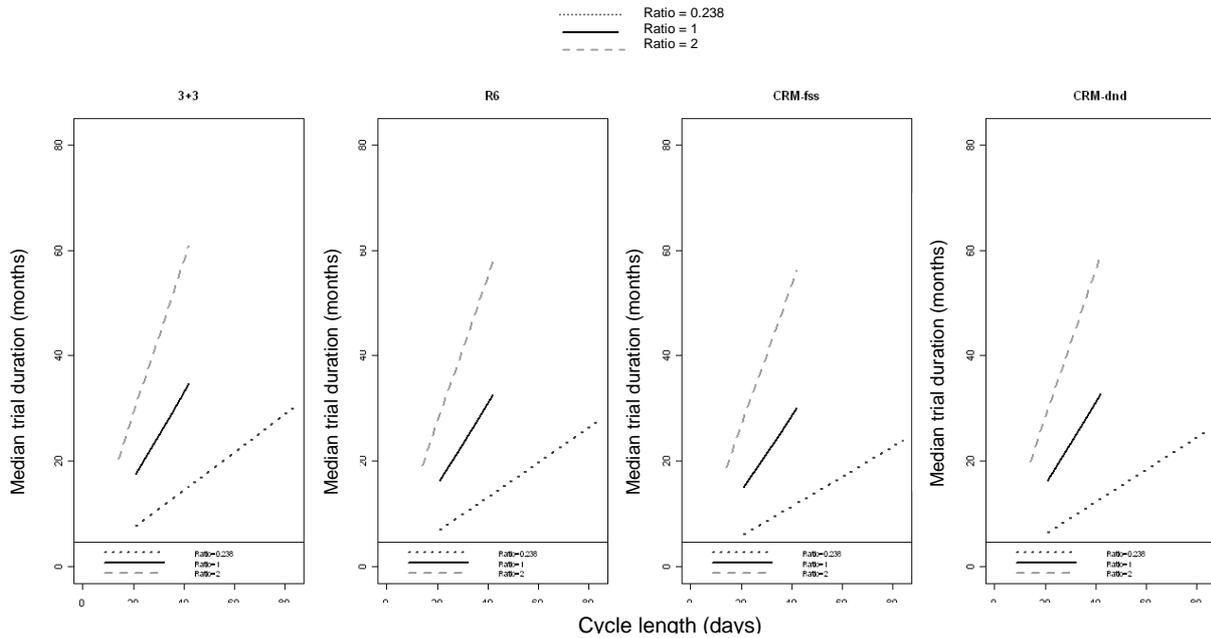
**Appendix G. Median (IQR) overall toxicity rate, according to various mean interpatient arrival times, in 1,000 simulated studies with 4 different designs for each situation.** Cycle length: 21 days, true RP2D at the 4th dose. RP2D: recommended phase 2 dose, R6: Rolling 6, CRM-f: CRM with fixed sample size stopping rule, CRM-d: CRM with defined number of patients at a dose stopping rule.



**Appendix H. Median (IQR) number of patients treated at dose 1 and 2, according to various mean interpatient arrival times, in 1,000 simulated studies with 4 different designs for each situation.** Cycle length: 21 days, true RP2D at the 4th dose. RP2D: recommended phase 2 dose, R6: Rolling 6, CRM-fss: CRM with fixed sample size stopping rule, CRM-dnd: CRM with defined number of patients at a dose stopping rule, IQR: interquartile range.



**Appendix I. Distribution of the recommended phase II dose (RP2D) over 1,000 simulations of each design.** Real toxicity corresponds to the probability of toxicity allocated to a dose, in our setting in which dose toxicity is known. Mean interpatient arrival time = 42 days, cycle length = 21 days, true RP2D at the 4th dose (20% toxicity dose). R6: Rolling 6, CRM-f: CRM with fixed sample size stopping rule, CRM-d: CRM with defined number of patients at a dose stopping rule. \*Interruption because of 2 DLT at the first level, \*\* interruption because the last level was reached and the design would require exploration of an upper level (3+3 and R6) or because no toxicity occurred up to the last level during initiation of CRM.



**Appendix J. Median trial duration, according to various cycle lengths (days) and according to different mean interpatient arrival time / cycle length ratios, in 1,000 simulated studies with 4 different designs for each situation.** True RP2D at the 4<sup>th</sup> dose.

### Annexe 3. Calcul de la pente pour le modèle logistique à un paramètre de la CRML

Soit  $TDL_i$  la variable correspondant à la survenue d'une TDL pour le patient  $i$ .

Le modèle utilisé par la CRML pour l'attribution des doses est le suivant :

$$\psi^{-1}(x_i, \theta) = \log\left(\frac{P(TDL_i = 0|x_i)}{1 - P(TDL_i = 0|x_i)}\right) = \alpha_0 - \beta_c x_i \quad \text{(Modèle 11)}$$

pour  $TDL_i = (0,1)$

Avec un paramètre à estimer :  $\theta = \alpha_0$ .

La méthode utilisée pour calculer  $\beta_c$  consiste à faire la moyenne des coefficients  $\beta$  qui correspondent à une variation de la probabilité de toxicité passant de 20 % à 35 %, entre deux niveaux de dose adjacents. Nous avons effectué cette moyenne pour les cinq dernières doses de la série des doses à explorer.

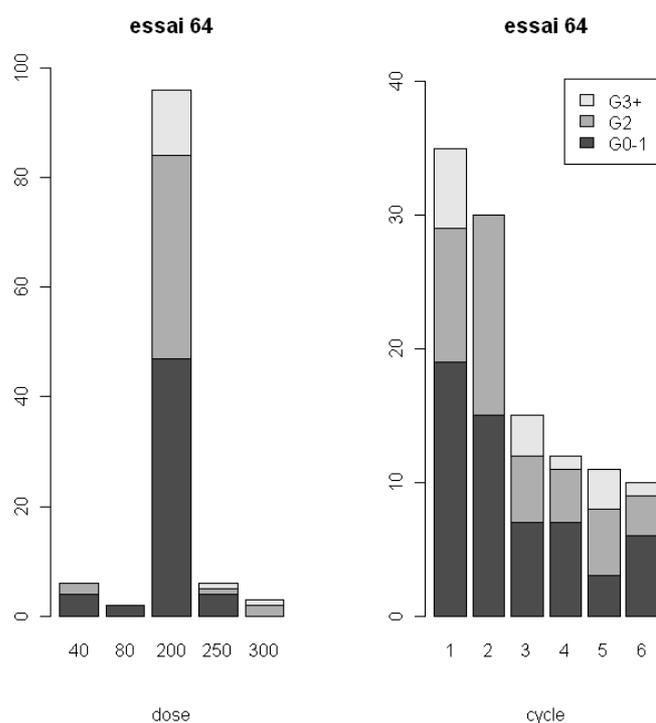
Pour les doses explorées dans la première étude de simulation, cela correspondait à une pente  $\beta_c = 0,0037$  (Doussau, Asselain et al. 2012). Pour les doses explorées dans la simulation étudiant le POMM, cela correspondait à une pente  $\beta_c = 2,42$  (Doussau, Thiebaut et al. 2013).

#### Annexe 4. Présentation des données d'un essai « essai 64 ». Syntaxe R, SAS et test de l'hypothèse de proportionnalité des cotes

Données d'un essai (« essai 64 »), issues de la base de données de deux centres d'essais de phase I (Postel-Vinay, Gomez-Roca et al. 2011).

113 cycles administrés sur 35 patients inclus (Figure 16), à 5 doses différentes (de 40 à 300, soit 0,4 à 3 unités décimales).

Grade maximum sur 113 cycles : G3+ (14 cycles), G2 (42 cycles).



**Figure 16. Répartition de distribution de la toxicité maximale par cycle, selon la dose (gauche) et le cycle (droite).** Données d'un essai issu de (Postel-Vinay, Gomez-Roca et al. 2011).

- **Syntaxe et estimations avec le logiciel R, fonction clmm du *package* Ordinal.**

Le modèle complet (modèle 7) s'écrit :

```
model_dose_temps <- clmm(factor(var_reponse) ~ temps + dose + (1|id),
                          data = data_modele)
summary(model_dose_temps)
```

La sortie resultant est :

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: factor(var\_reponse) ~ temps + dose + (1 | id)  
data: data\_modele

link	threshold	nobs	logLik	AIC	niter	max.grad	cond.H
logit	flexible	113	-102.39	214.78	17(564)	7.10e-06	5.6e+02

Random effects:

	Var	Std.Dev
id	1.531	1.237

Number of groups: id 35

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
temps	0.08609	0.13597	0.633	0.5266
dose	1.26262	0.76098	1.659	0.0971 .

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value
1 2	2.682	1.572	1.706
2 3	5.182	1.668	3.106

Les paramètres estimés sont ainsi :  $\hat{\theta} = (\hat{\alpha}_1 = 2,682, \hat{\alpha}_2 = 5,182, \hat{\beta}_1 = 1,263, \hat{\beta}_2 = 0,086, \hat{\sigma}_0^2 = 1,531)$ .

- **Estimations avec le logiciel SAS, procédure nlmixed**

On obtient des estimations quasi identiques du modèle 7 par les deux logiciels, à la troisième décimale près :  $\hat{\theta} = (\hat{\alpha}_1 = 2,688, \hat{\alpha}_2 = 5,188, \hat{\beta}_1 = 1,265, \hat{\beta}_2 = 0,086, \hat{\sigma}_0^2 = 1,534)$ .

L'hypothèse de proportionnalité des cotes sur la dose est respectée ( $p = 0,995$ ).

### Syntaxe avec le logiciel SAS

```
Proc NLMIXED data=tmp corr ecorr optcheck tech=congra;
parms b01=2.495 b02=-3.347 b1=0 b2=0 s20=0.035 ;
bounds b01>0, s20>=0;
eta1= b01+b02      + (b1)*dose + (b2)*temps  + u;
eta2= b02          + (b1)*dose + (b2)*temps  + u;
/* likelihood computation */
expeta1 = exp(eta1);
expeta2 = exp(eta2);
p0 = 1-(expeta1/(1+expeta1));
p1 = 1-(expeta2/(1+expeta2))-p0;
p2 = 1-p0-p1;
L = (p0**y0)*(p1**y1)*(p2**y2);
ll=log(L);
random u ~ normal(0,s20) subject=id;
model var_reponse ~ general(ll);run;
```

#### The NLMIXED Procedure

##### Spécifications

Data Set	WORK.TMP
Dependent Variable	var_reponse
Distribution for Dependent Variable	General
Random Effects	u
Distribution for Random Effects	Normal
Subject Variable	id
Optimization Technique	Conjugate-Gradient
Integration Method	Adaptive Gaussian Quadrature

##### Dimensions

Observations Used	113
Observations Not Used	0
Total Observations	113
Subjects	35
Max Obs Per Subject	6
Parameters	5
Quadrature Points	1

##### Paramètres

b01	b02	b1	b2	s20	NegLogLike
2.495	-3.347	0	0	0.035	122.159552

NOTE: GCONV critère de convergence exécuté.

Statistiques d'ajustement	
-2 Log Likelihood	204.8
AIC (smaller is better)	214.8
AICC (smaller is better)	215.3
BIC (smaller is better)	222.6

Résultats estimés des paramètres						
Paramètre	Estimation	Erreur standard	DF	Valeur du test t	Pr >  t	Alpha
b01	2.5001	0.3876	34	6.45	<.0001	0.05
b02	-5.1882	1.6783	34	-3.09	0.0040	0.05
b1	1.2652	0.7651	34	1.65	0.1074	0.05
b2	0.08597	0.1360	34	0.63	0.5314	0.05
s20	1.5337	0.9223	34	1.66	0.1055	0.05

Dans cette syntaxe utilisée avec le logiciel SAS, dérivée de celle proposée par (Pedrono, Thiebaut et al. 2009), la correspondance entre les *intercept* estimés par le modèle avec le logiciel R est la suivante :

```
b01+b02 = -(alpha_1|2)
b02 = - (alpha_2|3)
```

Cela est lié à l'écriture du modèle : d'une part, modélisation de logit ( $p \leq k$ ) avec R, versus logit ( $p \geq k$ ) avec SAS, et, d'autre part, formulation additive des *intercept* dans l'écriture du modèle avec SAS.

Au final, le modèle sous SAS estime les paramètres suivants :

$\theta = (\alpha_1 = 2,6881, \alpha_2 = 5,1882, \beta_1 = 1,2652, \beta_2 = 0,08597, \sigma_0^2 = 1,5337)$ , quasi identiques aux estimations faites avec le logiciel R.

### Test de l'hypothèse de proportionnalité des cotes : exemple du test sur la dose.

Modèle logistique généralisé ordonné pour la dose (Modèle 9), et à cotes proportionnelles pour le temps :

```
Proc NLMIXED data=tmp corr ecorr optcheck tech=congra;
parms b01=2.495 b02=-3.347 b1=0 b12=0 b2=0 s20=0.035 ;
bounds b01>0, s20>=0;
eta1= b01+b02 + (b1+b12)*dose + (b2)*temps + u;
eta2= b02 + (b1)*dose + (b2)*temps + u;
/* likelihood computation */
expeta1 = exp(eta1);
expeta2 = exp(eta2);
p0 = 1-(expeta1/(1+expeta1));
p1 = 1-(expeta2/(1+expeta2))-p0;
p2 = 1-p0-p1;
L = (p0**y0)*(p1**y1)*(p2**y2);
ll=log(L);
random u ~ normal(0,s20) subject=id;
model var_reponse ~ general(ll); run;
```

The NLMIXED Procedure  
Spécifications

Data Set	WORK.TMP
Dependent Variable	var_reponse
Distribution for Dependent Variable	General
Random Effects	u
Distribution for Random Effects	Normal
Subject Variable	id
Optimization Technique	Conjugate-Gradient
Integration Method	Adaptive Gaussian Quadrature

Dimensions

Observations Used	113
Observations Not Used	0
Total Observations	113
Subjects	35
Max Obs Per Subject	6
Parameters	6
Quadrature Points	1

Paramètres

b01	b02	b1	b12	b2	s20	NegLogLike
2.495	-3.347	0	0	0	0.035	122.159552

NOTE: GCONV critère de convergence exécuté.

Statistiques d'ajustement

-2 Log Likelihood	204.4
AIC (smaller is better)	216.4
AICC (smaller is better)	217.1
BIC (smaller is better)	225.7

Résultats estimés des paramètres

Paramètre	Estimation	Erreur standard	DF	Valeur du test t	Pr >  t	Alpha
b01	3.8242	2.1337	34	1.79	0.0820	0.05
b02	-6.2807	2.4315	34	-2.58	0.0143	0.05
b1	1.7877	1.1250	34	1.59	0.1213	0.05
<b>b12</b>	<b>-0.6468</b>	<b>1.0048</b>	<b>34</b>	<b>-0.64</b>	<b>0.5241</b>	<b>0.05</b>
b2	0.08975	0.1362	34	0.66	0.5143	0.05
s20	1.5199	0.9126	34	1.67	0.1050	0.05

Test de l'hypothèse de proportionnalité des cotes sur la dose, par le test du rapport de vraisemblance selon un  $\chi^2$ (1 degré de liberté).  $H_0 : b_{12} = 0$

-2LL GOMM = 204,4

-2LL POMM = 204,8

L'hypothèse de proportionnalité des cotes sur l'effet de la dose n'est pas rejetée,  $p = 0,995$ .

**Annexe 5. Communication affichée présentée au Congrès européen du cancer 2013  
(Doussau, Thiebaut et al. 2013)**

# A new approach to integrate the grade of toxicity and later cycles in the analysis and reporting of phase I dose finding trials

A Doussau<sup>1,2,3</sup>, R Thiébaud<sup>2,3</sup>, B Geoger<sup>4</sup>, MC Le Deley<sup>4</sup>, P Schöffski<sup>5</sup>, A Floquet<sup>6</sup>, S Mathoulin-Pélessier<sup>3,6</sup>, E Rizzo<sup>7</sup>, C Le Tourneau<sup>4</sup>, P Fumoleau<sup>8</sup>, X Paoletti<sup>1</sup>

1. Institut Curie, INSERM U900, Paris ; 2. INSERM U897, CIC-EC7, CHU de Bordeaux ; 3. Univ. Bordeaux Segalen, CIC-EC7, Bordeaux ; 4. Institut Gustave Roussy ; Univ. Paris-Sud XI, Villejuif ; 5. Univ. Hospitals Leuven, KU Leuven ; 6. Institut Bergonié, Bordeaux ; 7. EORTC, Bruxelles ; 8. Centre de Lutte contre le Cancer Georges-François Leclerc, Dijon.



## Background:

- The maximum tolerated dose (MTD) is commonly defined as: dose of a given treatment associated with a certain proportion of dose limiting toxicities (DLT), evaluated during a predefined time window (usually one or two cycles of treatment).
- Adverse events (AE) are assessed sequentially, mostly using NCI-CTCAE grading criteria.
- Working hypothesis: longitudinal assessment of AEs can provide relevant additional information for defining the optimal dose and schedule of a new treatment.
- A new approach was developed that uses all the data collected at the end of the trials (Doussau et al., 2013).
- **Aim: To illustrate our approach, we retrospectively reanalyzed three phase I clinical trials of anticancer agents.**

## Results:

The description of the three trials is summarized in table 1 and figure 1.

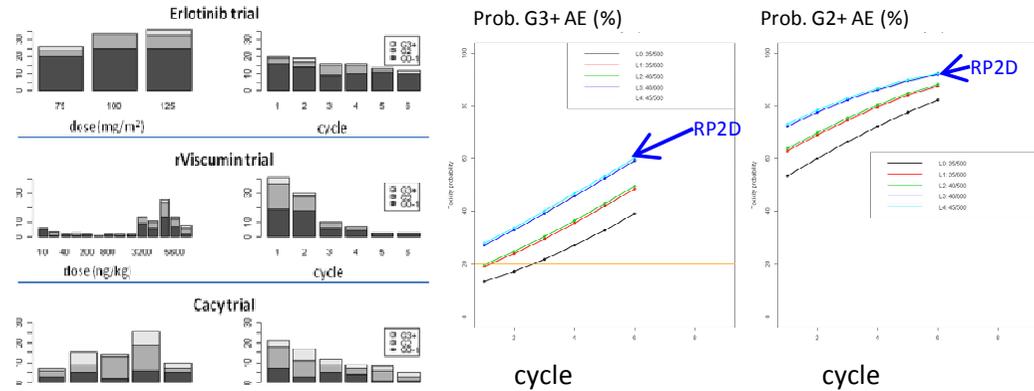
**Table 1. Description of the three trials.** G=grade, AE=adverse event, pts=patients, DLT=dose limiting toxicity.

Trial	Erlotinib	rViscumin	Cacy
Design	CRM	CRM	3+3
Nb pts included	20 pts	41 pts	21 pts
Nb levels explored	3 (75 to 125 mg/m <sup>2</sup> )	14 (10 to 6400 ng/kg)	5 levels *
Nb DLTs	2 DLTs	4 DLTs	2 DLTs
Nb cycles	96 cycles	94 cycles	72 cycles
Worst grade / cycle	7 G3+, 19 G2	7 G3, 34 G2	13 G3+, 33 G2
AE during DLT evaluation period	G3+ : fatal seizures (DLT) ; skin rash (DLT) ; vomiting ; G2 : digestive or cutaneous toxicities, asthenia	G3 : asthenia and liver toxicity (4 DLTs) ; G2 : hypersensitivity, gastrointestinal and liver toxicities, hypoalbuminemia, lymphopenia, and dyspnoea.	G3+ : asthenia, hematologic or gastro-intestinal toxicities (2 DLTs) ; G2 : idem, along with cutaneous toxicities
Later AE	G3+ : pneumopathy, surgical intervention, cutaneous ; G2 : cutaneous, digestive and asthenia	G3 : asthenia and liver toxicity ; G2 : asthenia, gastrointestinal and liver toxicities, hypoalbuminemia, lymphopenia, pain and increased blood creatinine.	idem DLT evaluation period.

\* Cacy trial explored the following combinations of doses : doxLipeg / cyclophosphamide mg/m<sup>2</sup> : 35/500, 35/600, 40/500, 40/600, 45/500.

## Material and methods:

- Individual patient data from 3 completed phase I clinical trials were retrieved (table 1) : erlotinib with radiotherapy for children with brainstem gliomas, rViscumine for adults with solid tumours, and combination of DoxLipeg (liposomal doxorubicin) + cyclophosphamide for women with ovarian carcinoma (Cacy trial).
- **Outcome** of interest: worst grade (G) of toxicity in each cycle, in three grading categories (G0-1/G2/G3-5, Figure 1). Analysis restricted to toxicities related to treatment, during cycles 1 to 6.
- We estimated the probability of G2-5 and G3-5 toxicity per cycle for each evaluated dose level (mixed effect **proportional odds model** to account for repeated **longitudinal** measurements, **LPOM**). We report : the risk of toxicity over time (=cycle), the probability of severe toxicity at the recommended phase II dose (RP2D) of the trial, with its 95% confidence interval (95%CI) and the probability of moderate or severe toxicity.



**Figure 1.** Distribution of the number of patients according to dose level (left part) or cycle (right part), and worst grade toxicity.

**Figure 2.** Estimated probability of G3+ AE (left part) or G2+ AE (right part) according to the cycle; longitudinal reanalysis of Cacy trial . RP2D = recommended phase II dose (dark blue line).

**Erlotinib**

Trial : RP2D = 125 mg/m<sup>2</sup> (3<sup>rd</sup> dose level)  
 Prob. DLT at RP2D : 16% (95%CI=4%–45%)

**Longitudinal reanalysis :**

No increased risk of toxicity with time.  
 Prob G3-5 per cycle at RP2D : 9% (95%CI=4%-19%)  
 Prob G2-5 per cycle at RP2D : 31%

LPOM

**rViscumin**

Trial : RP2D = 5600 ng/kg (13<sup>th</sup> dose level)  
 Prob. DLT at RP2D : 16% (7%–37%)

**Longitudinal reanalysis :**

No increased risk of toxicity with time.  
 Prob G3-5 per cycle at RP2D : 14% (95%CI= 6%-30%)  
 Prob G2-5 per cycle at RP2D : 61%

LPOM

**Cacy**

Trial : RP2D = 40/600 (4<sup>th</sup> dose level)

**Longitudinal reanalysis :**

Increased risk of toxicity with time (p=0.03). Figure 2 represents the estimated probability of graded toxicity of the different levels of doses explored, according to time (=cycle).

LPOM

## References:

- Doussau A, Thiébaud R, Paoletti X. Dose-finding design using mixed-effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials. Stat Med. 2013.
- Geoger B, et al. Innovative Therapies for Children with Cancer pediatric phase I study of erlotinib in brainstem glioma and relapsing/refractory brain tumors. Neuro Oncol. 2011.
- Schöffski P et al. Phase I trial of intravenous aviscumine (rViscumine) in patients with solid tumors. Ann Oncol. 2004.
- Floquet A, et al. Pegylated liposomal doxorubicin and cyclophosphamide in early recurrent ovarian carcinoma: Phase I dose-finding study (under publication).

**Acknowledgement:** INCa (grant for Optidose Research project), Institut Bergonié, EORTC, Institut Gustave Roussy (for data providing).

## Conclusion:

This reanalysis using the AE collected beside the DLT evaluation period, through **LPOM** approach:

- revealed an increased risk of toxicity with time in 1/3 trials (Cacy : liposomal doxorubicin + cyclophosphamide combination trial).
- enabled an accurate estimate of the per cycle burden of moderate or severe toxicity of the RP2D.
- the extra information used did not reveal any specific type of late toxicities.

Dedicated methods, whose operating characteristics were evaluated elsewhere (Doussau et al., 2013), allow analyzing toxic adverse events occurring during all cycles of treatment. They should be integrated in the recommended phase II dose assessment.

**Annexe 6. Chapitre de livre (Paoletti and Doussau 2014)**

P1

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

P2

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

P3

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

P4

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

P5

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

P6

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

P7

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p8

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p9

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p10

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p11

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p12

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p13

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p14

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p15

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p16

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p17

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p18

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p19

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p20

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p 21

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p22

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p23

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p24

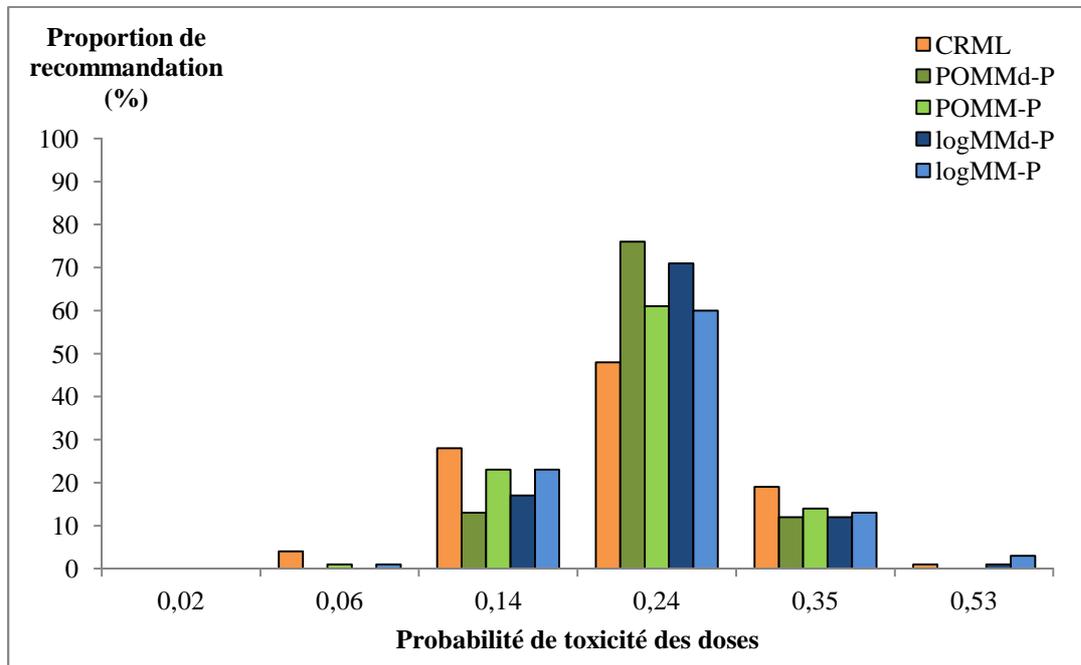
Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

p25

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

Paoletti X, Doussau A. *Dose Finding Methods in Oncology: From the Maximum Tolerated Dose to the Recommended Phase II Dose*. In: van Montfort K, Oud J, Ghidry W, editors. *Developments in Statistical Evaluation of Clinical Trials*. New York: Springer; 2014. p. 333-58. (in press)

**Annexe 7. Distribution des doses recommandées dans les nouveaux schémas adaptatifs utilisant les modèles mixtes logistique ou ordinal pour l'attribution des doses**



**Figure 17. Distribution des doses recommandées sur 1 000 échantillons d'essais simulés selon un schéma CRML, ou selon des schémas adaptatifs utilisant un modèle longitudinal pour l'attribution des doses (n = 30 patients, recevant 1 à 6 cycles de traitement, jusqu'à la première toxicité grave). CRML : méthode de réévaluation continue, par le maximum de vraisemblance, POMM : proportionnal odds mixed model, LogMM : modèle mixte logistique. Le d en indice signifie que le paramètre de la dose n'est pas estimé par le modèle. Le P signifie que les données longitudinales sont utilisées prospectivement pour l'attribution des doses, par opposition aux résultats de la Figure 13 13 qui concernent l'utilisation des modèles pour l'analyse rétrospective d'essais.**

**Annexe 8. Article comparant le modèle mixte longitudinal au modèle mixte à cotes proportionnelles, et l'impact de la simplification du modèle (soumis).**

P1

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

P2

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

P3

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

P4

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

P5

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

P6

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

P7

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

p8

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

p9

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

p10

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

p11

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

p12

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

p13

Paoletti X, Doussau A, Rizzo E, Thiébaud R. *Dose finding with longitudinal data: simpler models, richer outcomes* (soumis)

La sagesse est resplendissante, elle est inaltérable. Elle se laisse aisément contempler par ceux qui l'aiment, elle se laisse trouver par ceux qui la cherchent.

Elle devance leurs désirs en se montrant à eux la première.

Celui qui la cherche dès l'aurore ne se fatiguera pas : il la trouvera assise à sa porte.

(Sg 6, 12-14)

