



Estimation du bénéfice de survie à partir de méta-analyse sur données individuelles et évaluation économique.

Béranger Lueza

► To cite this version:

Béranger Lueza. Estimation du bénéfice de survie à partir de méta-analyse sur données individuelles et évaluation économique.. Santé publique et épidémiologie. Université Paris-Saclay, 2016. Français. NNT : 2016SACLS268 . tel-01396362

HAL Id: tel-01396362

<https://theses.hal.science/tel-01396362>

Submitted on 14 Nov 2016

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.

NNT : 2016SACLS268

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'UNIVERSITE PARIS-SUD

ECOLE DOCTORALE N°570
SANTÉ PUBLIQUE

Spécialité de doctorat : Biostatistiques

Par

M. Béranger LUEZA

Estimation du bénéfice de survie à partir de méta-analyse
sur données individuelles et évaluation économique.

Thèse présentée et soutenue à Villejuif, le 30 Septembre 2016 :

Composition du Jury :

M. Aurélien LATOUCHE
M. Roch GIORGI
M. Raphaël PORCHER
M. Bertrand BAUJAT
M. Jean-Pierre PIGNON
Mme. Julia BONASTRE

PU, Conservatoire National des Arts et Métiers
PU-PH, Université Aix-Marseille
MCU-PH, HDR, Université Paris Diderot
PU-PH, Université Pierre et Marie Curie
PhD, PH, HDR, Université Paris-Saclay
PhD, Université Paris-Saclay

Président
Rapporteur
Rapporteur
Examineur
Directeur de thèse
Co-directrice de thèse

Avant-propos

Cette thèse a été réalisée au sein du Service de Biostatistique et d'Epidémiologie (SBE), dirigé par le Dr Ellen Benhamou et au sein de l'unité Oncostat – INSERM U1018 du Centre de recherche en Epidémiologie et Santé des Populations (CESP), dirigée par Stefan Michiels. Ces deux services sont situés à Gustave Roussy, 114 rue Edouard Vaillant, 94805, Villejuif, France.

Une collaboration avec Oliver Rivero-Arias, *Assistant Professor* en économie de la santé à l'Université d'Oxford a été développée lors de cette thèse. Suite à un premier séjour à l'université d'Oxford de deux semaines lors de la deuxième année, un nouvel échange a eu lieu lors de la troisième année de thèse. Nous avons passé une semaine à Oxford et Mr Rivero-Arias a été accueilli une semaine à Gustave Roussy.

Pour le deuxième projet de cette thèse, nous avons collaboré avec Federico Rotolo et Stefan Michiels, tous deux biostatisticiens à Gustave Roussy (SBE et Oncostat – INSERM U1018 du CESP).

Enfin pour le troisième projet de cette thèse, nous avons collaboré avec les Dr Cécile Le Péchoux et Dr Rodrigo Arriagada, radiothérapeutes à Gustave Roussy, et le Dr Dirk de Ruyscher, radiothérapeute à la Maastric Clinic de Maastricht, Pays-Bas, ainsi qu'avec le *RadioTherapy Timing in Small Cell Lung Cancer Collaborative Group* lors de la réalisation d'une méta-analyse sur données individuelles.

Cette thèse a été financée par l'ITMO Cancer et l'IRESP dans le cadre du Plan Cancer 2009-2013, par le Programme Hospitalier de Recherche Clinique, et par la Ligue Nationale contre le Cancer que nous remercions.

Directeur de thèse : Jean-Pierre Pignon

Co-directrice de thèse : Julia Bonastre

Remerciements

Je tiens à remercier Pr Aurélien Latouche d'avoir accepté d'examiner ce travail de thèse et de présider mon jury de thèse. Merci à Pr Roch Giorgi et Dr Raphaël Porcher d'avoir accepté d'être rapporteurs de ma thèse. Je tiens à remercier également Pr Bertrand Baujat d'avoir accepté de faire partie du jury et d'apporter sa vision clinique. Je vous exprime à tous les quatre ma gratitude pour avoir pris le temps d'évaluer ce travail.

J'adresse mes très chers remerciements à Julia Bonastre et Jean-Pierre Pignon, un tandem de co-directeurs de thèse épatant et stimulant. Vous avez accepté de m'encadrer pendant cette thèse et m'avez fait confiance pendant 4 ans, me laissant une autonomie nouvelle et appréciable, tout en me guidant quand le besoin s'en faisait sentir. Je vous suis reconnaissant de l'aide que vous m'avez apportée pour appréhender chacun à votre manière deux thématiques qui me sont dorénavant chères : l'économie de la santé et la méta-analyse.

Julia : Merci pour cette discussion concernant une hypothétique thèse, un beau jour de janvier 2012 à Aix-en-Provence : je n'ai jamais regretté une seule seconde mon choix. Merci de m'avoir accompagné tout le long de cette route, en sachant quand me réconforter ou me recadrer. Au final, une thèse avec toi... ça déchire !

Jean-Pierre : Merci pour votre soutien, votre exigence, vos conseils avisés, votre regard à la fois méthodologique et clinique, et votre disponibilité à chaque instant. Vous m'avez fait découvrir le vaste monde de la méta-analyse et de la recherche clinique en général, et je vous en remercie.

Je remercie Dr Ellen Benhamou de m'avoir accueilli au sein du service de biostatistiques et d'épidémiologie, ainsi que Stefan Michiels au sein de l'équipe Oncostat - INSERM U1018 du CESP : il fait bon vivre sous votre toit !

Je remercie chaudement Federico Rotolo pour m'avoir *supporté* scientifiquement, R-codistiquement, et surtout amicalement ces dernières années. C'est un plaisir de passer chaque journée à tes côtés.

Stefan et Federico, je tiens à vous remercier d'avoir accepté de superviser mon deuxième projet de thèse. Je suis heureux d'avoir fait un bout de thèse avec vous et je suis très fier du fruit de notre collaboration. Merci à vous deux!

I would also like to thank Oliver Rivero-Arias for his help, his support and his kindness since the first time we met in Aix-en-Provence.

J'adresse mes remerciements aux Dr Cécile Le Péchoux, Dr Dirk De Ruyscher et Dr Rodrigo Arriagada pour leur collaboration sur notre méta-analyse.

Je tiens à remercier Jean Bouyer, directeur de l'Ecole Doctorale de Santé Publique, pour sa bienveillance permanente et son aide dans les derniers mois de cette thèse. Merci également à Audrey Bourgeois pour son accompagnement et le temps pris pour chaque doctorant.

Je me permets de remercier les investigateurs et les patients des essais inclus dans les méta-analyses. Un statisticien, bien qu'éloigné des salles de consultation, sait humblement ce que l'on doit aux patients ayant pris part à une étude de recherche clinique.

Mes remerciements vont aussi à mes collègues du Service de Biostatistiques et d'Epidémiologie de Gustave Roussy, et particulièrement à Nils Ternès pour avoir partagé avec moi les joies et les appréhensions de la thèse pendant 3 ans.

Je remercie ma famille et mes amis pour avoir pris des nouvelles de ma thèse ces quatre années durant et m'avoir encouragé. Merci à ma mère et mon père de m'avoir transmis les fibres scientifique d'une part et littéraire d'autre part, toutes deux nécessaires pour écrire ce manuscrit de thèse. Je remercie ma mère, ma sœur et mes beaux-parents qui nous ont permis de souffler ces dernières années, facilitant ainsi certains soirs l'écriture de mes articles et de mon manuscrit de thèse. Merci aussi à mes frères et mes grands-parents, passer des moments avec vous me ressourçait et me remplissait d'énergie.

Ett tusen sexton tack Josefin pour m'avoir incité à me lancer dans cette aventure, et avoir été là chaque jour à mes côtés. Tu es co-auteure de ma plus merveilleuse œuvre pendant cette thèse : Alva. Cette thèse je vous la dédie à toutes les deux, *mina älskelingar... Jag älskar er!*

Résumé

Le bénéfice de survie restreint à un horizon temporel donné a été proposé comme mesure alternative aux mesures relatives de l'effet d'un traitement notamment dans le cas de non proportionnalité des risques de décès. Le bénéfice de survie restreint correspond à la différence des survies moyennes entre deux bras de traitement et s'exprime en nombre d'années de vie gagnées. Dans la littérature, cette mesure est présentée comme plus intuitive que le *hazard ratio* et plusieurs auteurs ont proposé son utilisation pour le design et l'analyse d'un essai clinique. Toutefois, ce n'est pas actuellement la mesure qui est utilisée de manière courante dans les essais randomisés. Cette mesure s'applique quelle que soit la distribution des temps de survie et est adaptée si l'hypothèse des risques proportionnels n'est pas respectée. De plus, le bénéfice de survie restreint peut être utilisé en évaluation médico-économique où la mesure d'un effet absolu est nécessaire (nombre d'années de vie gagnées pondérées ou non par la qualité de vie). Si l'on souhaite estimer le bénéfice de survie restreint à partir d'une méta-analyse sur données individuelles, se pose alors la question de prendre en compte l'effet essai dû à la structure hiérarchique des données. L'objectif de cette thèse était de comparer des méthodes statistiques d'estimation du bénéfice de survie restreint à partir de données individuelles d'une méta-analyse d'essais cliniques randomisés. Le point de départ a été une étude de cas (étude coût-efficacité) réalisée à partir des données de la *Meta-Analysis of Radiotherapy in Lung Cancer*. Cette étude a montré que les cinq méthodes d'estimation étudiées conduisaient à des estimations différentes du bénéfice de survie et de son intervalle de confiance. Le choix de la méthode d'estimation avait également un impact sur les résultats de l'analyse coût-efficacité. Un second travail a consisté à mener une étude de simulation pour mieux comprendre les propriétés des méthodes d'estimation considérées en termes de biais moyen et d'erreur-type. Enfin, la dernière partie de la thèse a mis en application les enseignements de cette étude de simulation au travers de trois méta-analyses sur données individuelles dans le cancer du nasopharynx et dans le cancer du poumon à petites cellules.

Mots clés : analyse de survie, survie moyenne restreinte, méta-analyse sur données individuelles, hétérogénéité, essais cliniques randomisés, étude coût-efficacité, évaluation médico-économique.

Abstract

The survival benefit restricted up to a certain time horizon has been suggested as an alternative measure to the common relative measures used to estimate the treatment effect, especially in case of non-proportional hazards of death. The restricted survival benefit corresponds to the difference of the two restricted mean survival times estimated for each treatment arm, and is expressed in terms of life years gained. In the literature, this measure is considered as more intuitive than the hazard ratio and many authors have suggested its use for the design and the analysis of clinical trials. However, it is not currently the most used measure in randomized trials. This measure is valid under any distribution of the survival times and is adapted if the proportional hazards assumption does not hold. In addition, the restricted survival benefit can be used in medico-economic evaluation where an absolute measure of the treatment effect is needed (number of [quality adjusted] life years gained). If one wants to estimate the restricted survival benefit from an individual participant data meta-analysis, there is a need to take into account the trial effect due to the hierarchical structure of the data. The aim of this thesis was to compare statistical methods to estimate the restricted survival benefit from an individual participant data meta-analysis of randomized trials. The starting point was a case study (cost-effectiveness analysis) using data from the *Meta-Analysis of Radiotherapy in Lung Cancer*. This study showed that the five investigated methods yielded different estimates for the restricted survival benefit and its confidence interval. The choice of a method to estimate the survival benefit also impacted on cost-effectiveness results. Our second project consisted in a simulation study to have a better understanding of the properties of the investigated methods in terms of bias and standard error. Finally, the last part of the thesis illustrated the lessons learned from the simulation study through three examples of individual participant data meta-analysis in nasopharynx cancer and in small cell lung cancer.

Key words : survival analysis, restricted mean survival time, individual participant data meta-analysis, heterogeneity, randomized clinical trials, cost-effectiveness analysis, economic evaluation

Liste des productions scientifiques

Articles directement issus du travail de thèse

Lueza B, Mauguén A, Pignon JP, Rivero-Arias O, Bonastre J. Difference in Restricted Mean Survival Time for Cost-Effectiveness Analysis Using Individual Patient Data Meta-Analysis: Evidence from a Case Study. *PLoS One* 2016 11(3): e0150032. doi:[10.1186/1471-2288-14-72](https://doi.org/10.1186/1471-2288-14-72)

Lueza B, Rotolo F, Bonastre J, Pignon JP, Michiels S. Bias and precision of methods for estimating the difference in restricted mean survival time from an individual patient data meta-analysis. *BMC Med Res Meth* 2016 16(1):37. doi: [10.1186/s12874-016-0137-z](https://doi.org/10.1186/s12874-016-0137-z)

De Ruysscher D*, **Lueza B***, Le Péchoux C, Johnson DH, O'Brien M, Murray N, Spiro S, Wang X, Takada M, Lebeau B, Blackstock W, Skarlos D, Baas P, Choy H, Price A, Seymour L, Arriagada R, Pignon JP. Impact of thoracic radiotherapy timing in limited-stage small-cell lung cancer: usefulness of the individual patient data meta-analysis. *Annals of Oncology* 2016 27(10), 1818–1828. doi: [10.1093/annonc/mdw263](https://doi.org/10.1093/annonc/mdw263)

* co-premier auteur

Articles en rapport avec la thématique de la thèse

Ramaekers BLT, Joore M, **Lueza B**, Bonastre J, Mauguén A, Pignon JP, Le Péchoux C, De Ruysscher D, MAR-LC Collaborative Group, Grutters J. Cost Effectiveness of Modified Fractionation Radiotherapy versus Conventional Radiotherapy for Unresected Non–Small-Cell Lung Cancer Patients. *J Thorac Oncol* 2013; 8: 1295–307. doi: [10.1097/JTO.0b013e31829f6c55](https://doi.org/10.1097/JTO.0b013e31829f6c55)

Guyot P, Ades AE, Beasley M, **Lueza B**, Pignon JP, Welton NJ. Extrapolation of survival curves from cancer trials using external information. *Med Decis Making* 2016 (in press)

Communications orales

Lueza B, Rotolo F, Bonastre J, Rivero-Arias O, Pignon JP, Michiels S. Estimating the difference in restricted mean survival times from an individual patient data meta-analysis. *ISCB - Conference of the International Society for Clinical Biostatistics*, 2015 : poster presentation.

Lueza B, Rotolo F, Bonastre J, Pignon JP, Michiels S. Estimation du bénéfice de survie agrégé en méta-analyse sur données individuelles. *EPICLIN 9 – 22èmes Journées des Statisticiens des CLCC*, Montpellier, 2015 : présentation orale (Prix de la meilleure communication orale).

Lueza B. Estimating pooled survival benefit using individual patient-level data meta-analyses in cost-effectiveness analysis of health care technologies: case and simulation study. *Health Economics Research Centre Seminar*, Université d'Oxford, 2015 : séminaire invité.

Lueza B, Mauguén A, Pignon JP, Le Péchoux C, Bonastre J. Estimating the survival benefit from individual patient data meta-analysis: which methods and which impact on the cost effectiveness ratio? *European Conference of Health Economics*, Zürich, 2012 : présentation orale.

Lueza B, Mauguén A, Pignon JP, Le Péchoux C, Bonastre J. Evaluation économique à partir des données individuelles d'une méta-analyse. Comparaison de méthodes d'estimation du bénéfice de survie et impact sur le ratio coût-efficacité. *EPICLIN 6 – 19èmes Journées des Statisticiens des CLCC*, Lyon, 2012: présentation orale.

Lueza B, Le Teuff G, Pignon JP, De Ruyscher D, Le Péchoux C. Hétérogénéité inter et intra-sous-groupe d'essais pour explorer l'hétérogénéité globale: un exemple de méta-analyse sur données individuelles dans le traitement des cancers bronchiques. *EPICLIN 6 – 19èmes Journées des Statisticiens des CLCC*, Lyon, 2012: présentation orale.

Lueza B, Mauguén A, Pignon JP, Le Péchoux C, Bonastre J. Methods to estimate the survival benefit from individual patient data meta-analysis. Impact on the health economic outcome. *3rd Joint CES/HESG Workshop*, Aix-en-Provence, 2012: "working paper" discuté par Stirling Bryan.

Abréviations

Cont	Contrôle
CT	Chimiothérapie
ET	Erreur-type
Exp	Expérimental
GHM	Groupe Homogène de Malades
HAS	Haute Autorité de Santé
HR	<i>Hazard ratio</i> (rapport des risques instantanés)
IC	Intervalle de Confiance
ICER	Ratio incrémental coût-efficacité (<i>Incremental cost-effectiveness ratio</i>)
MAC-NPC	<i>Meta-Analysis of Chemotherapy in Nasopharynx Carcinoma</i>
MADI	Méta-analyse sur données individuelles
MADP	Méta-analyse sur données publiées
MAR-LC	<i>Meta-Analysis of Radiotherapy in Lung Cancer</i>
NICE	National Institute for Health and Clinical Excellence
OR	<i>Odds ratio</i> (rapport des cotes)
QALY	Année de vie pondérée par la qualité de vie (<i>Quality Adjusted Life Years</i>)
RMST	Survie moyenne restreinte (<i>Restricted mean survival time</i>)
rmstD	Bénéfice de survie restreint (<i>Difference in restricted mean survival time</i>)
rmstRD	Bénéfice de survie restreint relatif (<i>Relative difference in restricted mean survival time</i>)
rmstR	Ratio des survies moyennes restreintes (<i>Ratio of restricted mean survival time</i>)
RT	Radiothérapie
RTT-SCLC	<i>RadioTherapy Timing in Small Cell Lung Cancer</i>
T2A	Tarification A l'Activité

Table des matières

Avant-propos	iii
Remerciements.....	iv
Résumé	vi
Abstract.....	vii
Liste des productions scientifiques	viii
Abréviations.....	x
Table des matières	xi
Liste des figures.....	xiv
Liste des tableaux.....	xvi
 Chapitre 1. Introduction.....	 1
1.1. Contexte et objectifs de la thèse	1
1.2. Plan de la thèse.....	5
 Chapitre 2. Méthodes	 7
2.1. Bénéfice de survie.....	7
2.1.1. Définitions générales en analyse de survie.....	7
2.1.2. Bénéfice de survie et bénéfice de survie restreint.....	11
2.1.3. Choix de l'horizon temporel t^*	14
2.1.4. Extrapolation jusqu'à t^*	14
2.2. Méthodologie des méta-analyses	15

2.2.1.	Méta-analyses sur données individuelles et méta-analyses sur données publiées : définitions, avantages et inconvénients	15
2.2.2.	Méthodologie des méta-analyses sur données individuelles.....	16
2.3.	Estimation du bénéfice de survie en MADI.....	24
2.3.1.	Méthode « Naïve Kaplan-Meier »	24
2.3.2.	Méthode « Stewart-Parmar ».....	26
2.3.3.	Méthode « Peto ».....	27
2.3.4.	Méthode « Pooled Kaplan-Meier »	29
2.3.5.	Méthode « Pooled Exponentiel ».....	30
 Chapitre 3. Comparaison de méthodes d'estimation du bénéfice de survie et impact en évaluation médico-économique : une étude de cas		33
3.1.	Motivation.....	33
3.2.	Matériel : la méta-analyse Meta-Analysis of Radiotherapy in Lung Cancer	34
3.3.	Analyse coût-efficacité	39
3.3.1.	Estimation des coûts.....	41
3.3.2.	Estimation du bénéfice de survie restreint et du ratio coût-efficacité incrémental ..	44
3.4.	Discussion.....	53
 Chapitre 4. Comparaison de méthodes d'estimation du bénéfice de survie : une étude de simulation		57
4.1.	Motivation.....	57
4.2.	Etude de simulation.....	59
4.2.1.	Plan expérimental de l'étude de simulation	59
4.2.2.	Vrai bénéfice de survie restreint.....	62
4.2.3.	Scénarios des simulations	62
4.2.4.	Critères d'évaluation.....	64
4.2.5.	Résultats de l'étude de simulation	64

4.3. Mise en perspective des résultats de l'étude de cas Meta-Analysis of Radiotherapy in Lung Cancer	82
4.4. Discussion.....	84
Chapitre 5. Estimation du bénéfice de survie : application à trois méta-analyses sur données individuelles MAC-NPC 1, MAC-NPC 2 et RTT-SCLC	87
5.1. Meta-Analysis of Chemotherapy in Nasopharynx Carcinoma (MAC-NPC 1 et MAC-NPC 2).....	87
5.1.1. Résultats de l'analyse princeps avec le hazard ratio.....	88
5.1.2. Résultats de l'analyse secondaire avec le bénéfice de survie restreint.....	90
5.1.3. Conclusion des méta-analyses MAC-NPC 1 et MAC-NPC 2.....	100
5.2. RadioTherapy Timing in Small-Cell Lung Cancer (RTT-SCLC).....	101
5.2.1. Sélection des essais et stratégie de recherche	101
5.2.2. Résultats : analyse princeps avec le hazard ratio et le bénéfice de survie restreint	103
5.2.3. Conclusion de la méta-analyse RTT-SCLC	116
5.3. Comparaison du bénéfice de survie restreint et du hazard ratio à partir de MAR-LC, MAC-NPC 2 et RTT-SCLC.....	119
Chapitre 6. Discussion	123
6.1. Synthèse des résultats et apports de la thèse.....	123
6.2. Le bénéfice de survie restreint dans la littérature.....	125
6.3. Limites de la thèse	128
6.4. Conclusion et perspectives	129
Références	131
Annexes	138

Liste des figures

2.1 - Représentation graphique du bénéfice de survie restreint à $t^* = 5$ ans et $t^* = 10$ ans	13
2.2 - Exemple de forest plot : comparaison de la radiothérapie modifiée (RT Mod) à la radiothérapie conventionnelle (RT Conv)	20
2.3 - Exemple de forest plot avec sous-groupes d'essai : comparaison de la radiothérapie modifiée (RT Mod) à la radiothérapie conventionnelle (RT Conv)	23
3.1 - Classification des essais selon le type de la radiothérapie modifiée	37
3.2 - Courbes de survie estimées à l'aide des méthodes Naïve Kaplan-Meier et Stewart-Parmar pour la méta-analyse MAR-LC	44
3.3 - Courbes de survie estimées à l'aide des méthodes Peto-mois (panel A), Peto-année (panel B) et Peto-quintile (panel C) pour la méta-analyse MAR-LC	46
3.4 - Forest-plot des bénéfices de survie restreint exprimés en mois et estimés avec les méthodes Pooled Kaplan-Meier (carrés et losange gris) et Pooled Exponentiel (carrés et losange verts) pour la méta-analyse MAR-LC.	47
3.5 - Courbes d'acceptabilité montrant la probabilité que la radiothérapie modifiée soit coût-efficace pour différents seuils de disposition à payer pour une année de vie gagnée	51
4.1 - Illustration de l'hétérogénéité du risque de base entre trois essais simulés, à l'aide de l'effet aléatoire A_j pour un effet traitement fixé ($HR=0,5$)	60
4.2 - Illustration de l'hétérogénéité de l'effet traitement entre trois essais simulés, à l'aide de l'effet aléatoire B_j pour un risque de base fixé	61
4.3 - Comparaison graphique à $t^* = 5$ ans (panel A) et à $t^* = 10$ ans (panel B) des méthodes en termes de biais moyen et d'erreurs-types empirique et moyenne. Scénario avec proportionnalité des risques, 5 essais et 200 patients par essai et $\beta = -0,7$	72
4.4 - Comparaison graphique à $t^* = 5$ ans (panel A) et à $t^* = 10$ ans (panel B) des méthodes en termes de biais moyen et d'erreurs-types empirique et moyenne. Scénario avec non-proportionnalité des risques, 5 essais et 200 patients par essai et $\beta = -0,7$	73

5.1 - Forest-plot des <i>hazard ratios</i> (RT contre RT+CT) avec un modèle à effets aléatoires pour la méta-analyse MAC-NPC 1.....	92
5.2 - Forest-plot des bénéfices de survie restreints à $t^* = 10$ ans, exprimés en année et estimés avec la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires pour la méta-analyse MAC-NPC 1.	93
5.3 - Bénéfice de survie restreint (panel A) et bénéfice de survie restreint relatif (panel B) exprimés comme fonction de l'horizon temporel t^* dans la méta-analyse MAC-NPC 1.	95
5.4 - Forest-plot des <i>hazard ratios</i> (RT contre RT+CT) avec un modèle à effets aléatoires pour la méta-analyse MAC-NPC 2.....	98
5.5 - Forest-plot des bénéfices de survie restreints à $t^* = 10$ ans, exprimés en année et estimés avec la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires pour la méta-analyse MAC-NPC 2.	99
5.6 - Flow-chart des essais inclus dans la méta-analyse RTT-SCLC.....	103
5.7 - Forest-plot des <i>hazard ratios</i> pour la méta-analyse RTT-SCLC.	107
5.8 - Courbes de survie estimées à l'aide de la méthode Peto-année pour les sous-groupes d'essai avec une compliance à la chimiothérapie similaire ou différente entre les deux bras pour la méta-analyse RTT-SCLC.....	111
5.9 - Forest-plot des <i>hazard ratios</i> avec un modèle à effets fixes dans deux sous-groupes d'essais de la méta-analyse RTT-SCLC.....	112
5.10 - Forest-plot des bénéfices de survie restreints à $t^* = 5$ ans, exprimés en mois et estimés avec la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires dans deux sous-groupes d'essais de la méta-analyse RTT-SCLC.	113
5.11 - Bénéfice de survie restreint (panel A) et bénéfice de survie restreint relatif (panel B) exprimés comme fonction de l'horizon temporel t^* dans deux sous-groupes d'essais de la méta-analyse RTT-SCLC.....	115
5.12 - Estimation de l'effet traitement avec le hazard ratio et le bénéfice de survie restreint pour les 42 essais des méta-analyses MAR-LC, MAC-NPC 2 et RTT-SCLC.	121
A.1 - Publication de la méta-analyse RTT-SCLC acceptée à <i>Annals of Oncology</i>	141

Liste des tableaux

2.1 - Caractéristiques des méthodes d'estimation du bénéfice de survie restreint à partir de méta-analyses sur données individuelles	32
3.1 - Description des essais inclus dans la méta-analyse MAR-LC	35
3.2 - Estimation des différents coûts selon le type de radiothérapie	43
3.3 - Survies moyennes restreintes, bénéfice de survie restreint moyen et son erreur-type selon les sept méthodes d'estimation dans l'étude de cas MAR-LC	48
3.4 - ICER moyen selon les sept méthodes d'estimation du bénéfice de survie restreint dans l'étude de cas MAR-LC	50
3.5 - Probabilité que la radiothérapie modifiée soit coût-efficace pour différents seuils de disposition à payer et selon la méthode d'estimation du bénéfice de survie restreint.....	52
4.1 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint. Scénario avec 5 essais et 200 patients par essai, et avec proportionnalité des risques.	68
4.2 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint. Scénario avec 5 essais et 200 patients par essai, et avec non-proportionnalité des risques.	69
4.3 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint. Scénario avec 20 essais et 100 patients par essai, et avec proportionnalité des risques	75
4.4 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint. Scénario avec 5 essais et 200 patients par essai, avec proportionnalité des risques, et effet traitement délétère.	77
4.5 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint. Scénario avec 5 essais et 200 patients par essai, avec proportionnalité des risques, et avec corrélation entre les effets aléatoires.	79

4.6 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint. Scénario avec 5 essais et 200 patients par essai, avec proportionnalité des risques, et modèle à effets fixes.	81
4.7 - Estimation du bénéfice de survie restreint et de son erreur-type selon les différentes méthodes étudiées dans l'étude de cas et dans l'étude de simulation	83
5.1 - Résultats des méthodes d'estimation du bénéfice de survie restreint dans les méta-analyses MAC-NPC 1 et MAC-NPC 2.	91
5.2 - Description des essais inclus dans la méta-analyse RTT-SCLC	104
5.3 - Répartition des essais inclus dans la méta-analyse RTT-SCLC selon les sous-groupes étudiés	106
5.4 - <i>Hazard ratios</i> et hétérogénéité de l'effet de la radiothérapie expérimentale contre la radiothérapie contrôle sur la survie globale en fonction de caractéristiques d'essais	109
A.1 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint. Scénario avec 20 essais et 100 patients par essai, et avec non-proportionnalité des risques	139
A.2 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint. Scénario avec 5 essais et un nombre de patients par essai aléatoire ($N = 1\,000$), avec proportionnalité des risques, et avec $\beta = -0,7$	140

Chapitre 1. Introduction

1.1. Contexte et objectifs de la thèse

En recherche clinique, plusieurs mesures de l'effet d'un traitement peuvent être utilisées. Les mesures usuelles sont traditionnellement des mesures relatives. Dans le cas de données ne dépendant pas du temps ou pour des données catégorielles – pour des taux de décès ou des taux de toxicité par exemple – l'effet du traitement est quantifié à l'aide de rapport de cotes (*odds ratios* en anglais) ou de risque relatif. Pour des données dépendantes du temps avec ou sans censure, les mesures relatives s'expriment généralement sous la forme d'un *hazard ratio* (HR) qui est un rapport des risques instantanés de décès ou encore sous la forme du rapport des médianes de survie. Le *hazard ratio* est la mesure privilégiée pour quantifier l'effet relatif d'un traitement et est celle recommandée par les groupes CONSORT et Cochrane (Moher *et al.*, 2010; Higgins & Green, 2011). Des mesures absolues de l'effet traitement sont souvent associées en complément comme critères secondaires. Il peut s'agir par exemple de la différence des médianes de survie entre les deux bras d'un essai clinique ou de la différence des taux de survie à un moment donné.

Pour les évaluations médico-économiques, l'une des mesures de survie les plus utilisées pour estimer le critères d'efficacité est le bénéfice de survie restreint à un horizon temporel donné (Anell & Norinder, 2000; Gray *et al.*, 2011). L'étude coût-efficacité permet de mettre en regard les bénéfices de santé d'un nouveau traitement et les coûts supplémentaires qu'il engendre, de manière à disposer d'un indicateur permettant aux décideurs de réaliser des arbitrages concernant le remboursement ou la fixation du prix d'un nouveau traitement. Le résultat d'une étude coût-efficacité s'exprime généralement en coût par année de vie gagnée ou en coût par année de vie en parfaite santé (pondération des années de vie par la qualité de vie pour obtenir des QALYs [*quality adjusted life-years*]). Plus récemment, le bénéfice de survie restreint a été proposé en recherche clinique comme mesure alternative aux mesures relatives de l'effet d'un traitement notamment dans le cas de non proportionnalité des risques de décès, c'est-à-dire lorsque l'effet du traitement varie au cours du temps (Royston & Parmar, 2011).

Le bénéfice de survie restreint correspond à la différence des survies moyennes restreintes (*difference in restricted mean survival time* ou $rmstD(t^*)$) entre deux bras de traitement (Irwin, 1949; Zhao & Tsiatis, 1999, 1997; Murray & Tsiatis, 1999; Chen & Tsiatis, 2001; Andersen *et al.*, 2004; Royston & Parmar, 2011, 2013; Schaubel & Wei, 2011; Zhao *et al.*, 2012, 2016; Tian *et al.*, 2014; Uno *et al.*, 2014). Il permet de quantifier l'effet d'un traitement et s'exprime en années de vie gagnées. La survie moyenne restreinte (*restricted mean survival time* ou $RMST(t^*)$) correspond à l'espérance de vie jusqu'à un horizon temporel t^* et a été développée en recherche clinique afin de se limiter au suivi d'un essai clinique randomisé (Irwin, 1949; Karrison, 1987, 1997). Cette survie moyenne restreinte correspond graphiquement à l'aire sous une courbe de survie restreinte au temps t^* . Le bénéfice de survie restreint à un temps donné correspond ainsi à l'aire entre deux courbes de survie jusqu'à l'horizon temporel considéré t^* . Cette mesure est présentée comme plus intuitive que le *hazard ratio* et plusieurs auteurs ont proposé son utilisation pour le design et l'analyse d'un essai clinique (Royston & Parmar, 2011, 2013, 2016, Zhao *et al.*, 2012, 2016, Uno *et al.*, 2014, 2015). En effet, le bénéfice de survie restreint présente plusieurs avantages. C'est une mesure absolue qui combine le risque de décès de base et l'effet relatif du traitement, et s'interprète comme un allongement de l'espérance de vie d'un patient, ce qui est plus parlant qu'une réduction du risque de décès quantifiée par le *hazard ratio*. De plus, le bénéfice de survie restreint est une mesure qui s'applique quelle que soit la distribution des temps de survie et qui peut donc s'appliquer lorsque les risques ne sont pas proportionnels. Bien que le bénéfice de survie restreint ait fait l'objet de travaux méthodologiques ayant montré son intérêt pour mesurer l'effet d'un traitement (Royston & Parmar, 2011, 2013, Uno *et al.*, 2014, 2015; Wei *et al.*, 2015; Trinquart *et al.*, 2016), cette mesure n'est pas utilisée fréquemment dans les essais cliniques. Des études antérieures ont comparé des méthodes paramétriques d'estimation du bénéfice de survie afin d'étudier la question de l'extrapolation au-delà du suivi de l'essai (Jackson, 2010; Gong & Fang, 2012; Latimer, 2013; Lamb *et al.*, 2015). Cependant, ces études se limitaient au cas d'un seul essai clinique randomisé et n'abordaient pas la question de l'effet centre dans les essais multicentriques, ni la question de l'effet essai dans les méta-analyses.

Les méta-analyses d'essais cliniques randomisés représentent le plus haut niveau de preuve dans l'évaluation de l'efficacité d'un traitement (HAS, 2013). La méta-analyse consiste à recueillir les données provenant de tous les essais cliniques qui étudient la même question clinique. La méta-analyse permet ainsi d'estimer l'effet traitement global dans une analyse combinée des effets traitement estimés dans chaque essai. Elle permet aussi d'étudier l'hétérogénéité de l'effet traitement, c'est-à-dire la variation de cet effet traitement entre les essais.

La méta-analyse sur données individuelles (MADI), qui se fonde sur les données de chaque patient récupérées auprès des investigateurs des différents essais, est la méthode de référence par rapport à la méta-analyse sur données publiées (Riley *et al.*, 2010). Si l'on souhaite estimer le bénéfice de survie restreint à partir d'une MADI, se pose alors la question de prendre en compte l'effet essai dû à la structure hiérarchique des méta-analyses. En effet, dans une méta-analyse, il faut prendre en compte l'effet essai afin d'éviter un biais dans l'estimation de l'effet traitement connu sous le nom de paradoxe de Simpson (Altman & Deeks, 2002; Cates, 2002; Rücker & Schumacher, 2008). Cette question se pose de manière analogue dans un essai clinique multicentrique où l'on cherche à prendre en compte l'effet centre. Deux types de méthodes permettent d'estimer l'effet traitement global tout en permettant de prendre en compte cet effet essai (respectivement centre) : les méthodes dites à une étape et les méthodes à deux étapes. Les méthodes à une étape combinent les données de tous les essais (respectivement centres) dans une seule analyse stratifiée prenant en compte le niveau essai (respectivement centre). Les modèles de Cox stratifiés, où le risque de base est différent selon l'essai (respectivement centre), ou les modèles hiérarchiques, où l'effet essai (respectivement centre) est modélisé par un effet fixe ou un effet aléatoire, sont les méthodes à une étape les plus répandues. Les méthodes à deux étapes, qui sont les méthodes les plus utilisées en méta-analyse (Simmonds *et al.*, 2005), consistent à estimer l'effet traitement dans chaque essai (respectivement centre) dans un premier temps puis à estimer l'effet traitement global comme une moyenne pondérée des effets individuels. Là encore, les effets individuels peuvent être agrégés à l'aide d'un modèle à effets fixes ou d'un modèle à effets aléatoires en fonction de la présence ou non d'hétérogénéité.

Plusieurs études ont discuté et comparé différents modèles de Cox en une étape (stratifiés, avec effets fixes ou avec effets aléatoires) pour tenir compte de l'effet centre dans un essai multicentrique (Glidden & Vittinghoff, 2004; Legrand *et al.*, 2005; Munda & Legrand, 2014) ou de l'effet essai dans une méta-analyse (Michiels *et al.*, 2005a; Tudur Smith *et al.*, 2005; Rondeau *et al.*, 2008; Stewart *et al.*, 2012) en présence d'hétérogénéité entre les essais (respectivement centres) du risque de base et/ou de l'effet traitement. D'autres auteurs ont comparé à la fois des modèles de Cox en une étape et des modèles de Cox en deux étapes pour estimer le *hazard ratio* global à partir d'une MADI (Tudur Smith & Williamson, 2007; Bowden *et al.*, 2011). Toutes ces études se sont intéressées à l'estimation de l'effet traitement global en considérant le *hazard ratio* comme mesure relative de l'effet traitement et ont conclu que cette estimation était proche entre les modèles à une ou deux étapes dans la majorité des cas. A l'inverse, les études ayant également comparé ces modèles avec des modèles naïfs ne tenant pas compte de l'effet essai

(respectivement centre) ont conclu que l'estimation de l'effet relatif du traitement était biaisée si l'effet essai (respectivement centre) était ignoré dans l'analyse (Glidden & Vittinghoff, 2004; Munda & Legrand, 2014). Peu de recherches ont porté sur l'utilisation d'une mesure absolue de l'effet traitement pour l'analyse d'une MADI. En effet, jusqu'à présent, seule une étude récente s'est intéressée au bénéfice de survie restreint en MADI (Wei *et al.*, 2015). Dans cette étude, Wei et al. ont comparé trois méthodes à deux étapes afin d'estimer le bénéfice de survie restreint à partir d'une MADI. Dans leur étude, le bénéfice de survie restreint était estimé comme l'agrégation, à l'aide d'un modèle à effets fixes, des bénéfices de survie restreints estimés dans chacun des essais. Pour estimer le bénéfice de survie restreint dans chaque essai, Wei et al. ont comparé trois méthodes : deux méthodes non-paramétriques – l'une fondée sur les « pseudo-valeurs » (Andersen & Pohar Perme, 2010) et l'autre fondée sur l'estimateur de Kaplan-Meier – et une méthode de survie paramétrique flexible basée sur des splines (Royston & Parmar, 2002; méthode définie en Annexes). A travers une étude de simulation et deux études de cas en méta-analyse, Wei et al. ont montré que les trois méthodes conduisaient à des résultats similaires en terme de biais moyen, d'erreur quadratique moyenne et de probabilité de couverture des intervalles de confiance.

S'inscrivant dans la continuité des travaux de Wei et al., cette thèse étudie le bénéfice de survie restreint comme mesure alternative aux mesures relatives dans le contexte des MADI. L'objectif de cette thèse était d'étudier plusieurs méthodes d'estimation du bénéfice de survie restreint au travers d'une étude de cas, d'une étude de simulation et de plusieurs applications à partir de MADI. L'étude de Wei et al. a été publiée alors que nous réalisions l'étude de simulation. Nous avons souhaité étendre ces travaux en considérant un plus grand nombre de scénarios se fondant sur un cas de figure plus réaliste en méta-analyse et utilisant des effets aléatoires pour introduire de l'hétérogénéité sur le risque de base et sur l'effet traitement entre les essais. Parmi les trois méthodes à deux étapes étudiées par Wei et al., qui conduisaient à des résultats similaires, nous avons retenu celle fondée sur l'estimateur de Kaplan-Meier. Nous avons également choisi d'adapter d'autres méthodes non-paramétriques couramment utilisées en MADI. Nous nous sommes particulièrement intéressés à la manière dont ces méthodes prennent en compte la stratification sur l'essai, la non proportionnalité des risques et l'hétérogénéité du risque de base et de l'effet traitement. Une autre originalité de ce travail de thèse a été d'étudier le bénéfice de survie restreint dans le contexte d'une évaluation médico-économique réalisée à partir des données d'une MADI. Une précédente étude avait souligné l'importance du choix de la méthode d'analyse de survie dans les études coût-efficacité à partir des données d'un essai

clinique ou d'une MADI (Guyot *et al.*, 2011). Guyot et al. avaient ainsi précisé que la mesure de l'efficacité dans une étude coût-efficacité devrait être estimée avec la même méthode statistique que celle utilisée pour estimer l'efficacité dans l'analyse clinique princeps. C'est pourquoi nous avons choisi d'étudier des méthodes non-paramétriques d'estimation du bénéfice de survie restreint déjà utilisées en analyse de survie dans le domaine des MADI ainsi que des méthodes paramétriques qui sont utilisées classiquement en évaluation médico-économique. Notre question de recherche a été de déterminer si le choix d'une méthode d'estimation du bénéfice de survie restreint peut avoir un impact sur les résultats d'une étude coût-efficacité. Cette question est documentée pour les évaluations économiques à partir de données individuelles d'essais cliniques (Neymark *et al.*, 2001, 2002; Jackson, 2010; Connock *et al.*, 2011; Latimer, 2013), mais n'a pas été abordée dans le contexte des MADI. Si cette étude de cas a été le point de départ de notre recherche, nous avons découvert au fil du temps que l'estimation du bénéfice de survie restreint à partir de MADI posait des questions méthodologiques indépendantes du contexte particulier des études médico-économiques. C'est pourquoi une étude de simulation et d'autres applications ont été réalisées par la suite en se plaçant dans le contexte de l'analyse du critère de jugement principal d'une MADI.

1.2. Plan de la thèse

Le Chapitre 2 présente les principales méthodes utilisées en méta-analyse et en analyse de survie, ainsi que les méthodes d'estimation du bénéfice de survie restreint que nous avons étudiées. Le Chapitre 3 porte sur la comparaison des méthodes d'estimation du bénéfice de survie restreint dans le cadre d'une étude coût-efficacité sur données individuelles de méta-analyse d'essais cliniques randomisés. Cette étude de cas a été réalisée à partir des données de la *Meta-Analysis of Radiotherapy in Lung Cancer* et a abouti à la publication d'un article scientifique dans *PLoS ONE* (Lueza *et al.*, 2016a). Le Chapitre 4 présente l'étude de simulation qui a permis de mieux comprendre les propriétés des méthodes étudiées pour estimer le bénéfice de survie restreint à partir d'une méta-analyse sur données individuelles. L'article scientifique tiré de cette étude de simulation a été publié dans *BMC Medical Research Methodology* (Lueza *et al.*, 2016b). Le Chapitre 5 illustre les méthodes d'estimation du bénéfice de survie restreint au travers de trois méta-analyses sur données individuelles : la méta-analyse *Meta-Analysis of Chemotherapy in Nasopharynx Carcinoma* (MAC-NPC 1) et sa version actualisée MAC-NPC 2 dans le cancer du nasopharynx que nous avons ré-analysées en utilisant le bénéfice de survie restreint comme

c2015ritère de jugement, et la méta-analyse *RadioTherapy Timing in Small-Cell Lung Cancer* (RTT-SCLC) dans le cancer du poumon à petites cellules dont nous avons réalisé l'analyse princeps en utilisant à la fois le bénéfice de survie restreint et le *hazard ratio*. L'article correspondant à la méta-analyse sur données individuelles RTT-SCLC a été publié dans *Annals of Oncology* (De Ruyscher *et al.*, 2016) et est disponible en Annexes (Figure A.1). Enfin, une discussion et des perspectives de travail présentées dans le Chapitre 6 clôturent cette thèse.

Chapitre 2. Méthodes

2.1. Bénéfice de survie

2.1.1. Définitions générales en analyse de survie

Si l'on considère la variable aléatoire T symbolisant le temps de survenue du décès et $F(t)$ sa fonction de distribution, alors la fonction de survie au temps t est définie par :

$$S(t) = 1 - F(t) = \Pr(T > t) \quad (2.1)$$

La fonction de risque cumulée est définie par :

$$\Lambda(t) = -\log\{S(t)\} = \int_0^t \lambda(u) du \quad (2.2)$$

où $\lambda(t)$ est la fonction de risque instantané au temps t .

Dans le cas d'un essai clinique, les temps d'évènement peuvent être censurés dans le cas d'un patient « perdu de vue » ou « exclu vivant » : on parle alors de censure à droite et le temps de survie pris en compte dans l'analyse de survie sera celui à la date de dernières nouvelles (Klein & Moeschberger, 2003).

L'estimateur de Kaplan-Meier de la fonction de survie est défini par (Kaplan & Meier, 1958) :

$$\hat{S}_{KM}(t) = \prod_{t_i \leq t} \frac{Y_i - d_i}{d_i} \quad (2.3)$$

avec les E temps d'évènement distincts $t_1 < t_2 < \dots < t_E$, Y_i le nombre d'individus qui sont à risque de présenter l'évènement au temps t_i et d_i le nombre d'évènement au temps t_i .

Le test non-paramétrique du log-rank (Mantel, 1966) permet de comparer les fonctions de risque de deux bras de traitement (contrôle et expérimental). Les hypothèses testées sont les suivantes :

- H_0 : les fonctions de risque sont les mêmes dans les deux bras ;
- H_1 : les fonctions de risque diffèrent dans les deux bras.

A chaque temps d'évènement t_i , un tableau de contingence comme celui-ci peut être généré :

Etat	Bras contrôle	Bras expérimental	Total
Avec évènement	$O_{Cont,i}$	$O_{Exp,i}$	d_i
Sans évènement	$Y_{Cont,i} - O_{Cont,i}$	$Y_{Exp,i} - O_{Exp,i}$	$Y_i - d_i$
Total	$Y_{Cont,i}$	$Y_{Exp,i}$	Y_i

où :

- $O_{bras,i}$ = nombre d'évènements observés à l'instant t_i dans un bras donné;
- $Y_{bras,i}$ = nombre d'individus à risque à l'instant t_i dans un bras donné;
- $bras = Cont$ pour le bras contrôle et Exp pour le bras expérimental.

Le nombre d'évènements attendus à l'instant t_i dans un bras donné, sous l'hypothèse d'une absence de différence entre les deux bras, est calculé comme :

$$A_{bras,i} = \frac{Y_{bras,i} \times d_i}{Y_i} \quad (2.4)$$

et la variance de $(O_{bras,i} - A_{bras,i})$ est donnée par :

$$Var(O_{bras,i} - A_{bras,i}) = A_{bras,i} \times \left(1 - \frac{Y_{bras,i}}{Y_i}\right) \times \left(\frac{Y_i - d_i}{Y_i - 1}\right) \quad (2.5)$$

Ainsi, nous définissons :

- $O_{bras} = \sum_{i=1}^E O_{bras,i}$ le nombre d'évènements observés dans un bras donné ;
- $A_{bras} = \sum_{i=1}^E A_{bras,i}$ le nombre d'évènements attendus dans un bras donné sous l'hypothèse nulle ;
- $Var(O_{bras} - A_{bras}) = Var[\sum_{i=1}^E (O_{bras,i} - A_{bras,i})] = \sum_{i=1}^E Var(O_{bras,i} - A_{bras,i})$,
la variance de $(O_{bras} - A_{bras})$, définie grâce à l'indépendance entre deux temps distincts.

Finalement, la statistique du log-rank suivant une distribution Khi^2 à 1 degré de liberté sous l'hypothèse nulle est définie par :

$$LR = \frac{(O_{Exp} - A_{Exp})^2}{Var(O_{Exp} - A_{Exp})} \quad (2.6)$$

Le *hazard ratio* (HR) permet de quantifier l'effet d'un traitement par rapport à un autre et est défini comme le rapport des risques instantanés de survenue d'un évènement dans chacun des deux bras de traitement (noté Exp pour le bras expérimental et Cont pour le bras contrôle) :

$$HR(t) = \frac{\lambda_{Exp}(t)}{\lambda_{Cont}(t)} \quad (2.7)$$

Peto, dans une publication clinique de Yusuf et al. (Yusuf *et al.*, 1985), a proposé un estimateur pour le calcul d'un HR dans un essai basé sur la statistique du log-rank et qui permet de prendre en compte les temps d'évènement et de censure. Cet estimateur est souvent utilisé en méta-analyse et est défini comme suit :

$$\begin{aligned} \widehat{HR}_{Peto} &= \exp\left(\frac{O_{Exp} - A_{Exp}}{Var(O_{Exp} - A_{Exp})}\right) \\ \widehat{Var}[\log(\widehat{HR}_{Peto})] &= \frac{1}{Var(O_{Exp} - A_{Exp})} \end{aligned} \quad (2.8)$$

où :

- O_{Exp} = nombre d'évènements observés dans le bras expérimental ;
- A_{Exp} = nombre d'évènements attendus dans le bras expérimental sous l'hypothèse d'une absence de différence entre les deux bras.

Le HR peut également être estimé à l'aide du modèle de Cox (Cox, 1972). Ce modèle de survie semi-paramétrique à risques proportionnels définit la fonction de risque instantané d'un patient au temps t par :

$$\lambda(t; X) = \lambda_0(t) \cdot e^{\beta' X} \quad (2.9)$$

où $\lambda_0(t)$ est la fonction de risque de base, X le vecteur de variables explicatives et β le vecteurs des coefficients de régression associés.

Le modèle de Cox peut être stratifié sur une covariable, typiquement sur l'essai pour une méta-analyse, et devient :

$$\lambda(t; X) = \lambda_{0,j}(t) \cdot e^{\beta'x} \quad (2.10)$$

où j représente l'essai. Ce modèle stratifié permet que la fonction de risque de base soit différente pour chaque essai.

L'hypothèse des risques proportionnels considère que le HR entre les deux bras de traitement (ou une autre variable explicative du modèle de Cox) est constant dans le temps. Cette hypothèse peut être essentielle pour la validité de certains modèles de survie tels que le modèle de Cox. Différentes méthodes existent pour vérifier l'hypothèse des risques proportionnels, nous en présentons ici trois : une méthode graphique et deux tests statistiques. La méthode graphique consiste à vérifier que les courbes du graphe de $\log\{-\log[S(t)]\}$ en fonction de $\log(t)$ sont parallèles, ou tout du moins que l'écart entre ces courbes est constant. Un premier test a été proposé par Grambsch et Therneau qui teste si les résidus de Schoenfeld du coefficient de régression d'un modèle de Cox dépendent ou non du temps (Grambsch & Therneau, 1994). Un second test a été proposé par Royston et Parmar qui est basé sur un test de rapport de vraisemblance entre deux modèles de Cox avec et sans covariable dépendante du temps (Royston & Parmar, 2011). Dans cette thèse, nous utilisons le test de Grambsch et Therneau qui est disponible à l'aide de la fonction `cox.zph()` implémentée dans le package *survival* sous R.

Le test de non-proportionnalité des risques a été modifié par Wei et collègues pour tenir compte de la spécificité de la méta-analyse (Wei *et al.*, 2015). Ce test consiste en deux étapes :

1. Dans chaque essai j le test de Grambsch et Therneau est réalisé : on obtient alors J p-valeurs p_j .
2. Ces J p-valeurs sont agrégées à l'aide de la méthode de Fisher (Fisher, 1932).

La statistique du test global est alors donnée par :

$$\chi^2 = -2 \sum_{j=1}^J \log(p_j) \quad (2.11)$$

avec J le nombre d'essais et p_j la p-valeur du test de non-proportionnalité pour l'essai j . Cette statistique de test suit une distribution χ^2 à $2J$ degrés de liberté sous l'hypothèse nulle que les risques sont proportionnels dans tous les essais. Un test significatif (par exemple $p < 0,05$)

suggère que les risques dans au moins un essai ne sont pas proportionnels. Dans cette thèse, nous avons utilisé cette approche de Wei et al. pour tester globalement l'hypothèse de proportionnalité des risques dans une méta-analyse.

2.1.2. Bénéfice de survie et bénéfice de survie restreint

La survie moyenne, ou espérance de vie moyenne, μ peut être calculée à l'aide de la fonction de survie et d'un horizon temporel (Karrison, 1997) :

$$\mu = \int_0^{\infty} S(t)dt \quad (2.12)$$

Irwin a proposé une estimation de la survie moyenne, restreinte à un horizon temporel donné t^* (Irwin, 1949). Dans la littérature, cette survie moyenne restreinte est appelée *restricted mean survival time* et est notée $RMST(t^*)$ (Royston & Parmar, 2011), c'est cette notation que nous utilisons dans cette thèse :

$$RMST(t^*) = \mu(t^*) = \int_0^{t^*} S(t)dt \quad (2.13)$$

Graphiquement, la survie moyenne restreinte à l'horizon temporel t^* correspond à l'aire sous la courbe de survie définie par $S(t)$ entre 0 et t^* . Pour un horizon t^* fixé à 5 ans par exemple, le $RMST(t^*)$ peut être égal à 5 ans maximum dans le cas où aucun des patient ne décède.

Une mesure de l'effet traitement souvent utilisée en évaluation médico-économique est le bénéfice de survie. Il est défini comme la différence des survies moyennes des deux bras de traitement et peut s'exprimer en années de vie gagnées :

$$BS = \int_0^{\infty} S_{Exp}(t)dt - \int_0^{\infty} S_{Cont}(t)dt = \mu_{Exp} - \mu_{Cont} \quad (2.14)$$

où $S_{Exp}(t)$ et $S_{Cont}(t)$ sont les fonctions de survie respectivement des bras expérimental et contrôle.

De même, le bénéfice de survie peut être restreint à un horizon temporel t^* et est défini comme la différence des survies moyennes restreintes estimées dans les deux bras de traitement (Royston & Parmar, 2011) :

$$rmstD(t^*) = \int_0^{t^*} S_{Exp}(t)dt - \int_0^{t^*} S_{Cont}(t)dt = RMST_{Exp} - RMST_{Cont} \quad (2.15)$$

Nous adoptons dans cette thèse l'acronyme *rmstD* pour *difference in restricted mean survival time* qui est le terme employé dans la littérature (Royston & Parmar, 2011, 2013; Wei *et al.*, 2015; Trinquart *et al.*, 2016). La différence de survie moyenne restreinte $rmstD(t^*)$ peut correspondre à un gain maximal de t^* années de vie gagnée, et est égal à 0 en cas d'équivalence des deux bras de traitement. A l'inverse, le $rmstD(t^*)$ peut également être négatif et correspondre à une perte maximale de t^* années de vie. Nous adopterons cependant le terme de bénéfice de survie restreint pour évoquer le $rmstD(t^*)$ et signalerons le cas où le bras expérimental est moins efficace que le bras contrôle.

Graphiquement, le bénéfice de survie restreint $rmstD(t^*)$ est calculé comme l'aire entre les deux courbes de survie définies par $S_{Cont}(t)$ pour le bras contrôle et $S_{Exp}(t)$ pour le bras expérimental entre 0 et t^* . Pour un traitement expérimental qui serait toujours plus efficace que le traitement du bras contrôle, cette aire sera d'autant plus grande que l'horizon temporel t^* est éloigné.

La Figure 2.1 reflète les données d'un essai clinique fictif où le traitement expérimental, représenté par la courbe de survie rouge, est toujours plus efficace que le traitement du bras contrôle, représenté par la courbe de survie bleue, avec un *hazard ratio* fixé à $HR = 0,63$. Ces courbes exponentielles reflètent un risque de décès constant $\lambda = \log(2)/5$ pour le bras expérimental, ce qui correspond à une médiane de survie égale à 5 ans, et $\lambda = \log(3)/5$ pour le bras contrôle.

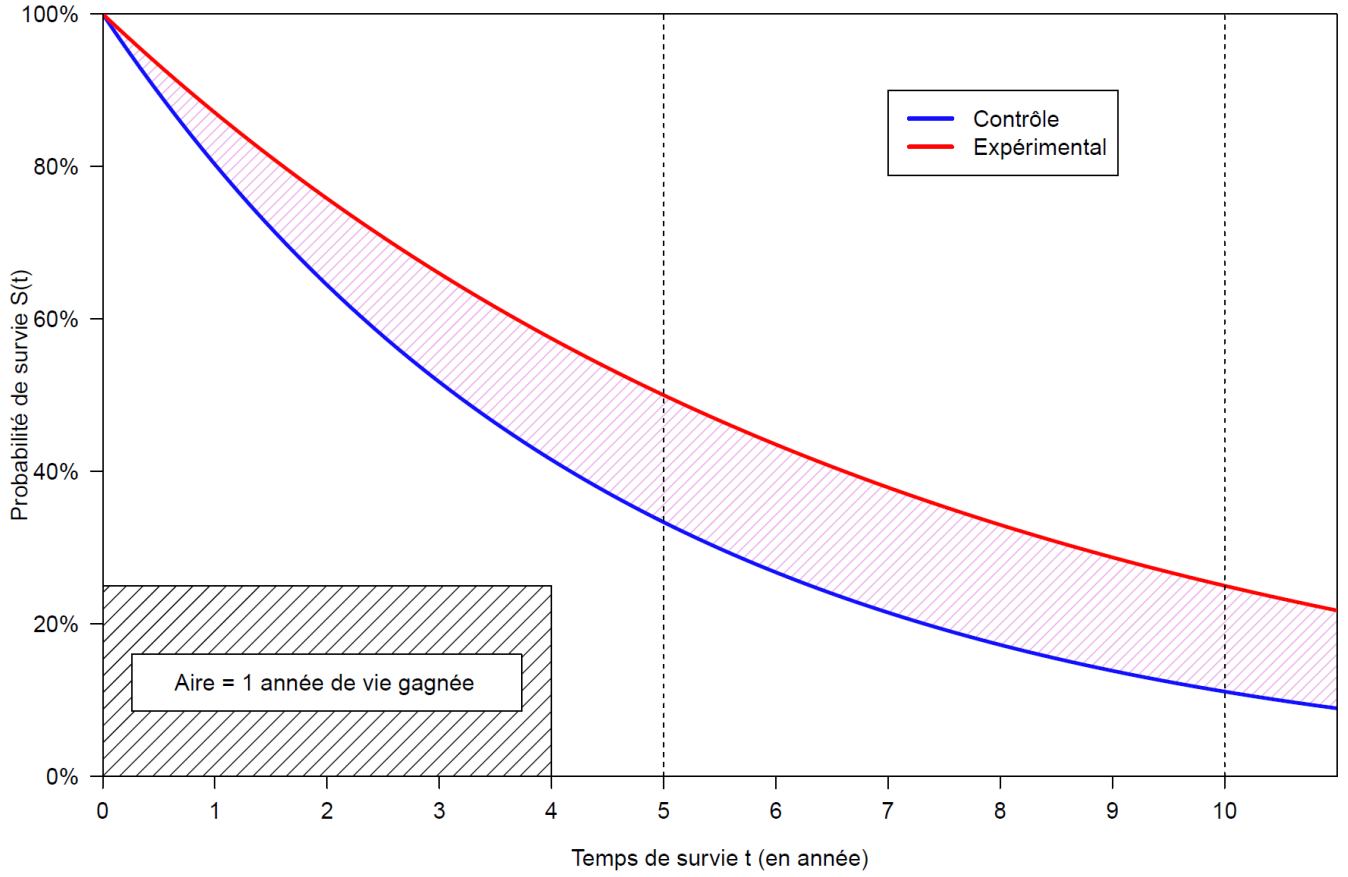


Figure 2.1 - Représentation graphique du bénéfice de survie restreint à $t^* = 5$ ans et $t^* = 10$ ans

L'aire entre les deux courbes de survie, hachurée en rose, correspond au bénéfice de survie. Si l'on restreint l'estimation du bénéfice de survie à 5 ans, l'aire rose de 0 à 5 ans vaut $rmstD(t^* = 5 \text{ ans}) = 6,9$ mois en faveur du traitement expérimental. En d'autres termes, le traitement expérimental prolonge la survie moyenne (espérance de vie) pendant les 5 premières années de 6,9 mois. Le bénéfice de survie restreint à 10 ans, équivalent à l'aire rose de 0 à 10 ans, vaut lui $rmstD(t^* = 10 \text{ ans}) = 16,4$ mois. L'aire du rectangle hachuré en noir sur la Figure 2.1 représente à titre illustratif un bénéfice de survie d'un an.

D'après Karrison, la variance du bénéfice de survie restreint $\widehat{Var}(rmstD(t^*))$ peut être estimée comme suit (Karrison, 1997) :

$$\widehat{Var}(rmstD(t^*)) = \widehat{Var}(RMST_{Exp}(t^*)) + \widehat{Var}(RMST_{Cont}(t^*)) \quad (2.16)$$

Les variances $\widehat{Var}(RMST_{Cont}(t^*))$ et $\widehat{Var}(RMST_{Exp}(t^*))$ sont définies dans la section 2.3 pour chacune des méthodes étudiées dans cette thèse.

2.1.3. Choix de l'horizon temporel t^*

L'estimation du bénéfice de survie restreint $rmstD(t^*)$ est intrinsèquement dépendante du choix de l'horizon temporel t^* . Karrison a recommandé de choisir un horizon temporel maximum t^* tel que l'erreur-type de $\hat{S}_{KM}(t)$ soit inférieure à une valeur seuil choisie au préalable dans le cadre d'un essai clinique (Karrison, 1987, 1997). Dans le cas plus particulier d'une méta-analyse sur données individuelles, les essais peuvent avoir différents suivis. Il y a donc un compromis à réaliser entre choisir de petites valeurs de t^* qui reviendrait à censurer de nombreuses données avec une forte perte d'information, et de plus grandes valeurs de t^* qui nécessiteraient d'extrapoler la fonction de survie pour certains essais qui n'auraient plus de patient à risque à l'horizon t^* (voir paragraphe 2.1.4 ci-après).

Royston et Parmar, et Wei et al. ont précisé que le choix de t^* doit également être guidé par l'intérêt clinique et être défini au préalable par exemple lors de la rédaction du protocole d'un essai clinique (Royston & Parmar, 2011, 2013; Wei *et al.*, 2015). Wei et al. ont également suggéré de présenter le graphique du bénéfice de survie restreint $rmstD(t^*)$ en fonction de l'horizon t^* . En effet, une analyse de sensibilité en variant l'horizon t^* doit être réalisée pour juger de la variation de l'effet traitement au cours du temps. Une telle analyse est particulièrement utile dans le cas où par exemple les risques de décès s'inverseraient après t^* et où l'estimation du bénéfice de survie restreint après un délai trop court serait biaisée.

2.1.4. Extrapolation jusqu'à t^*

En cas d'une différence de suivis entre les essais inclus dans une méta-analyse, certains essais peuvent ne plus présenter de patient à risque à l'horizon temporel choisi t^* . Afin d'extrapoler la fonction de survie au-delà du dernier temps d'évènement observé t_{max} jusqu'à t^* nous utilisons la méthode d'extrapolation de Brown et al. (Brown *et al.*, 1974). Cette méthode d'extrapolation consiste à compléter la queue de la courbe de survie de Kaplan-Meier par une courbe exponentielle. La fonction de survie pour $t > t_{max}$ est :

$$\hat{S}_{Brown}(t) = \exp\{t \times \log[\hat{S}_{KM}(t_{max})]/t_{max}\} \quad (2.17)$$

où $\hat{S}_{KM}(t)$ est l'estimateur de Kaplan-Meier de la fonction de survie.

Lamb et al. ont montré que le bénéfice de survie estimé à l'aide de cette méthode d'extrapolation est moins biaisée que la survie moyenne restreinte au temps de dernier événement observé (Lamb *et al.*, 2015).

2.2. Méthodologie des méta-analyses

2.2.1. Méta-analyses sur données individuelles et méta-analyses sur données publiées : définitions, avantages et inconvénients

Une méta-analyse est la synthèse quantitative et exhaustive de l'ensemble des résultats des essais randomisés étudiant une même question (Pignon & Hill, 2001). On tient compte dans une méta-analyse des résultats d'une analyse combinée, mais aussi, si besoin, de l'hétérogénéité des résultats des études individuelles. Deux grands types de méta-analyse existent :

- la méta-analyse sur données individuelles (MADI), réalisée à partir des données individuelles de chaque patient, qui est la méthode de référence (Stewart & Tierney, 2002)
- la méta-analyse sur données publiées (MADP), réalisée à partir de données agrégées extraites de publications

Pour une MADI, les données de chaque patient sont obtenues à partir des bases de données des essais fournies par les investigateurs. Une MADI inclut des essais quel que soit leur statut de publication (publié ou non-publié, résumé dans un congrès, enregistrement dans un registre d'essai). Elle comporte une vérification complète des données et de la qualité des essais (randomisation, suivi des patients), et une analyse standardisée des différents critères de jugement (Stewart & Clarke, 1995). Elle permet également d'actualiser le suivi des patients, ainsi qu'une analyse plus approfondie des données comme l'utilisation de modèle d'analyse de survie à partir des temps de survie individuels par exemple (Stewart & Tierney, 2002). La méta-analyse sur données individuelles peut ainsi répondre à une question avec une meilleure puissance statistique qu'une méta-analyse sur données publiées (Whitehead, 2002). Un autre avantage de ce type de méta-analyse est de pouvoir explorer plus en détail les origines de l'hétérogénéité entre les essais. Puisque des informations concernant les patients sont recueillies auprès des investigateurs des essais cliniques (âge, sexe, histologie, stade de la maladie, par exemple), il est possible d'étudier les interactions entre les caractéristiques des patients ou des essais, et l'effet du traitement (Stewart & Clarke, 1995). Le recueil de facteurs pronostiques permet ainsi d'étudier ces interactions avec une

puissance suffisante pour identifier les patients bénéficiant plus du traitement ou ceux en bénéficiant moins.

Les méta-analyses sur données publiées sont exposées au risque de biais de publication et de sélection des résultats rapportés (Pignon & Hill, 2001; Dwan *et al.*, 2013). En effet, en général les essais publiés sont souvent des essais « positifs », i.e. montrant un bénéfice du traitement. De plus, ces essais « positifs » sont souvent plus rapidement publiés, dans de meilleures revues scientifiques et plus souvent en anglais. Ainsi, lorsque les deux types de méta-analyse ont été utilisés pour étudier une même question, les méta-analyses basées sur la littérature ont surestimé le bénéfice associé au traitement expérimental dans la plupart des cas (Stewart & Parmar, 1993). Il faut donc veiller à récupérer également tous les essais « négatifs », y compris les essais interrompus prématurément : ceci est réalisé en méta-analyse sur données individuelles.

Enfin, le fait que certains patients randomisés aient été exclus de l'analyse de certains essais randomisés ne permet pas de réaliser une analyse en « intention de traiter » (c'est-à-dire sur l'ensemble des sujets tirés au sort, en les analysant dans leurs groupes initiaux) dans le cas de la méta-analyse sur données publiées. A l'inverse, l'analyse en « intention de traiter » est possible lorsque l'on dispose des données individuelles, en réclamant les données des patients exclus aux investigateurs.

2.2.2. Méthodologie des méta-analyses sur données individuelles

2.2.2.a. Stratification sur l'essai

Pour éviter un biais connu sous le nom de paradoxe de Simpson (Altman & Deeks, 2002; Cates, 2002; Rücker & Schumacher, 2008) qui serait dû à une comparaison directe de patients n'ayant pas les mêmes risques de base, les analyses statistiques des critères de jugement principal et secondaires sont stratifiées sur l'essai dans une méta-analyse. Ainsi, les patients du bras expérimental d'un essai sont uniquement comparés à des patients du bras contrôle du même essai. Chaque essai correspond ainsi à une strate de l'analyse finale : c'est ce que l'on appelle la stratification sur l'essai. Certaines méthodes à une étape, telles que les modèles de Cox stratifiés (équation (2.10)), permettent de tenir compte du facteur essai, en modélisant un risque de décès de base différent pour chaque essai. Pour les méthodes à deux étapes, l'effet traitement global est alors une combinaison des estimations des effets traitements dans chaque essai.

2.2.2.b. Estimation de l'effet traitement

Pour l'analyse d'évènement dépendant du temps, le *hazard ratio* (HR) est la mesure la plus souvent utilisée pour quantifier l'effet d'un traitement par rapport à un autre (Cox, 1972; Moher *et al.*, 2010; Higgins & Green, 2011). Il correspond au rapport du risque instantané de survenue d'un évènement dans le groupe expérimental et du risque dans le bras contrôle (équation (2.7)). Ce HR est souvent estimé en méta-analyse soit à l'aide de l'estimateur de Peto (équation (2.8)) soit à l'aide du modèle de Cox à risques proportionnels (équation (2.9); Cox, 1972).

En méta-analyse, le HR global, qui mesure l'effet traitement global, est estimé comme la moyenne pondérée des HR estimés dans chacun des essais de la méta-analyse. Cette agrégation est réalisée à l'aide d'un modèle à effets fixes ou d'un modèle à effets aléatoires (DerSimonian & Laird, 1986). Ces deux types de modèle sont définis dans le paragraphe suivant.

2.2.2.c. Estimation de l'effet traitement global par modèle à effets fixes et à effets aléatoires

L'effet traitement global (le terme « poolé » est parfois utilisé dans la littérature) $\hat{\theta}_{global}$ peut être obtenu par une moyenne pondérée des effets traitements $\hat{\theta}_j$ estimés dans chacun des J essais à l'aide soit d'un modèle à effets fixes, soit d'un modèle à effets aléatoires (Whitehead, 2002) :

$$\hat{\theta}_{global} = \frac{\sum_{j=1}^J \hat{\theta}_j w_j}{\sum_{j=1}^J w_j} \quad \text{et} \quad \widehat{Var}[\theta_{global}] = \frac{1}{\sum_{j=1}^J w_j} \quad (2.18)$$

où les poids sont estimés par $w_j = \frac{1}{\widehat{var}(\hat{\theta}_j)}$ pour le modèle à effets fixes et par $w_j = \frac{1}{\widehat{var}(\hat{\theta}_j) + \tau^2}$ pour le modèle à effets aléatoires. Le paramètre τ^2 évalue l'hétérogénéité inter-essai de l'effet traitement $\hat{\theta}_j$ et est estimé par la méthode des moments de DerSimonian-Laird (DerSimonian & Laird, 1986). Lorsque c'est le modèle à effets fixes qui est utilisé, on parle parfois de moyenne pondérée par l'inverse de la variance (*inverse variance weighted average*) puisque le poids accordé à chaque essai est alors égal à l'inverse de la variance de $\hat{\theta}_j$ estimé dans chaque essai. Ainsi plus cette variance est faible, pour un essai de grande taille par exemple, plus le poids accordé à cet essai dans l'estimation de l'effet global est important.

La mesure de l'effet traitement θ doit suivre une distribution de loi normale pour que l'agrégation entre les essais soit licite (Whitehead, 2002), d'où la nécessité d'une transformation

logarithmique pour les HR et les odds ratio (OR). Ainsi, en méta-analyse, les mesures de l'effet traitement $\hat{\theta}_j$ estimées dans chaque essai peuvent prendre la forme :

- $\hat{\theta}_j = \log(\widehat{HR}_j)$ pour des données de survie ;
- $\hat{\theta}_j = \hat{\mu}_j$ pour des données quantitatives (bénéfice de survie, moyenne, etc.) ;
- $\hat{\theta}_j = \log(\widehat{OR}_j)$ pour des données catégorielles.

Il est à noter, que si l'on utilise l'estimateur de Peto $\widehat{HR}_{Peto}$ défini dans l'équation (2.8), alors le calcul du *hazard ratio* global calculé par effets fixes se simplifie :

$$\widehat{HR}_{global,Peto} = \exp\left(\frac{\sum_{j=1}^J (O_{Exp,j} - A_{Exp,j})}{\sum_{j=1}^J Var(O_{Exp,j} - A_{Exp,j})}\right) \quad (2.19)$$

$$\widehat{Var}[\log(\widehat{HR}_{global,Peto})] = \frac{1}{\sum_{j=1}^J Var(O_{Exp,j} - A_{Exp,j})}$$

où :

- J = nombre d'essais cliniques ;
- $O_{Exp,j}$ = nombre d'événements observés dans le bras expérimental de l'essai j ;
- $A_{Exp,j}$ = nombre d'événements attendus dans le bras expérimental de l'essai j , estimé par la statistique du log-rank, sous l'hypothèse d'une absence de différence entre les deux bras.

Avec l'estimateur de Peto (équation (2.8)), la variance $Var(O_{Exp,j} - A_{Exp,j})$ est approximativement égale à 1/4 du nombre de décès total dans l'essai j (Parmar *et al.*, 1998).

2.2.2.d. Forest plot

Dans une méta-analyse, l'un des graphiques les plus utilisés pour représenter les résultats est le « forest plot » (Whitehead, 2002). Pour des données de survie, il représente les HR de chaque essai, le HR global ainsi que les intervalles de confiances correspondants. On peut ainsi détecter visuellement l'hétérogénéité de l'effet traitement existant entre les essais.

La Figure 2.2 est un exemple de forest plot qui décrit les données de la méta-analyse MAR-LC (Meta-Analysis of Radiotherapy in Lung Cancer) de dix essais randomisés comparant la radiothérapie (RT) thoracique conventionnelle à la radiothérapie avec un fractionnement modifié (dite RT modifiée) chez les patients ayant un cancer bronchique non à petites cellules (Mauguen *et al.*, 2012).

Chaque ligne représente un essai, et cette ligne comporte le nombre d'événements et l'effectif de chaque bras de l'essai, la différence entre le nombre de cas observés et le nombre de cas attendus (O-A), sa variance et le HR. Le HR est représenté par un carré dont la taille est inversement proportionnelle à la variance de (O-A), et les segments partant de chaque côté du carré représentent l'intervalle de confiance à 95 % du HR. Le HR global est représenté par un losange dont l'axe horizontal correspond à l'intervalle de confiance à 95 %. Un HR de 1 équivaut à l'équivalence entre les deux bras de traitement. Si le HR est à gauche de la ligne verticale ($HR < 1$), le traitement expérimental est plus efficace que le traitement du bras contrôle (réduction du risque de décès). Si l'intervalle de confiance (segment pour les HR de chaque essai ou losange pour le HR global) n'inclut pas le chiffre 1, le traitement est significativement efficace ou délétère.

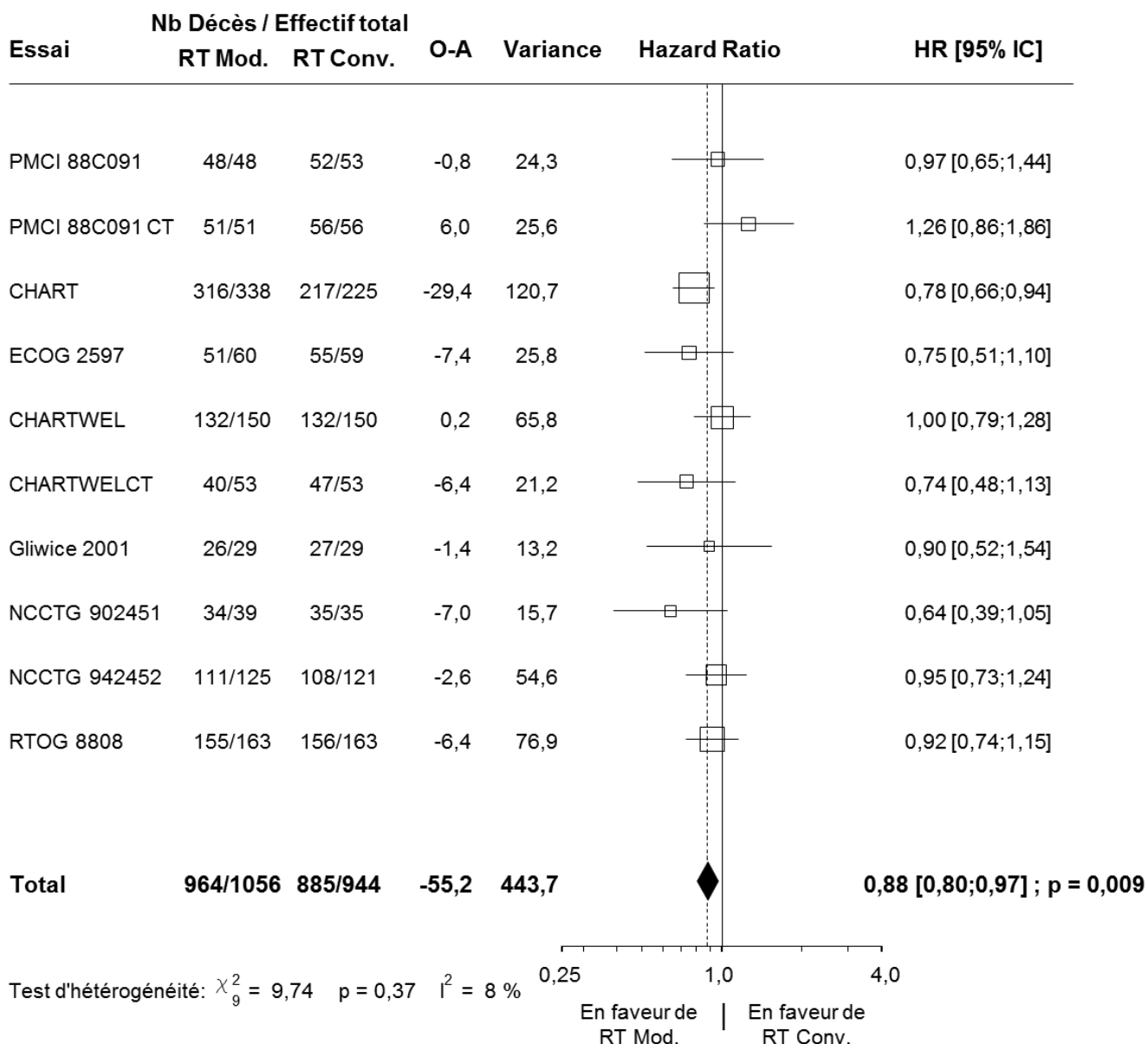


Figure 2.2 - Exemple de forest plot : comparaison de la radiothérapie modifiée (RT Mod) à la radiothérapie conventionnelle (RT Conv)

Légende de la Figure 2.2 :

Le *hazard ratio* global a été estimé par un modèle à effets fixes.

CHART: Continuous Hyperfractionated Accelerated Radiation Therapy; CHARTWEL: CHART Week-End Less; CT: Chimiothérapie; ECOG: Eastern Cooperative Oncology Group; HR: *Hazard ratio*; IC 95%: Intervalle de confiance à 95 %; NCCTG: North Central Cancer Treatment Group; O-A : Nombre d'événements observés moins ceux attendus dans le bras RT modifiée; PCMI: Peter MacCallum Institute; RTOG: Radiation Therapy Oncology Group; RT: Radiothérapie

2.2.2.e. Tests d'hétérogénéité

Le test d'hétérogénéité (appelé aussi test Q de Cochran) teste si les résultats de tous les essais peuvent être considérés comme similaires (Cochran, 1954 ; Whitehead, 2002). C'est l'hypothèse d'homogénéité (H_0) ; l'estimation de l'effet traitement global est alors plus facile à interpréter. Si le test d'hétérogénéité est significatif, il existe au moins un essai dont le résultat ne peut pas être considéré comme comparable aux autres. Les hypothèses testées sont donc les suivantes :

- H_0 : l'effet du traitement est homogène dans tous les essais ;
- H_1 : l'effet du traitement dans au moins un essai diffère de l'effet global.

La statistique de test est la suivante :

$$Q = \sum_{j=1}^J w_j (\hat{\theta}_j - \hat{\theta}_{global, fixe})^2 \quad (2.20)$$

qui suit une loi du Khi^2 à $J-1$ degrés de liberté, avec J essais et où $w_j = \frac{1}{\text{Var}(\hat{\theta}_j)}$ et $\hat{\theta}_{global, fixe}$ est l'estimateur de l'effet traitement global par un modèle à effets fixes (équation (2.18)). En utilisant l'estimateur $\widehat{HR}_{Peto}$ défini dans l'équation (2.8), cette formule peut se réécrire :

$$Q_{Peto} = \sum_{j=1}^J \left[\frac{(O_{Exp,j} - A_{Exp,j})^2}{\text{Var}(O_{Exp,j} - A_{Exp,j})} \right] - \frac{[\sum_{j=1}^J (O_{Exp,j} - A_{Exp,j})]^2}{\sum_{j=1}^J \text{Var}(O_{Exp,j} - A_{Exp,j})} \quad (2.21)$$

qui suit une loi du Khi^2 à $J-1$ degrés de liberté, avec J essais. Le test d'hétérogénéité étant peu puissant, on choisit le plus souvent pour ce test un seuil de significativité de 10 %.

La statistique I^2 définie par Higgins permet de quantifier l'hétérogénéité de l'effet traitement et se calcule de la manière suivante (Higgins *et al.*, 2003) :

$$I^2 = \frac{Q - (J - 1)}{Q} \times 100 \quad (2.22)$$

avec Q la statistique du test de Cochran et J le nombre d'essais. Cette statistique de test donne le pourcentage de variabilité de l'effet traitement, qui est dû à l'hétérogénéité plutôt qu'à l'erreur d'échantillonnage. Si ce pourcentage est compris entre 25% et 50%, il reflète une hétérogénéité notable. Une valeur supérieure à 50 % traduit une hétérogénéité importante.

Le test d'hétérogénéité Q peut être également utilisé pour tester la variation de l'effet du traitement selon des sous-groupes d'essais (par exemple, le type de chimiothérapie ou de

radiothérapie). Dans ce cas, on l'appelle le plus souvent test d'interaction. Le Khi^2 d'hétérogénéité s'obtient en calculant O-A et $\text{Var}(\text{O-A})$ dans chaque sous-groupe d'essais, puis en les sommant selon la formule (2.21) définissant la statistique Q présentée auparavant. Cette statistique de test suit une loi du Khi^2 à $K-1$ degrés de liberté avec K le nombre de sous-groupes.

Formellement, l'hétérogénéité totale entre les essais est la somme de l'hétérogénéité inter-sous-groupe et l'hétérogénéité (résiduelle) intra-sous-groupe :

$$\chi_{J-1}^2(\text{Het}_{\text{totale}}) = \chi_{K-1}^2(\text{Het}_{\text{Inter}}) + \chi_{J-K}^2(\text{Het}_{\text{Intra}}) \quad (2.23)$$

avec J le nombre d'essais de la méta-analyse et K le nombre de sous-groupes.

Le test associé à l'hétérogénéité inter-sous-groupe correspond au test d'interaction des sous-groupes avec l'effet traitement. Plus l'hétérogénéité inter-sous-groupe est importante, plus l'hétérogénéité totale entre les essais est expliquée par cette classification en sous-groupes.

La Figure 2.3 présente un test d'interaction avec $K = 4$ sous-groupes d'essai : ce test correspond à tester que les quatre HR (illustrés par les losanges blancs) sont égaux. Ce test étant non-significatif ($p = 0,98$ avec $K-1=3$ degrés de liberté pour le Khi^2), nous en concluons que la radiothérapie modifiée est plus efficace que la radiothérapie conventionnelle quel que soit le sous-groupe de radiothérapie modifiée considéré.

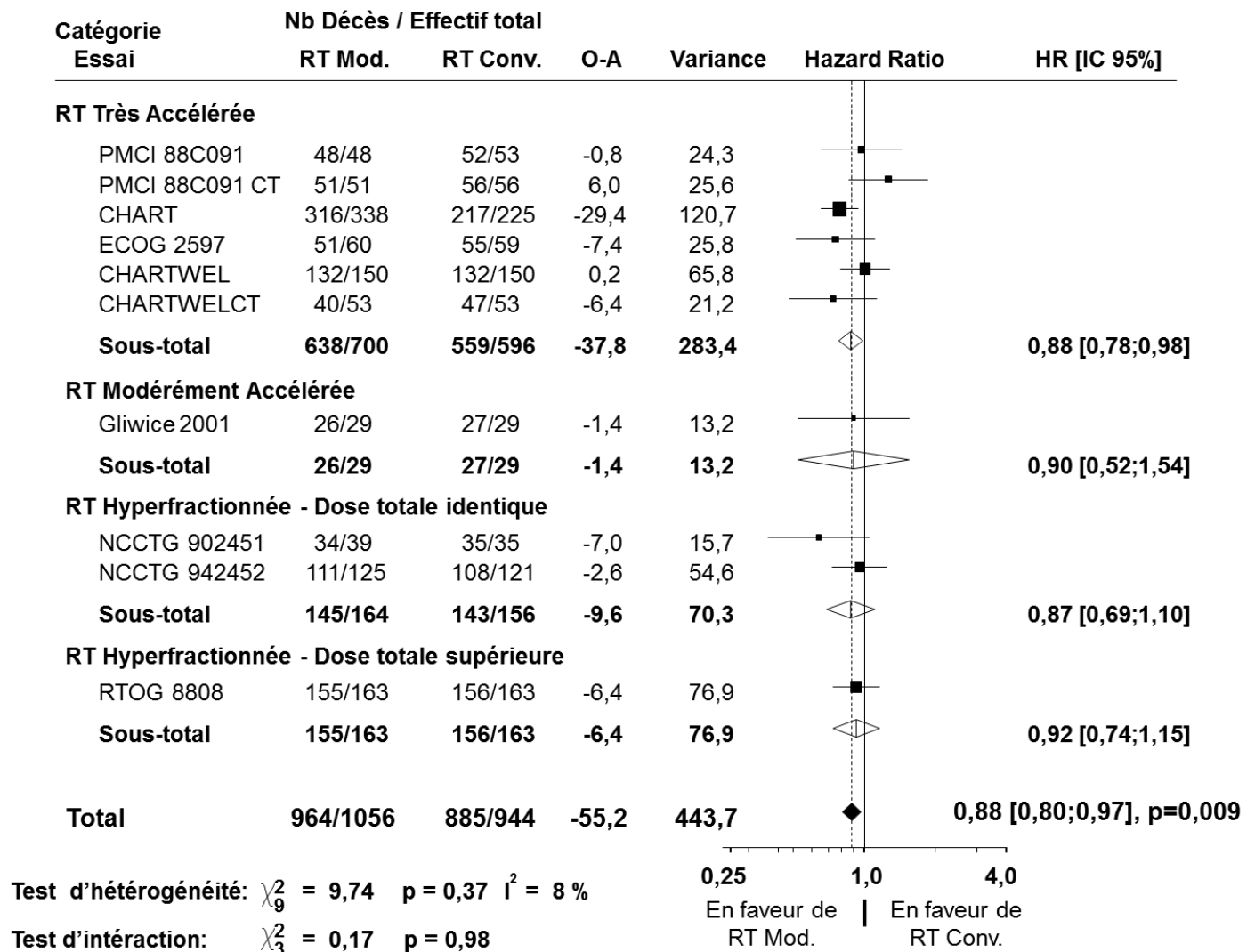


Figure 2.3 - Exemple de forest plot avec sous-groupes d'essai : comparaison de la radiothérapie modifiée (RT Mod) à la radiothérapie conventionnelle (RT Conv)

Légende de la Figure 2.3 :

Les *hazard ratio* de chaque sous-groupe ont été estimés par un modèle à effets fixes.

HR: *Hazard ratio*; IC 95%: Intervalle de confiance à 95 %; O-A : Nombre d'évènements observés moins ceux attendus dans le bras RT modifiée; RT: Radiothérapie

2.3. Estimation du bénéfice de survie en MADI

Afin d'estimer le bénéfice de survie restreint $rmstD(t^*)$ à partir d'une méta-analyse sur données individuelles (MADI), nous avons considéré des méthodes dites à deux étapes (*two-stage*) que Wei et al. ont utilisées dans le même contexte (Wei *et al.*, 2015). Parmi les trois méthodes étudiées par Wei et al, qui conduisaient à des résultats similaires, nous avons retenu celle basée sur l'estimateur de Kaplan-Meier. Nous avons également choisi d'adapter d'autres méthodes non-paramétriques utilisées en MADI. Ces différentes méthodes peuvent être classées en deux approches.

Dans la première approche, le $rmstD(t^*)$ est directement estimé comme l'aire entre les deux courbes de survie agrégées tenant compte ou non de l'effet essai. Les méthodes de cette approche utilisées dans cette thèse sont les méthodes dites « Naïve Kaplan-Meier », « Stewart-Parmar » et « Peto » ; nous les détaillons dans les paragraphes suivants.

Dans la deuxième approche, les méthodes à deux étapes estiment le $rmstD(t^*)$ comme la moyenne pondérée des $rmstD_j(t^*)$ estimés dans chaque essai j . Dans la première étape, le $rmstD(t^*)$ est estimé dans chaque essai clinique comme l'aire entre les deux courbes de survie. Dans la deuxième étape, les $rmstD_j(t^*)$ sont ensuite agrégés à l'aide d'un modèle à effets fixes ou un modèle à effets aléatoires. C'est cette deuxième approche qui avait été retenue par Wei et al. (Wei *et al.*, 2015). Dans cette thèse, les $rmstD_j(t^*)$ sont estimés dans chaque essai j comme l'aire entre les deux courbes de survie définies soit par Kaplan-Meier (méthode dite « Pooled Kaplan-Meier ») soit par un modèle exponentiel (méthode dite « Pooled Exponentiel »). Ces deux méthodes sont également détaillées dans les paragraphes suivants.

2.3.1. Méthode « Naïve Kaplan-Meier »

La méthode la plus simple pour estimer le bénéfice de survie restreint $rmstD(t^*)$ est de considérer la méta-analyse comme un essai unique de grande taille. Le $rmstD(t^*)$ est alors estimé comme l'aire entre les deux courbes de survie de Kaplan-Meier des bras expérimental et contrôle, obtenu en agrégeant les données de tous les essais, et ainsi en ignorant l'effet essai (Durand-Zaleski *et al.*, 1997) :

$$\begin{aligned} \widehat{rmstD}_{KM}(t^*) = & \sum_{i=1}^{E_{Exp}-1} \hat{S}_{KM,Exp}(t_{Exp,i})(t_{Exp,i+1} - t_{Exp,i}) \\ & - \sum_{i=1}^{E_{Cont}-1} \hat{S}_{KM,Cont}(t_{Cont,i})(t_{Cont,i+1} - t_{Cont,i}) \end{aligned} \quad (2.24)$$

avec $\hat{S}_{KM,bras}$ l'estimateur de Kaplan-Meier défini pour les bras expérimental et contrôle (notés Exp et Cont), $t_{bras,0} = 0$, $\hat{S}_{KM,bras}(t_{bras,0}) = 1$, et E_{bras} le nombre de temps d'évènement distincts ($t_{bras,1} < t_{bras,2} < \dots < t_{bras,E_{bras}}$).

Afin d'estimer la variance du bénéfice de survie, nous utilisons la méthode delta qui est dérivée des deux premiers termes d'une approximation des séries de Taylor (Hosmer *et al.*, 2008a). Cette méthode permet d'estimer la variance de $f(X)$ fonction d'une variable aléatoire X :

$$Var[f(X)] = Var[X] \left[\frac{\partial f(X)}{\partial X} \right]^2 \quad (2.25)$$

A partir de la méthode delta, nous pouvons définir la formule dite « de Greenwood » pour estimer la variance de la fonction de survie de Kaplan-Meier (Hosmer *et al.*, 2008b) :

$$\widehat{Var}[\hat{S}_{KM}(t)] = [\hat{S}_{KM}(t)]^2 \sum_{t_i \leq t_E} \frac{d_i}{Y_i(Y_i - d_i)} \quad (2.26)$$

En appliquant la formule de Greenwood à la survie moyenne restreinte estimée par Kaplan-Meier

$\widehat{RMST}_{KM}(t^*) = \int_0^{t^*} \hat{S}_{KM}(t) dt$ nous obtenons un estimateur de $\widehat{Var}[\widehat{RMST}_{KM}(t^*)]$ (Klein & Moeschberger, 2003; Hosmer *et al.*, 2008b):

$$\widehat{Var}[\widehat{RMST}_{KM}(t^*)] = \sum_{i=1}^E \left[\int_{t_i}^{t^*} \hat{S}_{KM}(t) dt \right]^2 \frac{d_i}{Y_i(Y_i - d_i)} \quad (2.27)$$

En remplaçant $\hat{S}_{KM}(t)$ par l'estimateur de Kaplan-Meier de l'équation (2.3), nous obtenons (Hosmer *et al.*, 2008b) :

$$\widehat{Var}(\widehat{RMST}_{KM}(t^*)) = \frac{m}{m-1} \sum_{i=1}^{E-1} \frac{a_i^2}{Y_i(Y_i - d_i)} \quad (2.28)$$

où $a_i = \sum_{l=i}^{E-1} \hat{S}_{KM}(t_l)(t_{l+1} - t_l)$ et $m = \sum_{l=1}^E d_l$.

La variance du bénéfice de survie restreint $\widehat{Var}[rmstD_{KM}(t^*)]$ est estimée par ré-échantillonnage non-paramétrique (*bootstrap*) stratifié sur l'essai – i.e. en conservant le nombre de patients inclus dans chaque essai de l'échantillon initial – sur 1 000 échantillons dans le Chapitre 3 et de façon analytique à l'aide des formules (2.16) et (2.28) dans les Chapitre 4 et Chapitre 5.

Le désavantage majeur de la méthode dite « Naïve Kaplan-Meier » est de ne pas tenir compte ni de la stratification sur l'essai ni de l'hétérogénéité potentielle de l'effet traitement puisque tous les patients du bras contrôle (et du bras expérimental) sont regroupés quel que soit leur essai d'origine et donc quel que soit leur risque de base.

2.3.2. Méthode « Stewart-Parmar »

La méthode de Stewart et Parmar est une méthode non-paramétrique dérivée de celle de Kaplan-Meier (Stewart & Parmar, 1993). La méthode originelle de Stewart et Parmar permet d'estimer la probabilité de survie à un temps t donné dans le bras expérimental d'une méta-analyse à partir de deux paramètres :

- la probabilité de survie au temps t dans le bras contrôle estimée par Kaplan-Meier,
- le *hazard ratio* global HR_{global} de la méta-analyse.

Nous avons appliqué cette méthode dite « Stewart-Parmar » à tous les temps d'évènement afin d'estimer la courbe de survie agrégée du bras expérimental $\hat{S}_{SP,Exp}(t)$ à partir :

- de la courbe de survie de Kaplan-Meier $\hat{S}_{KM,Cont}(t)$ du bras contrôle, en agrégeant les données de tous les essais sans tenir compte de l'effet essai,
- du *hazard ratio* global HR_{global} de la méta-analyse.

La courbe de survie agrégée du bras contrôle $\hat{S}_{SP,Cont}(t)$ est définie égale à celle de Kaplan-Meier $\hat{S}_{KM,Cont}(t)$:

$$\hat{S}_{SP,Cont}(t) = \hat{S}_{KM,Cont}(t) \quad (2.29)$$

$$\hat{S}_{SP,Exp}(t) = [\hat{S}_{KM,Cont}(t)]^{HR_{global}} \quad (2.30)$$

A noter que la formule (2.30) repose sur l'hypothèse de proportionnalité des risques.

Le bénéfice de survie restreint $\widehat{rmstD}_{SP}(t^*)$ est alors estimé comme l'aire entre les deux courbes de survie définies par $\hat{S}_{SP,Cont}(t)$ et $\hat{S}_{SP,Exp}(t)$:

$$\widehat{rmstD}_{SP}(t^*) = \int_0^{t^*} \hat{S}_{SP,Exp}(t) dt - \int_0^{t^*} \hat{S}_{SP,Cont}(t) dt \quad (2.31)$$

La variance du bénéfice de survie restreint $\widehat{Var}[\widehat{rmstD}_{SP}(t^*)]$ est estimée par *bootstrap* non-paramétrique sur 1 000 échantillons dans le Chapitre 3.

L'inconvénient majeur de la méthode Stewart-Parmar est de ne tenir compte de la stratification sur l'essai seulement dans le calcul du *hazard ratio* global HR_{global} . L'effet essai n'est cependant pas pris en compte pour le risque de base puisque, similairement à la méthode Naïve Kaplan-Meier, tous les patients du bras contrôle sont regroupés quel que soit leur essai d'origine et donc quel que soit leur risque de base. De plus, cette méthode repose par définition sur l'hypothèse des risques proportionnels. L'hétérogénéité potentielle de l'effet traitement entre les essais peut, en revanche, être prise en compte dans le calcul du HR_{global} utilisé dans l'équation (2.30).

2.3.3. Méthode « Peto »

La méthode actuarielle dite « Peto » a été développée par Richard Peto afin d'implémenter des courbes de survie pour les méta-analyse sur données individuelles en tenant compte de l'effet essai (Early Breast Cancer Trialists' Collaborative Group, 1990, 2012; Rotolo & Michiels, 2014). Dans cette méthode, les probabilités de survie dans les deux bras de traitement sont estimées à la fin d'intervalle de temps prédéterminés i , à partir de la probabilité de survie p_i de l'ensemble des patients des deux bras confondus et du *hazard ratio* global $HR_{i,global}$.

La probabilité de survie $\hat{p}_i$ de l'ensemble de la population est estimée à la fin de l'intervalle de temps i par :

$$\hat{p}_i = \exp\left(-\frac{D_i}{PI_i}\right) \quad (2.32)$$

où D_i est le nombre de décès dans l'intervalle de temps i et PI_i est le nombre total de personnes-intervalle dans l' i -ème intervalle. Une personne-intervalle est équivalent à une personne-année lorsque l'intervalle de temps choisi est l'année. Le *hazard ratio* global $\widehat{HR}_{i,global}$ dans l'intervalle i est estimé soit par un modèle à effets fixes (Chapitre 3, Chapitre 4 et Chapitre 5) soit par un

modèle à effets aléatoires (Chapitre 4 et Chapitre 5) pour combiner les différents $\widehat{HR}_{i,j}$ estimés dans chaque essai j avec l'estimateur de Peto (défini à l'équation (2.8)). Les probabilités de survie à la fin de chaque intervalle i dans le bras contrôle ($\hat{p}_{Cont,i}$) et dans le bras expérimental ($\hat{p}_{Exp,i}$) sont estimés comme suit :

$$\hat{p}_{Cont,i} = \hat{p}_i - [0,5\hat{p}_i(\hat{p}_i - 1)\log(\widehat{HR}_{i,global})] \quad (2.33)$$

$$\hat{p}_{Exp,i} = \hat{p}_i + [0,5\hat{p}_i(\hat{p}_i - 1)\log(\widehat{HR}_{i,global})] \quad (2.34)$$

La fonction de survie au temps t dans chaque bras est le produit des probabilités de survie sur les intervalles de temps n_i jusqu'à t :

$$\hat{S}_{Peto,Cont}(t) = \prod_{i=1}^{n_i} \hat{p}_{Cont,i} \text{ et } \hat{S}_{Peto,Exp}(t) = \prod_{i=1}^{n_i} \hat{p}_{Exp,i} \quad (2.35)$$

Le bénéfice de survie restreint $\widehat{rmstD}_{Peto}(t^*)$ est estimé comme l'aire entre les deux courbes de survie définies par $\hat{S}_{Peto,Cont}(t)$ et $\hat{S}_{Peto,Exp}(t)$:

$$\begin{aligned} \widehat{rmstD}_{Peto}(t^*) = \sum_{i=0}^{n_i} \frac{(t_{i+1} - t_i)}{2} & \left[(\hat{S}_{Peto,Exp}(t_{i+1}) - \hat{S}_{Peto,Cont}(t_{i+1})) \right. \\ & \left. + (\hat{S}_{Peto,Exp}(t_i) - \hat{S}_{Peto,Cont}(t_i)) \right] \end{aligned} \quad (2.36)$$

où $t_0 = 0$ et $(t_1, \dots, t_{n_i} = t^*)$ représentent les intervalles de temps prédéterminés.

La méthode Peto a été utilisée originellement avec un modèle à effets fixes pour l'estimation du $HR_{i,global}$ et des intervalles de temps d'un an (Early Breast Cancer Trialists' Collaborative Group, 1990). Nous avons cependant considéré trois définitions d'intervalle de temps différentes : une année (« Peto-année »), un mois (« Peto-mois ») et un intervalle de temps basé sur les quintiles du nombre total de décès survenant avant t^* dans la méta-analyse (« Peto-quintile »). Les méthodes Peto-année et Peto-mois sont utilisées dans le Chapitre 3 et la méthode Peto-quintile est utilisée dans les Chapitre 3 et Chapitre 4.

La variance $\widehat{Var}(\widehat{rmstD}_{Peto}(t^*))$ est estimée par *bootstrap* stratifié sur l'essai sur 1 000 échantillons dans le Chapitre 3 et par *bootstrap* « classique » sur 50 échantillons dans le Chapitre 4 et le Chapitre 5. Le faible nombre (50) d'échantillons simulés dans les Chapitre 4 et Chapitre 5 est à mettre en parallèle du nombre de simulations (1 000) réalisées dans ces chapitres.

Les avantages de cette méthode sont de tenir compte de la potentielle hétérogénéité de l'effet traitement dans le calcul du $HR_{i,global}$ et de permettre que le *hazard ratio* varie d'une période i à une autre et donc de relâcher l'hypothèse de proportionnalité des risques. De plus, la méthode Peto permet de prendre en compte l'hétérogénéité de l'effet traitement dans le calcul du

$HR_{i,global}$. Cependant, un inconvénient de la méthode Peto est de ne pas prendre en compte l'hétérogénéité du risque de base, égal à $(1 - \text{la probabilité de survie } \hat{p}_i)$, qui est estimé tous essais et tous bras confondus (équation (2.32)).

2.3.4. Méthode « Pooled Kaplan-Meier »

La méthode dite « Pooled Kaplan-Meier », est basée sur l'agrégation des bénéfices de survie restreints $\widehat{rmstD}_{KM,j}$ estimés dans chaque essai j comme l'aire entre les courbes de survie de Kaplan-Meier des deux bras de traitement :

$$\begin{aligned} \widehat{rmstD}_{KM,j}(t^*) = & \sum_{i=1}^{E_{j,Exp}-1} \hat{S}_{KM,j,Exp}(t_{j,Exp,i})(t_{j,Exp,i+1} - t_{j,Exp,i}) \\ & - \sum_{i=1}^{E_{j,Cont}-1} \hat{S}_{KM,j,Cont}(t_{j,Cont,i})(t_{j,Cont,i+1} - t_{j,Cont,i}) \end{aligned} \quad (2.37)$$

avec pour chaque essai j : $\hat{S}_{KM,j,bras}$ l'estimateur de Kaplan-Meier défini pour les bras expérimental et contrôle (notés Exp et Cont), $t_{j,bras,0} = 0$, $\hat{S}_{KM,j,bras}(t_{j,bras,0}) = 1$, et $E_{j,bras}$ le nombre de temps d'évènement distincts ($t_{j,bras,1} < t_{j,bras,2} < \dots < t_{j,bras,E_{j,bras}}$).

Le bénéfice de survie restreint global $\widehat{rmstD}_{KM,global}(t^*)$ est estimé soit par un modèle à effets fixes (Chapitre 3, Chapitre 4 et Chapitre 5) soit par un modèle à effets aléatoires (Chapitre 4 et Chapitre 5) pour combiner les différents $\widehat{rmstD}_{KM,j}(t^*)$ estimés dans chaque essai j . Ces deux types de modèle d'agrégation sont définis à l'équation (2.18).

L'estimation de la variance du bénéfice de survie restreint dans chaque essai j $\widehat{Var}[\widehat{rmstD}_{KM,j}(t^*)]$ est basée sur les formules (2.16) et (2.28) dans les Chapitre 3, Chapitre 4 et Chapitre 5. La variance du bénéfice de survie restreint global pour la méthode Pooled Kaplan-Meier $\widehat{Var}[\widehat{rmstD}_{KM,global}(t^*)]$ est estimée par *bootstrap* stratifié sur l'essai sur 1 000 échantillons dans le Chapitre 3 et analytiquement à l'aide de la formule (2.18) dans les Chapitre 4 et Chapitre 5.

La méthode Pooled Kaplan-Meier permet de tenir compte de la stratification sur l'essai et de la potentielle hétérogénéité de l'effet traitement dans le modèle à deux étapes utilisé pour agréger les $\widehat{rmstD}_{KM,j}(t^*)$ estimés dans chaque essai j . De plus, cette méthode s'applique si les risques ne sont pas proportionnels.

2.3.5. Méthode « Pooled Exponentiel »

De manière similaire à la méthode Pooled Kaplan-Meier, la méthode dite « Pooled Exponentiel » est basée sur l'agrégation des $\widehat{rmstD}_{Exponentiel,j}(t^*)$ estimés dans chaque essai j comme l'aire entre les courbes de survie exponentielles des deux bras de traitement :

$$\begin{aligned}\widehat{rmstD}_{Exponentiel,j} &= \int_0^{t^*} e^{-\hat{\lambda}_{j,Exp}t} dt - \int_0^{t^*} e^{-\hat{\lambda}_{j,Cont}t} dt \\ &= \frac{1 - e^{-\hat{\lambda}_{j,Exp}t^*}}{\hat{\lambda}_{j,Exp}} - \frac{1 - e^{-\hat{\lambda}_{j,Cont}t^*}}{\hat{\lambda}_{j,Cont}}\end{aligned}\quad (2.38)$$

où pour chaque essai j , $\hat{\lambda}_{j,bras}$ est le paramètre d'échelle de la distribution exponentielle estimé par la méthode du maximum de vraisemblance pour le bras contrôle et le bras expérimental (notés Cont et Exp).

Le bénéfice de survie restreint global $\widehat{rmstD}_{Exponentiel,global}(t^*)$ est estimé soit par un modèle à effets fixes (Chapitre 3, Chapitre 4 et Chapitre 5) soit par un modèle à effets aléatoires (Chapitre 4 et Chapitre 5) pour combiner les différents $\widehat{rmstD}_{Exponentiel,j}(t^*)$ estimés dans chaque essai j . Ces deux types de modèle d'agrégation sont définis à l'équation (2.18).

La variance de la survie moyenne restreinte $\widehat{RMST}_{Exponentiel,j,bras}(t^*)$ pour un bras donné et un essai j donné est :

$$\widehat{Var}[\widehat{RMST}_{Exponentiel,j,bras}(t^*)] = Var\left[\frac{1 - e^{-\hat{\lambda}_{j,bras}t^*}}{\hat{\lambda}_{j,bras}}\right]\quad (2.39)$$

En appliquant la méthode delta (formule (2.25)) à l'équation (2.39) ci-dessus, nous obtenons :

$$\widehat{Var}[\widehat{RMST}_{Exponentiel,j,bras}(t^*)] = \widehat{Var}[\hat{\lambda}_{j,bras}] \left[\frac{\partial f(\hat{\lambda}_{j,bras})}{\partial \hat{\lambda}_{j,bras}} \right]^2\quad (2.40)$$

En dérivant cette formule (2.40), nous obtenons l'estimateur de la variance de la survie moyenne restreinte suivant :

$$\widehat{Var}[\widehat{RMST}_{Exponentiel,j,bras}(t^*)] = \widehat{Var}[\hat{\lambda}_{j,bras}] \left[\frac{\hat{\lambda}_{j,bras} \times t^* \times e^{-\hat{\lambda}_{j,bras}t^*} - (1 - e^{-\hat{\lambda}_{j,bras}t^*})}{\hat{\lambda}_{j,bras}^2} \right]^2\quad (2.41)$$

L'estimation de la variance du bénéfice de survie restreint dans chaque essai j $\widehat{Var}[\widehat{rmstD}_{Exponentiel,j}(t^*)]$ est basée sur cette dernière formule (2.41) et la formule (2.16), pour les Chapitre 3, Chapitre 4 et Chapitre 5. La variance du bénéfice de survie restreint global pour la méthode Pooled Exponentiel $\widehat{Var}[\widehat{rmstD}_{Exponentiel,global}(t^*)]$ est estimée par *bootstrap* stratifié sur l'essai sur 1 000 échantillons dans le Chapitre 3 et analytiquement à l'aide de la formule (2.18) dans les Chapitre 4 et Chapitre 5.

Tout comme la méthode Pooled Kaplan-Meier, la méthode Pooled Exponentiel permet de tenir compte de la stratification sur l'essai et de la potentielle hétérogénéité de l'effet traitement lors de l'agrégation des $\widehat{rmstD}_{KM,j}(t^*)$ estimés dans chaque essai j . L'inconvénient de la méthode Pooled Exponentiel est qu'elle repose sur la distribution exponentielle qui est un modèle paramétrique à risques proportionnels et qui repose sur l'hypothèse d'un risque instantané constant.

Les méthodes d'estimation du bénéfice de survie restreint à partir d'une méta-analyse sur données individuelles et leurs caractéristiques sont résumées dans le Tableau 2.1.

Tableau 2.1 - Caractéristiques des méthodes d'estimation du bénéfice de survie restreint à partir de méta-analyses sur données individuelles

Méthode d'estimation	Prend en compte			Bénéfice de survie restreint
	Stratification sur l'essai	Hétérogénéité potentielle de l'effet traitement	Risques non-proportionnels	
Naïve Kaplan-Meier	Non	Non	Oui	
Stewart-Parmar	Non [*]	Oui [§]	Non ^μ	$rmstD(t^*) = \int_0^{t^*} S_{Exp}(t)dt - \int_0^{t^*} S_{Cont}(t)dt$
Peto	Non [*]	Oui [§]	Oui ^ℓ	
Pooled Kaplan-Meier	Oui	Oui [§]	Oui	Moyenne pondérée à l'aide d'un modèle à effets fixes ou à effets aléatoires pour agréger les $rmstD_j(t^*)$ estimés dans chaque essai j
Pooled Exponentiel	Oui	Oui [§]	Non	$rmstD_j(t^*) = \int_0^{t^*} S_{j,Exp}(t)dt - \int_0^{t^*} S_{j,Cont}(t)dt$

rmstD: différence de survie moyenne restreinte; S_{Exp} : fonction de survie du bras expérimental; S_{Cont} : fonction de survie du bras contrôle; $S_{j,Exp}$: fonction de survie du bras expérimental de l'essai j ; $S_{j,Cont}$: fonction de survie du bras contrôle de l'essai j ;

^{*} Stratification sur l'essai seulement prise en compte dans le calcul de l'effet traitement global ; [§] L'hétérogénéité de l'effet traitement peut être prise en compte dans le modèle à deux étapes utilisé pour agréger les HR_j (méthodes Stewart-Parmar et Peto) ou les $rmstD_j(t^*)$ (Pooled Kaplan-Meier et Pooled Exponentiel) estimés dans chaque essai ; ^μ La non-proportionnalité des risques peut cependant être prise en compte en utilisant un HR différent en fonction du temps ; ^ℓ Le $HR_{i,global}$ peut varier entre les intervalles de temps.

Chapitre 3. Comparaison de méthodes d'estimation du bénéfice de survie et impact en évaluation médico-économique : une étude de cas

3.1. Motivation

Dans ce chapitre, nous illustrons les différentes méthodes d'estimation du bénéfice de survie restreint à travers une étude de cas et dans le contexte d'une analyse coût-efficacité à partir de la méta-analyse sur données individuelles *Meta-Analysis of Radiotherapy in Lung Cancer* (MAR-LC). La méta-analyse MAR-LC avait déjà fait l'objet d'une analyse coût-efficacité dans le contexte des Pays-Bas à l'aide d'un modèle de Markov. Nous avons participé à la conception, l'analyse et l'interprétation de cette évaluation médico-économique (Ramaekers *et al.*, 2013). Dans ce chapitre, nous présentons les résultats de l'analyse coût-efficacité que nous avons réalisé à l'aide de l'approche « aire sous la courbe » (*area under the curve approach* en anglais) (Glasziou *et al.*, 1990; Rochau *et al.*, 2015) à partir des données individuelles de MAR-LC dans le contexte français.

L'objectif de ce Chapitre 3 est d'étudier si le choix d'une méthode d'estimation du bénéfice de survie restreint a un impact sur les résultats d'une étude coût-efficacité. Aucune méthode d'estimation du critère d'efficacité pour une analyse coût-efficacité à partir d'une méta-analyse sur données individuelles n'ayant été proposée par la Haute Autorité de Santé (HAS), ni par l'agence similaire en Grande-Bretagne (*National Institute for Health and Clinical Excellence* - NICE) ni même dans la littérature, il nous semblait opportun de comparer et d'étudier différentes méthodes d'estimation sur une étude de cas. Nous avons donc estimé le $rmstD(t^*)$ avec les sept méthodes d'estimation présentées dans le Chapitre 2 : Naïve Kaplan-Meier, Stewart-Parmar, Peto-mois, Peto-année, Peto-quintile, Pooled Kaplan-Meier, et Pooled Exponentiel. Dans ce Chapitre 3, nous présentons tout d'abord la méta-analyse MAR-LC, puis nous définissons la méthodologie et les résultats de l'analyse coût-efficacité.

3.2. Matériel : la méta-analyse *Meta-Analysis of Radiotherapy in Lung Cancer*

La méta-analyse sur données individuelles *Meta-Analysis of Radiotherapy in Lung Cancer* (MAR-LC) comparait la radiothérapie thoracique modifiée (hyper-fractionnée et/ou accélérée) à la radiothérapie conventionnelle chez des patients atteints de cancer bronchique non à petites cellules non métastatique (Mauguen *et al.*, 2012).

La radiothérapie conventionnelle pour ces patients consiste généralement à donner une dose total de 60 Gray (Gy), principalement aux USA, à 70 Gy, en Europe, à raison de 1,8 à 2 Gy par fraction, avec 5 fractions par semaine étalées sur une période de 6 à 7 semaines.

La radiothérapie hyper-fractionnée consiste à réduire la dose par fraction en augmentant le nombre de fractions. La même dose est alors administrée en 2 à 3 fractions par jour, au lieu d'une seule dans la radiothérapie conventionnelle (essais NCCTG).

La radiothérapie accélérée consiste à réduire la durée globale de la radiothérapie. Ceci est réalisé généralement en traitant les patients soit le week-end ou le samedi en plus des 5 autres jours de la semaine soit plusieurs fois par jour avec l'hyper-fractionnement.

Deux essais inclus dans MAR-LC comportaient également de la chimiothérapie (CT) en plus de la radiothérapie : CT concomitante pour l'essai PMCI 88C091 et néo-adjuvante pour l'essai CHARTWEL. L'essai PMCI 88C091 avait un plan factoriel 2x2 : RT modifiée vs. RT conventionnelle ; CT concomitante vs pas de CT concomitante. Dans l'essai CHARTWEL, la randomisation a été stratifiée sur l'administration d'une CT néo-adjuvante. Le design de ces deux essais a donc permis de les diviser en deux essais séparés, avec et sans CT dans les deux bras (PMCI 88C091 et PMCI 88C091 CT ; CHARTWEL et CHARTWEL CT). Au final, dix essais cliniques randomisés de phase III ont été inclus dans la méta-analyse MAR-LC, réalisés entre 1989 et 2006 (cf. Tableau 3.1).

Tableau 3.1 - Description des essais inclus dans la méta-analyse MAR-LC

Catégorie	Essais	Période d'inclusion	Radiothérapie Dose totale/ Nb fractions/ Durée	Caractéristiques des patients	Nombre de patients randomisés	Suivi médian (années)	Survie globale HR [IC 95%]
RT très accélérée	PMCI 88C091	1989-1995	Conventionnelle: 60 Gy / 30 fr / 6 s Modifiée: 60 Gy / 30 fr / 3 s bid	PS 0-1 Stade I-III	101	Non atteint	0,97 [0,65-1,44]
	PMCI 88C091 CT	1989-1995	Conventionnelle: 60 Gy / 30 fr / 6 s Modifiée: 60 Gy / 30 fr / 3 s bid	PS 0-1 Stade I-III	107	Non atteint	1,26 [0,86-1,86]
	CHART	1990-1995	Conventionnelle: 60 Gy / 30 fr / 6 s Modifiée: 54 Gy / 36 fr / 1,5 s tid	PS 0-1 Stade I-III	563	6,9	0,78 [0,66-0,94]
	ECOG 2597	1998-2001	Conventionnelle: 64 Gy / 32 fr / 6,5 s Modifiée: 57,6 Gy / 36 fr / 2,5 s tid	PS 0-1 Stade III	119	6,7	0,75 [0,51-1,10]
	CHARTWEL	1997-2005	Conventionnelle: 66 Gy / 33 fr / 6,5 s Modifiée: 60 Gy / 40 fr / 2,5 s tid	PS 0-1 Stade I-III	300	4,9	1,00 [0,79-1,28]
	CHARTWEL CT	1997-2005	Conventionnelle: 66 Gy / 33 fr / 6,5 s Modifiée: 60 Gy / 40 fr / 2,5 s tid	PS 0-1 Stade I-III	106	3,5	0,74 [0,48-1,13]
RT modérément accélérée	Gliwice 2001	2001-2006	Conventionnelle: 72 Gy / 40 fr / 8 s Modifiée: 72 Gy / 40 fr / 5,5 s	PS 0-1 Stade II-III	58	5,3	0,90 [0,52-1,54]
RT hyper-fractionnée – dose totale identique	NCCTG 902451	1992-1993	Conventionnelle: 60 Gy / 30 fr / 6 s Modifiée: 60 Gy / 40 fr / 6 s bid	PS 0-2 Stade III	74	8,1	0,64 [0,39-1,05]
	NCCTG 942452	1994-1999	Conventionnelle: 60 Gy / 30 fr / 6 s Modifiée: 60 Gy / 40 fr / 6 s bid	PS 0-1 Stade III	246	7,3	0,95 [0,73-1,24]
RT hyper-fractionnée – dose totale supérieure	RTOG 8808	1989-1992	Conventionnelle: 60 Gy / 30 fr / 6 s Modifiée: 69,6 Gy / 58 fr / 6 s bid	KPS $\geq$ 70 Stade II-III	326	6,8	0,92 [0,74-1,15]
Total		1989-2006	NA	NA	2 000	6,9	0,88 [0,80-0,97]

bid: Radiothérapie donnée deux fois par jour; CHART: Continuous Hyperfractionated Accelerated Radiation Therapy; CHARTWEL: CHART Week-End Less; CT: chimiothérapie; ECOG: Eastern Cooperative Oncology Group; fr: fraction; Gy: Gray; HR: Hazard ratio RT modifiée contre RT conventionnelle; IC: Intervalle de confiance; (K)PS: (Karnofsky) Performance Status; MAR-LC: Meta-Analysis of Radiotherapy in Lung Cancer; NCCTG: North Central Cancer Treatment Group; PCMI: Peter MacCallum Institute; RTOG: Radiation Therapy Oncology Group; RT: Radiothérapie; s: semaine ; tid: Radiothérapie donnée trois fois par jour.

Dans la méta-analyse MAR-LC, les essais ont été classés en 4 sous-groupes (cf. Figure 2.3, Figure 3.1 et Tableau 3.1) :

- (1) Six essais avec une « Radiothérapie très accélérée », avec un raccourcissement de la durée globale du traitement par radiothérapie de 50 % ou plus comparé au bras contrôle, utilisant une dose totale inférieure ou égale à celle du bras conventionnel (PMCI 88C091, PMCI 88C091 CT, CHART, ECOG 2597, CHARTWELL et CHARTWELL CT; 1 296 patients),
- (2) Un essai avec une « Radiothérapie modérément accélérée », avec un raccourcissement de la durée globale entre 15 et 50 %, utilisant une dose totale égale à celle du bras conventionnel (Gliwice 2001; 58 patients),
- (3) Deux essais avec une « Radiothérapie hyper-fractionnée et utilisant une dose totale identique » à celle du bras conventionnel (NCCTG 902451, NCCTG 942452; 320 patients),
- (4) Un essai avec une « Radiothérapie hyper-fractionnée et utilisant une dose totale supérieure » (RTOG 8808; 326 patients).

Dose totale		Inférieure		Egale (± 5%)		Supérieure	
A C C E L E R A T I O N	0-13%			NCCTG 902451 NCCTG 942452 (3)	RTOG 8808 (4)		
	14-49%			Gliwice 2001 (2)			
	≥ 50%	(1)	CHART CHARTWEL CHARTWEL-CT ECOG 2597	PMCI 88C091 PMCI 88C091 CT			
Dose par fraction (Gy)		Hyper-fractionné Normal		Hyper-fractionné Normal		Hyper-fractionné Normal	
		< 1,25 1,25-1,75	1,8-2	< 1,25 1,25-1,75	1,8-2	< 1,25 1,25-1,75	1,8-2

Figure 3.1 - Classification des essais selon le type de la radiothérapie modifiée

La méta-analyse MAR-LC comportait 2 000 patients, inclus dans 10 essais, avec un suivi médian de 6,9 années. Les données disponibles étaient :

- les caractéristiques du patient à l'inclusion (âge, sexe, score d'état général de l'OMS ou de Karnofsky) ;
- les caractéristiques de la tumeur (histologie, stade de la maladie) ;
- les dates de randomisation, date et cause du décès, date de rechute, date de dernières nouvelles ;
- le bras de traitement (radiothérapie modifiée ou conventionnelle) et des informations sur le traitement effectivement reçu (nombre de fractions, dose totale, date du premier et du dernier jour du traitement par radiothérapie) ;
- des informations sur les toxicités aiguës et tardives (les toxicités dont le grade était supérieur ou égal à 3 ont été considérées comme sévères) ;
- raison de l'exclusion pour les patients exclus de l'analyse.

Le critère de jugement principal de la méta-analyse était la survie globale, définie comme le délai entre la randomisation et le décès. Les patients vivants ont été censurés à la date de dernières nouvelles. Deux critères secondaires ont également été analysés :

- La survie sans rechute, définie comme le délai entre la randomisation et la date de rechute (locale ou à distance) ou de décès. Les patients n'ayant pas rechuté ou décédé en l'absence de rechute ont été censurés à la date de dernières nouvelles.
- Les toxicités, mesurées au moyen des échelles classiques NCI-CTC, OMS, RTOG/EORTC et définies comme :
 - aiguës lors des six premiers mois du traitement,
 - tardives, au-delà de six mois après le traitement.

Les résultats de l'analyse de la survie globale se fondaient sur 2 000 patients, dont 1 849 étaient décédés. Le risque de décès était significativement réduit de 12% avec la radiothérapie modifiée (HR=0,88 ; IC 95 % : [0,80 ; 0,97] ; p=0,009). Il n'y avait pas d'hétérogénéité de l'effet du traitement entre les essais (test d'hétérogénéité, p=0,37, I²=8%) et l'effet de la radiothérapie modifiée ne différait pas entre les quatre sous-groupes de radiothérapie (test d'interaction, p=0,98) (Figure 2.3, (Mauguen *et al.*, 2012)).

Les résultats de l'analyse de la survie sans rechute se fondaient sur 2 000 patients, dont 1 926 avaient rechuté ou étaient décédés. Aucun bénéfice significatif de la radiothérapie modifiée sur la survie sans rechute n'a été observé (HR=0,94 ; IC 95% : [0,86 ; 1,03] ; p=0,19). Il n'y avait pas d'hétérogénéité de l'effet du traitement entre les essais (p=0,28 ; I²=18%), ni d'interaction entre les quatre sous-groupes de radiothérapie (p=0,85). Concernant les toxicités, seul le risque de toxicité œsophagienne aiguë sévère était significativement augmenté avec la radiothérapie modifiée (OR=2,44 ; IC 95 % : [1,90 ; 3,14], p<0,001).

3.3. Analyse coût-efficacité

L'objectif de l'évaluation économique d'une action de santé est de mettre en regard son efficacité avec les coûts qu'elle engendre, afin d'optimiser l'allocation des ressources (HAS, 2011). La méta-analyse MAR-LC ayant démontré l'efficacité de la radiothérapie modifiée, il était également intéressant d'étudier si la radiothérapie modifiée (bras expérimental) est plus coût-efficace que la radiothérapie conventionnelle (bras contrôle).

Le critère de résultat d'une analyse coût-efficacité est le rapport coût-efficacité incrémental (*Incremental Cost Effectiveness Ratio* en anglais : ICER) qui représente le coût à consentir pour une unité de santé supplémentaire et qui est défini par :

$$ICER = \frac{\Delta C}{\Delta E} = \frac{\bar{C}_{Exp} - \bar{C}_{Cont}}{\bar{E}_{Exp} - \bar{E}_{Cont}} \quad (3.1)$$

où :

- ΔC est la différence entre le coût moyen du bras expérimental $\bar{C}_{Exp}$ et le coût moyen du bras contrôle $\bar{C}_{Cont}$,
- ΔE est la différence des effets entre le bras expérimental et le bras contrôle ; cette mesure d'efficacité est généralement exprimée soit en année de vie, soit en année de vie pondérée par la qualité de vie liée à la santé (QALY pour *Quality Adjusted Life Years*) (HAS, 2011).

Dans une étude coût-efficacité, les résultats de santé et les coûts sont considérés sur une même période. L'horizon temporel de notre étude a été fixé à $t^* = 5$ ans, qui était l'horizon temporel pour le critère de jugement principal de la méta-analyse MAR-LC : en effet, Mauguen et al. ont comparé les taux de survie des deux types de radiothérapie à 3 et à 5 ans (Mauguen *et al.*,

2012). De plus, ce choix de $t^* = 5$ ans était cohérent avec à la fois le faible taux de survie des patients à 5 ans (9 %) et le suivi des essais cliniques inclus dans MAR-LC (Tableau 3.1).

Dans notre étude de cas, le critère d'efficacité était le nombre d'années de vie gagnées grâce au traitement. La différence des effets ΔE correspondait donc au bénéfice de survie restreint à l'horizon temporel $t^* = 5$ ans ($rmstD(t^* = 5)$). Ainsi dans MAR-LC, l'ICER s'interprète comme le coût par année de vie gagnée de la radiothérapie modifiée et est défini par :

$$ICER = \frac{\bar{C}_{RT\ Mod} - \bar{C}_{RT\ Conv}}{rmstD(t^* = 5)} \quad (3.2)$$

Plus l'ICER est faible, plus le coût supplémentaire à investir dans la radiothérapie modifiée pour une année de vie gagnée est faible, et donc plus la radiothérapie modifiée est dite « coût-efficace ».

Dans certains pays, une règle de décision permet de situer l'ICER par rapport à un seuil indiquant la disposition à payer de la collectivité. Ainsi, en Grande-Bretagne, le NICE utilise un seuil entre 20k£ et 30k£ (entre 25k€ et 40k€) par QALY, et 70k£/QALY (90k€) pour un traitement de fin de vie (NICE, 2013). Ceci signifie que le système de santé britannique est disposé à financer un traitement dont le ratio coût-efficacité est inférieur à 20k£ (ou 30k£) par année de vie gagnée en parfaite santé. En France, la Haute Autorité de Santé (HAS) n'utilise pas de seuil de disposition à payer explicite pour fonder ses recommandations. Dans notre étude, différents seuils de disposition à payer ont été étudiés.

Afin de clarifier le message de notre étude de cas et de nous focaliser sur l'estimation du bénéfice de survie restreint à partir d'une méta-analyse sur données individuelles, nous n'avons pas tenu compte de la qualité de vie liée à la santé, et n'avons donc pas exprimé notre mesure d'efficacité en QALY. De même, nous n'avons pas étendu l'horizon temporel de l'analyse coût-efficacité à la vie entière comme cela est classiquement fait en évaluation médico-économique, mais nous nous sommes limités à la durée du suivi des essais. En tenant compte du fait que la probabilité de survie des patients atteints d'un cancer bronchique non à petites cellules était de 9 % à 5 ans, l'extrapolation de la survie au-delà de cet horizon n'était pas susceptible d'affecter nos résultats de manière importante. La HAS recommande également une actualisation des coûts et des résultats futurs afin de les ramener à leur valeur présente (HAS, 2011). Cependant, dans notre étude ni les coûts ni les résultats n'ont été actualisés puisque la médiane de survie des patients était seulement de 14 mois.

Les coûts moyens, les bénéfices de survie restreints moyens et les ICERs moyens, ainsi que leurs intervalles de confiance à 95 % ont été estimés par ré-échantillonnage non-paramétrique

(*bootstrap*) sur 1 000 échantillons. Ce *bootstrap* était stratifié sur l'essai afin de conserver le nombre de patients inclus dans chaque essai de l'échantillon initial de MAR-LC. Pour chaque échantillon du *bootstrap*, nous avons estimé le coût moyen par patient pour chaque type de radiothérapie, sept bénéfices de survie restreints – pour chacune des sept méthodes d'analyse de survie – et donc sept ICERs (défini à l'équation (3.2)).

3.3.1. Estimation des coûts

La perspective retenue dans l'évaluation des coûts dans notre analyse était celle de l'Assurance Maladie et les coûts étaient exprimés en euros 2012. Les coûts ont été estimés à partir des données de consommation de ressources disponibles dans la méta-analyse MAR-LC. Les quatre postes de ressource valorisés étaient le traitement par radiothérapie, les transports sanitaires pour la radiothérapie, le traitement de la rechute et le traitement de l'œsophagite aiguë. L'estimation de chacun de ces quatre postes de coût est détaillée dans les paragraphes suivants.

3.3.1.a. Coût de traitement par radiothérapie

Le coût de traitement par radiothérapie a été calculé, pour chaque patient, à partir du nombre de fractions reçues et des tarifs de la tarification à l'activité (T2A) par Groupe Homogène de Malades (GHM) pour la dosimétrie et l'irradiation. Pour chaque patient p , le coût de la radiothérapie $C_{\text{Traitement RT},p}$ par patient était donc ainsi calculé :

$$C_{\text{Traitement RT},p} = 1\,004 + n_{\text{fractions},p} \times 138 \quad (3.3)$$

avec $n_{\text{fractions},p}$ le nombre total de fractions reçues par le patient p .

3.3.1.b. Coût de transport pour la radiothérapie

La radiothérapie modifiée augmentant le nombre de fractions reçues par patient et donc, potentiellement, les coûts liés aux transports pouvaient être plus élevés dans ce bras de traitement. Il nous apparaissait important de valoriser ce poste de coût. Afin d'estimer les coûts de transport pour la radiothérapie pour les patients traités pour un cancer bronchique, nous avons utilisé un coût unitaire issu de la littérature (Martin, 2003). Cet article étudiait le coût global de la radiothérapie et proposait un coût de 80€ (soit 89,68 euros 2012) pour un aller/retour en

véhicule sanitaire léger. D'après cet article, les patients utilisaient fréquemment ce type de transport. Dans notre analyse, nous avons supposé que le taux de recours des patients aux transports sanitaires était de 50 % et le coût de transport pour la radiothérapie était donc estimé pour chaque patient p par :

$$C_{Transport\ RT,p} = 44,84 \times n_{fractions,p} \quad (3.4)$$

3.3.1.c. Coût de traitement de la rechute

Dans la méta-analyse MAR-LC, les durées de survie après rechute étaient disponibles pour chaque patient ayant rechuté. Le coût unitaire de traitement de la rechute a lui été estimé à partir des données d'un article d'A.C. Braud. Cet article étudiait une population de patients similaire à celle de la MAR-LC, dans un contexte français (Braud *et al.*, 2003). Dans cet article, les coûts directs de traitement ont été estimés à partir des données d'une série de patients atteints d'un cancer bronchique non à petite cellule entre la première rechute et le décès. Nous avons calculé un coût mensuel de traitement de la rechute égal à 3 073 euros 2012. Nous avons ensuite appliqué ce coût mensuel aux délais observés pour les patients de la méta-analyse MAR-LC. Finalement, le coût de traitement de la rechute $C_{Rechute}$ pour un patient ayant rechuté était estimé par :

$$C_{Rechute} = 3\,073 \times T_{post-rechute} \quad (3.5)$$

où $T_{post-rechute}$ est la durée de survie après rechute exprimé en mois.

3.3.1.d. Coût de traitement de l'œsophagite aigüe

Nous avons valorisé le coût de traitement de l'œsophagite aigüe puisque le risque de toxicité œsophagienne aigüe sévère était significativement augmenté avec la radiothérapie modifiée (OR=2,44 ; IC 95 % : [1,90 ; 3,14], $p < 0,001$). Nous avons fait l'hypothèse que la présence d'une toxicité œsophagienne aigüe sévère nécessitait une hospitalisation. Le coût de l'œsophagite aigüe a été valorisé à l'aide du classement par GHM des séjours de la base PMSI pour l'œsophagite (code K20 de la 10^{ème} Classification Internationale des Maladies [CIM10]). Nous avons retenu les trois principaux GHM dans lesquels les séjours comportant ce diagnostic principal sont classés et nous avons calculé un coût moyen pondéré par les effectifs de la base PMSI 2012 égal à 1 745€. Finalement, le coût de traitement de l'œsophagite aigüe était :

$$C_{Oesophagite} = 1\,745 \quad (3.6)$$

si le patient avait subi une toxicité aigüe œsophagienne sévère.

Les méthodes d'estimation et les sources des quatre postes de coût qui ont été valorisés sont résumées dans le Tableau 3.2.

Tableau 3.2 - Estimation des différents coûts selon le type de radiothérapie

Poste de coût	Ressource MAR-LC	Coût unitaire	Source
Traitement par radiothérapie	Nombre total de fractions reçues	1 004€ dosimétrie + 138€/fraction d'irradiation	PMSI 2012
Transports pour la radiothérapie	Nombre total de fractions reçues	40€/fraction d'irradiation	Martin, 2003
Traitement de la rechute	Temps de survie après rechute	3 073€/mois	Braud, 2003
Traitement de l'œsophagite	Toxicité œsophagienne aigüe sévère (oui ou non)	1 745€ si présence de toxicité	PMSI 2012

Finalement, pour chaque bras de radiothérapie, le coût total moyen par patient a été estimé comme la somme des 4 coûts moyens par patient pour chacun des postes de coût.

Dans la méta-analyse MAR-LC, le coût moyen par patient était de 25 331€ (IC 95 % : [23 630€ ; 27 115€]) pour la radiothérapie conventionnelle et de 29 659€ (IC 95 % : [27 845€ ; 31 507€]) pour la radiothérapie modifiée, ce qui correspondait à un coût moyen incrémental de 4 328€ (IC 95% : [1 830€ ; 6 804€]). Le coût moyen du traitement de la rechute et le coût moyen du traitement par radiothérapie représentaient environ respectivement 70 % et 20 % du coût total moyen par patient.

3.3.2. Estimation du bénéfice de survie restreint et du ratio coût-efficacité incrémental

Le critère d'efficacité de l'analyse coût-efficacité était défini comme le bénéfice de survie restreint à 5 ans. Nous avons donc estimé le $rmstD(t^*)$ avec les sept méthodes d'estimation présentées dans le Chapitre 2 : Naïve Kaplan-Meier, Stewart-Parmar, Peto-mois, Peto-année, Peto-quintile, Pooled Kaplan-Meier, et Pooled Exponentiel.

Les courbes de survie estimées pour les deux bras de traitement de la méta-analyse MAR-LC à l'aide des méthodes Naïve Kaplan-Meier et Stewart-Parmar, et des méthodes Peto-mois, Peto-année, et Peto-quintile sont respectivement présentées dans les Figure 3.2 et Figure 3.3.

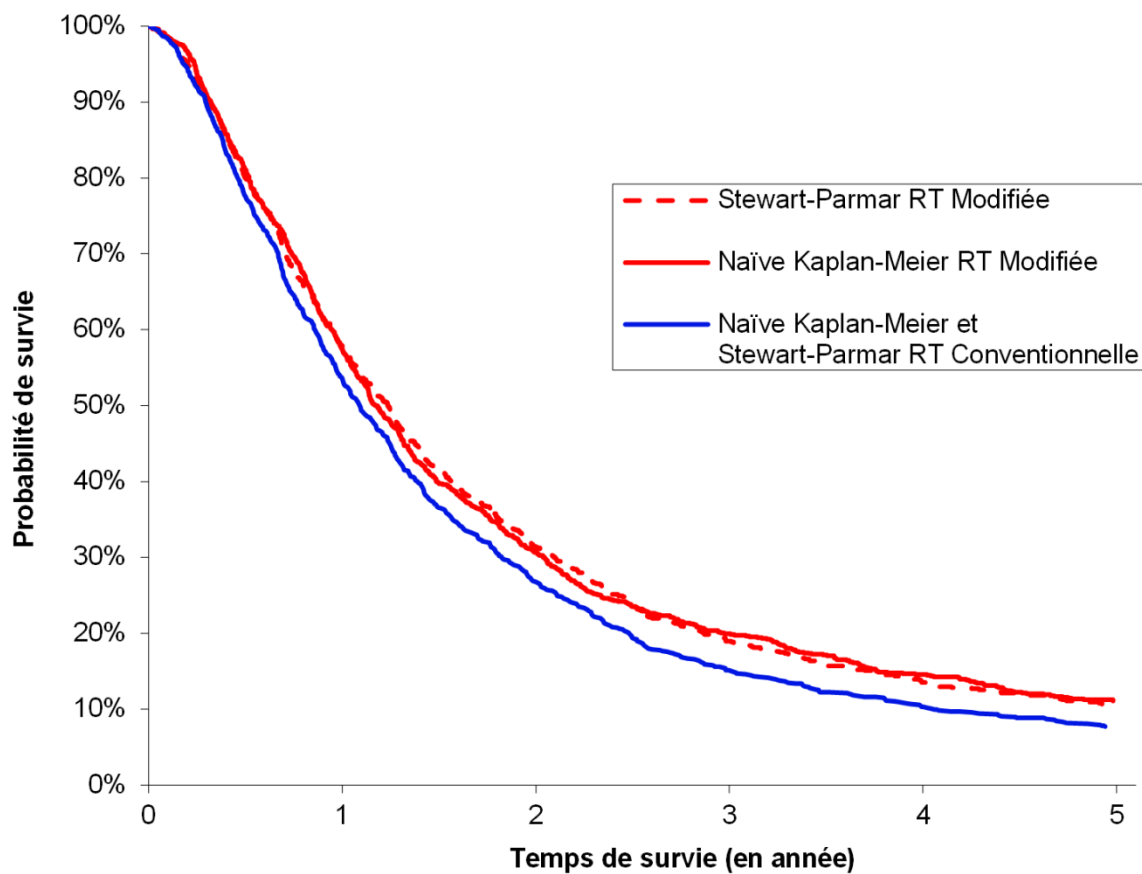
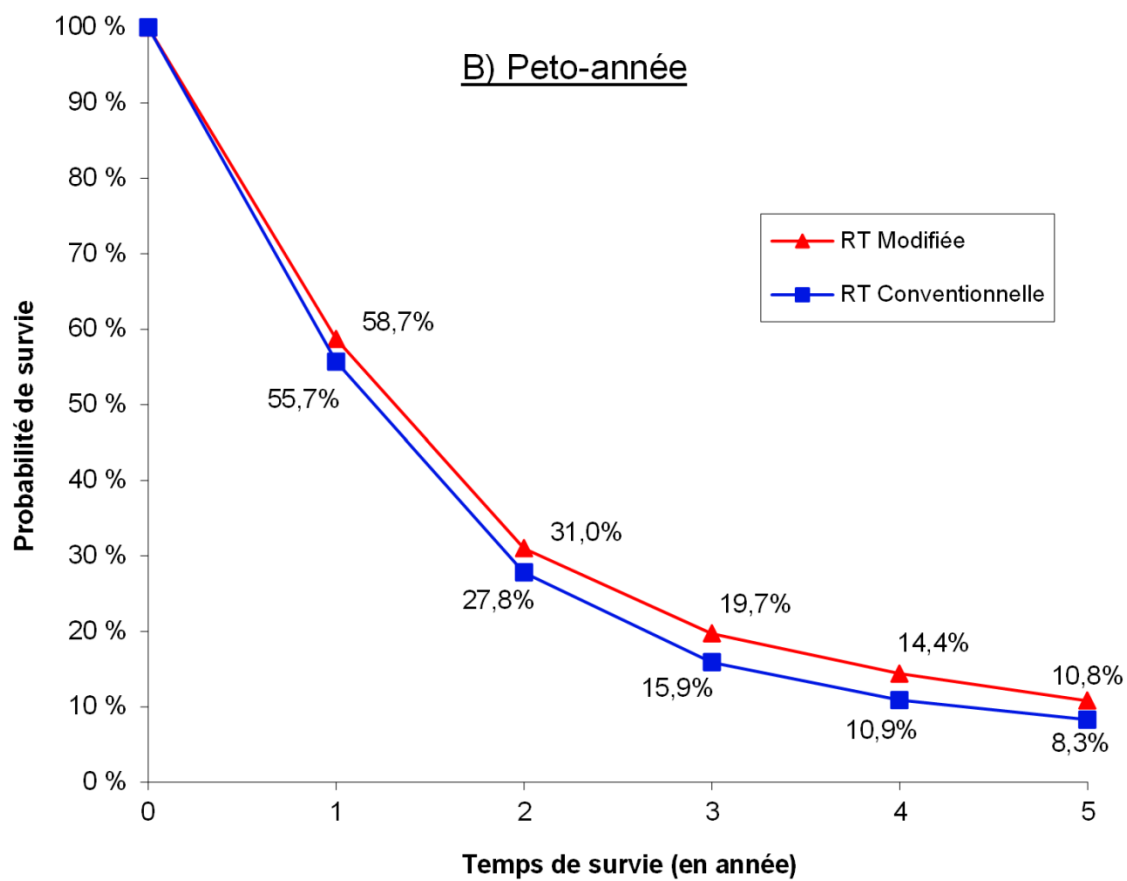
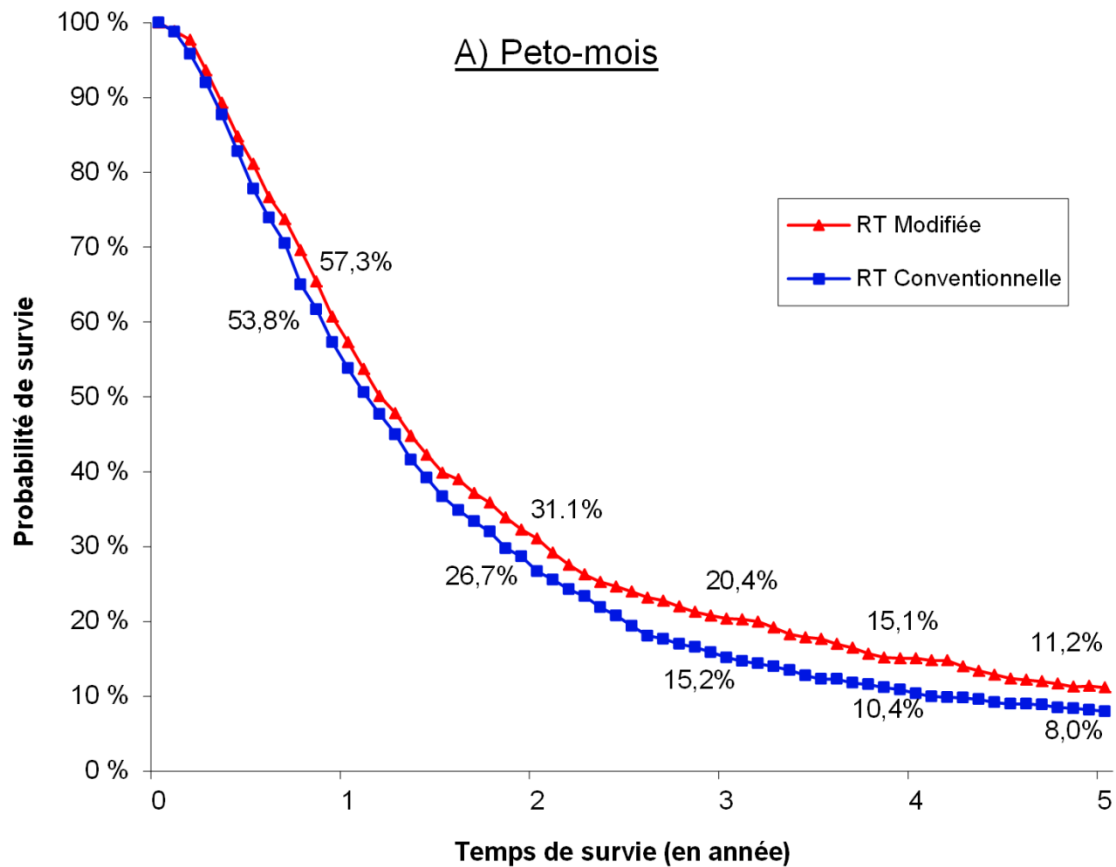


Figure 3.2 - Courbes de survie estimées à l'aide des méthodes Naïve Kaplan-Meier et Stewart-Parmar pour la méta-analyse MAR-LC.



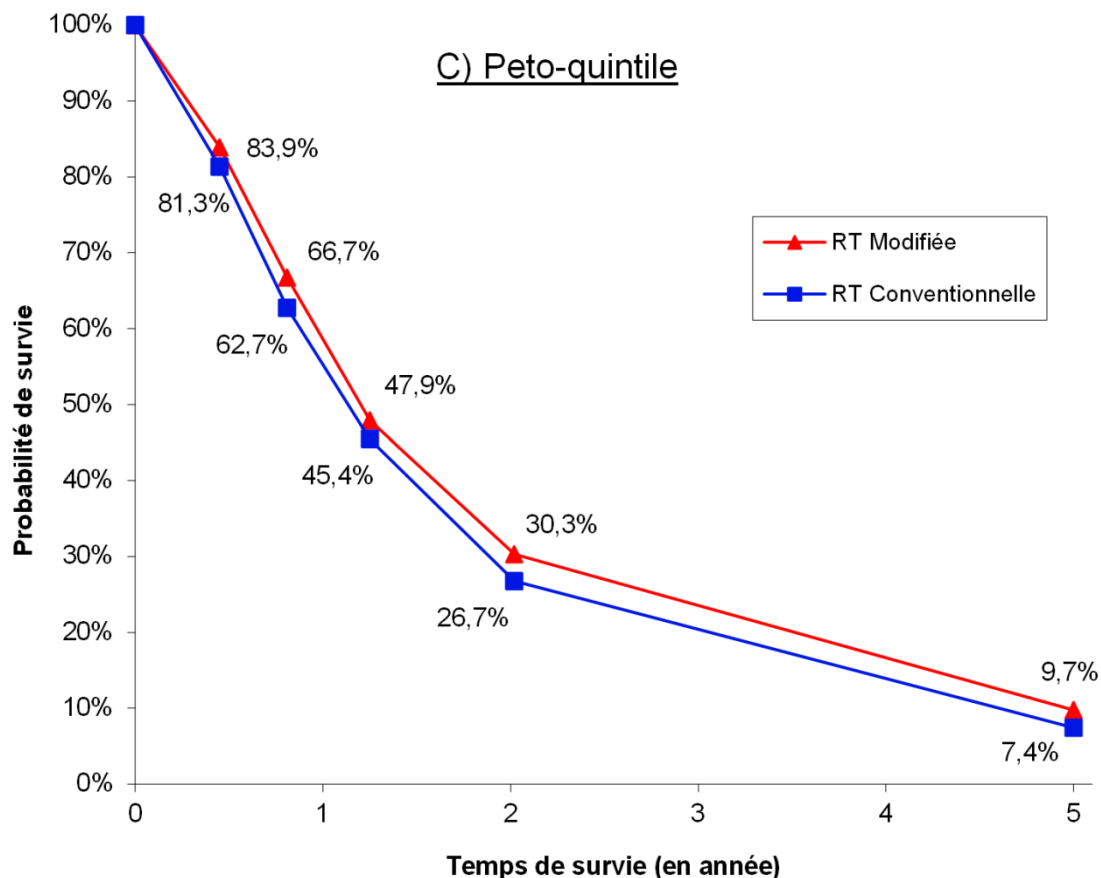


Figure 3.3 - Courbes de survie estimées à l'aide des méthodes Peto-mois (panel A), Peto-année (panel B) et Peto-quintile (panel C) pour la méta-analyse MAR-LC.

Les méthodes Naïve Kaplan-Meier et Stewart-Parmar estimaient la même courbe de survie, par définition (Equation (2.29)), pour le bras RT conventionnelle, et des courbes très similaires pour le bras RT modifiée (Figure 3.2). Les courbes de survie des méthodes Peto-mois, Peto-année et Peto-quintile différaient puisqu'elles ne reposaient pas sur les mêmes intervalles de temps (Figure 3.3).

La Figure 3.4 illustre le forest-plot des bénéfices de survie restreints estimés par Kaplan-Meier ou par le modèle exponentiel pour chacun des dix essais de MAR-LC, indiquant aucune hétérogénéité entre les essais (Cochran- $p=0,47$, I^2 d'Higgins =0% pour Pooled Kaplan-Meier et $p=0,31$, $I^2=15\%$ pour Pooled Exponentiel). Sur cette figure, chaque essai est représenté par un carré, dont le centre indique le bénéfice de survie restreint (rmstD) pour l'essai en question, et avec la ligne horizontale représentant l'IC à 95 %. La taille du carré est directement proportionnelle à la quantité d'information apportée par l'essai. Les losanges représentent les rmstD globaux, dont le centre indique la valeur du rmstD et les extrémités l'IC à 95%.

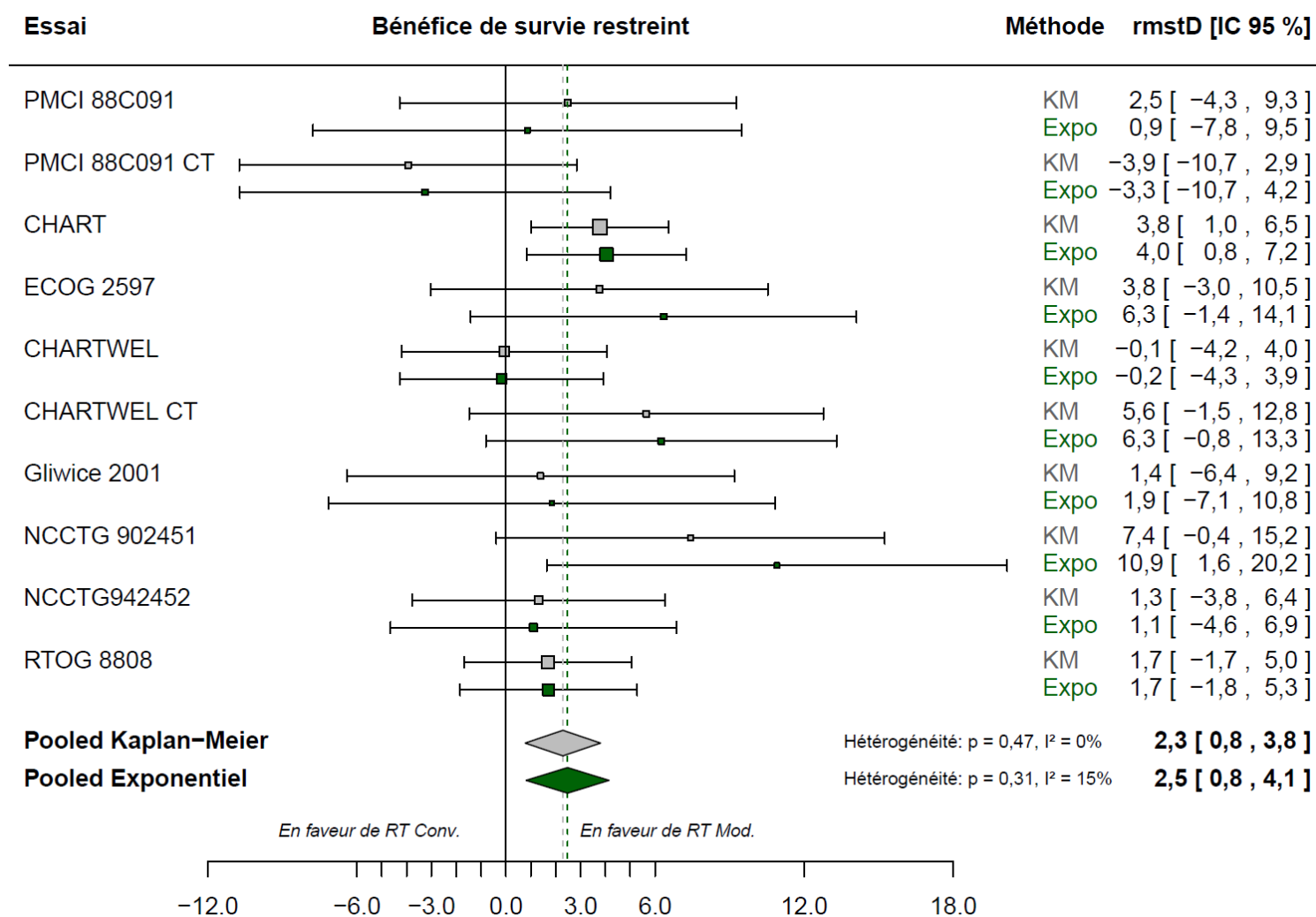


Figure 3.4 - Forest-plot des bénéfices de survie restreint exprimés en mois et estimés avec les méthodes Pooled Kaplan-Meier (carrés et losange gris) et Pooled Exponentiel (carrés et losange verts) pour la méta-analyse MAR-LC.

Légende de la Figure 3.4 :

Un modèle à effets fixes a été utilisé pour agréger les $rmstD(t^*)$ estimés dans chaque essai.

CHART: Continuous Hyperfractionated Accelerated Radiation Therapy; CHARTWEL: CHART Week-End Less; CT: Chimiothérapie; ECOG: Eastern Cooperative Oncology Group; Expo: Exponentiel; IC: Intervalle de confiance; KM: Kaplan-Meier; MAR-LC: Meta-Analysis of Radiotherapy in Lung Cancer; NCCTG: North Central Cancer Treatment Group; PCMI: Peter MacCallum Institute; rmstD: bénéfice de survie restreint; RTOG: Radiation Therapy Oncology Group; RT: Radiothérapie; t^* : horizon temporel.

Les survies moyennes restreintes ($RMST(t^* = 5)$) des deux bras de traitement par RT ainsi que les bénéfices de survie restreints moyens pour chacune des sept méthodes d'estimation sont présentés dans le Tableau 3.3. Les estimations moyennes de ces paramètres, leurs intervalles de confiance et l'erreur-type du $rmstD(t^* = 5)$ étaient calculés à l'aide des 1 000 échantillons du *bootstrap*.

Tableau 3.3 - Survies moyennes restreintes, bénéfice de survie restreint moyen et son erreur-type selon les sept méthodes d'estimation dans l'étude de cas MAR-LC

Méthode d'estimation	Survie moyenne restreinte RT Modifiée (en mois) [IC 95 %] [§]	Survie moyenne restreinte RT Conventionnelle (en mois) [IC 95 %] [§]	Bénéfice de survie restreint (en mois) [IC 95 %] [§]	Erreur-type du bénéfice de survie restreint [§]
Peto-quintile	22,0 [21,0 ; 23,1]	20,3 [19,3 ; 21,3]	1,7 [0,4 ; 3,1]	0,7
Peto-année	21,5 [20,5 ; 22,5]	19,7 [18,8 ; 20,7]	1,8 [0,5 ; 3,0]	0,6
Stewart-Parmar	21,1 [20,0 ; 22,2]	19,0 [17,9 ; 20,1]	2,1 [0,6 ; 3,5]	0,7
Naïve Kaplan-Meier	21,2 [20,1 ; 22,3]	19,0 [17,9 ; 20,1]	2,2 [0,6 ; 3,7]	0,8
Peto-mois	21,3 [20,1 ; 22,4]	19,0 [18,0 ; 20,0]	2,3 [0,7 ; 3,9]	0,8
Pooled Kaplan-Meier	20,7 [19,5 ; 21,8]	18,0 [16,7 ; 19,2]	2,3 [0,7 ; 3,8]	0,8
Pooled Exponentiel	21,6 [20,3 ; 22,9]	18,8 [17,6 ; 20,1]	2,5 [0,7 ; 4,2]	0,9

[§] IC 95% et erreur-type estimés à l'aide d'un *bootstrap* non-paramétrique avec 1 000 échantillons

Selon les méthodes, la $RMST(t^* = 5)$ variait de 20,7 mois à 22,0 mois pour la RT modifiée et de 18,0 à 20,3 mois pour la RT conventionnelle. Les $rmstD(t^* = 5)$ variaient de 1,7 mois pour la méthode Peto-quintile à 2,5 mois pour la méthode Pooled Exponentiel. Pour toutes les méthodes sauf Pooled Kaplan-Meier et Pooled Exponentiel, la différence des moyennes de la $RMST(t^* = 5)$ de la RT modifiée (1^{ère} colonne du Tableau 3.3) moins la $RMST(t^* = 5)$ de la RT conventionnelle (2^{ème} colonne) était égale à la moyenne des 1 000 estimations du $rmstD(t^* = 5)$ (3^{ème} colonne). La différence des moyennes ($RMST(t^* = 5)$) était donc bien égale à la moyenne des différences ($rmstD(t^* = 5)$). A l'inverse, ces deux estimations étaient différentes pour les

méthodes Pooled Kaplan-Meier et Pooled Exponentiel. Pour un type de RT donné, la $RMST(t^* = 5)$ était estimée comme la moyenne pondérée des $RMST_j(t^* = 5)$ calculées dans chaque essai de MAR-LC. De même pour le $rmstD(t^* = 5)$ estimé comme la moyenne pondérée des $rmstD_j(t^* = 5)$. Ainsi, la différence des moyennes pondérées n'était pas égale à la moyenne pondérée des différences, ce qui était probablement dû à la différence de poids accordée à chaque essai pour ces deux estimations.

Pour les méthodes Peto, le découpage en intervalles d'un mois impliquait un plus faible nombre d'évènements dans chaque intervalle de temps, ce qui expliquerait la plus grande erreur-type du $rmstD(t^* = 5)$ pour la méthode Peto-mois comparée à celles de Peto-quintile et Peto-année. Les méthodes basées sur l'estimateur de Kaplan-Meier (Naïve Kaplan-Meier, Stewart-Parmar, et Pooled Kaplan-Meier) et la méthode Pooled Exponentiel avaient également des erreurs-types plus grandes et donc des intervalles de confiance plus larges par rapport aux méthodes Peto-quintile et Peto-année.

Il est à noter que l'estimation du $rmstD(t^*)$ global provient directement de la méta-analyse MAR-LC avec une erreur-type estimée analytiquement pour la Figure 3.4 alors qu'elle provient de l'estimation moyenne des 1 000 échantillons du *bootstrap*, avec une erreur-type estimée comme l'écart type, pour le Tableau 3.3. Ceci explique les faibles différences observées entre la Figure 3.4 et le Tableau 3.3 à la fois pour la méthode Pooled Kaplan-Meier (2,3 [0,8 ; 3,8] contre 2,3 [0,7 ; 3,8]) et pour la méthode Pooled Exponentiel (2,5 [0,8 ; 4,1] contre 2,5 [0,7 ; 4,2]).

La légère différence entre la méthode Pooled Kaplan-Meier et la méthode Pooled Exponentiel dans les estimations du $rmstD(t^*)$ et de son erreur-type peut s'expliquer par le fait que la distribution exponentielle n'était pas parfaitement adaptée à la distribution des temps de survie des patients dans chacun des 10 essais de la méta-analyse MAR-LC. Bien que dans chaque essai la proportionnalité des risques ait été vérifiée et que les tests de rapport de vraisemblance guidaient vers le choix d'une distribution exponentielle plutôt que la distribution de Weibull, on remarque une différence dans l'estimation du $rmstD(t^*)$ pour la majorité des essais dans la Figure 3.4 entre les méthodes Kaplan-Meier et exponentielle.

Il est à noter que de faibles différences entre les $rmstD(t^*)$ entraînaient cependant des différences plus importantes entre les ratios coût-efficacité incrémentaux (ICER) moyens pour chaque méthode (Tableau 3.4). En effet, les ICERs moyens variaient de 24 299€, pour la méthode Pooled Exponentiel, à 34 934€, pour la méthode Peto-quintile, pour une année de vie gagnée.

Tableau 3.4 - ICER moyen selon les sept méthodes d'estimation du bénéfice de survie restreint dans l'étude de cas MAR-LC

Méthode d'estimation	ICER moyen (coût par année de vie gagnée) [IC 95 %] [§]
Peto-quintile	34 934€ [12 506 ; 98 066]
Peto-année	33 387€ [13 512 ; 75 753]
Stewart-Parmar	29 017€ [11 822 ; 68 206]
Naïve Kaplan-Meier	26 848€ [11 152 ; 68 297]
Peto-mois	28 022€ [10 563 ; 61 608]
Pooled Kaplan-Meier	25 527€ [10 355 ; 66 289]
Pooled Exponentiel	24 299€ [9 584 ; 59 119]

[§] IC 95% et erreur-type estimés à l'aide d'un *bootstrap* non-paramétrique avec 1 000 échantillons

Dans chacun des 1 000 échantillons du *bootstrap* la radiothérapie modifiée était à la fois plus efficace, quelle que soit la méthode d'estimation du bénéfice de survie restreint utilisée, et plus chère que la radiothérapie conventionnelle.

En plus de l'ICER moyen et de son intervalle de confiance, une des représentations usuelles permettant de juger l'incertitude entourant les résultats d'une analyse coût-efficacité est la courbe d'acceptabilité. Nous avons représenté ces courbes d'acceptabilité à l'aide des 1 000 ICERs estimés pour chacune des sept méthodes d'analyse de survie à partir des échantillons du *bootstrap*. Pour une méthode d'estimation du $rmstD(t^*)$ donnée, la RT modifiée est considérée

coût-efficace si l'ICER estimé est inférieur à la disposition à payer pour une année de vie gagnée. Ainsi les courbes d'acceptabilité de la Figure 3.5 représentent la proportion des 1 000 échantillons pour lesquels la RT modifiée est coût-efficace et ce pour différents seuils de disposition à payer.

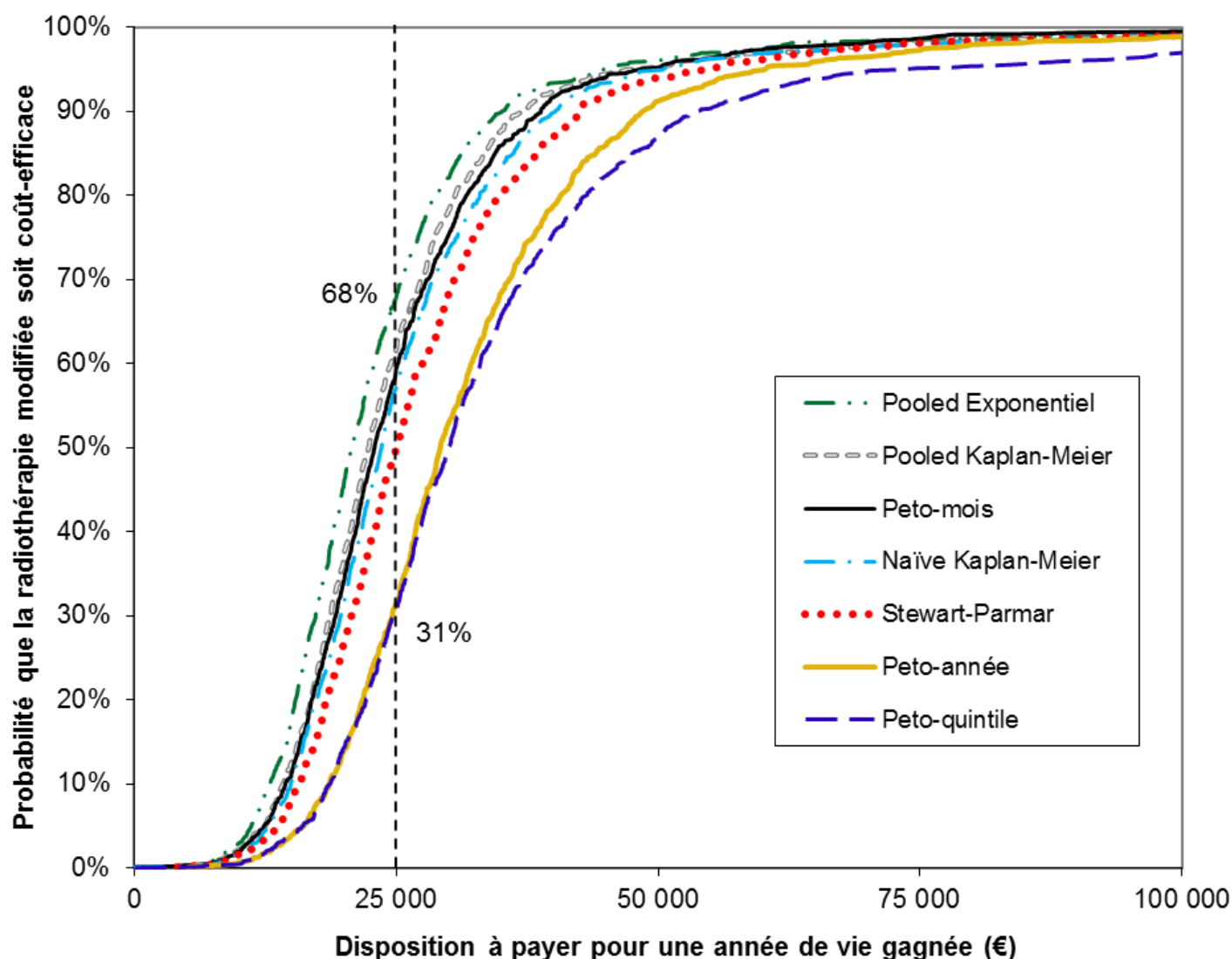


Figure 3.5 - Courbes d'acceptabilité montrant la probabilité que la radiothérapie modifiée soit coût-efficace pour différents seuils de disposition à payer pour une année de vie gagnée.

Note de lecture de la Figure 3.5:

Par exemple au seuil de 25 000€ par année de vie gagnée : la probabilité que la RT modifiée soit coût-efficace est de 68% lorsque le bénéfice de survie restreint était estimé par la méthode Pooled Exponentiel. En effet, 68% des 1 000 ICERs estimés avec Pooled Exponentiel dans les échantillons du *bootstrap* étaient inférieurs à 25 000€.

La courbe d'acceptabilité de la méthode Pooled Exponentiel était toujours au-dessus de celles des six autres méthodes, reflétant ainsi une probabilité plus grande que la RT modifiée soit coût-efficace lorsque le $rmstD(t^*)$ était estimé par la méthode Pooled Exponentiel (Figure 3.5). Les courbes d'acceptabilité des méthodes Pooled Kaplan-Meier, Peto-mois, Naïve Kaplan-Meier, et Stewart-Parmar étaient relativement proches, tandis que les courbes des méthodes Peto-année et Peto-quintile étaient particulièrement en-dessous de toutes les autres.

Le Tableau 3.5 complète les courbes d'acceptabilité de la Figure 3.5 et présente les probabilités que la radiothérapie modifiée soit coût-efficace pour trois seuils de disposition à payer pour une année de vie gagnée (25 000€, 50 000€ et 75 000€), et selon la méthode d'estimation du $rmstD(t^*)$.

Tableau 3.5 - Probabilité que la radiothérapie modifiée soit coût-efficace pour différents seuils de disposition à payer et selon la méthode d'estimation du bénéfice de survie restreint.

Méthode d'estimation	Seuil d'acceptabilité		
	25 000€	50 000€	75 000€
Peto-quintile	31 %	87 %	95 %
Peto-année	31 %	91 %	97 %
Stewart-Parmar	50 %	94 %	98 %
Naïve Kaplan-Meier	57 %	95 %	98 %
Peto-mois	59 %	95 %	99 %
Pooled Kaplan-Meier	62 %	95 %	99 %
Pooled Exponentiel	68 %	96 %	99 %

Avec un seuil de disposition à payer pour une année de vie gagnée supérieur à 50 000€, toutes les méthodes concluaient à ce que la RT modifiée était coût-efficace avec une probabilité d'environ 90 % voire plus (Figure 3.5 et Tableau 3.5). Cependant, les conclusions économiques seraient différentes en dessous de ce seuil. Avec un seuil de 25 000€ par année de vie gagnée, la

probabilité que la RT modifiée soit coût-efficace variait de 31 % avec les méthodes Peto-quintile et Peto-année à 68 % avec la méthode Pooled Exponentiel.

3.4. Discussion

A travers cette étude de cas, nous avons montré que différentes méthodes d'estimation du bénéfice de survie restreint à partir d'une méta-analyse sur données individuelles (MADI) peuvent conduire à des résultats économiques différents. Ces différences peuvent s'expliquer par deux facteurs.

En premier lieu, les méthodes d'estimation du $rmstD(t^*)$ ont différentes caractéristiques pour tenir compte de la spécificité de la structure hiérarchique d'une méta-analyse sur données individuelles (Tableau 2.1). La méthode Pooled Kaplan-Meier est la seule à tenir compte à la fois de la stratification sur l'essai, de l'hétérogénéité de l'effet traitement et de la non-proportionnalité des risques. Les méthodes Pooled Kaplan-Meier et Pooled Exponentiel permettent également d'étudier l'hétérogénéité du bénéfice de survie restreint entre les essais (Figure 3.4).

Une deuxième explication est liée à la définition de l'intervalle de temps considéré dans chacune des méthodes. Les probabilités de survie pour les méthodes Naïve Kaplan-Meier, Pooled Kaplan-Meier, et Stewart-Parmar, trois méthodes dérivées de Kaplan-Meier, sont estimées après chaque évènement. Pour la méthode Peto-mois, les probabilités de survie sont estimées chaque mois qui est un découpage de temps proche d'une estimation à chaque évènement. De même, pour la méthode Pooled Exponentiel, tous les temps de décès sont utilisés pour estimer le paramètre d'échelle de la distribution exponentielle. Ces cinq méthodes conduisaient aux courbes d'acceptabilité les plus optimistes vis-à-vis de la probabilité de la radiothérapie modifiée d'être coût-efficace (Figure 3.5). A l'inverse, les méthodes Peto-année et Peto-mois conduisaient à des résultats différents dû au fait que les intervalles de temps considérés étaient plus larges. Cela induisait une moins grande variabilité dans l'estimation du bénéfice de survie pour ces deux méthodes, possiblement au prix d'une estimation biaisée du $rmstD(t^*)$: 1,7 et 1,8 mois respectivement pour Peto-quintile et Peto-année contre 2,1 à 2,5 mois pour les autres méthodes (Tableau 3.3).

Les résultats de notre étude coût-efficacité utilisant un modèle dit d'aire sous la courbe pouvaient être comparés avec ceux de l'étude coût-efficacité dans le contexte hollandais fondée sur un modèle de Markov (Ramackers *et al.*, 2013). Bien que ces deux types d'analyse aient

certaines différences structurelles (Minacori *et al.*, 2015; Rochau *et al.*, 2015), Goere *et al.* ont montré dans une étude de cas que les résultats des deux analyses étaient similaires lorsque les sources de données étaient les mêmes (Goeree *et al.*, 2016). L'analyse coût-efficacité hollandaise ayant également utilisé MAR-LC comme source principale de données, il semblait donc intéressant de comparer les résultats en termes de coût, d'efficacité clinique et d'efficience. Une différence notable entre nos deux études était que l'étude hollandaise avait étudié séparément chacun des 4 sous-groupes de la RT modifiée définis dans le Tableau 3.1 et la Figure 3.1, alors que nous les avons considérés comme une seule stratégie de traitement dans notre étude de cas. Cependant, les estimations étaient assez proches entre les deux études en termes de coût, d'efficacité clinique et de coût-efficacité bien que ces deux études fussent réalisées dans des contextes français et hollandais.

Le coût total par patient était de 25k€ pour la RT conventionnelle et de 30k€ pour la RT modifiée dans notre étude de cas, contre 24k€ pour la RT conventionnelle et entre 26k€ (RT très accélérée) et 30k€ (RT hyper-fractionnée avec dose total supérieure) pour la RT modifiée dans le modèle de Markov hollandais (Ramaekers *et al.*, 2013). La survie moyenne était également du même ordre de grandeur entre les des deux études. Elle variait entre 18 et 20 mois pour la RT conventionnelle et entre 21 et 22 mois pour la RT modifiée dans notre étude de cas (Tableau 3.3) contre 20 mois pour la RT conventionnelle et environ 22-23 mois pour la RT modifiée dans l'étude de Ramaekers *et al.* Enfin, à un seuil de disponibilité à payer supérieur à 50k€, la probabilité que la RT conventionnelle soit coût-efficace était inférieure à 5% dans les deux études.

Notre étude de cas a montré que le choix d'une méthode d'estimation du bénéfice de survie restreint à partir d'une MADI avait un impact sur les résultats d'une étude coût-efficacité. Il apparaissait que la méthode Pooled Kaplan-Meier présentait de nombreux avantages. Tout d'abord, elle permet de tenir compte de la stratification sur l'essai, de l'hétérogénéité de l'effet traitement et de risques non proportionnels. De plus, cette méthode ne nécessite pas de prédéfinir des intervalles de temps contrairement à la méthode actuarielle de Peto. Enfin, la méthode Pooled Kaplan-Meier permet d'étudier la potentielle hétérogénéité du $rmstD(t^*)$ entre les essais et avait été prouvée comme étant non-biaisée et avec une bonne probabilité de couverture des intervalles de confiance dans une précédente étude de simulation (Wei *et al.*, 2015).

L'une des limites inhérente aux études de cas était que les résultats de cette étude dépendaient des caractéristiques des données cliniques utilisées. Toutefois, même sans

hétérogénéité de l'effet traitement entre les essais et avec des risques proportionnels dans la méta-analyse MAR-LC, nous avons noté une différence dans les ICERs moyens estimés par les méthodes étudiées. Cette différence aurait pu être encore plus importante dans le cas d'une hétérogénéité de l'effet traitement ou de risques non proportionnels. La suite logique de ce travail était donc d'étudier les propriétés des méthodes d'estimation du $rmstD(t^*)$ à travers différents scénarios (introduction d'hétérogénéité de l'effet traitement, risques non proportionnels) dans une étude de simulation.

Chapitre 4. Comparaison de méthodes d'estimation du bénéfice de survie : une étude de simulation

4.1. Motivation

L'étude de cas à partir de la méta-analyse sur données individuelles MAR-LC précédemment présentée nous a permis de constater que les différentes méthodes étudiées pouvaient conduire à des estimations différentes du bénéfice de survie restreint et de son intervalle de confiance, et potentiellement à des conclusions différentes pour l'évaluation économique selon le seuil de disposition à payer considéré. Afin de mieux comprendre les propriétés des méthodes étudiées nous avons réalisé une étude de simulation : c'est l'objet de ce Chapitre 4.

Pour estimer le bénéfice de survie restreint ($rmstD(t^*)$) à partir d'une méta-analyse sur données individuelles, Wei et al. avaient déjà réalisé une étude de simulation pour comparer différentes méthodes d'estimation. Dans leur étude, le $rmstD(t^*)$ était estimé comme l'agrégation, à l'aide d'un modèle à effets fixes, des $rmstD_j(t^*)$ estimés dans chacun des essais. Pour estimer le bénéfice de survie restreint $rmstD_j(t^*)$ dans chaque essai, Wei et al. avaient comparé trois méthodes : deux méthodes non-paramétriques – l'une fondée sur les « pseudo-valeurs » (Andersen & Pohar Perme, 2010) et l'autre fondée sur l'estimateur de Kaplan-Meier – et une méthode de survie paramétrique flexible basée sur des splines (Royston & Parmar, 2002; méthode définie en Annexes). Il est à noter que la méthode fondée sur l'estimateur de Kaplan-Meier de Wei et al. correspond à la méthode que nous désignons par Pooled Kaplan-Meier avec effets fixes présentée dans le Chapitre 2. Les auteurs avaient simulé des données de survie censurées de méta-analyse à l'aide d'une distribution Weibull et avaient fait varier les paramètres de la distribution Weibull, le taux de censure, et la taille de l'effet traitement ($HR=0,7$ et $HR=1,2$). Deux tailles d'échantillon ($N=250$ et $N=500$) avaient également été simulées, mais il n'était toutefois pas spécifié combien d'essais ni combien de patients par essai étaient simulés. A travers ces simulations et deux études de cas en méta-analyse, Wei et al. ont montré que les trois

méthodes conduisaient à des résultats similaires en terme de biais moyen, d'erreur quadratique moyenne et de probabilité de couverture des intervalles de confiance.

Nous nous sommes appuyés sur les résultats de cet article qui a été publié en cours de réalisation de notre étude de simulation. Ainsi, nous avons étendu les résultats de l'étude de Wei et al. en comparant de nouvelles méthodes d'estimation et à travers un plus grand nombre de scénarios. De plus, notre étude de simulation s'est fondée sur un cas de figure plus réaliste en méta-analyse avec l'utilisation d'effets aléatoires pour introduire de l'hétérogénéité sur le risque de base et sur l'effet traitement entre les essais.

Parmi les sept méthodes présentées dans les chapitres précédents, nous avons décidé de comparer quatre méthodes dans l'étude de simulation : les méthodes non-paramétriques Naïve Kaplan-Meier, Peto-quintile, et Pooled Kaplan-Meier, et la méthode paramétrique Pooled Exponentiel. Nous avons choisi de restreindre le nombre de méthodes pour plus de clarté d'interprétation de l'étude de simulation. Parmi les trois intervalles de temps considérés pour la méthode Peto (mois, année et quintile) dans l'étude de cas, nous avons décidé de conserver l'intervalle de temps défini sur les quintiles des temps de décès (Peto-quintile) qui nous semblait plus intéressant car fondé sur la distribution des temps de décès. Nous n'avons pas retenu la méthode Stewart-Parmar, pour laquelle nous appliquons le *hazard ratio* global de la méta-analyse aux patients du bras contrôle pour estimer la survie du bras expérimental. Cette méthode ne prenait donc pas en compte l'hétérogénéité du risque de base entre les essais, et reposait sur l'hypothèse forte des risques proportionnels. Enfin, parmi les méthodes étudiées par Wei et al, nous n'avons pas inclus d'autres méthodes que la méthode Pooled Kaplan-Meier, puisque les résultats entre les trois méthodes de leur étude de simulation étaient similaires (Wei *et al.*, 2015).

Dans nos simulations, nous avons fait varier la taille de l'effet traitement global et l'horizon temporel t^* , comme étudiés précédemment par Wei et al. Nous avons également étudié l'hétérogénéité inter-essai du risque de base et de l'effet traitement, la corrélation entre les deux effets aléatoires, le nombre d'essais et le nombre de patients par essai, l'utilisation de modèles à effets fixes ou à effets aléatoires de DerSimonian-Laird (DerSimonian & Laird, 1986), et la proportionnalité des risques.

4.2. Etude de simulation

4.2.1. Plan expérimental de l'étude de simulation

Nous avons simulé des données individuelles de méta-analyse à partir d'un modèle de survie avec des effets aléatoires au niveau essai sur le risque de base et sur l'effet traitement. Ainsi, nous avons simulé des temps de survie censurés pour N patients inclus dans J essais, chacun de taille n_j avec $\sum_{j=1}^J n_j = N$. La fonction de risque pour un essai $j \in \{1, \dots, J\}$ était ainsi définie par :

$$\lambda(t; x | A_j = a_j, B_j = b_j) = \lambda_0(t) \exp\{a_j + (\beta + b_j)x\} \quad (4.1)$$

où

- $\lambda_0(t)$ est la fonction de risque de base ;
- A_j est une quantité aléatoire qui affecte le risque de base au niveau essai avec $\text{Var}(A_j) = \sigma^2$;
- B_j est une quantité aléatoire qui affecte l'effet traitement au niveau essai avec $\text{Var}(B_j) = \tau^2$;
- β est l'effet traitement global ;
- x est la variable de traitement binaire, codée -1/2 pour le bras contrôle et +1/2 pour le bras expérimental afin d'obtenir une hétérogénéité égale dans les deux bras (Tudur Smith & Williamson, 2007; Rondeau *et al.*, 2008; Bowden *et al.*, 2011).

Il est à noter que la covariance entre les deux effets aléatoires est définie par $\text{cov}(A_j, B_j) = \rho\sigma\tau$ avec ρ la corrélation entre A_j et B_j .

Nous avons introduit de l'hétérogénéité sur le risque de base et sur l'effet traitement entre les essais à l'aide des deux variables aléatoires A_j et B_j suivant une distribution binomiale. Le fait d'utiliser une distribution discrète avec la loi binomiale nous permettait de dériver de manière directe le vrai bénéfice de survie sous-jacent à chacun de nos scénarios de simulation. Les variables aléatoires binomiales pouvaient prendre chacune 51 valeurs, et étaient centrées et réduites afin d'obtenir les variances σ^2 et τ^2 souhaitées :

$$\begin{aligned} A &\sim [\text{Bin}(n = 50, p = 0,5) - 25] \cdot \sigma / \sqrt{12,5} \\ B &\sim [\text{Bin}(n = 50, p = 0,5) - 25] \cdot \tau / \sqrt{12,5} \end{aligned} \quad (4.2)$$

Le rationnel du choix de $n = 50$ pour les deux variables binomiales était que cette distribution permettait de s'approcher d'une distribution continue comme la loi normale par exemple, tout en permettant un calcul immédiat du vrai bénéfice de survie restreint $rmstD(t^*)$.

Pour la fonction de risque de base, nous avons utilisé une distribution exponentielle avec $\lambda_0(t) = \log(2)/5$, ce qui correspond à fixer un temps de survie médian de 5 ans. La courbe de survie du bras expérimental de la Figure 2.1 permet d'illustrer le risque de base utilisé dans ces simulations. L'essai 2 présenté dans les Figure 4.1 et Figure 4.2 permet également d'illustrer le risque de base global simulé ; pour ces figures-là la médiane de survie n'est égale à 5 ans pour aucun des bras, puisque l'effet traitement codé $\pm 1/2$ et égal à $(\beta + b_j)$ augmente ou diminue le risque de décès. Le risque de base simulé est commun aux deux bras : la courbe de survie virtuelle du risque de base global correspondant aux deux bras confondus se trouverait entre les deux courbes de survie.

La Figure 4.1 illustre trois essais de 200 patients chacun simulés avec les paramètres de l'étude de simulation, avec le même effet traitement ($\beta = -0,7$ correspondant à un HR = 0,5), mais avec une hétérogénéité du risque de base entre les essais : un risque de base diminué dans l'essai 1 ($a_j = -1$), un risque de base égal au risque de base global dans l'essai 2 ($a_j = 0$), et un risque de base augmenté dans l'essai 3 ($a_j = 1$). Plus le risque de base augmente, plus les courbes de survie des deux bras sont tirées vers le bas. On remarque que bien que l'effet traitement relatif soit égal pour les 3 essais fictifs simulés (HR = 0,5), le bénéfice de survie restreint à 5 ans, estimé comme l'aire entre les deux courbes de survie, augmente lorsque le risque de base augmente également : 5 mois pour l'essai 1, 9 mois pour l'essai 2, et 12 mois pour l'essai 3.

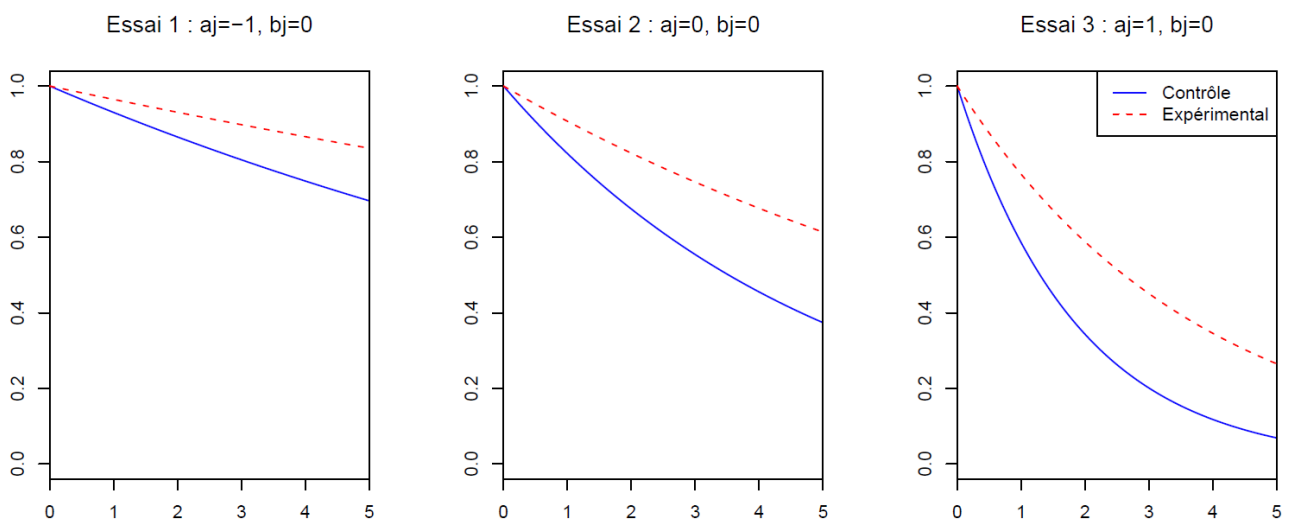


Figure 4.1 - Illustration de l'hétérogénéité du risque de base entre trois essais simulés, à l'aide de l'effet aléatoire A_j pour un effet traitement fixé (HR=0,5)

La Figure 4.2 illustre trois essais de 200 patients chacun simulés avec les paramètres de l'étude de simulation, avec le même risque de base ($a_j = 0$), mais avec une hétérogénéité de l'effet traitement entre les essais : un effet traitement augmenté dans l'essai 1 ($b_j = -1$, $HR = \exp(-1,7) = 0,2$), un effet traitement égal à l'effet traitement global dans l'essai 2 ($b_j = 0$, $HR = \exp(-0,7) = 0,5$), et un effet traitement diminué dans l'essai 3 ($b_j = 1$, $HR = \exp(0,3) = 1,3$). Plus l'effet traitement diminue, plus l'écart entre les deux courbes de survie diminue, voire la courbe de survie du bras contrôle passe au-dessus de celle du bras expérimental dans le cas d'un effet traitement délétère dans l'essai 3 avec un HR égal à 1,3. Comme expliqué dans les paragraphes précédents, le risque de base est commun aux deux bras. Ainsi, bien que le risque de base ne soit pas hétérogène entre ces trois essais ($a_j = 0$), la courbe de survie du bras contrôle est différente pour ces trois essais, car modifiée par l'effet traitement codé $+1/2$ pour ce bras.

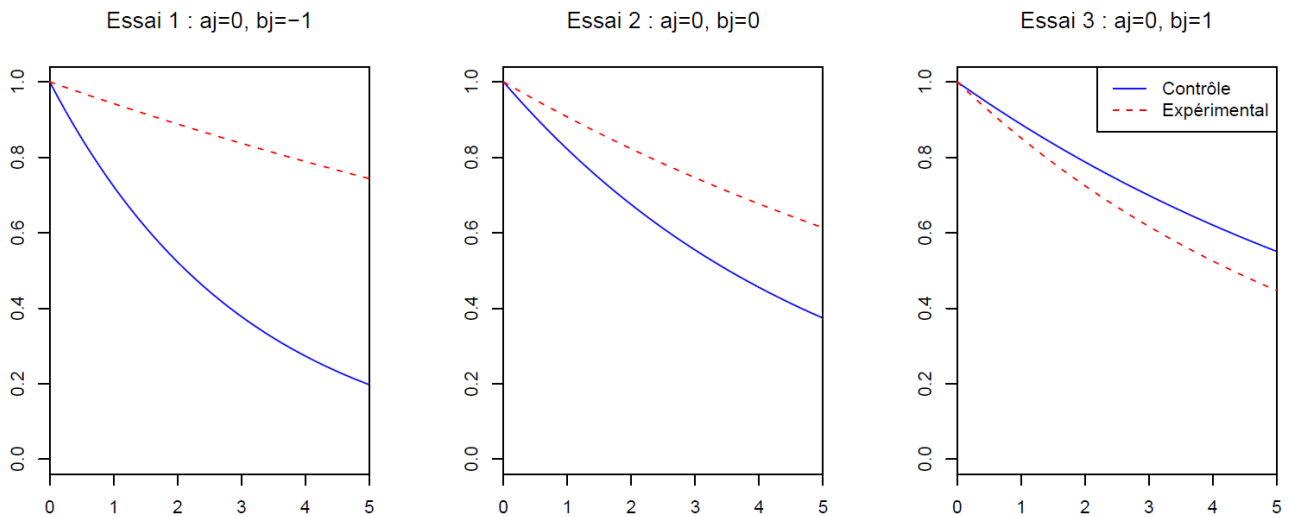


Figure 4.2 - Illustration de l'hétérogénéité de l'effet traitement entre trois essais simulés, à l'aide de l'effet aléatoire B_j pour un risque de base fixé

Nous avons également introduit de la censure à droite, indépendante et non-informative, en fixant le même temps de recrutement égal à 3 ans pour tous les essais et en variant le temps de suivi maximum entre les essais. Le temps de suivi maximum était choisi uniformément entre 2 et 9 ans afin de répliquer la différence des temps de suivi observés entre les essais inclus dans une méta-analyse sur données individuelles.

4.2.2. Vrai bénéfice de survie restreint

A partir de notre modèle de simulation de l'équation (4.1), le bénéfice de survie restreint est défini par :

$$\begin{aligned} rmstD(t^*) &= \int_0^{t^*} S\left(t; x = \frac{1}{2}\right) dt - \int_0^{t^*} S\left(t; x = -\frac{1}{2}\right) dt \\ &= \int_0^{t^*} \left[\int_{\mathcal{K}} S\left(t; x = \frac{1}{2} | a_j, b_j\right) dF_{A,B}(a, b) \right] dt - \int_0^{t^*} \left[\int_{\mathcal{K}} S\left(t; x = -\frac{1}{2} | a_j, b_j\right) dF_{A,B}(a, b) \right] dt \end{aligned} \quad (4.3)$$

où $\mathcal{K} = \{(a_k, b_k)\}_{k=1, \dots, K}$ est le support de la variable bivariable (A, B) . La distribution jointe $F_{A,B}(a, b)$ est définie par les probabilités $p_k = \mathbb{P}(A = a_k, B = b_k)_{k=1, \dots, K}$ de tous les $K = 51^2$ couples (a_k, b_k) potentiels. La distribution de probabilité jointe $F_{A,B}(a, b)$ ayant été définie discrète, l'intégrale sur $\mathcal{K}$ de l'équation (4.3) se transforme en une somme sur les K points du support $\mathcal{K}$. Ainsi, le vrai bénéfice de survie restreint $rmstD(t^*)$ correspond à une somme des K $rmstD_k(t^*)$ pondérée par les probabilités p_k :

$$rmstD(t^*) = \sum_{k=1}^K p_k rmstD_k(t^*), \quad (4.4)$$

avec le bénéfice de survie restreint conditionnel $rmstD_k(t^*)$ défini pour un couple (a_k, b_k) par :

$$\begin{aligned} rmstD_k(t^*) &= \int_0^{t^*} \exp\left(-\frac{\log(2)}{5} \cdot t \cdot e^{a_k + 1/2(\beta + b_k)}\right) dt \\ &\quad - \int_0^{t^*} \exp\left(-\frac{\log(2)}{5} \cdot t \cdot e^{a_k - 1/2(\beta + b_k)}\right) dt \end{aligned} \quad (4.5)$$

4.2.3. Scénarios des simulations

A travers différents scénarios, nous avons fait varier la force de l'hétérogénéité inter-essai du risque de base (faible avec $\sigma^2 = 0,01$ et forte avec $\sigma^2 = 0,10$) et de l'effet traitement ($\tau^2 = 0,01$; $0,10$). Les valeurs de ces deux hétérogénéités étaient fondées sur des études de simulation en méta-analyse précédemment publiées (Tudur Smith & Williamson, 2007; Rondeau *et al.*, 2008; Bowden *et al.*, 2011). Notre analyse principale était conduite avec des effets aléatoires A_j et B_j non corrélés ($\rho = 0$), cependant, nous avons étudié l'impact d'une corrélation négative ($\rho = -0,8$) dans

une étude de sensibilité. Une corrélation négative entre le risque de base et l'effet traitement dans notre étude de simulation signifiait schématiquement que le traitement expérimental serait plus efficace dans les essais avec un plus mauvais pronostic vital de base. Une telle corrélation négative avait été observée et étudiée par Rondeau et al. dans leur étude de simulation (Rondeau *et al.*, 2008). Nous avons considéré différentes valeurs pour le nombre d'essais J et le nombre de patients par essai : deux scénarios avec des tailles d'essai fixées ($J = 5$, $n_j = 200$; $J = 20$, $n_j = 100$), et un scénario avec un nombre d'essais ($J = 5$) et un nombre total de patients ($N = 1\,000$) fixés mais avec un nombre de patients par essai tiré aléatoirement à l'aide d'une distribution de Poisson. Avec les paramètres de la distribution de Poisson que nous avons choisis, 95% des essais avaient un nombre de patients $n_j \in [57 ; 376]$. Nous avons également étudié différentes valeurs pour la taille de l'effet traitement global ($\beta = 0$; $\pm 0,2$; $\pm 0,7$). Nous avons envisagé quatre valeurs pour l'effet traitement global lors de nos premières simulations ($\beta = 0$; $-0,2$; $-0,5$; $-0,7$ correspondant respectivement à des valeurs de HR = 1 ; 0,8 ; 0,6 ; et 0,5) mais les premiers résultats montraient une trop grande similitude avec $\beta = -0,5$ et $\beta = -0,7$: nous avons donc préféré ne garder que trois valeurs reflétant une absence d'effet traitement ($\beta = 0$), un effet traitement modéré (traitement expérimental bénéfique $\beta = -0,2$ ou délétère $\beta = +0,2$) et un fort effet traitement ($\beta = \pm 0,7$).

Nous avons également étudié l'impact du choix de l'horizon temporel de restriction avec $t^* = 5$ ans (le temps de survie médian) et $t^* = 10$ ans. Nous avons choisi ces deux valeurs afin d'illustrer un scénario dans lequel tous les essais présenteraient des patients encore à risque de décès à t^* (5 ans) et un scénario dans lequel certains essais auraient un suivi plus court que le temps de restriction ($t^* = 10$ ans). Le taux moyen de censures administratives variait entre les scénarios de 49 % à 52 % à $t^* = 5$ ans et de 38 % à 40 % à $t^* = 10$ ans. Dans le cas d'absence d'effet traitement global ($\beta = 0$, $\tau^2 = 0$) et sans hétérogénéité du risque de base ($\sigma^2 = 0$), la survie moyenne restreinte dans les deux bras était égale à 3,6 ans à $t^* = 5$ ans et à 5,4 ans à $t^* = 10$ ans.

Enfin, l'influence de la non-proportionnalité des risques a été étudiée à l'aide d'une distribution exponentielle par morceaux (*piecewise*) en simulant un effet du traitement expérimental délétère ($\beta' = -\beta$, avec $\beta \leq 0$) les 2 premières années et bénéfique (β) ensuite.

4.2.4. Critères d'évaluation

Nous avons simulé 1 000 méta-analyses pour chacun des scénarios envisagés. Les trois critères de comparaison de nos quatre méthodes d'estimation du bénéfice de survie restreint ($rmstD(t^*)$) étaient :

- le biais moyen, défini comme la moyenne des 1 000 $\widehat{rmstD}(t^*)$ estimés dans chacune des méta-analyses simulées moins la vraie valeur du $rmstD(t^*)$ pour le scénario considéré ;
- l'erreur-type empirique (ETE), définie comme l'écart-type des 1 000 $\widehat{rmstD}(t^*)$ estimés (*empirical standard error* en anglais) ;
- l'erreur-type moyenne (ETM), définie comme la moyenne des 1 000 erreurs-types $\widehat{ET}(rmstD(t^*))$ estimées dans chacune des méta-analyses simulées (*average standard error* en anglais).

Un biais moyen proche de 0 indiquerait que le bénéfice de survie restreint global estimé ($\widehat{rmstD}(t^*)$) est proche de la vraie valeur du $rmstD(t^*)$ simulée, et donc que la méthode étudiée estime de manière non biaisée le bénéfice de survie restreint global ; on parle alors d'une bonne exactitude (*accuracy* en anglais) de la méthode d'estimation.

Egalement, plus l'erreur-type empirique (ETE) est faible, plus la précision – faible variabilité de notre estimation (*precision* en anglais) – de notre méthode d'estimation est élevée. Enfin, plus l'erreur-type moyenne (ETM) est proche de l'erreur-type empirique (ETE), plus l'estimateur de l'erreur-type du bénéfice de survie restreint ($\widehat{ET}(rmstD(t^*))$) est non biaisé. Ainsi, une méthode d'estimation avec $ETM < ETE$ sous-estimerait la véritable variabilité dans l'estimation du bénéfice de survie restreint et l'intervalle de confiance pour le $\widehat{rmstD}(t^*)$ serait alors trop étroit et donc trop optimiste.

4.2.5. Résultats de l'étude de simulation

Les tableaux suivants (Tableau 4.1 - Tableau 4.6) présentent la vraie valeur du bénéfice de survie restreint simulé, ainsi que le biais moyen et les erreurs-types empirique et moyenne des quatre méthodes d'estimation, pour les différents scénarios étudiés dans nos simulations. Pour chacun des tableaux, les quatre scénarios relatifs à l'hétérogénéité de l'effet traitement (τ^2) et du risque de base (σ^2) sont représentés en ligne : les scénarios avec une forte hétérogénéité de l'effet traitement entre les essais ($\tau^2 = 0,10$) correspondent aux 2^{ème} et 4^{ème} blocs horizontaux, et les

scénarios avec une forte hétérogénéité du risque de base entre les essais ($\sigma^2 = 0,10$) correspondent aux 3^{ème} et 4^{ème} blocs horizontaux. En colonne sont représentés les résultats pour un effet traitement simulé nul ($\beta = 0$), modéré ($\beta = -0,2$) et fort ($\beta = -0,7$). Les résultats correspondant à un horizon temporel $t^* = 5$ ans et $t^* = 10$ ans sont rapportés pour chacun de ces effets traitement.

Le Tableau 4.1 correspond à des méta-analyses simulées de 5 essais et 200 patients par essai avec proportionnalité des risques, sans corrélation entre les deux effets aléatoires, et avec l'utilisation d'un modèle à effets aléatoires pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-Quintile. Les autres tableaux permettent d'étudier les différents scénarios envisagés : la non-proportionnalité des risques (Tableau 4.2), des méta-analyses de 20 essais et 50 patients par essai (Tableau 4.3), des méta-analyses de 5 essais et un nombre de patients par essai aléatoire (Tableau A.2), un effet traitement délétère (Tableau 4.4), une corrélation négative entre les deux effets aléatoires (Tableau 4.5), et l'utilisation d'un modèle à effets fixes pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-Quintile (Tableau 4.6).

Les termes soulignés dans le texte et les titres des tableaux correspondent aux paramètres qui varient par rapport au scénario « standard » présenté dans le Tableau 4.1. Les chiffres grisés dans ces tableaux correspondent à des résultats notables et qui sont détaillés dans les paragraphes respectifs.

Les Figure 4.3 et Figure 4.4 permettent d'illustrer graphiquement le biais moyen, et les erreurs-types moyenne et empirique des méthodes d'estimation, dans le cas d'un scénario avec des méta-analyses simulées de 5 essais et 200 patients par essai, et un fort effet traitement bénéfique ($\beta = -0,7$). La Figure 4.3 (respectivement Figure 4.4) illustre les résultats avec des risques simulés proportionnels (respectivement des risques non-proportionnels) et correspond aux chiffres donnés dans les deux derniers blocs de colonnes du Tableau 4.1 (respectivement du Tableau 4.2). Sur ces deux figures, les carrés noirs représentent le biais moyen du $rmstD(t^*)$ estimé par chaque méthode pour un scénario particulier. La droite rouge horizontale correspond à l'absence de biais dans l'estimation du $rmstD(t^*)$: plus le carré noir est proche de cette droite rouge, plus la méthode étudiée estime de façon non-biaisée le bénéfice de survie restreint. Les barres verticales noires et violettes représentent les intervalles de confiance à 95% du biais estimés respectivement avec l'erreur-type empirique et moyenne : deux barres de même longueur signifient que l'erreur-type empirique est égale à l'erreur-type moyenne et donc que l'erreur-type du bénéfice de survie restreint ($\widehat{ET}(rmstD(t^*))$) est non biaisée pour la méthode en question. Les scénarios avec une forte hétérogénéité de l'effet traitement entre les essais ($\tau^2 = 0,10$)

correspondent aux 2^{ème} et 4^{ème} colonnes. Les scénarios avec une forte hétérogénéité du risque de base entre les essais ($\sigma^2 = 0,10$) correspondent aux 3^{ème} et 4^{ème} colonnes.

Globalement, pour tous les scénarios étudiés, les méthodes ne présentaient quasiment aucun biais en cas d'absence d'effet traitement ($\beta = 0$). Lorsque l'effet traitement simulé était bénéfique ($\beta = -0,2$ ou $-0,7$), la méthode Peto-quintile sous-estimait le bénéfice de survie restreint $rmstD(t^*)$ (Figure 4.3 et Figure 4.4) particulièrement sur le long terme ($t^* = 10$ ans). La méthode Pooled Exponentiel et, dans une bien moindre mesure, la méthode Pooled Kaplan-Meier étaient biaisées dans le cas de risques non-proportionnels (Figure 4.4). Dans tous ces cas, lorsqu'une méthode d'estimation du $rmstD(t^*)$ était biaisée, le biais augmentait avec $|\beta|$ (Tableau 4.1 et Tableau 4.2). Pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile un modèle à effets aléatoires était utilisé pour tenir compte de l'hétérogénéité de l'effet traitement (Figure 4.3 et Figure 4.4, Tableau 4.1 et Tableau 4.2).

Tableau 4.1 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint.
Scénario avec 5 essais et 200 patients par essai, et avec proportionnalité des risques.

Scénario hétérogénéité	Méthodes	$\beta = 0$						$\beta = -0,2$						$\beta = -0,7$					
		$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$		
		Vrai $rmstD = 0$			Vrai $rmstD = 0$			Vrai $rmstD = 0,2$			Vrai $rmstD = 0,6$			Vrai $rmstD = 0,8$			Vrai $rmstD = 2,0$		
		Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM
$(\sigma^2, \tau^2) = (0,01;0,01)$	Naïve Kaplan-Meier	0,00	0,12	0,11	0,00	0,28	0,24	0,00	0,12	0,11	0,00	0,29	0,24	0,01	0,12	0,11	0,01	0,28	0,23
	Pooled Kaplan-Meier	0,00	0,12	0,13	0,00	0,29	0,30	0,00	0,12	0,13	0,01	0,31	0,30	0,01	0,12	0,13	0,02	0,30	0,29
	Pooled Exponentiel	0,00	0,10	0,11	0,00	0,27	0,28	0,00	0,11	0,11	0,00	0,28	0,28	0,01	0,11	0,11	0,02	0,27	0,27
	Peto-quintile	0,00	0,11	0,10	0,00	0,24	0,21	-0,01	0,12	0,10	-0,07	0,25	0,21	-0,04	0,12	0,10	-0,21	0,26	0,23
$(\sigma^2, \tau^2) = (0,01;0,10)$	Naïve Kaplan-Meier	0,00	0,19	0,11	0,01	0,49	0,24	0,00	0,19	0,11	0,01	0,48	0,24	0,00	0,19	0,11	0,00	0,48	0,23
	Pooled Kaplan-Meier	0,00	0,20	0,18	0,01	0,50	0,47	0,00	0,19	0,19	0,00	0,49	0,47	0,00	0,20	0,18	0,01	0,48	0,44
	Pooled Exponentiel	0,00	0,19	0,17	0,01	0,48	0,45	0,00	0,18	0,17	0,01	0,47	0,46	0,00	0,19	0,17	0,01	0,47	0,44
	Peto-quintile	0,00	0,18	0,10	0,00	0,43	0,22	-0,01	0,18	0,10	-0,07	0,42	0,22	-0,04	0,19	0,10	-0,20	0,46	0,23
$(\sigma^2, \tau^2) = (0,10;0,01)$	Naïve Kaplan-Meier	-0,01	0,12	0,11	-0,01	0,28	0,24	0,00	0,12	0,11	0,01	0,27	0,24	0,00	0,14	0,11	0,01	0,28	0,24
	Pooled Kaplan-Meier	-0,01	0,12	0,13	0,00	0,28	0,29	0,00	0,12	0,13	0,00	0,29	0,29	0,00	0,14	0,13	0,01	0,28	0,29
	Pooled Exponentiel	0,00	0,10	0,11	-0,01	0,26	0,28	0,00	0,10	0,11	0,01	0,26	0,28	0,00	0,12	0,12	0,02	0,27	0,27
	Peto-quintile	-0,01	0,11	0,10	-0,01	0,23	0,22	-0,01	0,12	0,10	-0,06	0,24	0,21	-0,03	0,13	0,11	-0,15	0,26	0,23
$(\sigma^2, \tau^2) = (0,10;0,10)$	Naïve Kaplan-Meier	0,00	0,18	0,11	0,00	0,45	0,24	-0,01	0,19	0,11	0,00	0,48	0,24	0,01	0,19	0,11	0,01	0,44	0,24
	Pooled Kaplan-Meier	0,00	0,18	0,18	0,00	0,45	0,45	-0,01	0,19	0,18	0,00	0,47	0,45	0,01	0,19	0,18	0,02	0,44	0,43
	Pooled Exponentiel	0,00	0,17	0,17	0,00	0,44	0,44	0,00	0,18	0,17	0,00	0,47	0,44	0,01	0,19	0,18	0,02	0,44	0,42
	Peto-quintile	0,00	0,17	0,10	0,00	0,40	0,22	-0,01	0,18	0,10	-0,06	0,42	0,22	-0,02	0,19	0,11	-0,13	0,43	0,24

Un modèle à effets aléatoires a été utilisé pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile.

β : Taille de l'effet traitement ($=\log(HR)$); σ^2 : hétérogénéité du risque de base; τ^2 : hétérogénéité de l'effet traitement; ETE: erreur-type empirique; ETM: erreur-type moyenne; $rmstD$: bénéfice de survie restreint; t^* : horizon temporel.

Tableau 4.2 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint.
Scénario avec 5 essais et 200 patients par essai, et avec non-proportionnalité des risques.

Scénario hétérogénéité	Méthodes	$\beta = 0$						$\beta = -0,2$						$\beta = -0,7$					
		$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$		
		Vrai $rmstD = 0$			Vrai $rmstD = 0$			Vrai $rmstD = -0,1$			Vrai $rmstD = 0,1$			Vrai $rmstD = -0,3$			Vrai $rmstD = 0,3$		
		Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM
$(\sigma^2, \tau^2) = (0,01;0,01)$	Naïve Kaplan-Meier	0,00	0,12	0,11	0,00	0,28	0,24	0,00	0,11	0,11	0,01	0,27	0,24	0,00	0,12	0,11	0,00	0,28	0,24
	Pooled Kaplan-Meier	0,00	0,12	0,13	0,01	0,29	0,30	0,00	0,12	0,13	-0,01	0,29	0,30	0,01	0,12	0,13	-0,05	0,29	0,29
	Pooled Exponentiel	0,00	0,10	0,11	0,00	0,27	0,28	0,12	0,10	0,11	0,03	0,27	0,28	0,40	0,12	0,12	0,07	0,30	0,30
	Peto-quintile	0,00	0,11	0,10	0,00	0,24	0,21	0,01	0,11	0,10	-0,05	0,24	0,21	0,02	0,11	0,10	-0,19	0,24	0,21
$(\sigma^2, \tau^2) = (0,01;0,10)$	Naïve Kaplan-Meier	0,00	0,19	0,11	0,02	0,50	0,24	-0,01	0,19	0,11	-0,02	0,48	0,24	0,00	0,19	0,11	0,00	0,48	0,24
	Pooled Kaplan-Meier	0,01	0,19	0,18	0,02	0,50	0,46	-0,01	0,19	0,18	-0,03	0,49	0,46	0,00	0,19	0,18	-0,05	0,48	0,46
	Pooled Exponentiel	0,00	0,19	0,17	0,01	0,49	0,45	0,11	0,18	0,17	0,00	0,48	0,45	0,39	0,20	0,18	0,05	0,50	0,47
	Peto-quintile	0,00	0,18	0,10	0,01	0,43	0,22	0,00	0,18	0,10	-0,07	0,41	0,22	0,02	0,18	0,10	-0,18	0,42	0,22
$(\sigma^2, \tau^2) = (0,10;0,01)$	Naïve Kaplan-Meier	0,00	0,12	0,11	0,00	0,28	0,24	0,01	0,12	0,11	0,01	0,28	0,24	0,00	0,12	0,11	-0,01	0,27	0,24
	Pooled Kaplan-Meier	0,00	0,12	0,13	0,00	0,29	0,29	0,01	0,12	0,12	-0,01	0,29	0,29	0,00	0,12	0,13	-0,07	0,28	0,29
	Pooled Exponentiel	0,00	0,10	0,11	0,00	0,27	0,28	0,12	0,10	0,11	0,03	0,27	0,27	0,38	0,12	0,12	0,04	0,30	0,31
	Peto-quintile	0,00	0,11	0,10	0,00	0,24	0,21	0,01	0,12	0,10	-0,05	0,24	0,22	0,00	0,11	0,10	-0,21	0,25	0,22
$(\sigma^2, \tau^2) = (0,10;0,10)$	Naïve Kaplan-Meier	-0,01	0,19	0,11	-0,01	0,46	0,24	0,00	0,18	0,11	-0,01	0,46	0,24	0,00	0,20	0,11	0,01	0,49	0,24
	Pooled Kaplan-Meier	-0,01	0,19	0,18	-0,02	0,47	0,45	0,00	0,18	0,18	-0,03	0,46	0,45	0,00	0,20	0,18	-0,05	0,48	0,45
	Pooled Exponentiel	-0,01	0,18	0,17	-0,01	0,45	0,44	0,11	0,18	0,17	0,00	0,45	0,44	0,38	0,20	0,18	0,05	0,50	0,46
	Peto-quintile	-0,01	0,18	0,10	-0,01	0,41	0,22	0,00	0,18	0,11	-0,06	0,41	0,22	0,01	0,19	0,10	-0,18	0,44	0,22

Un modèle à effets aléatoires a été utilisé pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile.

β : Taille de l'effet traitement ($=\log(HR)$); σ^2 : hétérogénéité du risque de base; τ^2 : hétérogénéité de l'effet traitement; ETE: erreur-type empirique;

ETM: erreur-type moyenne; $rmstD$: bénéfice de survie restreint; t^* : horizon temporel.

Dans les scénarios avec une forte hétérogénéité inter-essai de l'effet traitement ($\tau^2 = 0,10$), toutes les méthodes présentaient une plus grande erreur-type empirique (ETE), comme montré dans les Figure 4.3 et Figure 4.4, et les Tableau 4.1 et Tableau 4.2. L'erreur-type était correctement estimée (erreur-type moyenne ETM = ETE) seulement pour les méthodes Pooled Kaplan-Meier et Pooled Exponentiel (avec effets aléatoires). L'erreur-type était généralement sous-estimée (ETM < ETE) pour les deux autres méthodes: l'ETM était par exemple deux fois plus petite que l'ETE pour les méthodes Naïve Kaplan-Meier et Peto-quintile dans le cas où $\tau^2 = 0,10$. Aucun impact n'a été noté ni sur le biais ni sur l'erreur-type, lorsque l'on faisait varier l'hétérogénéité du risque de base entre les essais (σ^2).

A la fois avec des risques proportionnels et non-proportionnels, la méthode Peto-quintile était peu biaisée à $t^* = 5$ ans mais davantage à $t^* = 10$ ans avec un biais jusqu'à 0,21 ans dans le cas d'un fort effet traitement avec $\beta = -0,7$ (Figure 4.3-B et Figure 4.4-B). Dans le cas de risques non-proportionnels, qui étaient incorporés à l'aide d'une distribution exponentielle par morceaux avec un effet traitement délétère les 2 premières années et un effet traitement bénéfique par la suite, la méthode Pooled Exponentiel était fortement biaisée à $t^* = 5$ ans, avec un biais de presque 0,40 ans comparé à un vrai bénéfice de survie restreint $rmstD(t^* = 5) = -0,30$ ans (Figure 4.4-A et Tableau 4.2). Ce biais suggérait que la méthode Pooled Exponentiel ne permet pas de refléter la distribution exponentielle par morceaux avec $\beta' = -\beta$ ($\beta \leq 0$) pour $t \in [0-2]$ ans et β pour $t > 2$ ans. Cependant, ce biais disparaissait à $t^* = 10$ ans (Figure 4.4-B), probablement parce que les risques étaient proportionnels entre 2 et 10 ans dans les données simulées et le différentiel de l'effet traitement des 2 premières années était ainsi atténué. Nous avons également noté un léger biais pour la méthode Pooled Kaplan-Meier à $t^* = 10$ ans lorsque les risques étaient non-proportionnels : un biais d'environ 0,05 ans à comparer avec un vrai bénéfice de survie restreint $rmstD(t^* = 10) = 0,30$ ans (Figure 4.4-B et Tableau 4.2).

Concernant le choix du temps de restriction t^* , quelle que soit la proportionnalité des risques, les erreurs-types empirique et moyenne étaient plus petites à $t^* = 5$ ans (Figure 4.3-A et Figure 4.4-A) qu'à $t^* = 10$ ans (Figure 4.3-B et Figure 4.4-B) ce qui peut s'expliquer de manière intuitive par le fait que nous avons de moins en moins de patients à risque au fur et à mesure des années et donc une moins grande précision dans notre estimation du $rmstD(t^*)$.

Légende des Figure 4.3 et Figure 4.4 :

Un modèle à effets aléatoires a été utilisé pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile.

β : Taille de l'effet traitement ($=\log(HR)$); σ^2 : hétérogénéité du risque de base; τ^2 : hétérogénéité de l'effet traitement; IC 95%: Intervalle de confiance à 95 %; $rmstD$: bénéfice de survie restreint; t^* : horizon temporel.

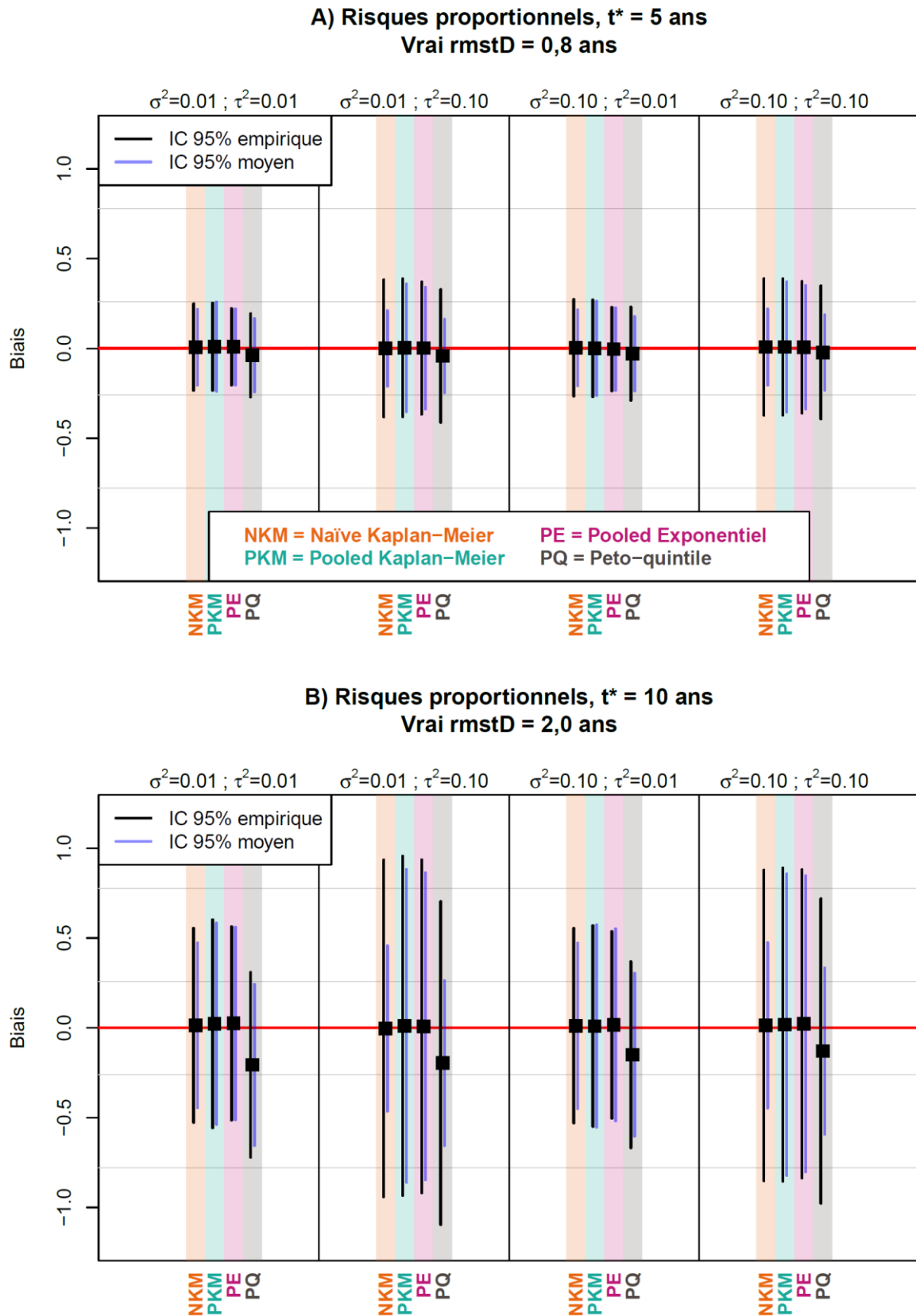


Figure 4.3 - Comparaison graphique à $t^* = 5$ ans (panel A) et à $t^* = 10$ ans (panel B) des méthodes en termes de biais moyen et d'erreurs-types empirique et moyenne. Scénario avec proportionnalité des risques, 5 essais et 200 patients par essai et $\beta = -0,7$.

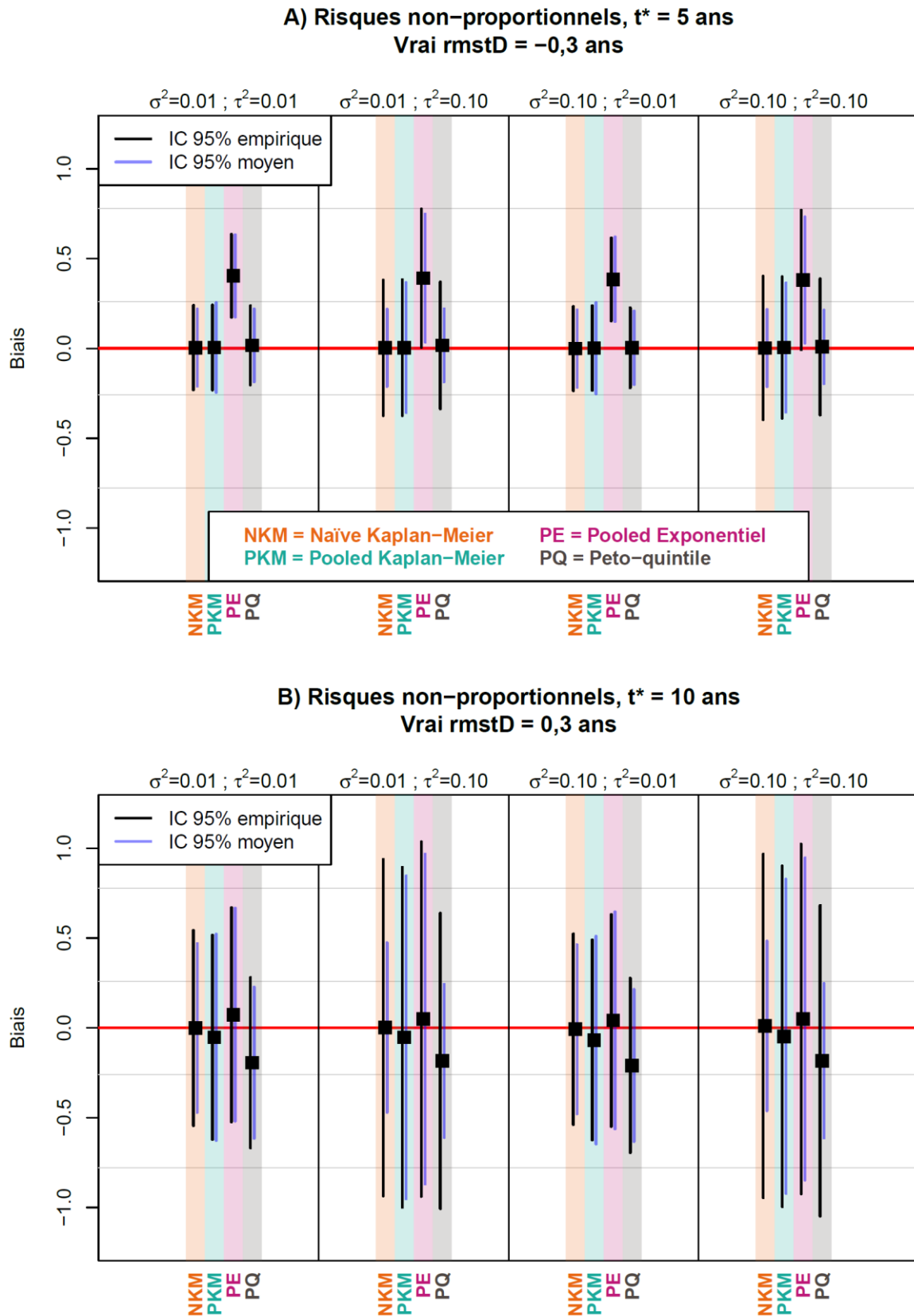


Figure 4.4 - Comparaison graphique à $t^* = 5$ ans (panel A) et à $t^* = 10$ ans (panel B) des méthodes en termes de biais moyen et d'erreurs-types empirique et moyenne. Scénario avec non-proportionnalité des risques, 5 essais et 200 patients par essai et $\beta = -0,7$.

Le nombre d'essais et le nombre de patients par essai n'avaient pas d'impact en termes de biais. Pour les méta-analyses simulées de $J = 20$ essais et $n_j = 100$ patients par essai (Tableau 4.3 pour les scénarios avec risques proportionnels et Tableau A.1 pour les scénarios avec risques non-proportionnels), toutes les méthodes avaient de plus petites erreurs-types empirique et moyenne que pour les méta-analyses de $J = 5$ essais et $n_j = 200$ patients par essai (Tableau 4.1 et Tableau 4.2).

Avec un nombre d'essais ($J = 5$) et un nombre total de patients ($N = 1\,000$) fixés, et un fort effet traitement bénéfique ($\beta = -0,7$), les résultats en termes de biais et d'erreur-types étaient similaires que le nombre de patients par essai soit fixé ($n_j = 200$; Tableau 4.1) ou tiré aléatoirement à l'aide d'une distribution de Poisson (Tableau A.2).

Tableau 4.3 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint.
Scénario avec 20 essais et 100 patients par essai, et avec proportionnalité des risques

Scénario hétérogénéité	Méthodes	$\beta = 0$						$\beta = -0,2$						$\beta = -0,7$					
		$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$		
		Vrai $rmstD = 0$			Vrai $rmstD = 0$			Vrai $rmstD = 0,2$			Vrai $rmstD = 0,6$			Vrai $rmstD = 0,8$			Vrai $rmstD = 2,0$		
		Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM
$(\sigma^2, \tau^2) = (0,01;0,01)$	Naïve Kaplan-Meier	0,00	0,08	0,08	0,00	0,18	0,17	0,00	0,08	0,08	0,01	0,19	0,17	0,00	0,08	0,08	0,00	0,17	0,16
	Pooled Kaplan-Meier	0,00	0,08	0,08	0,00	0,20	0,20	0,00	0,08	0,08	0,02	0,21	0,20	0,01	0,08	0,08	0,02	0,19	0,19
	Pooled Exponentiel	0,00	0,07	0,08	0,00	0,19	0,19	0,01	0,07	0,08	0,01	0,19	0,19	0,01	0,07	0,08	0,03	0,17	0,18
	Peto-quintile	0,00	0,08	0,07	0,00	0,16	0,15	-0,01	0,08	0,07	-0,07	0,16	0,15	-0,05	0,07	0,07	-0,22	0,16	0,16
$(\sigma^2, \tau^2) = (0,01;0,10)$	Naïve Kaplan-Meier	0,00	0,11	0,08	0,01	0,27	0,17	0,00	0,11	0,08	0,01	0,27	0,17	0,00	0,11	0,08	0,00	0,26	0,16
	Pooled Kaplan-Meier	0,00	0,11	0,11	0,01	0,28	0,28	0,01	0,12	0,11	0,01	0,28	0,27	0,01	0,11	0,11	0,03	0,26	0,26
	Pooled Exponentiel	0,00	0,11	0,10	0,01	0,27	0,27	0,01	0,11	0,10	0,01	0,28	0,27	0,01	0,10	0,10	0,03	0,26	0,26
	Peto-quintile	0,00	0,11	0,07	0,00	0,24	0,15	-0,01	0,11	0,07	-0,06	0,24	0,15	-0,04	0,11	0,07	-0,19	0,25	0,17
$(\sigma^2, \tau^2) = (0,10;0,01)$	Naïve Kaplan-Meier	0,00	0,08	0,08	0,01	0,18	0,17	0,00	0,08	0,08	0,00	0,18	0,17	0,00	0,09	0,08	0,00	0,18	0,17
	Pooled Kaplan-Meier	0,00	0,08	0,08	0,00	0,20	0,19	0,00	0,08	0,08	0,01	0,20	0,19	0,00	0,09	0,09	0,03	0,19	0,19
	Pooled Exponentiel	0,00	0,07	0,08	0,00	0,18	0,18	0,00	0,07	0,08	0,01	0,19	0,18	0,00	0,08	0,08	0,03	0,18	0,18
	Peto-quintile	0,00	0,08	0,07	0,00	0,16	0,15	-0,01	0,08	0,07	-0,05	0,17	0,16	-0,03	0,08	0,08	-0,14	0,17	0,17
$(\sigma^2, \tau^2) = (0,10;0,10)$	Naïve Kaplan-Meier	0,00	0,11	0,08	0,00	0,27	0,17	0,00	0,11	0,08	0,00	0,26	0,17	0,00	0,11	0,08	0,01	0,26	0,17
	Pooled Kaplan-Meier	0,00	0,11	0,11	0,00	0,27	0,27	0,00	0,11	0,11	0,00	0,26	0,26	0,01	0,11	0,11	0,02	0,27	0,26
	Pooled Exponentiel	0,00	0,10	0,10	0,00	0,27	0,26	0,00	0,10	0,10	0,00	0,26	0,26	0,00	0,11	0,10	0,03	0,26	0,25
	Peto-quintile	0,00	0,11	0,07	0,00	0,24	0,16	-0,01	0,10	0,07	-0,06	0,23	0,16	-0,03	0,11	0,08	-0,13	0,26	0,17

Un modèle à effets aléatoires a été utilisé pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile.

β : Taille de l'effet traitement ($=\log(HR)$); σ^2 : hétérogénéité du risque de base; τ^2 : hétérogénéité de l'effet traitement; ETE: erreur-type empirique; ETM: erreur-type moyenne; rmstD: bénéfice de survie restreint; t^* : horizon temporel.

Nous avons également considéré l'impact d'un effet traitement délétère ($\beta = +0,2 ; +0,7$) mais, les résultats n'étaient pas affectés : les méthodes biaisées avaient des biais inversés, et les erreurs-types empirique et moyenne restaient inchangées (Tableau 4.4). Le biais de la méthode Peto-quintile par exemple était de 0,22 an à $t^* = 10$ ans pour le scénario $\sigma^2 = \tau^2 = 0,01$ et $\beta = +0,7$ par rapport à un vrai $rmstD(t^* = 10)$ égal à -2,0 années. Cette méthode estimait donc un $rmstD(t^* = 10)$ plus proche de 0 que la vraie valeur ; c'était déjà le cas avec un effet traitement bénéfique ($\beta = -0,7$), avec un biais de -0,21 an par rapport à un vrai $rmstD(t^* = 10)$ de 2,0 années (Tableau 4.1). La symétrie des résultats par rapport aux scénarios avec un effet traitement bénéfique (Tableau 4.1) était intuitivement attendue au vu du codage utilisé pour la variable représentant le bras de traitement ($\pm 1/2$) (équation (4.1)) et de la distribution symétrique des deux effets aléatoires (équation (4.2)).

Tableau 4.4 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint.
Scénario avec 5 essais et 200 patients par essai, avec proportionnalité des risques, et effet traitement délétère.

Hétérogénéité scenario	Méthodes	$\beta = 0,2$						$\beta = 0,7$					
		$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$		
		Vrai $rmstD = -0,2$			Vrai $rmstD = -0,6$			Vrai $rmstD = -0,8$			Vrai $rmstD = -2,0$		
		Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM
$(\sigma^2, \tau^2) = (0,01;0,01)$	Naïve Kaplan-Meier	0,00	0,12	0,11	0,00	0,28	0,24	0,00	0,12	0,11	0,01	0,28	0,23
	Pooled Kaplan-Meier	0,00	0,12	0,13	-0,01	0,29	0,30	0,00	0,12	0,13	0,00	0,30	0,29
	Pooled Exponentiel	0,00	0,10	0,11	-0,01	0,27	0,28	0,00	0,11	0,11	0,00	0,28	0,28
	Peto-quintile	0,02	0,11	0,10	0,08	0,24	0,21	0,04	0,12	0,10	0,22	0,27	0,23
$(\sigma^2, \tau^2) = (0,01;0,10)$	Naïve Kaplan-Meier	0,00	0,19	0,11	0,01	0,49	0,24	0,00	0,19	0,11	0,00	0,48	0,23
	Pooled Kaplan-Meier	0,00	0,20	0,18	0,00	0,50	0,46	0,00	0,20	0,18	0,00	0,48	0,45
	Pooled Exponentiel	0,00	0,19	0,17	0,00	0,48	0,45	0,00	0,19	0,17	-0,01	0,47	0,44
	Peto-quintile	0,01	0,18	0,10	0,07	0,43	0,22	0,04	0,19	0,11	0,20	0,46	0,23
$(\sigma^2, \tau^2) = (0,10;0,01)$	Naïve Kaplan-Meier	-0,01	0,12	0,11	-0,01	0,27	0,24	0,00	0,13	0,11	0,00	0,27	0,23
	Pooled Kaplan-Meier	0,00	0,12	0,13	-0,01	0,28	0,29	0,00	0,13	0,13	0,00	0,29	0,29
	Pooled Exponentiel	0,00	0,10	0,11	-0,01	0,26	0,28	0,01	0,12	0,12	-0,01	0,26	0,27
	Peto-quintile	0,01	0,11	0,10	0,05	0,23	0,22	0,03	0,13	0,11	0,16	0,26	0,23
$(\sigma^2, \tau^2) = (0,10;0,10)$	Naïve Kaplan-Meier	0,00	0,18	0,11	-0,01	0,44	0,24	0,00	0,20	0,11	0,00	0,46	0,24
	Pooled Kaplan-Meier	0,00	0,18	0,18	-0,01	0,45	0,45	0,00	0,20	0,19	0,00	0,45	0,44
	Pooled Exponentiel	0,00	0,17	0,17	0,00	0,44	0,44	0,00	0,18	0,18	-0,01	0,44	0,43
	Peto-quintile	0,01	0,17	0,10	0,06	0,40	0,22	0,03	0,19	0,11	0,14	0,45	0,24

Un modèle à effets aléatoires a été utilisé pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile.

β : Taille de l'effet traitement ($=\log(HR)$); σ^2 : hétérogénéité du risque de base; τ^2 : hétérogénéité de l'effet traitement; ETE: erreur-type empirique; ETM: erreur-type moyenne; $rmstD$: bénéfice de survie restreint; t^* : horizon temporel.

L'introduction d'une corrélation négative entre les deux effets aléatoires A_j et B_j n'avait pas d'impact majeur en termes de biais moyen et d'estimation de l'erreur-type, à l'exception du scénario avec une forte hétérogénéité du risque de base et de l'effet traitement ($\sigma^2 = \tau^2 = 0,10$) et avec $\beta = -0,7$ pour lequel les erreurs-types moyenne et empirique étaient légèrement plus importantes pour toutes les méthodes (Tableau 4.5) par rapport aux scénarios sans corrélation. Par exemple, à $t^* = 10$ ans, l'erreur-type empirique était de 0,50 pour les méthodes Naïve Kaplan-Meier, Pooled Kaplan-Meier et Pooled Exponentiel et 0,52 pour la méthode Peto-quintile avec corrélation, contre respectivement 0,44 et 0,43 pour un scénario sans corrélation entre les deux effets aléatoires (Tableau 4.1).

Tableau 4.5 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint.
Scénario avec 5 essais et 200 patients par essai, avec proportionnalité des risques, et avec corrélation entre les effets aléatoires.

Scénario hétérogénéité	Méthodes	$\beta = 0$						$\beta = -0,2$						$\beta = -0,7$					
		$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$		
		Vrai $rmstD = 0$			Vrai $rmstD = 0$			Vrai $rmstD = 0,2$			Vrai $rmstD = 0,6$			Vrai $rmstD = 0,8$			Vrai $rmstD = 2,0$		
		Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM
$(\sigma^2, \tau^2) = (0,01;0,01)$	Naïve Kaplan-Meier	0,00	0,12	0,11	0,00	0,28	0,24	0,00	0,12	0,11	0,00	0,29	0,24	0,01	0,13	0,11	0,01	0,27	0,23
	Pooled Kaplan-Meier	0,00	0,12	0,13	0,00	0,29	0,30	0,00	0,13	0,13	0,00	0,30	0,30	0,01	0,13	0,13	0,02	0,30	0,29
	Pooled Exponentiel	-0,01	0,10	0,11	0,00	0,27	0,28	0,00	0,11	0,11	0,00	0,28	0,28	0,01	0,11	0,11	0,03	0,27	0,28
	Peto-quintile	0,00	0,11	0,10	0,00	0,24	0,21	-0,01	0,12	0,10	-0,07	0,25	0,21	-0,04	0,12	0,11	-0,20	0,26	0,23
$(\sigma^2, \tau^2) = (0,01;0,10)$	Naïve Kaplan-Meier	0,00	0,19	0,11	0,02	0,47	0,24	0,00	0,19	0,11	0,00	0,46	0,24	0,01	0,19	0,11	0,02	0,46	0,23
	Pooled Kaplan-Meier	0,00	0,19	0,18	0,02	0,47	0,45	-0,01	0,19	0,19	0,00	0,48	0,46	0,01	0,19	0,19	0,03	0,47	0,45
	Pooled Exponentiel	0,00	0,18	0,17	0,02	0,46	0,44	0,00	0,18	0,17	0,00	0,46	0,45	0,01	0,19	0,18	0,03	0,46	0,43
	Peto-quintile	0,00	0,18	0,10	0,02	0,41	0,22	-0,02	0,18	0,10	-0,07	0,41	0,22	-0,03	0,19	0,11	-0,16	0,44	0,24
$(\sigma^2, \tau^2) = (0,10;0,01)$	Naïve Kaplan-Meier	0,00	0,12	0,11	-0,01	0,27	0,24	0,00	0,13	0,11	0,01	0,28	0,24	0,01	0,15	0,11	0,01	0,29	0,23
	Pooled Kaplan-Meier	-0,01	0,12	0,13	0,01	0,28	0,29	0,00	0,13	0,13	0,02	0,30	0,30	0,01	0,15	0,15	0,03	0,30	0,29
	Pooled Exponentiel	-0,01	0,10	0,11	0,00	0,27	0,28	0,00	0,11	0,11	0,02	0,27	0,28	0,00	0,14	0,13	0,03	0,28	0,28
	Peto-quintile	0,00	0,11	0,10	0,01	0,24	0,22	-0,01	0,12	0,10	-0,04	0,25	0,22	-0,02	0,15	0,11	-0,12	0,28	0,23
$(\sigma^2, \tau^2) = (0,10;0,10)$	Naïve Kaplan-Meier	0,00	0,19	0,11	0,02	0,46	0,24	-0,02	0,21	0,11	-0,03	0,49	0,24	0,02	0,24	0,11	0,04	0,50	0,24
	Pooled Kaplan-Meier	0,00	0,19	0,19	0,01	0,47	0,46	-0,02	0,21	0,20	-0,03	0,49	0,46	0,01	0,24	0,23	0,04	0,50	0,47
	Pooled Exponentiel	0,00	0,18	0,18	0,02	0,46	0,45	-0,02	0,20	0,19	-0,03	0,49	0,45	0,01	0,24	0,22	0,04	0,50	0,46
	Peto-quintile	0,00	0,18	0,11	0,04	0,42	0,22	-0,03	0,20	0,11	-0,06	0,45	0,23	-0,01	0,25	0,11	-0,04	0,52	0,25

Un modèle à effets aléatoires a été utilisé pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile.

β : Taille de l'effet traitement ($=\log(HR)$); σ^2 : hétérogénéité du risque de base; τ^2 : hétérogénéité de l'effet traitement; ETE: erreur-type empirique;

ETM: erreur-type moyenne; $rmstD$: bénéfice de survie restreint; t^* : horizon temporel.

Dans le cas d'une forte hétérogénéité de l'effet traitement ($\tau^2 = 0,10$), les méthodes d'estimation Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile présentaient un biais légèrement plus important lorsqu'un modèle à effets fixes était utilisé (Tableau 4.6) comparé aux résultats avec un modèle à effets aléatoires de DerSimonian-Laird utilisé dans le Tableau 4.1, et ce particulièrement pour $\beta = -0,7$ et à $t^* = 10$ ans. Si l'on regarde par exemple le quatrième bloc horizontal, correspondant au scénario $\sigma^2 = \tau^2 = 0,10$, et la dernière colonne ($\beta = -0,7$ et $t^* = 10$ ans) le biais était de 0,02 pour Pooled Kaplan-Meier et Pooled Exponentiel, et -0,13 an pour Peto-quintile avec l'utilisation d'un modèle à effets aléatoires (Tableau 4.1), contre 0,05 et -0,15 (Tableau 4.6) avec un modèle à effets fixes.

Concernant l'estimation de l'erreur-type, toujours en regardant les mêmes scénarios ($\sigma^2 = \tau^2 = 0,10$; $\beta = -0,7$ et $t^* = 10$ ans) les valeurs d'erreur-type empirique (ETE) étaient très proches entre l'utilisation d'un modèle à effets aléatoires avec une ETE de 0,44 pour Pooled Kaplan-Meier et Pooled Exponentiel, et 0,43 pour Peto-quintile (Tableau 4.1), et l'utilisation d'un modèle à effets fixes avec une ETE respectivement de 0,45 et 0,43 (Tableau 4.6). Par contre, les erreurs-types moyennes (ETM) étaient plus faibles avec un modèle à effets fixes : l'ETM valait 0,22 pour Pooled Kaplan-Meier, 0,22 pour Pooled Exponentiel (Tableau 4.6) contre respectivement 0,43, et 0,42 avec un modèle à effets aléatoires (Tableau 4.1). Des résultats similaires pouvaient être observés dans le scénario avec une forte hétérogénéité de l'effet traitement ($\tau^2 = 0,10$) et une faible hétérogénéité du risque de base ($\sigma^2 = 0,01$) (Tableau 4.6).

Ainsi, l'utilisation d'un modèle à effets fixes entraînait une sous-estimation de l'erreur-type (ETM $\ll$ ETE) dans le cas d'une forte hétérogénéité de l'effet traitement ($\tau^2 = 0,10$), la particularité d'un modèle à effets aléatoires permettant justement de prendre en compte la variabilité de l'effet traitement inter-essai (τ^2) dans l'estimation de l'erreur-type de l'effet global. Cette sous-estimation de l'erreur-type du $rmstD(t^*)$ était cependant moins marquée dans les deux scénarios avec une faible hétérogénéité de l'effet traitement ($\tau^2 = 0,01$). Il est à noter que pour la méthode Peto-quintile, l'ETM était sous-estimée que ce soit avec un modèle à effets fixes (ETM = 0,24 contre ETE = 0,43 ; Tableau 4.6) ou avec un modèle à effets aléatoires (ETM = 0,23 contre ETE = 0,43 ; Tableau 4.1).

Tableau 4.6 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint.
Scénario avec 5 essais et 200 patients par essai, avec proportionnalité des risques, et modèle à effets fixes.

Scénario hétérogénéité	Méthodes	$\beta = 0$						$\beta = -0,2$						$\beta = -0,7$					
		$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$		
		Vrai $rmstD = 0$			Vrai $rmstD = 0$			Vrai $rmstD = 0,2$			Vrai $rmstD = 0,6$			Vrai $rmstD = 0,8$			Vrai $rmstD = 2,0$		
		Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM
$(\sigma^2, \tau^2) = (0,01;0,01)$	Naïve Kaplan-Meier	0,00	0,12	0,11	0,00	0,28	0,24	0,00	0,12	0,11	0,00	0,29	0,24	0,01	0,12	0,11	0,01	0,28	0,23
	Pooled Kaplan-Meier	0,00	0,12	0,11	0,00	0,29	0,23	0,00	0,12	0,11	0,01	0,31	0,23	0,01	0,12	0,11	0,03	0,30	0,23
	Pooled Exponentiel	0,00	0,10	0,09	0,00	0,27	0,24	0,00	0,11	0,09	0,01	0,28	0,24	0,01	0,11	0,09	0,03	0,27	0,23
	Peto-quintile	0,00	0,11	0,10	0,00	0,24	0,21	-0,01	0,12	0,10	-0,07	0,25	0,21	-0,04	0,12	0,10	-0,21	0,26	0,22
$(\sigma^2, \tau^2) = (0,01;0,10)$	Naïve Kaplan-Meier	0,00	0,19	0,11	0,01	0,49	0,24	0,00	0,19	0,11	0,01	0,48	0,24	0,00	0,19	0,11	0,00	0,48	0,23
	Pooled Kaplan-Meier	0,00	0,20	0,11	0,01	0,51	0,23	0,00	0,19	0,11	0,02	0,50	0,23	0,01	0,20	0,11	0,06	0,49	0,23
	Pooled Exponentiel	0,00	0,19	0,09	0,01	0,50	0,24	0,00	0,18	0,09	0,02	0,49	0,23	0,00	0,19	0,09	0,05	0,49	0,23
	Peto-quintile	0,00	0,18	0,10	0,00	0,42	0,21	-0,01	0,18	0,10	-0,07	0,42	0,21	-0,04	0,19	0,10	-0,21	0,45	0,22
$(\sigma^2, \tau^2) = (0,10;0,01)$	Naïve Kaplan-Meier	-0,01	0,12	0,11	-0,01	0,28	0,24	0,00	0,12	0,11	0,01	0,27	0,24	0,00	0,14	0,11	0,01	0,28	0,24
	Pooled Kaplan-Meier	-0,01	0,12	0,11	-0,01	0,28	0,23	0,00	0,12	0,11	0,00	0,29	0,23	0,00	0,14	0,11	0,02	0,29	0,22
	Pooled Exponentiel	0,00	0,10	0,09	-0,01	0,26	0,23	0,00	0,10	0,09	0,01	0,26	0,23	-0,01	0,12	0,09	0,02	0,27	0,22
	Peto-quintile	-0,01	0,11	0,10	-0,01	0,23	0,21	-0,01	0,12	0,10	-0,06	0,24	0,21	-0,03	0,13	0,10	-0,15	0,26	0,23
$(\sigma^2, \tau^2) = (0,10;0,10)$	Naïve Kaplan-Meier	0,00	0,18	0,11	0,00	0,45	0,24	-0,01	0,19	0,11	0,00	0,48	0,24	0,01	0,19	0,11	0,01	0,44	0,24
	Pooled Kaplan-Meier	0,00	0,18	0,11	0,00	0,46	0,23	-0,01	0,19	0,11	0,01	0,49	0,23	0,01	0,20	0,10	0,05	0,45	0,22
	Pooled Exponentiel	0,00	0,17	0,09	0,00	0,45	0,23	0,00	0,18	0,09	0,01	0,48	0,23	0,00	0,19	0,09	0,05	0,45	0,22
	Peto-quintile	0,00	0,17	0,10	0,00	0,40	0,21	-0,02	0,18	0,10	-0,06	0,42	0,21	-0,02	0,19	0,11	-0,15	0,43	0,23

Un modèle à effets fixes est utilisé pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile.

β : Taille de l'effet traitement ($=\log(HR)$); σ^2 : hétérogénéité du risque de base; τ^2 : hétérogénéité de l'effet traitement; ETE: erreur-type empirique;

ETM: erreur-type moyenne; $rmstD$: bénéfice de survie restreint; t^* : horizon temporel.

4.3. Mise en perspective des résultats de l'étude de cas *Meta-Analysis of Radiotherapy in Lung Cancer*

Les résultats de l'étude de simulation permettent de mieux appréhender les résultats obtenus dans l'étude de cas avec la méta-analyse sur données individuelles MAR-LC (*Meta-Analysis of Radiotherapy in Lung Cancer*) du Chapitre 3.

Le *hazard ratio* global dans MAR-LC était de 0,88 (IC 95 % : [0,80 ; 0,97]), ce qui correspond à un effet traitement global $\beta = \log(\text{HR}) = -0,13$. L'hypothèse de la proportionnalité des risques était vérifiée globalement dans la méta-analyse MAR-LC ($p = 0,91$) d'après la méthodologie de Wei et al. (Wei *et al.*, 2015). Il n'y avait pas d'hétérogénéité significative de l'effet du traitement entre les essais (Q-Cochran $p = 0,37$, $I^2 = 8\%$) et nous avons donc utilisé un modèle à effets fixes pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile. Au vu de ces éléments, les résultats obtenus dans le Chapitre 3 sont à comparer avec les résultats de l'étude de simulation, avec la proportionnalité des risques et l'utilisation d'un modèle à effets fixes, décrits dans le Tableau 4.6. Plus précisément, les résultats les plus comparables sont ceux donnés pour un effet traitement modéré ($\beta = -0,2$) et pour le scénario avec une très faible hétérogénéité de l'effet traitement et du risque de base ($\sigma^2 = \tau^2 = 0,01$) qui correspondent au premier bloc horizontal du Tableau 4.6.

De plus, le risque de décès de base était plus élevé dans la méta-analyse MAR-LC avec une médiane de survie globale de 1,2 an, comparé au risque de base défini dans l'étude de simulation avec une médiane de survie égale à 5 ans. L'horizon temporel dans MAR-LC était $t^* = 5$ ans, ce qui correspondait à un temps où seuls 10 % des patients dans MAR-LC étaient encore vivants. Les risques de base étant différents entre ces deux études, les résultats de l'estimation du bénéfice de survie restreint dans MAR-LC à $t^* = 5$ ans doivent donc être comparés à ceux obtenus dans l'étude de simulation à $t^* = 10$ ans, où environ 20 % des patients étaient encore vivants, plutôt qu'à ceux obtenus à 5 ans qui était la médiane de survie simulée avec 50 % de patients encore vivants.

Nous avons ainsi résumé dans un même tableau, les résultats de l'étude de cas extraits du Tableau 3.3 et ceux de l'étude de simulation extraits du Tableau 4.6. Afin de pouvoir comparer directement les résultats de ces deux études, nous avons exprimé les bénéfices de survie en mois.

Tableau 4.7 - Estimation du bénéfice de survie restreint et de son erreur-type selon les différentes méthodes étudiées dans l'étude de cas et dans l'étude de simulation

Méthode d'estimation	Etude de cas MAR-LC		Etude de simulation	
	$t^* = 5$ ans		$(\beta = -0,2 \text{ et } \sigma^2 = \tau^2 = 0,01)$	
			$t^* = 10$ ans	
	rmstD moyen (en mois) [IC 95 %] [§]	Erreur-type du rmstD [§]	rmstD moyen (en mois) [IC 95 %] ^ℓ	Erreur-type du rmstD ^ℓ
Peto-quintile	1,7 [0,4 ; 3,1]	0,7	6,1 [1,1 ; 11,0]	2,5
Naïve Kaplan-Meier	2,2 [0,6 ; 3,7]	0,8	7,0 [1,3 ; 12,7]	2,9
Pooled Kaplan-Meier	2,3 [0,7 ; 3,8]	0,8	7,1 [1,6 ; 12,6]	2,8
Pooled Exponentiel	2,5 [0,7 ; 4,2]	0,9	7,1 [1,5 ; 12,6]	2,9

[§] : IC 95% et erreur-type estimés à l'aide d'un *bootstrap* non-paramétrique avec 1 000 échantillons

^ℓ : IC 95% et erreur-type basés sur l'erreur-type moyenne estimée à l'aide des 1 000 méta-analyses simulées

Lorsque l'on compare les résultats issus de l'étude de cas et ceux de l'étude de simulation, nous pouvons tout d'abord remarquer que les quatre méthodes peuvent être classées dans le même ordre croissant en termes d'estimation du bénéfice de survie restreint et de son erreur-type. Ainsi, la méthode Peto-quintile se démarque des trois autres méthodes d'estimation : le bénéfice de survie restreint et son erreur-type estimés par cette méthode sont inférieurs à ceux des autres méthodes. On retrouve dans l'étude de cas, le fait que le bénéfice de survie restreint de la méthode Peto-quintile est plus faible que celui des autres méthodes, bien que cette différence soit non statistiquement significative, puisque les intervalles de confiance se chevauchent entre les quatre méthodes.

Les trois autres méthodes donnent des résultats similaires entre elles que ce soit pour l'étude de cas ou pour l'étude de simulation. La seule différence que nous pourrions noter est entre les méthodes Pooled Kaplan-Meier et Pooled Exponentiel pour l'étude de cas (2,3 mois contre 2,5 mois), alors que les résultats pour l'étude de simulation étaient identiques (7,1 mois).

L'étude de simulation semble donc infirmer la surestimation du $rmstD(t^*)$ et de son erreur-type par la méthode Pooled Exponentiel par rapport à la méthode Pooled Kaplan-Meier, ce qui était suggéré dans l'étude de cas du Chapitre 3. Néanmoins cette différence est légère et peut s'expliquer par un problème d'adéquation (*fit* en anglais) du modèle exponentiel aux données. Dans l'étude de simulation, les données étaient simulées à l'aide d'une distribution exponentielle, le modèle exponentiel était donc par nature le meilleur candidat parmi les différents modèles de survie paramétriques. Dans l'étude de cas sur la méta-analyse MAR-LC, le modèle exponentiel semblait bien correspondre à la distribution des données de survie dans chaque essai d'après les tests du rapport de vraisemblance effectués par rapport à un modèle de Weibull. Cependant, un modèle paramétrique plus flexible tel que le modèle Gamma ou log-normal aurait pu être plus approprié pour certains essais. Cela peut expliquer la légère différence entre ces deux méthodes à la fois dans l'estimation du $rmstD(t^*)$, mais aussi dans l'estimation de son erreur-type, différence que l'on peut également étudier graphiquement pour chacun des essais d'après le forest-plot de la Figure 3.4. Concernant l'erreur-type du $rmstD(t^*)$, une autre explication à la différence d'estimation entre les méthodes Pooled Kaplan-Meier et Pooled Exponentiel serait que l'erreur-type était estimée par *bootstrap* dans l'étude de cas et analytiquement dans l'étude de simulation.

4.4. Discussion

Dans cette étude de simulation, nous avons comparé quatre méthodes pour estimer le bénéfice de survie restreint $rmstD(t^*)$ à partir d'une méta-analyse sur données individuelles (MADI) à travers différents scénarios étudiant plusieurs paramètres clé en méta-analyse.

Nous avons montré que la méthode Pooled Kaplan-Meier était très rarement biaisée, à l'exception d'un léger biais dans le cas de risques non-proportionnels avec un temps de restriction à 10 ans. De même, la méthode Naïve Kaplan-Meier était non biaisée pour tous les scénarios envisagés, tandis que la méthode Pooled Exponentiel présentait un léger biais avec des risques non-proportionnels à $t^* = 10$ ans, et un biais plus important à $t^* = 5$ ans. Enfin, la méthode Peto-quintile sous-estimait le $rmstD(t^*)$, à l'exception d'une estimation à $t^* = 5$ ans dans le cas de non-proportionnalité des risques.

En présence d'une forte hétérogénéité de l'effet traitement, seules les méthodes Pooled Kaplan-Meier et Pooled Exponentiel avec un modèle à effets aléatoires de DerSimonian-Laird estimaient correctement l'erreur-type du $rmstD(t^*)$, puisque pour ces méthodes l'erreur-type moyenne était égale à l'erreur-type empirique.

La méthode Naïve Kaplan-Meier, n’offrant pas la possibilité de tenir compte de l’hétérogénéité de l’effet traitement, sous-estimait l’erreur-type du $rmstD(t^*)$ en cas d’hétérogénéité même faible de l’effet traitement (Figure 4.3 et Figure 4.4). Cette méthode n’était pas biaisée dans notre étude de simulation qui comportait des essais simulés avec un ratio d’allocation 1:1 entre les deux bras de traitement. Cependant, il est probable qu’un biais lié au paradoxe de Simpson apparaîtrait si au sein d’une méta-analyse le ratio d’allocation de certains essais était variable et différent de 1:1.

La méthode Peto-quintile sous-estimait le $rmstD(t^*)$ dans la grande majorité des scénarios. Dans les scénarios avec une forte hétérogénéité de l’effet traitement, la méthode Peto-quintile, malgré l’utilisation d’un modèle à effets aléatoires dans l’estimation du HR global, sous-estimait l’erreur-type du $rmstD(t^*)$.

La méthode Pooled Exponentiel était fortement biaisée à $t^* = 5$ ans dans le cas de risque non-proportionnels, ce qui était attendu puisque ces scénarios étaient simulés à l’aide d’un modèle exponentiel par morceaux (Figure 4.4-A et Tableau 4.2).

Par conséquent, parmi les méthodes étudiées dans cette étude de simulation, la méthode Pooled Kaplan-Meier – avec un modèle à effets aléatoires de DerSimonian-Laird pour agréger les $rmstD_j(t^*)$ estimés dans chacun des essais – était la méthode présentant les meilleures performances pour estimer le bénéfice de survie restreint à partir d’une méta-analyse sur données individuelles. Cette méthode permettait d’obtenir le meilleur compromis en termes de biais moyen et d’erreur-type pour estimer le $rmstD(t^*)$ à travers les scénarios investigués dans notre étude de simulation. C’est cette méthode que nous avons privilégiée dans les trois exemples d’utilisation du bénéfice de survie restreint à partir de MADI décrits dans le chapitre suivant.

Chapitre 5. Estimation du bénéfice de survie : application à trois méta-analyses sur données individuelles MAC-NPC 1, MAC-NPC 2 et RTT-SCLC

Dans ce chapitre, nous présentons trois exemples de méta-analyses sur données individuelles (MADI) pour lesquelles nous avons estimé le bénéfice de survie restreint. Les deux premiers exemples correspondent à la *Meta-Analysis of Chemotherapy in Nasopharynx Carcinoma* (MAC-NPC 1 ; (Baujat *et al.*, 2006)) et sa version actualisée MAC-NPC 2 (Blanchard *et al.*, 2015). Ces deux méta-analyses ont déjà été publiées, et avaient choisi un *hazard ratio* (HR) global estimé par un modèle à effets fixes comme mesure de l'effet traitement dans l'analyse princeps. En analyse secondaire, nous avons estimé le HR global par un modèle à effets aléatoires et avons également estimé le bénéfice de survie restreint ($rmstD(t^*)$) pour MAC-NPC 1 et MAC-NPC 2 avec les quatre méthodes étudiées dans les simulations (Lueza *et al.*, 2016b). Le troisième exemple, *RadioTherapy Timing in Small-Cell Lung Cancer* (RTT-SCLC) est une MADI originale pour laquelle l'analyse princeps a été réalisée à l'aide à la fois du HR et du $rmstD(t^*)$ estimé par la méthode Pooled Kaplan-Meier.

5.1. Meta-Analysis of Chemotherapy in Nasopharynx Carcinoma (MAC-NPC 1 et MAC-NPC 2)

Nous avons illustré les quatre méthodes d'estimation du bénéfice de survie restreint $rmstD(t^*)$, précédemment étudiées dans le Chapitre 4, à partir des données individuelles de la méta-analyse *Meta-Analysis of Chemotherapy in Nasopharynx Carcinoma* (MAC-NPC 1) et de sa version actualisée MAC-NPC 2. Ces deux MADI étudiaient l'ajout de la chimiothérapie (CT) à la radiothérapie (RT) dans le cancer du nasopharynx, et différaient en termes de significativité de l'effet traitement global et d'hétérogénéité de l'effet traitement. Nous présentons les estimations

de l'effet traitement dans ces deux MADI mesuré à l'aide à la fois du *hazard ratio* et du *rmstD(t*)*. Les résultats obtenus avec le *hazard ratio* global estimé par un modèle à effets fixes provenaient d'analyses antérieures à la thèse et avaient été publiés dans les articles cliniques correspondant aux méta-analyses MAC-NPC 1 et MAC-NPC 2. Dans le cadre de cette thèse, nous avons ré-estimé le HR global à l'aide d'un modèle à effets aléatoires, et nous avons également estimé le *rmstD(t*)* avec les quatre méthodes étudiées dans les simulations (Lueza *et al.*, 2016b).

5.1.1. Résultats de l'analyse princeps avec le hazard ratio

La méta-analyse MAC-NPC 1 incluait 1 753 patients dans 8 essais avec un suivi médian de 6 ans, variant de 3 à 9 ans selon les essais (Baujat *et al.*, 2006). Du fait du plan factoriel 2x2 d'un essai, 11 comparaisons étaient finalement incluses correspondant à un total de 1 975 patients puisque certains patients avaient été dupliqués dans l'essai comportant un plan factoriel.

Le *hazard ratio* global estimé par un modèle à effets fixes était de 0,82 (IC 95 % : [0,71 ; 0,94]), indiquant une amélioration significative ($p=0,006$) de la survie globale avec l'ajout de la CT à la RT comparé à la RT seule. L'hétérogénéité de l'effet traitement était significative (Q-test : $p=0,03$; I^2 d'Higgins = 50% ; Figure 5.1) et était expliquée par le timing de la CT. Pour tenir compte de cette hétérogénéité, nous avons recalculé le *hazard ratio* global $HR_{RT+CT \text{ vs } RT}$ avec un modèle à effets aléatoires de DerSimonian-Laird qui était toujours de 0,82 mais non significatif (IC 95% : [0,66 ; 1,02]; $p = 0,08$; le $HR_{RT \text{ vs } RT+CT}$ est donné dans la Figure 5.1). Nous avons également testé l'hypothèse globale de proportionnalité des risques à l'aide de la méthodologie décrite par Wei et al. qui consiste à agréger les p-valeurs du test de Grambsch-Therneau estimées pour chaque essai (Wei *et al.*, 2015). Cette hypothèse était vérifiée dans la méta-analyse MAC-NPC 1 à un seuil de significativité de 5 % ($p = 0,09$).

La méta-analyse MAC-NPC 2 correspondait à la version actualisée de MAC-NPC 1 et incluait de nouveaux essais mais également la mise à jour du suivi des essais déjà inclus dans MAC-NPC 1 pour un total de 5 028 patients inclus dans 23 comparaisons et un suivi médian de 8 ans, variant de 3 à 17 ans selon les essais (Blanchard *et al.*, 2015).

La survie globale était significativement améliorée avec l'ajout de la chimiothérapie à la radiothérapie avec un $HR_{RT+CT \text{ vs } RT}$ de 0,79 (IC 95 % [0,73 ; 0,86 ; $p<0,001$) estimée par un modèle à effets fixes. L'hétérogénéité de l'effet traitement était moins grande dans MAC-NPC2 que dans MAC-NPC 1 (Q-test : $p=0,09$; $I^2 = 30\%$; Figure 5.4). Nous avons ré-estimé le HR

global à l'aide d'un modèle à effets aléatoires de DerSimonian-Laird : $HR_{RT+CT \text{ vs } RT} = 0,79$ (IC 95 % : [0,70 ; 0,87] ; $p < 0,001$; le $HR_{RT \text{ vs } RT+CT}$ est donné dans la Figure 5.4). L'hypothèse des risques proportionnels était vérifiée globalement ($p = 0,16$).

5.1.2. Résultats de l'analyse secondaire avec le bénéfice de survie restreint

Pour l'estimation du $rmstD(t^*)$, deux temps de restriction, $t^* = 5$ ans et $t^* = 10$ ans, ont été retenus. Ces deux horizons étaient deux temps d'intérêt clinique dans les publications de MAC-NPC 1 et MAC-NPC 2 : en effet, Blanchard et al. ont par exemple comparé les taux de survie des deux bras de traitement à 5 et à 10 ans. Nous avons utilisé les quatre méthodes comparées dans l'étude de simulation à la fois avec un modèle à effets fixes et avec un modèle à effets aléatoires : Naïve Kaplan-Meier, Pooled Kaplan-Meier, Pooled Exponentiel, et Peto-quintile.

Pour la méta-analyse MAC-NPC 1, le $rmstD(t^*)$ variait entre les méthodes de 0,17 an à 0,23 an à $t^* = 5$ ans, et de 0,46 à 0,55 an à $t^* = 10$ ans (Tableau 5.1). Le $rmstD(t^*)$ n'était pas statistiquement différent de 0 avec les méthodes Pooled Kaplan-Meier et Pooled Exponentiel avec un modèle à effets aléatoires (p-valeurs grisées dans le Tableau 5.1). Puisqu'il y avait une hétérogénéité importante dans MAC-NPC 1 ($I^2 = 50\%$), un modèle à effets aléatoires de DerSimonian-Laird était le plus approprié pour agréger les $rmstD(t^*)$ estimés dans chaque essai. En effet, comme précédemment vu dans l'étude de simulation du Chapitre 4, un modèle à effets fixes sous-estime la variance du $rmstD(t^*)$ global. Comme pour l'étude de simulation, le $rmstD(t^*)$ et sa variance estimés à $t^* = 10$ ans étaient plus grands qu'à $t^* = 5$ ans (Tableau 5.1).

Pour la méta-analyse MAC-NPC 2, le $rmstD(t^*)$ variait entre les méthodes de 0,17 à 0,21 an à $t^* = 5$ ans et de 0,53 à 0,59 an à $t^* = 10$ ans (Tableau 5.1). Le $rmstD(t^*)$ était significativement différent de 0 et en faveur de la combinaison radiothérapie plus chimiothérapie (RT+CT) pour toutes les méthodes. En comparaison avec les résultats de la méta-analyse MAC-NPC 1, pour toutes les méthodes, l'erreur-type du $rmstD(t^*)$ était plus petite dans MAC-NPC 2, avec un $rmstD(t^*)$ similaire. Ces résultats étaient cohérents avec ceux de l'étude de simulation, puisqu'il y avait plus d'essais et un plus grand nombre global de patients dans MAC-NPC 2.

Dans la suite de cette section, nous avons retenu comme méthode de référence pour estimer le $rmstD(t^*)$ global la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires, qui présentait le meilleur compromis en termes de biais moyen et d'erreur-type dans notre étude de simulation.

Tableau 5.1 - Résultats des méthodes d'estimation du bénéfice de survie restreint dans les méta-analyses MAC-NPC 1 et MAC-NPC 2.

Modèle à deux étapes	Méthodes	MAC-NPC 1						MAC-NPC 2					
		$t^* = 5$ ans			$t^* = 10$ ans			$t^* = 5$ ans			$t^* = 10$ ans		
		<i>rmstD</i>	ET	p-valeur	<i>rmstD</i>	ET	p-valeur	<i>rmstD</i>	ET	p-valeur	<i>rmstD</i>	ET	p-valeur
Effets aléatoires	Naïve Kaplan-Meier	0,20	0,08	0,008	0,51	0,19	0,006	0,17	0,04	<0,001	0,54	0,11	<0,001
	Pooled Kaplan-Meier	0,17	0,11	0,106	0,49	0,28	0,081	0,20	0,05	<0,001	0,59	0,13	<0,001
	Pooled Exponentiel	0,17	0,09	0,076	0,51	0,29	0,078	0,17	0,03	<0,001	0,55	0,11	<0,001
	Peto-quintile	0,23	0,09	0,007	0,55	0,22	0,011	0,21	0,04	<0,001	0,59	0,12	<0,001
Effets fixes	Pooled Kaplan-Meier	0,20	0,07	0,005	0,52	0,18	0,004	0,18	0,04	<0,001	0,59	0,10	<0,001
	Pooled Exponentiel	0,18	0,06	0,003	0,55	0,18	0,002	0,17	0,03	<0,001	0,56	0,09	<0,001
	Peto-quintile	0,20	0,07	0,006	0,46	0,16	0,004	0,18	0,04	<0,001	0,53	0,09	<0,001

MAC-NPC: Meta-Analysis of Chemotherapy in Nasopharynx Carcinoma; *rmstD*: bénéfice de survie restreint; ET: erreur-type; t^* : horizon temporel.

Les Figure 5.1 et Figure 5.2 représentent les forests plots respectivement des *hazard ratios* et des bénéfices de survie restreints à $t^* = 10$ ans estimés pour chaque essai et globalement par un modèle à effets aléatoires de DerSimonian-Laird (DerSimonian & Laird, 1986) pour la méta-analyse MAC-NPC 1. Les HR donnés dans la Figure 5.1 ont été recalculés par rapport à ceux donnés dans le paragraphe 5.1.1 en prenant cette fois-ci comme référence la combinaison RT+CT. Ainsi, un $HR > 1$ ou un $rmstD(t^*) > 0$ illustre la supériorité du bras RT+CT par rapport au bras RT.

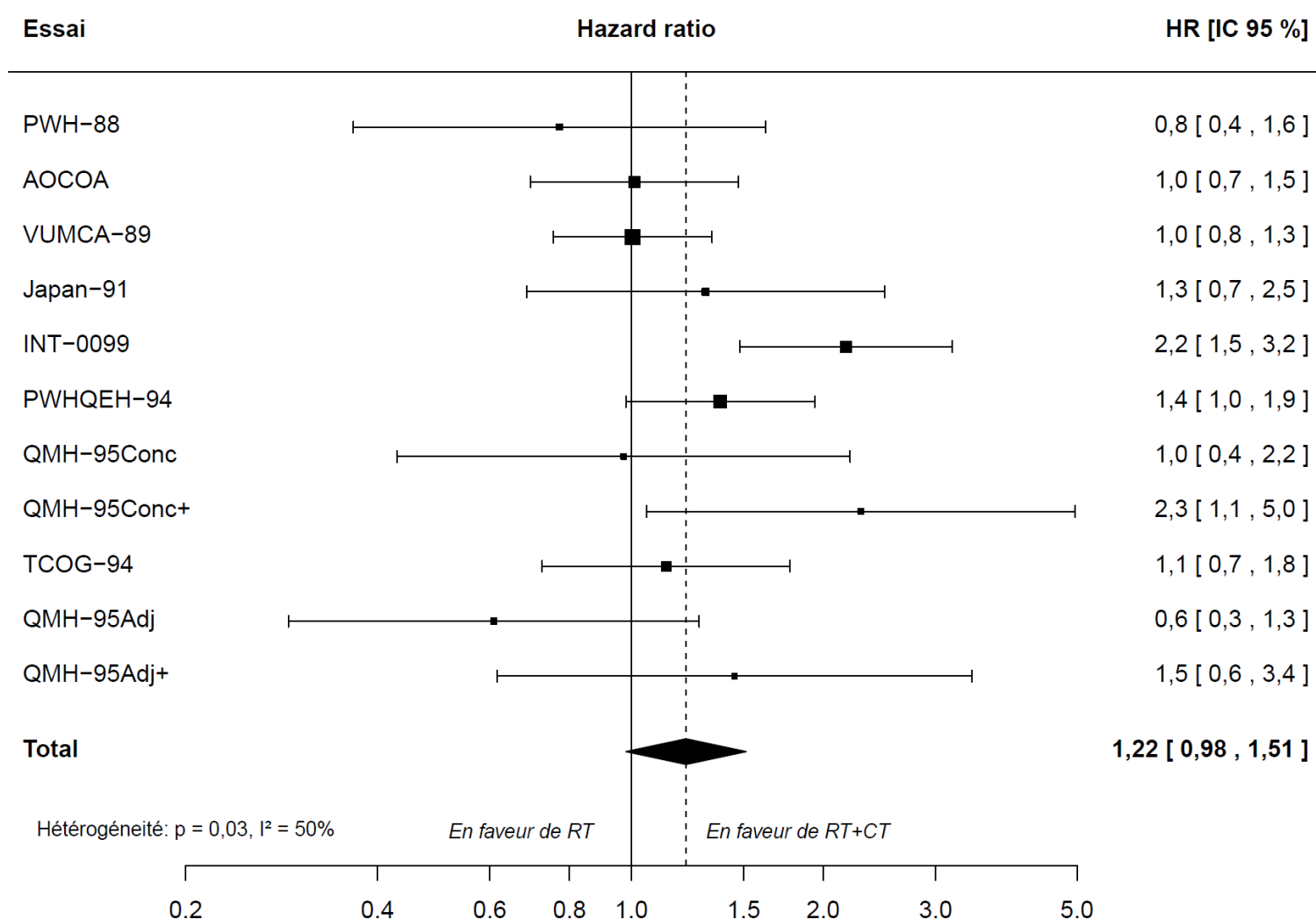


Figure 5.1 - Forest-plot des *hazard ratios* (RT contre RT+CT) avec un modèle à effets aléatoires pour la méta-analyse MAC-NPC 1.

Légende des Figure 5.1 et Figure 5.2:

Adj: RT vs. RT + CT adjuvante; Adj+: RT + CT concomitante vs. RT + CT concomitante + CT adjuvante; AOCOA: Asian-Oceanian Clinical Oncology Association; Conc: RT vs. RT + CT concomitante; Conc+: RT + CT adjuvante vs. RT + CT adjuvante+ CT concomitante; CT: Chimiothérapie; HR: *Hazard ratio* du risque de décès du bras RT vs. RT+CT; IC 95 %: Intervalle de confiance à 95 %; INT: Intergroup study; PWHQMH: Prince of Wales Hospital, Queen Mary Hospital; QMH: Queen Mary Hospital; rmstD: bénéfice de survie restreint; RT: Radiothérapie; TCOG: Taiwan Cooperative Oncology Group; VUMCA: International Nasopharynx Cancer Study Group;

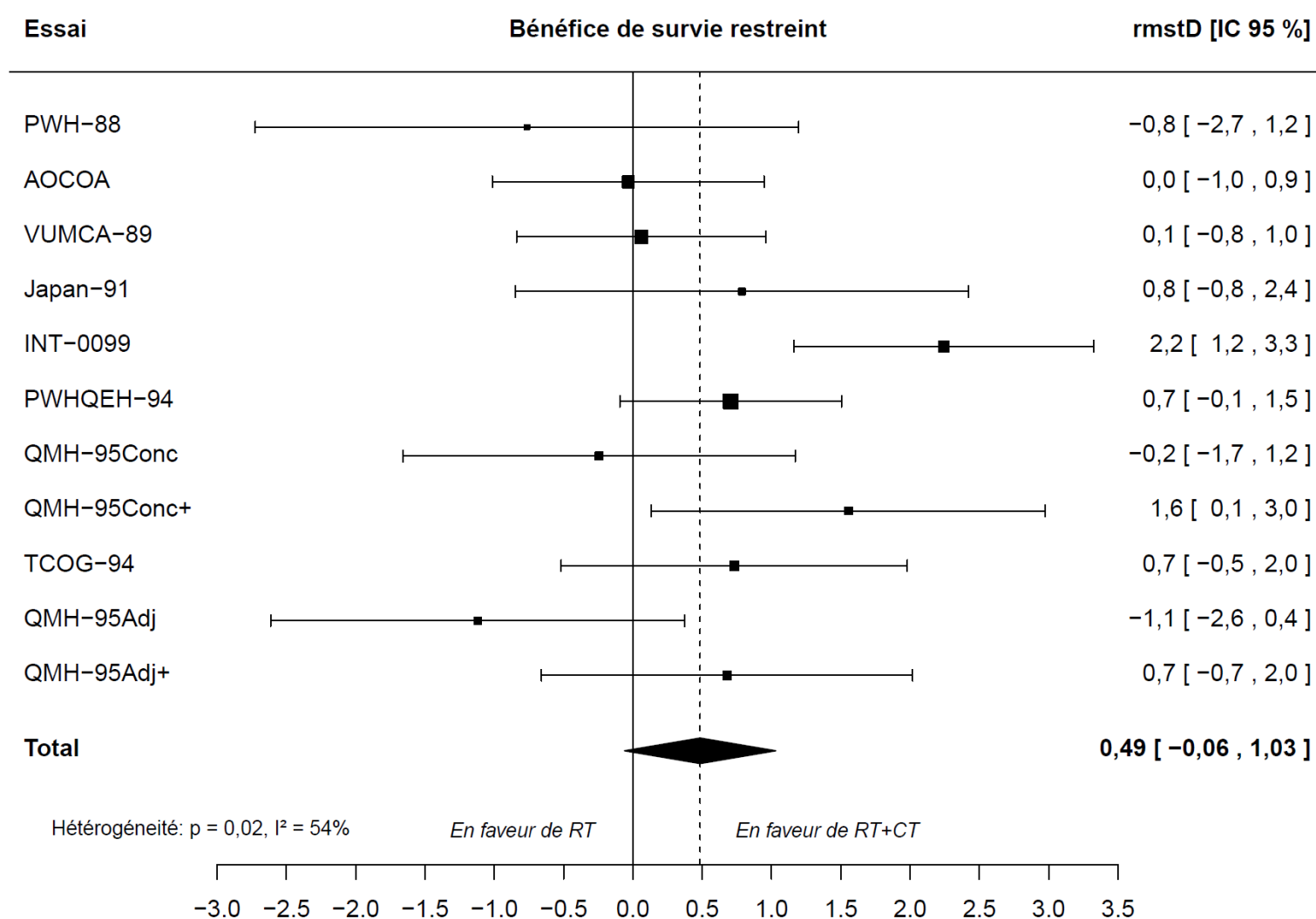


Figure 5.2 - Forest-plot des bénéfices de survie restreints à $t^* = 10$ ans, exprimés en année et estimés avec la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires pour la méta-analyse MAC-NPC 1.

Dans MAC-NPC 1, nous pouvons remarquer que pour chaque essai, la significativité et le sens de l'effet traitement étaient similaires que l'effet traitement fût mesuré à l'aide du HR (Figure 5.1) ou du $rmstD(t^* = 10)$ (Figure 5.2). De même, avec ces deux mesures l'effet traitement global indiquait une diminution du risque de décès ($HR_{RT \text{ vs } RT+CT} = 1,22 = 1/0,82$; IC 95 % : $[0,98 ; 1,51]$) ou une augmentation de la survie moyenne restreinte à 10 ans ($rmstD(t^* = 10) = 0,49$; IC 95 % : $[-0,06 ; 1,03]$), en faveur du bras RT+CT. Dans les deux cas l'effet du traitement global était non significatif à un seuil de 5 % avec une p-valeur qui valait 0,08 pour à la fois le HR et le $rmstD(t^*)$.

Afin de s'affranchir du fait que le $rmstD(t^*)$ s'interprète relativement à l'horizon t^* , certains auteurs ont proposé l'utilisation du bénéfice de survie restreint relatif, noté $rmstRD(t^*)$ (pour *relative difference in restricted mean survival time*) (Royston & Parmar, 2011; Zhao *et al.*, 2012; Wei *et al.*, 2015). Le $rmstRD(t^*)$ est présenté en termes relatifs comme un pourcentage de t^* et est défini par :

$$rmstRD(t^*) = rmstD(t^*)/t^* \quad (5.1)$$

Pour un effet traitement bénéfique ($rmstD(t^*) > 0$), le $rmstRD(t^*)$ varie entre 0 et 1 et s'interprète comme un pourcentage égal à l'espérance de vie gagnée grâce au traitement expérimental par rapport à l'espérance de vie maximale possible. Sa variance est estimée par $Var[rmstD(t^*)]/(t^*)^2$.

La Figure 5.3 représente le bénéfice de survie restreint global (panel A) et le bénéfice de survie restreint relatif (panel B) estimés dans MAC-NPC 1 par la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires de DerSimonian-Laird en faisant varier l'horizon temporel t^* . La courbe noire représente le bénéfice de survie restreint (relatif) comme fonction de l'horizon temporel t^* et la droite horizontale en pointillée indique l'absence d'effet traitement ($rmstD(t^*) = 0$ ou $rmstRD(t^*) = 0$). Les droites verticales rouges correspondent aux horizons temporels 5 et 10 ans et la zone grisée correspond à la région de confiance à 95% point par point.

La Figure 5.3-A montre que le bénéfice de survie restreint augmentait lorsque l'horizon temporel t^* augmentait, indiquant un prolongement de l'effet bénéfique du traitement expérimental RT+CT même après 5 ans. La Figure 5.3-B montre que le bénéfice de survie restreint relatif augmentait les 10 premières années et semblait se stabiliser autour de 5 % après 10 ans. A noter que ni le $rmstD(t^*)$, et donc par définition ni le $rmstRD(t^*)$, n'étaient significativement différents de 0 pour $t^* \in [0-10]$ ans.

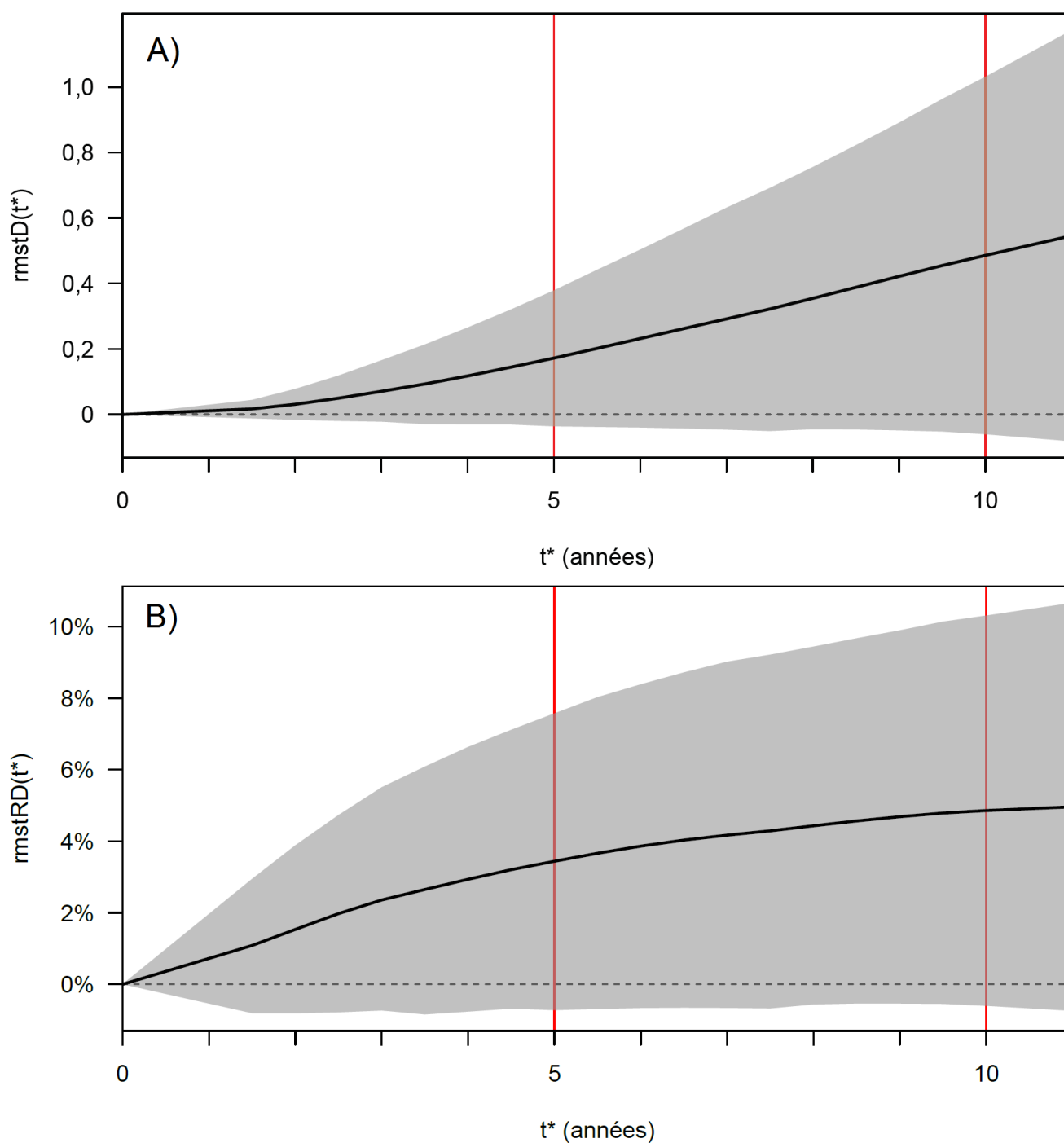


Figure 5.3 - Bénéfice de survie restreint (panel A) et bénéfice de survie restreint relatif (panel B) exprimés comme fonction de l'horizon temporel t^* dans la méta-analyse MAC-NPC 1.

Légende de la Figure 5.3 :

MAC-NPC: Meta-Analysis of Chemotherapy in Nasopharynx Carcinoma; $rmstD$: bénéfice de survie restreint ; $rmstRD$: bénéfice de survie restreint relatif

Le $rmstD(t^*)$, exprimés en années, et le $rmstRD(t^*)$ sont estimés à l'aide de la méthode Pooled Kaplan-Meier avec effets aléatoires.

Nous avons également comparé les estimations de l'effet traitement obtenues avec le *hazard ratio* ou le bénéfice de survie restreint à $t^* = 10$ ans pour la méta-analyse actualisée MAC-NPC 2. Les Figure 5.4 et Figure 5.5 représentent les forests plots respectivement des HR et des $rmstD(t^* = 10)$ estimés pour chaque essai et globalement par un modèle à effets aléatoires de DerSimonian-Laird (DerSimonian & Laird, 1986) pour MAC-NPC 2. Similairement, à la méta-analyse MAC-NPC 1, les *hazard ratios* donnés dans la Figure 5.4 ont été recalculés par rapport à ceux donnés dans le paragraphe 5.1.1 en prenant cette fois-ci comme référence la combinaison RT+CT. Ainsi, un $HR > 1$ ou un $rmstD(t^*) > 0$ illustre la supériorité du bras RT+CT par rapport au bras RT.

Pour la plupart des essais de MAC-NPC 2, la significativité et le sens de l'effet traitement étaient similaires que l'effet traitement fût mesuré à l'aide du HR (Figure 5.4) ou du $rmstD(t^* = 10)$ (Figure 5.5). Il y avait cependant trois essais qui faisaient exception :

- l'essai PWHQEH-94 pour lequel le HR n'était pas significatif ($HR = 1,2$; IC 95 % : [0,9 ; 1,6]) mais le $rmstD(t^* = 10)$ l'était ($rmstD(t^* = 10) = 0,8$ an [0,0 ; 1,5])
- les essais NPC-9901 et Guangzhou 2002-01 pour lesquels le HR était significatif ($HR = 1,4$ [1,0 ; 1,9] et $HR = 1,5$ [1,0 ; 2,1]) mais pas le $rmstD(t^* = 10)$ ($rmstD(t^* = 10) = 0,5$ an [-0,2 ; 1,3] et $rmstD(t^* = 10) = 0,6$ an [-0,3 ; 1,5]).

Avec ces deux mesures l'effet traitement global reflétait une diminution du risque de décès ($HR_{RT \text{ vs } RT+CT} = 1,26 = 1/0,79$; IC 95 % : [1,13 ; 1,41]) ou une augmentation de la survie moyenne restreinte à 10 ans ($rmstD(t^* = 10) = 0,59$ an ; IC 95 % : [0,34 ; 0,84]), en faveur du bras RT+CT. Dans les deux cas, l'effet du traitement global était significatif à un seuil de 5 % avec une p-valeur qui était inférieure à 0,001 pour à la fois le HR et le $rmstD(t^*)$.

Légende des Figure 5.4 et Figure 5.5 :

Adj: RT vs. RT + CT adjuvante; Adj+: RT + CT concomitante vs. RT + CT concomitante + CT adjuvante; AF: Fractionnement accéléré; AOCOA: Asian-Oceanian Clinical Oncology Association; CF: Fractionnement conventionnel; Conc: RT vs. RT + CT concomitante; Conc+: RT + CT adjuvante vs. RT + CT adjuvante+ CT concomitante; CT: Chimiothérapie; HeCOG: Hellenic Cooperative Oncology Group; HR: *Hazard ratio* du risque de décès du bras RT vs. RT+CT; IC 95 %: Intervalle de confiance à 95 % ; INT: Intergroup study; NPC: Nasopharyngeal Carcinoma; PWHQMH: Prince of Wales Hospital, Queen Mary Hospital; QMH: Queen Mary Hospital; rmstD: bénéfice de survie restreint exprimé en années; RT: Radiothérapie; SQNP: Singapore Naso-Pharynx; TCOG: Taiwan Cooperative Oncology Group; VUMCA: International Nasopharynx Cancer Study Group;

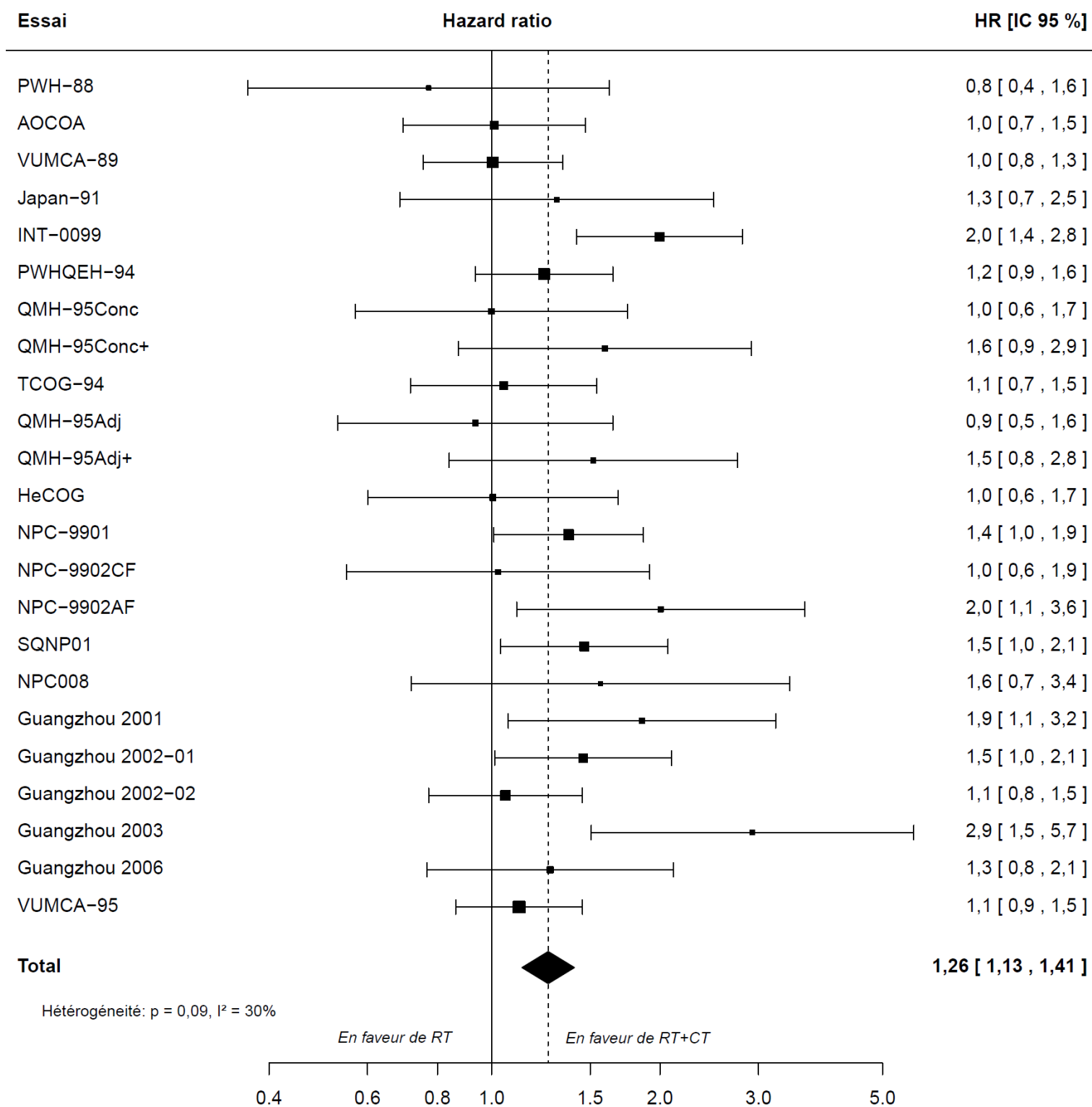


Figure 5.4 - Forest-plot des *hazard ratios* (RT contre RT+CT) avec un modèle à effets aléatoires pour la méta-analyse MAC-NPC 2.

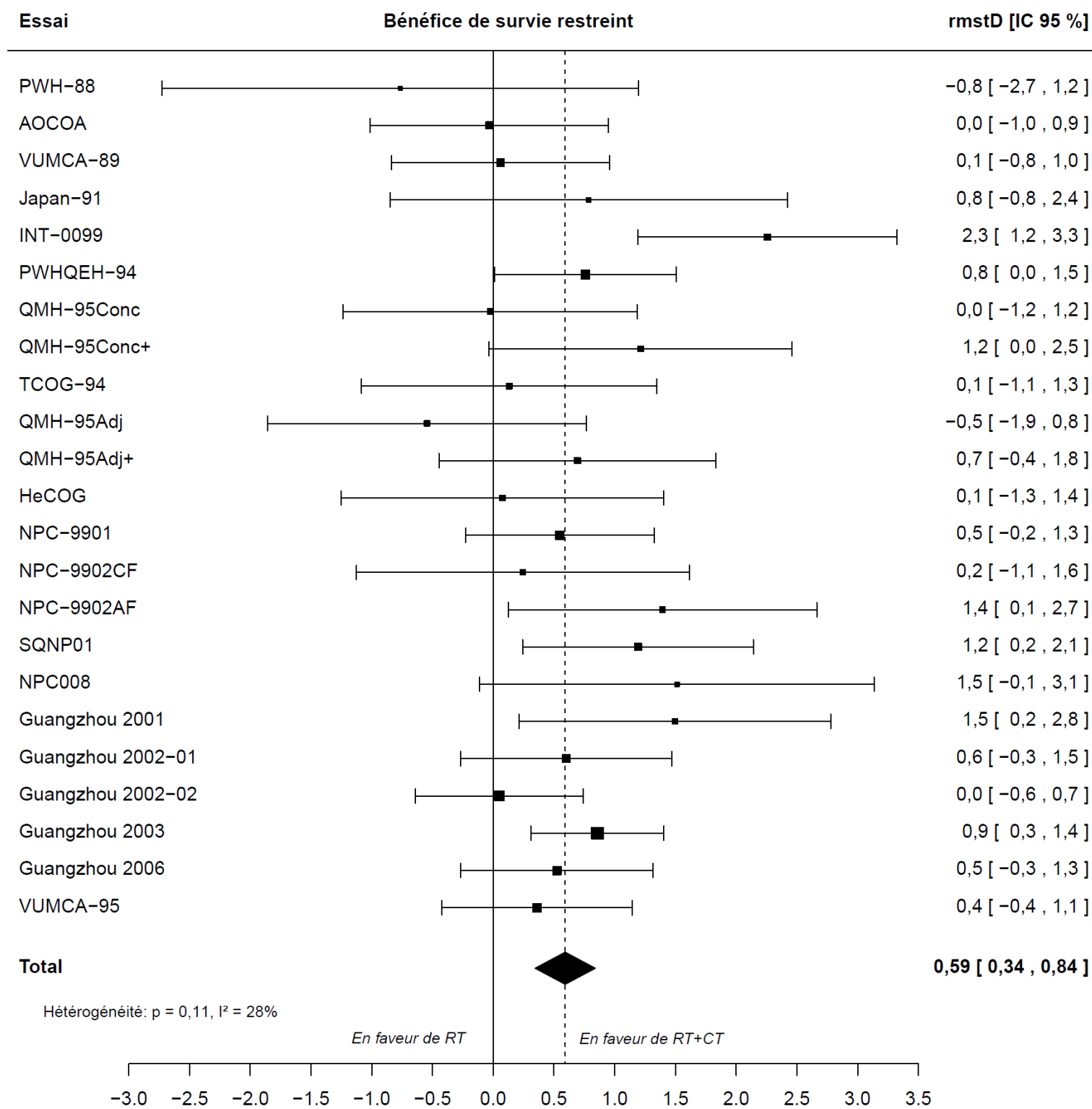


Figure 5.5 - Forest-plot des bénéfices de survie restreints à $t^* = 10$ ans, exprimés en année et estimés avec la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires pour la méta-analyse MAC-NPC 2.

5.1.3. Conclusion des méta-analyses MAC-NPC 1 et MAC-NPC 2

Nous avons estimé le bénéfice de survie restreint à 5 et à 10 ans dans deux méta-analyses sur données individuelles (MADI) et avons comparé les estimations de l'effet traitement à l'aide du *hazard ratio* et du $rmstD(t^*)$. Ces deux mesures donnaient des estimations très proches pour tous les essais et pour l'estimation de l'effet traitement global, ce qui s'explique principalement par le fait que l'hypothèse des risques proportionnels était vérifiée dans chacun des essais de MAC-NPC 1 et de MAC-NPC 2.

Le $rmstD(t^*)$ global estimé par la méthode de Pooled Kaplan-Meier avec un modèle à effets aléatoires était de 0,49 an (IC 95 % : [-0,06 ; 1,03], $p=0,08$) dans MAC-NPC 1 et de 0,59 an (IC 95 % : [0,34 ; 0,84], $p<0,001$) dans la version actualisée MAC-NPC 2. En d'autres termes, dans MAC-NPC 2, l'ajout de la chimiothérapie à la radiothérapie prolonge la survie moyenne restreinte à 10 ans de 7,1 mois (IC 95% : [4,1 ; 10,1]). Nous pensons que l'interprétation clinique du $rmstD(t^*)$ est plus intuitive que celle du *hazard ratio* global égal à 0,79 (IC 95 % : [0,70 ; 0,87]) ou que celle dérivée de la différence des taux de survie à 5 ans égale à 6,3% en faveur du bras avec chimiothérapie (IC 95 % : [3,5 ; 9,1]) (Blanchard *et al.*, 2015).

Il est à noter que dans ces deux MADI, l'hétérogénéité entre les $rmstD_j(t^*)$ estimés dans chaque essai était semblable à celle entre les HR_j . Pour MAC-NPC 1, l'hétérogénéité était similaire avec la p-valeur du test Q de Cochran égale à 0,03 et l' I^2 d'Higgins égal à 50 % pour le HR (Figure 5.1), comparés à $p = 0,02$ et $I^2 = 54$ % pour le $rmstD(t^* = 10)$ (Figure 5.2). Les valeurs d'hétérogénéité étaient également semblables entre les deux mesures dans MAC-NPC 2, avec $p = 0,09$ et $I^2 = 30$ % pour le HR (Figure 5.4) contre $p = 0,11$, $I^2 = 28$ % pour le $rmstD(t^* = 10)$ (Figure 5.5).

Dans la section suivante, nous présentons les résultats d'une MADI que nous avons réalisé à Gustave Roussy et pour laquelle l'analyse de la survie globale, qui était le critère de jugement principal, a été faite avec à la fois le *hazard ratio* (mesure principale) et le bénéfice de survie restreint (mesure secondaire).

5.2. RadioTherapy Timing in Small-Cell Lung Cancer (RTT-SCLC)

La méta-analyse sur données individuelles RTT-SCLC (RadioTherapy Timing in Small-Cell Lung Cancer) s'intéressait au timing de la radiothérapie (RT) thoracique dans le traitement du cancer bronchique à petites cellules non métastatique. En effet, l'association de la RT à une chimiothérapie (CT) est devenue le traitement standard dans cette pathologie (Pignon *et al.*, 1992; Van Meerbeeck *et al.*, 2011), cependant le timing optimal de la RT par rapport à la CT est toujours débattu. Plusieurs méta-analyses sur données publiées ont étudié cette question, mais aucune n'a pu conclure en faveur soit d'une RT précoce, initiée dans les 8 premières semaines après le début de la chimiothérapie, ou d'une RT tardive (Fried *et al.*, 2004; Spiro *et al.*, 2006; Pijls-Johannesma *et al.*, 2007). Une méta-analyse a cependant suggéré que la RT précoce serait plus efficace lorsque tous les cycles de CT prévus par le protocole étaient administrés (Spiro *et al.*, 2006), ce qui impliquerait que la question du timing optimal de la RT ne pourrait être abordée qu'en étudiant également la compliance à la CT pour chaque patient. De telles informations ne pouvant être connues qu'à l'aide des données individuelles des patients, nous avons donc réalisé une méta-analyse sur données individuelles (MADI) avec comme objectif de définir la meilleure approche pour combiner la RT thoracique à la CT. L'originalité de cette MADI est que la survie globale, qui était le critère de jugement principal, a été analysée à la fois avec le *hazard ratio* (mesure principale) et avec le bénéfice de survie restreint (mesure secondaire).

5.2.1. Sélection des essais et stratégie de recherche

Afin d'être éligibles, les essais devaient comparer deux timings de RT thoracique, précoce ou tardive, chez des patients atteints d'un cancer bronchique à petites cellules non métastatique et traités par chimio-radiothérapie. La RT était définie comme précoce si elle était initiée dans les 8 premières semaines après la randomisation, et comme tardive si initiée lors de la 9^{ème} semaine ou au-delà. Les essais comparant deux durées de traitement par RT, i.e. une RT courte contre une RT longue, avec au moins une différence d'au moins 2 semaines entre les bras étaient également éligibles. Nous utiliserons le terme de RT expérimentale pour les bras avec une RT précoce ou courte, et le terme de RT contrôle pour les bras avec une RT tardive ou longue.

Les essais devaient également avoir commencé après 1969, s'être terminés avant décembre 2006, et être correctement randomisés. Le protocole de CT prévu (drogue, dose, nombre de cycle) devait être le même dans les deux bras, mais les modalités de RT pouvaient varier. La dose totale de RT devait être d'au moins 30 Gray (Gy). La radiothérapie par

orthovoltage était un critère d'exclusion. Les patients éligibles devaient avoir un score d'état général défini par l'OMS (Organisation Mondiale de la Santé) ou un autre score équivalent, compris entre 0 et 1, et devaient ne pas avoir reçu de traitement préalable pour leur cancer.

Afin de limiter le biais de publication, nous avons recherché à la fois les essais publiés et non publiés, sans restriction de la langue, à l'aide des sources suivantes :

- des bases électroniques : PubMed Medline et le registre centralisé d'essais contrôlés de la Cochrane. La stratégie de recherche utilisée dans Medline est disponible en Annexes (*Web-Appendix 1* du manuscrit de RTT-SCLC, Figure A.1),
- des registres d'essais : UKCCCR Trials Register, ClinicalTrials.gov, Physicians Data Query Protocols et le Current Controlled Trials 'metaRegister' of controlled trials,
- une recherche manuelle dans les proceedings de conférences (ASCO, WLCC, ESMO, ECCO) et la bibliographie des essais et revues d'essai déjà identifiés,
- les investigateurs des essais inclus dans la méta-analyse RTT-SCLC à qui nous avons demandé de compléter la liste provisoire des essais.

Une première recherche avait été réalisée sur PubMed MEDLINE le 11 janvier 2007 et avait abouti à 905 résultats, auxquelles s'ajoutent 138 références provenant d'autres sources décrites au-dessus. Le flow-chart de la Figure 5.6 décrit ces 1 043 références. Une recherche actualisée a été réalisée le 21 février 2014, conduisant à 1 285 résultats sur Pubmed (*Web-Appendix 1* du manuscrit de RTT-SCLC, Figure A.1). Un seul nouvel essai a été considéré comme éligible à notre méta-analyse et a été inséré dans une analyse de sensibilité avec le HR extrait de la publication (Sun *et al.*, 2013).

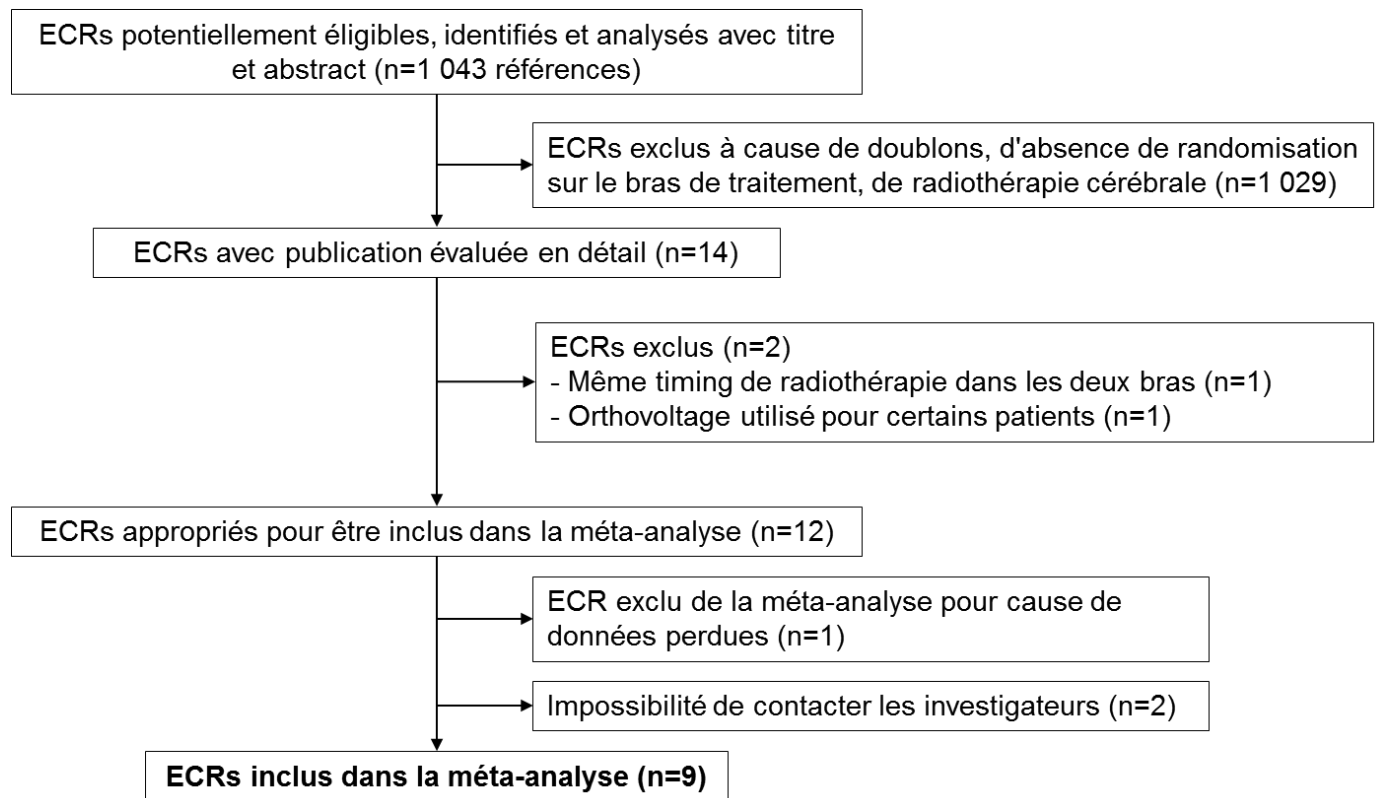


Figure 5.6 - Flow-chart des essais inclus dans la méta-analyse RTT-SCLC

Légende de la Figure 5.6 :

ECR: Essai contrôlé randomisé ; RTT-SCLC: RadioTherapy Timing in Small-Cell Lung Cancer.

5.2.2. Résultats : analyse princeps avec le hazard ratio et le bénéfice de survie restreint

Douze essais incluant 2 688 patients étaient éligibles pour notre méta-analyse RTT-SCLC. Nous n'avons pas pu récupérer les données de trois essais : l'un pour cause de données perdues et les deux autres par faute de contact avec les investigateurs (Figure 5.6). Les caractéristiques de ces trois essais sont décrites en Annexes du manuscrit (Tableau S2, Figure A.1).

Finalement, nous avons récupéré les données individuelles concernant 9 essais et 2 305 patients, correspondant à 86% du nombre total de patients éligibles : ces 9 essais sont décrits dans le Tableau 5.2. Le suivi médian était de 10 ans, variant de 5 à 17 ans selon les essais. La distribution des caractéristiques des patients entre les deux bras de RT était comparable.

Tableau 5.2 - Description des essais inclus dans la méta-analyse RT^T-SCLC

Essais	Période d'inclusion	Début de la radiothérapie	Dose RT / Nb fraction / Durée totale	Nb cycles CT	Nb de patients randomisés	Suivi médian (années)
CALGB8083	1981-84	Exp : Jour 1 Cont: Jour 64	50 Gy / 24 fr / 5 s	26 cycles environ Exp : 2 cycles pendant RT, jusqu'à 24 cycles après RT Cont : 3 cycles avant RT, 2 cycles pendant RT, jusqu'à 21 cycles après RT	292	17,2
BR.6	1985-88	Exp : Jour 22 Cont: Jour 106	40 Gy / 15 fr / 3 s	Exp : 6 cycles (1 avant RT, 1 pendant RT, 4 après RT) Cont : 6 cycles (5 avant RT, 1 pendant RT)	332	11,2
CCCWFU62286	1987-92	Exp* : Jour 1 Cont* : Jour 8	Exp* : 50 Gy / 25 fr / 5 s Cont* : 20 Gy / 8 fr / 2 s + pause 1 s + 20 Gy / 8 fr / 2 s + pause 1 s + 10 Gy / 4 fr / 1 s	Exp* : 6 cycles (2 pendant RT, 4 après RT) Cont* : 6 cycles (3 en alternance avec RT, 3 après RT)	114	17,3
03PCL88	1988-94	Exp* : Jour 30 Cont* : Jour 36	Exp* : 50 Gy / 20 fr / 5 s Cont* : 20 Gy / 8 fr / 2 s + pause 2s + 20 Gy / 8 fr / 2 s + pause 2s + 15 Gy / 6 fr / 1,5 s	Exp* : 6 cycles (2 avant RT, 1 pendant RT, 3 après RT) Cont* : 6 cycles (2 avant RT, 2 en alternance avec RT, 2 après RT)	164	6,5
ECOG3588	1989-92	Deux bras : Jour 1	Exp* : 45 Gy / 30 fr / 3 s bid Cont* : 45 Gy / 25 fr / 5 s	4 cycles (2 pendant RT, 2 après RT)	417	13,0
EORTC08877	1989-95	Exp : Jour 43 Cont: Jour 99	Exp : 12,5 Gy / 5 fr / 1 s + pause 3s + 12,5 Gy / 5 fr / 1 s + pause 3s + 12,5 Gy / 5 fr / 1 s + pause 3s + 12,5 Gy / 5 fr / 1 s Cont: 50 Gy / 20 fr / 4 s	Exp : 5 cycles (2 avant RT, 3 en alternance avec RT) Cont : 5 cycles (5 avant RT)	349	7,2
JCOG9104	1991-95	Exp : Jour 2 Cont: Jour 85	45 Gy / 30 fr / 3 s bid	Exp : 4 cycles (1 pendant RT, 3 après RT) Cont : 4 cycles (4 avant RT)	231	6,8
LLCG93	1993-99	Exp : Jour 22 Cont: Jour 106	40 Gy / 15 fr / 3 s	Exp : 6 cycles (1 avant RT, 1 pendant RT, 4 après RT) Cont : 6 cycles (5 avant RT, 1 pendant RT)	325	5,3
HeCOG93	1993-99	Exp : Jour 1 Cont: Jour 57	45 Gy / 30 fr / 3 s bid	Exp : 6 cycles (1 pendant RT, 5 après RT) Cont : 6 cycles (3 avant RT, 1 pendant RT, 2 après RT)	81	11,8
Total	1981-1999	NA	NA	NA	2 305	10,4

Légende du Tableau 5.2 :

Les essais sont classés chronologiquement selon leur période d'inclusion.

bid: Radiothérapie donnée deux fois par jour; BR: Bronchus; CALGB: Cancer and Leukaemia Group B; CCCWFU: Comprehensive Cancer Centre of Wake Forest University; Cont: RT contrôle (RT tardive); Cont*: RT contrôle (RT longue); CT: chimiothérapie; ECOG: Eastern Cooperative Oncology Group; EORTC: European Organisation for Research and Treatment of Cancer; Exp: RT expérimentale (RT précoce); Exp*: RT expérimentale (RT courte); fr: fraction; Gy: Gray; HeCOG: Hellenic Cooperative Oncology Group; JCOG: Japan Clinical Oncology Group; LLCG: London Lung Cancer Group; PCL: Petites Cellules Limitées; RT: Radiothérapie; s: semaine.

Plusieurs analyses de sous-groupes en fonction de caractéristiques d'essais étaient prévues dans le protocole : la compliance à la CT entre les deux bras (similaire/différente), la dose de RT par fraction ($< 1,8 \text{ Gy}$ / $1,8 - 2,4 \text{ Gy}$ / $> 2,4 \text{ Gy}$), le type de RT (hyper-fractionnée/standard), la durée totale de RT (> 30 jours dans les 2 bras / un bras ≤ 30 jours et un > 30 jours / ≤ 30 jours dans les 2 bras), la présence de CT à base de platine pendant la RT dans les deux bras (Oui/Non), la présence de CT concurrente à la RT dans les deux bras (Oui/Non), et si la RT était administrée de la même façon dans les deux bras (Oui/Non). La répartition des essais selon ces sous-groupes est donnée dans le Tableau 5.3.

Dans cette étude, un patient était défini comme compliant à la chimiothérapie s'il/elle avait reçu 100 % ou plus du nombre de cycles de CT prévus par le protocole, excepté pour l'essai CALGB8083 pour lequel les patients ayant reçu 6 cycles ou plus étaient considérés compliants. Un essai était considéré avec une compliance différente entre les 2 bras RT expérimentale et RT contrôle, si la différence entre les 2 bras de la proportion de patients compliants à la CT était $\geq 10\%$ et avec une compliance similaire entre les 2 bras si la différence était $< 10\%$ (Spiro *et al.*, 2006). Nous n'avions pas de données individuelles ni de données publiées sur la compliance à la CT disponibles pour l'essai CCWFU62286, cet essai était donc exclu de l'analyse en sous-groupe sur la compliance à la CT. Trois essais (BR6, ECOG3588, JCOG9104) étaient classés comme ayant une compliance à la CT similaire entre les bras, et avaient une proportion au moins égale à 79 % de patients compliants à la CT (Tableau 5.3). Cinq essais (CALGB8083, 03PCL88, EORTC08877, LLCG93, HeCOG93) étaient classés comme ayant une compliance différente entre les bras; pour chacun de ces essais la proportion de patients compliants à la CT était plus faible dans le bras RT expérimentale.

Tableau 5.3 - Répartition des essais inclus dans la méta-analyse RTT-SCLC selon les sous-groupes étudiés

Trial	Compliance à la CT entre les 2 bras ^s	Dose RT par fraction	Type de RT	Durée totale de la RT	CT à base de platine pendant la RT dans les 2 bras	CT concurrente dans les 2 bras	Même RT dans les 2 bras
CALGB8083	Différente	Cont: 73 Exp: 52	1,8 – 2,4 Gy	Standard	> 30 jours dans les deux bras	Non	Oui
BR6	Similaire	Cont: 80 Exp: 79	> 2,4 Gy	Standard	≤ 30 jours dans les deux bras	Oui	Oui
CCWFU62286	ND		1,8 – 2,4 Gy	Standard	> 30 jours dans les deux bras	No	Non
03PCL88	Différente	Cont: 85 Exp: 75	> 2,4 Gy	Standard	> 30 jours dans les deux bras	Non	Non
ECOG3588	Similaire	Cont: 91 Exp: 85	< 1,8 Gy	Hyperfr	Un bras ≤ 30 jours, un > 30 jours	Oui	Oui
EORTC08877	Différente	Cont: 90 Exp: 80	> 2,4 Gy	Standard	Un bras ≤ 30 jours, un > 30 jours	Non	Non
JCOG9104	Similaire	Cont: 85 Exp: 87	< 1,8 Gy	Hyperfr	≤ 30 jours dans les deux bras	Non	Oui
LLCG93	Différente	Cont: 80 Exp: 69	> 2,4 Gy	Standard	≤ 30 jours dans les deux bras	Oui	Oui
HeCOG93	Différente	Cont: 95 Exp: 83	< 1,8 Gy	Hyperfr	≤ 30 jours dans les deux bras	Oui	Oui

[§]: Proportion de patients conformes à la CT: si la différence des 2 bras ≥10% alors l'essai est considéré avec une conformité « Différente »

CT: Chimiothérapie; Gy: Gray; Hyperfr: Hyper-fractionnée; ND: Non Disponible; RT: Radiothérapie.

L'hypothèse de proportionnalité des risques était globalement vérifiée dans la méta-analyse à un seuil de significativité de 5 % ($p = 0,16$) (Wei *et al.*, 2015). La Figure 5.7 décrit les HRs des 9 essais inclus dans RTT-SCLC, et le HR global estimé par un modèle à effets fixes : la RT expérimentale n'avait pas d'impact significatif sur la survie globale ($HR_{RT\text{ Exp vs RT Cont}} = 0,99$, IC 95 % : $[0,91; 1,08]$, $p=0,78$). L'hétérogénéité de l'effet traitement entre les HRs estimés dans chaque essai était significative (Q-test : $p = 0,006$, $I^2 = 63\%$) et nous verrons dans le paragraphe suivant, à travers des analyses de sous-groupes, que cette hétérogénéité pouvait être expliquée par différentes caractéristiques d'essais.

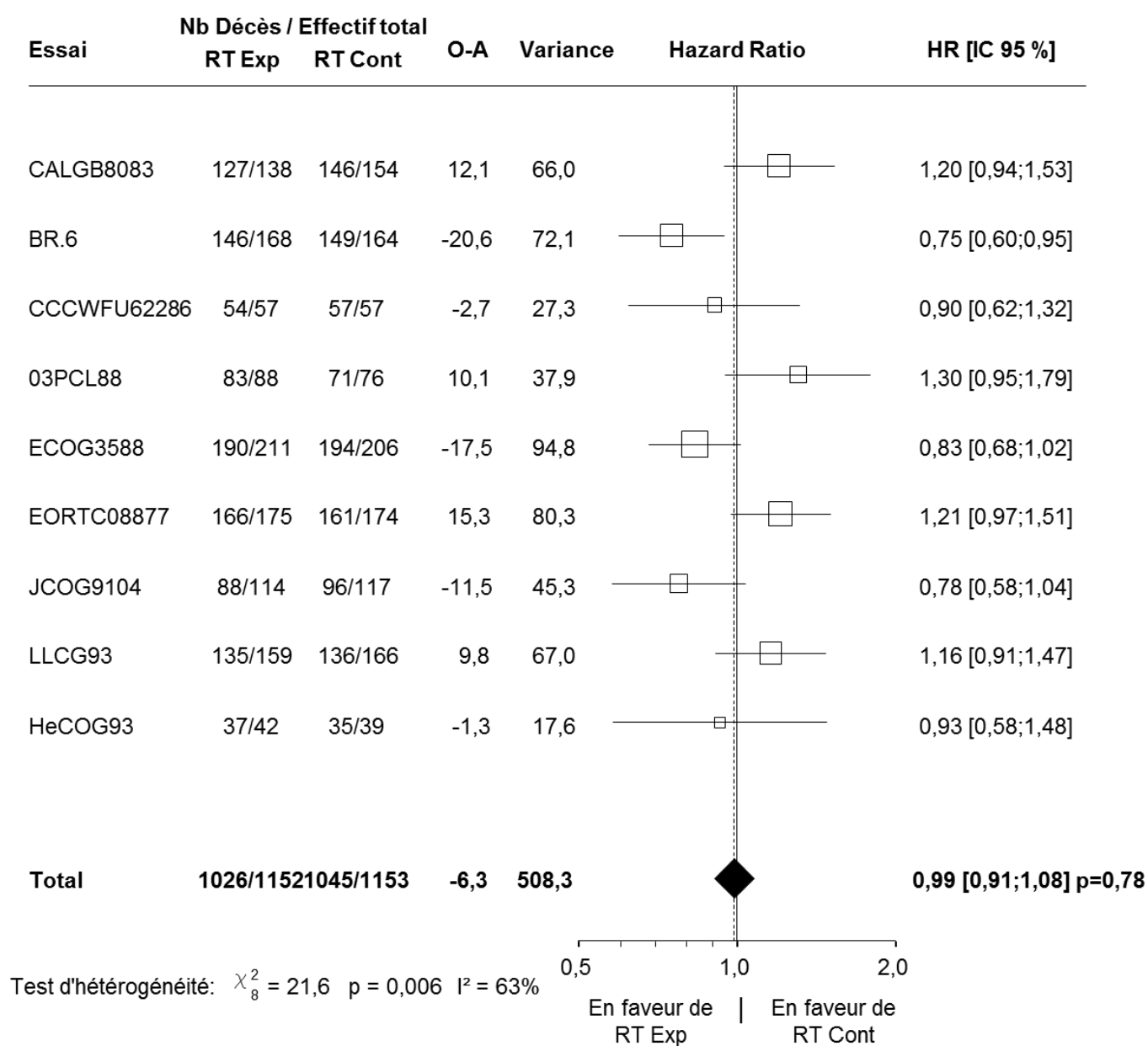


Figure 5.7 - Forest-plot des *hazard ratios* pour la méta-analyse RTT-SCLC.

Le HR global estimé par un modèle à effets aléatoires était également non significatif ($HR_{RT\text{ Exp vs RT Cont}} = 0,99$, IC 95 % : [0,85 ; 1,15], $p = 0,90$). Le bénéfice de survie restreint global à $t^* = 5$ ans et estimé avec la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires était également non significatif et égal à -0,1 mois (IC 95 % : [-2,7 ; 2,5], $p = 0,95$).

Le Tableau 5.4 synthétise les résultats des analyses de sous-groupes en fonction des caractéristiques d'essais et montre les HRs pour la survie globale selon les sous-groupes d'essais. Pour chacun des sous-groupes, l'hétérogénéité globale peut être divisée en hétérogénéités inter et intra sous-groupes, comme définies dans l'équation (2.23) du Chapitre 2. Plus l'hétérogénéité inter-sous-groupe est élevée, plus la caractéristique étudiée explique l'hétérogénéité globale. Dans le Tableau 5.4 les sous-groupes d'essai sont classés par ordre décroissant d'hétérogénéité inter-sous-groupes. Ainsi cinq caractéristiques d'essais étaient associées à un test d'interaction ou un test de tendance significatif avec le traitement reçu. Pour chacun de ces sous-groupes, une catégorie était associée à une amélioration de la survie globale avec la RT expérimentale. Cette amélioration de la survie était significative pour trois caractéristiques (une compliance à la CT similaire entre les deux bras, une dose par fraction inférieure à 1,8 Gy, et une RT hyper-fractionnée) et non-significative pour les deux autres (une durée totale de traitement inférieure à 30 jours, et la présence de CT à base de platine durant la RT dans les deux bras). Il est à noter que les essais utilisant de la RT hyper-fractionnée délivraient également des fractions inférieures à 1,8 Gy, et dont la durée totale de traitement était inférieure à 30 jours.

La compliance à la CT est la caractéristique d'essais qui avait l'hétérogénéité inter-sous-groupe la plus élevée et expliquait donc le mieux l'hétérogénéité globale (p -valeur du test d'interaction $< 0,0001$). Pour les essais avec une compliance à la CT similaire entre les deux bras, la RT expérimentale avait un effet significativement bénéfique sur la survie globale par rapport à la RT contrôle ($HR_{RT\text{ Exp vs RT Cont}} = 0,79$, IC 95 % : [0,69 ; 0,91]). A l'inverse, pour les essais avec une compliance à la CT différente entre les deux bras, la RT expérimentale avait un effet significativement délétère sur la survie globale ($HR_{RT\text{ Exp vs RT Cont}} = 1,19$, IC 95 % : [1,05 ; 1,34] ; Tableau 5.4).

Tableau 5.4 - *Hazard ratios* et hétérogénéité de l'effet de la radiothérapie expérimentale contre la radiothérapie contrôle sur la survie globale en fonction de caractéristiques d'essais

Caractéristiques essais		HR [IC 95 %]	Hétérogénéité [§]	
			Inter-sous-groupe	Intra-sous-groupe
Compliance à la CT entre les bras	Similaire	0,79 [0,69 ; 0,91]	19,5***	1,9
	Différente	1,19 [1,05 ; 1,34]		
Dose RT par fraction	< 1,8 Gy	0,82 [0,71 ; 0,96]	7,5* (p = 0,02) ^{§§}	14,1
	1,8- 2,4 Gy	1,11 [0,90 ; 1,35]		
	> 2,4 Gy	1,06 [0,94 ; 1,20]		
Type de RT	Hyper-fractionnée	0,82 [0,71 ; 0,96]	7,4**	14,2
	Standard	1,07 [0,96 ; 1,19]		
Durée totale de la RT	≤ 30 jours dans les deux bras	0,89 [0,78 ; 1,02]	5,6 (p = 0,02) ^{§§}	16,0
	Un bras ≤ 30 jours, un > 30 jours	0,99 [0,85 ; 1,15]		
	> 30 jours dans les deux bras	1,16 [0,98 ; 1,38]		
CT à base de platine pendant la RT dans les deux bras	Oui	0,89 [0,79 ; 1,01]	5,5**	16,1
	Non	1,09 [0,97 ; 1,24]		
CT concurrente dans les deux bras	Oui	0,95 [0,85 ; 1,06]	1,5	20,1
	Non	1,06 [0,92 ; 1,22]		
Même RT dans les deux bras	Oui	0,96 [0,85 ; 1,08]	0,5	21,1
	Non	1,02 [0,90 ; 1,16]		

Légende du Tableau 5.4 :

* : p<0,05 ; ** : p<0,01 ; *** : p<0,001.

[§] L'hétérogénéité totale est égale à 21,6 (9 essais analysés) excepté pour l'observance à la CT où elle vaut 21,4 (8 essais analysés : CCWFU62286 exclu de l'analyse).^{§§} Test de tendance;CT: Chimiothérapie; IC: Intervalle de Confiance; HR: *Hazard ratio* comparant la RT expérimentale à la RT contrôle; RT: Radiothérapie.

Cette interaction entre le traitement par RT et la compliance à la CT pouvait également s'apprécier sur les courbes de survies de Peto données dans la Figure 5.8 pour chaque bras de traitement et chacun des sous-groupes d'essai. Nous retrouvions ainsi le fait que dans les essais avec une compliance à la CT similaire entre les deux bras (courbes bleues), la RT expérimentale comparée à la RT contrôlée augmentait significativement le taux de survie de 5,7 % à 3 ans (de 24,4 % à 30,1 %) et de 7,7 % à 5 ans (de 16,5 % à 24,2 %). A l'inverse, dans les essais avec une compliance différente (courbes rouges/roses) la RT expérimentale diminuait le taux de survie de 3,8 % à 3 ans (de 16,1 % à 12,3 %) et de 2,2 % à 5 ans (de 10,5 % à 8,3 %). Nous pouvions également remarquer que les deux courbes bleues du sous-groupe avec une compliance à la CT similaire entre les deux bras étaient systématiquement au-dessus des deux courbes du sous-groupe avec une compliance à la CT différente.

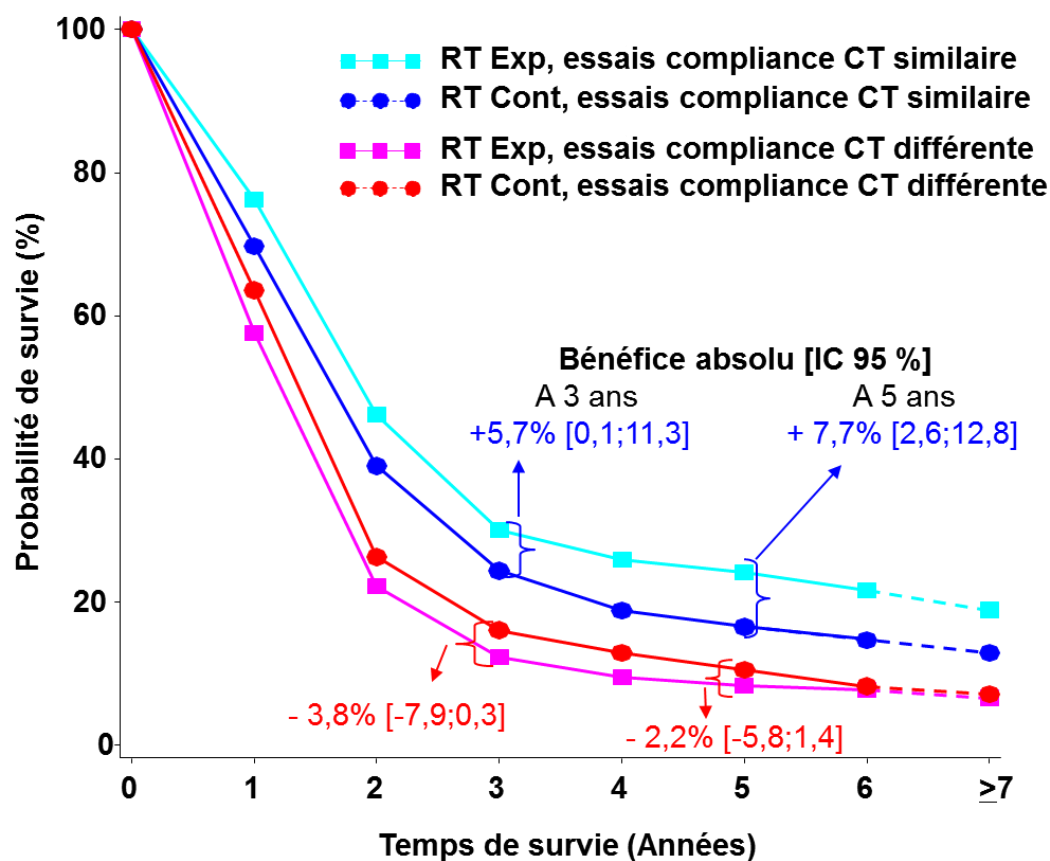


Figure 5.8 - Courbes de survie estimées à l'aide de la méthode Peto-année pour les sous-groupes d'essai avec une compliance à la chimiothérapie similaire ou différente entre les deux bras pour la méta-analyse RTT-SCLC

Légende de la Figure 5.8 :

Cont: Contrôle; CT: Chimiothérapie; Exp: Expérimentale; IC: Intervalle de Confiance; RT: Radiothérapie.

Afin de pouvoir comparer les forest-plots des HRs et des $rmstD(t^* = 5)$ donnés respectivement en Figure 5.9 et Figure 5.10 pour les deux sous-groupes d'essais définis par la compliance à la CT, nous avons recalculé les HRs donnés précédemment en inversant le traitement de référence. Ainsi, $HR_{RT\text{ Cont vs RT Exp}} = 1/0,79 = 1,26$ (IC 95 % : [1,10 ; 1,45]) pour les essais avec une compliance à la CT similaire entre les deux bras, et $HR_{RT\text{ Cont vs RT Exp}} = 1/1,19 = 0,84$ ([0,75 ; 0,95]) pour les essais avec une compliance à la CT différente.

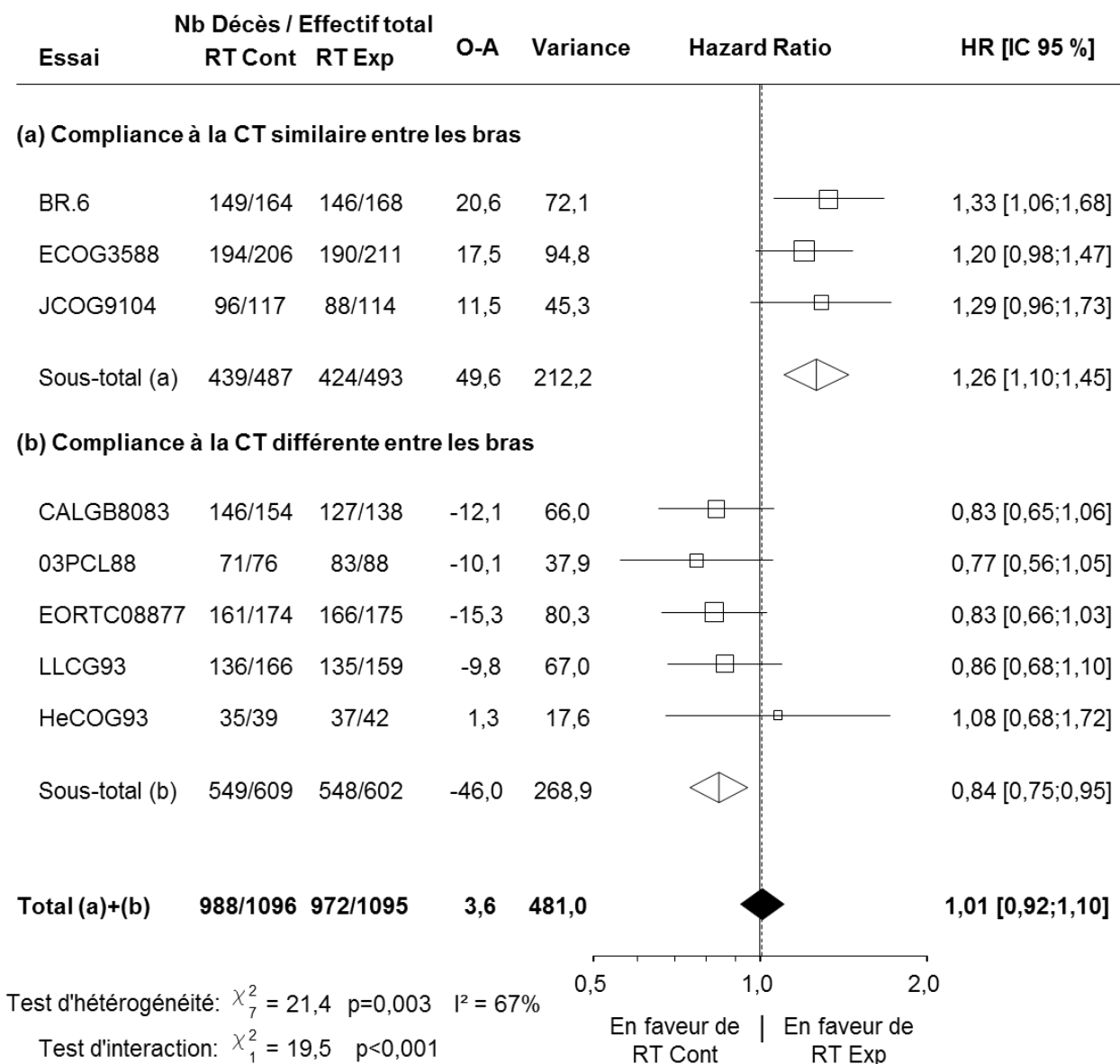


Figure 5.9 - Forest-plot des *hazard ratios* avec un modèle à effets fixes dans deux sous-groupes d'essais de la méta-analyse RTT-SCLC.

La Figure 5.10 indique que l'interaction entre le traitement reçu et la compliance à la CT était significative également avec le $rmstD(t^* = 5)$ ($p < 0,001$). La RT expérimentale augmentait la survie moyenne restreinte à 5 ans de 4,2 mois (IC 95 % : [1,8 ; 6,7]) pour les essais avec une compliance à la CT similaire entre les bras, et la diminuait de 3,1 mois ([1,3 ; 4,9]) pour les essais avec une compliance à la CT différente, par rapport à la RT contrôle. La significativité et le sens de l'effet traitement étaient similaires que l'effet traitement fût mesuré à l'aide du HR ou du $rmstD(t^* = 5)$ excepté pour l'essai CALGB8083 pour lequel seul le $rmstD(t^* = 5)$ était significatif ($rmstD(t^* = 5) = -4,0$ mois [-7,8 ; -0,2] et HR = 0,83 [0,65 ; 1,06]).

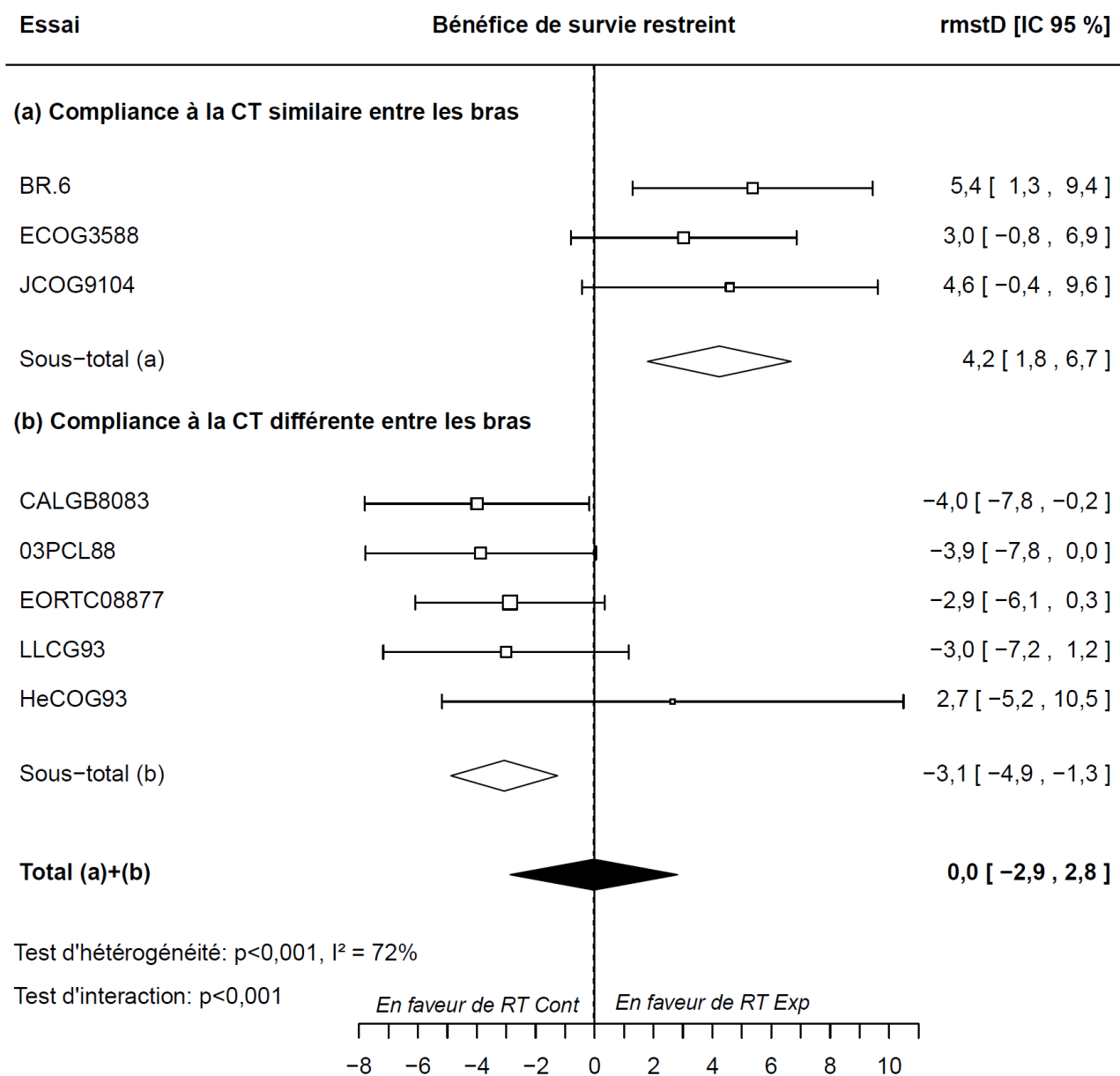


Figure 5.10 - Forest-plot des bénéfices de survie restreints à $t^* = 5$ ans, exprimés en mois et estimés avec la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires dans deux sous-groupes d'essais de la méta-analyse RTT-SCLC.

La Figure 5.11 représente le bénéfice de survie restreint (panel A) et le bénéfice de survie restreint relatif (égal au $rmstD(t^*)$ divisé par t^* ; équation (5.1)) estimés pour les deux sous-groupes définis par la compliance à la CT. Le $rmstD(t^*)$ et le $rmstRD(t^*)$ étaient estimés par la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires de DerSimonian-Laird en faisant varier l'horizon temporel t^* . Les courbes bleues et magenta représentent le bénéfice de survie restreint (relatif) pour le sous-groupe avec une compliance à la CT respectivement similaire ou différente entre les bras. La droite horizontale en pointillée indique l'absence d'effet traitement ($rmstD(t^*) = 0$ ou $rmstRD(t^*) = 0$). Les droites verticales rouges correspondent aux horizons temporels 5 et 10 ans et la zone grisée correspond à la région de confiance à 95% point par point.

La Figure 5.11-A montre que le bénéfice de survie était significativement différent de 0 et ne cessait d'augmenter pour $t^* \in [0-10]$ ans que ce soit en faveur de la RT expérimentale ou de la RT contrôle pour les sous-groupes avec une compliance à la CT respectivement similaire (courbe bleue) ou différente (courbe magenta) entre les bras.

La Figure 5.11-B indique que le bénéfice de survie relatif était maximal autour de 2 ans et valait environ 6 % de t^* que ce soit en faveur de la RT expérimentale ou de la RT contrôle pour les sous-groupes avec une compliance à la CT respectivement similaire (courbe bleue) ou différente (courbe magenta) entre les bras. Pour les essais avec une compliance à la CT différente, le bénéfice de survie restreint relatif en faveur de la RT contrôle diminuait après 2 ans en passant de 6 % à 2 ans à 3 % à 10 ans. Ce résultat indique un effet délétère précoce de la RT expérimentale.

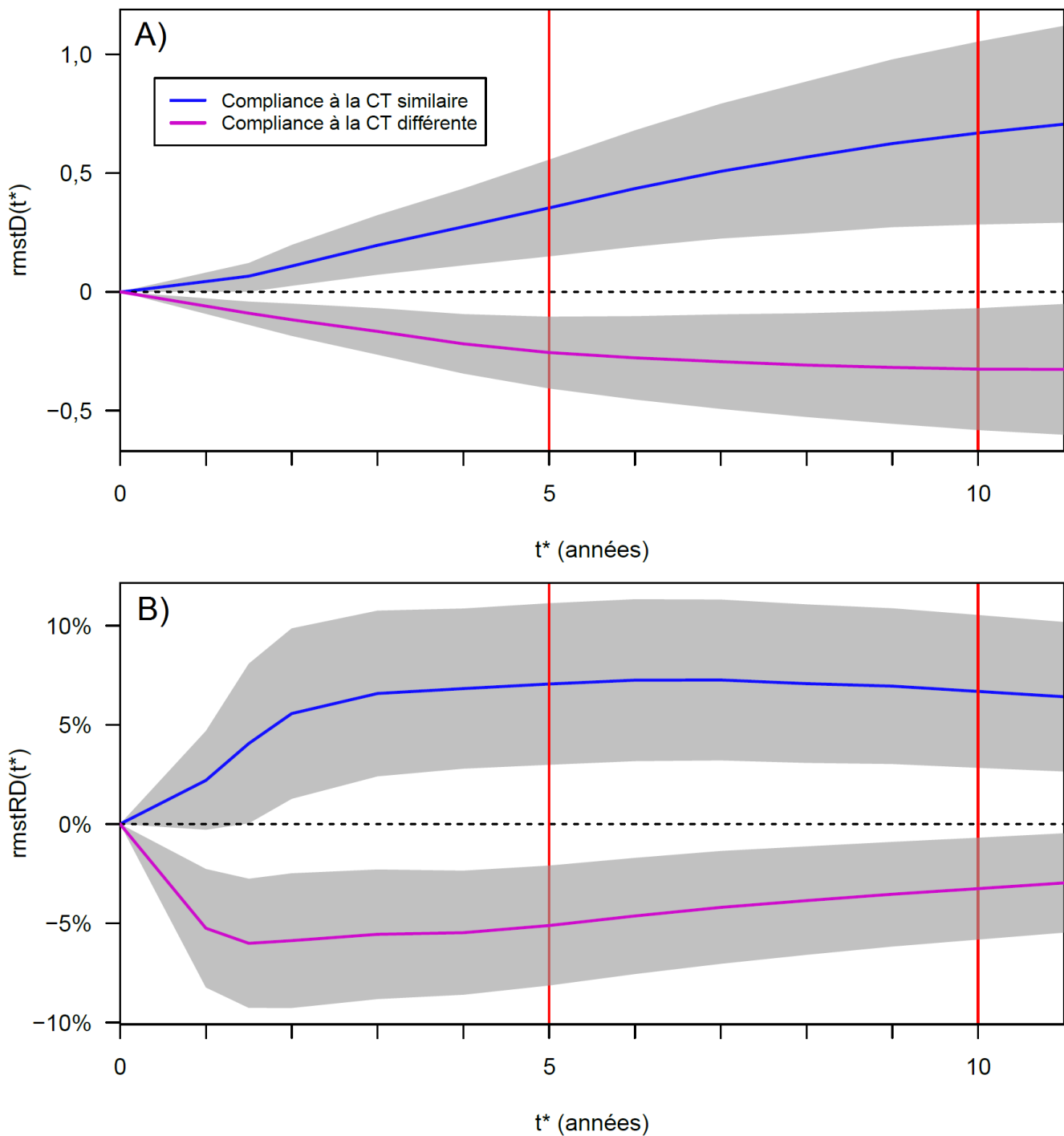


Figure 5.11 - Bénéfice de survie restreint (panel A) et bénéfice de survie restreint relatif (panel B) exprimés comme fonction de l'horizon temporel t^* dans deux sous-groupes d'essais de la méta-analyse RTT-SCLC.

Légende de la Figure 5.11 :

$rmstD$: bénéfice de survie restreint ; $rmstRD$: bénéfice de survie restreint relatif

Le $rmstD(t^*)$ et $rmstRD(t^*)$ sont estimés à l'aide de la méthode Pooled Kaplan-Meier avec effets aléatoires. Les courbes bleue et magenta représentent les sous-groupes d'essais avec une compliance à la chimiothérapie respectivement similaire ou différente entre les deux bras.

5.2.3. Conclusion de la méta-analyse RTT-SCLC

La méta-analyse sur données individuelles RTT-SCLC incluait neuf essais de phase III randomisés et étudiait le timing optimal de la radiothérapie thoracique dans le cancer du poumon à petites cellules. Cette MADI a montré que globalement il n'y avait pas de différence statistiquement significative en termes de survie globale entre la RT expérimentale (précoce ou courte) et la RT contrôle (tardive ou longue) ($HR=0,99$; $p=0,78$) et qu'il y avait une forte hétérogénéité de l'effet traitement entre les essais (Q-test : $p = 0,006$; $I^2 = 63\%$). Cependant, les analyses de sous-groupes en fonction de caractéristiques d'essais prévues dans le protocole de RTT-SCLC ont permis de mettre en évidence plusieurs caractéristiques ayant une interaction significative avec le bras de traitement reçu.

La compliance à la chimiothérapie, définie comme la proportion de patients ayant reçu le nombre de cycles de CT prévu, était la caractéristique d'essais qui expliquait le mieux l'hétérogénéité globale (test d'interaction : $p < 0,0001$). Ainsi, pour les essais avec une compliance à la CT différente entre les bras, pour lesquels une plus faible compliance était toujours observée dans le bras RT expérimentale, la RT expérimentale avait un effet délétère sur la survie globale comparée à la RT contrôle ($HR_{RT\ Exp\ vs\ RT\ Cont} = 1,19$; IC 95 % : [1,05 ; 1,34]). A l'inverse, pour les essais avec une compliance à la CT similaire (et bonne, avec au moins 79% de patients compliants dans chaque bras) entre les deux bras, la RT expérimentale avait un effet significativement bénéfique sur la survie globale ($HR_{RT\ Exp\ vs\ RT\ Cont} = 0,79$; IC 95 % : [0,69 ; 0,91]). Cette amélioration de la survie correspondait à un bénéfice de survie restreint à $t^* = 5$ ans de 4,2 mois en faveur de la RT expérimentale (IC 95 % : [1,8 ; 6,7]) et de 3,1 mois en faveur de la RT contrôle (IC 95 % : [1,3 ; 4,9]) pour les sous-groupes avec une compliance à la CT respectivement similaire ou différente entre les bras (Figure 5.10).

L'une des limites de cette MADI est que certains essais n'ont pas pu être inclus : trois essais éligibles et décrits dans le flow-chart de la Figure 5.6 mais pour lesquels nous n'avons pas pu récupérer les données individuelles, et un quatrième essai trop récent pour être inclus dans notre MADI. Cependant nous avons réalisé une analyse de sensibilité sur les sous-groupes d'essai définis par la compliance à la CT en incluant trois de ces quatre essais à l'aide des *hazard ratios* extraits des publications respectives. Deux de ces essais furent inclus dans le groupe avec une compliance à la CT similaire entre les bras et le troisième fut inclus dans le groupe avec une compliance à la CT différente. Pour cette analyse *post-hoc*, nous avons retrouvé les mêmes résultats sur la survie globale : un HR égal à 0,81 (IC 95 % : [0,72 ; 0,90]) en faveur de la RT expérimentale contre un HR égal à 1,18 (IC 95 % : [1,06 ; 1,32]) en faveur de la RT contrôle

respectivement pour le sous-groupe avec une compliance à la CT similaire ou différente. Les résultats étaient également similaires pour la survie sans progression mais ne sont pas présentés ici.

Une autre limite de cette MADI est que nous n'avons pas pu reproduire au niveau individuel, à l'aide d'un modèle de Cox multivarié, la valeur prédictive de la compliance à la CT trouvée au niveau essai ($p = 0,21$). Ce résultat est probablement lié à un manque de puissance et à la qualité des données reçues, puisque nous avons obtenu seulement le nombre de cycles de CT et non la dose effectivement reçus pour chaque patient. Nous avons par contre montré à l'aide d'une analyse multivariée que la compliance à la CT était pronostique d'une meilleure survie globale au niveau individuel avec un hazard ratio égal à 0,56 [0,49–0,64] en faveur des patients compliants à la CT comparés à ceux non-compliants (cf. Tableau S8 donné en Annexes du manuscrit RTT-SCLC ; Figure A.1). Enfin, notre conclusion est basée sur une étude de l'interaction entre l'effet traitement et une caractéristique d'essais, il s'agit cependant des meilleures données disponibles actuellement.

L'originalité de cette MADI était de mesurer l'effet traitement à la fois à l'aide du *hazard ratio* et du bénéfice de survie restreint. Quelque que soit la mesure utilisée, les résultats concluaient que la RT expérimentale lorsque associée à une CT à base de sels de platine et donnée avec une bonne compliance était le meilleur traitement puisque les trois essais avec une compliance à la CT similaire utilisaient ce régime. Ce résultat était en accord avec les méta-analyses sur données publiées précédentes (Fried et al., 2004; Spiro et al., 2006; Pijls-Johannesma et al., 2007) et en particulier avec celle de Spiro et al. qui avait également étudié la compliance à la CT (Spiro *et al.*, 2006).

5.3. Comparaison du bénéfice de survie restreint et du hazard ratio à partir de MAR-LC, MAC-NPC 2 et RTT-SCLC

A travers les différents exemples de méta-analyse sur données individuelles de cette thèse, nous avons constaté que les conclusions relativement à l'efficacité d'un traitement fondées sur le *hazard ratio* ou sur le bénéfice de survie restreint étaient similaires dans la grande majorité des cas. Nous retrouvions ainsi un résultat déjà mis en évidence par deux études antérieures (Wei *et al.*, 2015; Trinquart *et al.*, 2016).

Wei *et al* ont comparé les estimations de l'effet traitement avec le HR ou avec le bénéfice de survie restreint à 5 ans, à travers un exemple de méta-analyse sur données individuelles incluant 12 essais dans le cancer du poumon non à petites cellules. Dans chaque essai, le HR était estimé à l'aide d'un modèle de Cox, et le $rmstD(t^*)$ était estimé à l'aide du modèle paramétrique flexible de Royston et Parmar (Royston & Parmar, 2002; modèle défini en Annexes). Les HR et $rmstD(t^*)$ globaux étaient estimés par une moyenne pondérée à l'aide d'un modèle à effets fixes. La p-valeur globale du test de proportionnalité des risques de Grambsch-Therneau ($p = 0,03$) indiquait que cette hypothèse n'était pas vérifiée à un seuil de 5 %. Cependant, la direction et la significativité des estimations de l'effet traitement avec les deux mesures (HR et $rmstD(t^*)$) étaient similaires dans chaque essai et pour l'effet global ($(rmstD(t^*) = 0,07$ année $[-0,07 ; 0,21]$ et $HR = 0,9$ $[0,8 ; 1,0]$).

Trinquart *et al.* ont étudié les estimations du $rmstD(t^*)$ et du HR sur plusieurs essais de phase III avec des données individuelles reconstruites à l'aide des courbes de Kaplan-Meier publiées et d'un algorithme défini par Guyot *et al.* (Guyot *et al.*, 2012; Trinquart *et al.*, 2016). A l'aide d'une revue de la littérature d'articles publiés dans 5 revues majeures de recherche clinique (*New England Journal of Medicine*, *Lancet*, *Journal of the American Medical Association*, *Journal of Clinical Oncology*, et *Lancet Oncology*), les auteurs ont identifié 54 essais cliniques en oncologie publiés dans les 6 derniers mois de 2014. Dans chaque essai, le HR était estimé à l'aide d'un modèle de Cox, et le $rmstD(t^*)$ était estimé à l'aide des courbes de Kaplan-Meier. Sur les 54 essais, le HR et le $rmstD(t^*)$ estimaient un effet traitement d'une direction différente seulement pour 4 essais. Trinquart *et al.* soulignaient que dans 13 des 54 essais (24 %), l'hypothèse des risques proportionnels n'était pas vérifiée (test de Grambsch-Therneau), et donc que l'utilisation du $rmstD(t^*)$ dans les analyses et les publications des essais cliniques était fortement recommandée (Trinquart *et al.*, 2016).

La Figure 5.12 reprend un graphique présenté dans la publication de Trinquart et al, sur lequel pour chaque essai est représenté le HR en abscisses et le $rmstD(t^*)$ en ordonnées. Sur ce graphique, nous avons représenté les estimations de l'effet traitement pour les 42 essais inclus dans les méta-analyses MAR-LC, MAC-NPC 2 et RT^T-SCLC. Dans chaque essai, le HR était estimé à l'aide de l'estimateur de Peto, défini à l'équation (2.8), et le $rmstD(t^*)$ était estimé comme l'aire entre les courbes de survie de Kaplan-Meier, en extrapolant éventuellement jusqu'à t^* à l'aide de la méthode de Brown définie à l'équation (2.17). L'horizon temporel t^* était de 5 ans pour les méta-analyses MAR-LC et RT^T-SCLC et de 10 ans pour MAC-NPC 2. Parmi les 42 essais étudiés dans la Figure 5.12, seuls 3 essais (7 %) avaient des risques non-proportionnels. La direction de l'effet traitement selon qu'il est estimé par le HR ou le $rmstD(t^*)$ était différente pour 2 essais (5%), le HR indiquant un effet bénéfique du traitement expérimental et le $rmstD(t^*)$ un effet délétère. Cependant, pour ces 2 essais les estimations étaient non-significatives. Pour l'un de ces 2 essais, celui pour laquelle la différence était la plus marquée ($rmstD(t^*) = -0,7$ mois $[-7,9 ; 6,5]$ et HR = 1,1 $[0,8 ; 1,6]$) les risques étaient non-proportionnels ($p = 0,04$), ce qui pouvait expliquer la divergence entre le HR et le $rmstD(t^*)$. Pour 4 autres essais la direction de l'effet traitement était similaire mais pas sa significativité : pour 2 essais le HR était significatif mais pas le $rmstD(t^*)$ (losange noir) et inversement pour 2 autres essais (triangle noir).

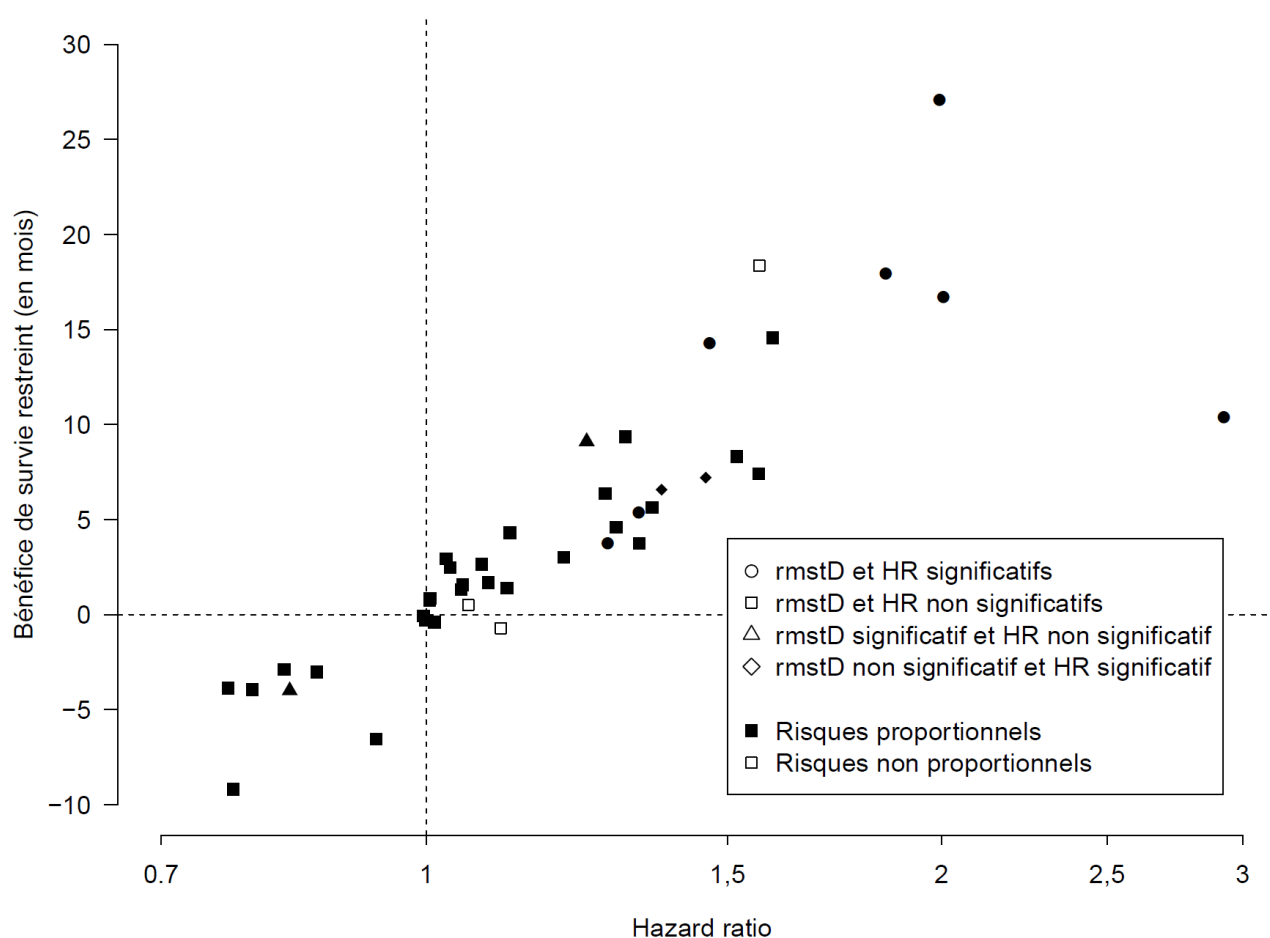


Figure 5.12 - Estimation de l'effet traitement avec le hazard ratio et le bénéfice de survie restreint pour les 42 essais des méta-analyses MAR-LC, MAC-NPC 2 et RTT-SCLC.

Légende de la Figure 5.12 :

Le bénéfice de survie restreint correspond à la différence de la survie moyenne restreinte du bras expérimental moins celle du bras contrôle. Le *hazard ratio* avait le bras expérimental comme bras de référence. Ainsi, un $rmstD(t^*) > 0$ et un $HR > 1$ indiquaient un effet bénéfique du traitement expérimental. L'horizon temporel t^* était de 5 ans pour les méta-analyses MAR-LC et RTT-SCLC et de 10 ans pour MAC-NPC 2.

HR: *Hazard ratio* du risque de décès du bras contrôle vs. bras expérimental ; rmstD: bénéfice de survie restreint exprimé en mois;

Chapitre 6. Discussion

6.1. Synthèse des résultats et apports de la thèse

L'objectif de cette thèse était d'étudier différentes méthodes d'estimation du bénéfice de survie à partir de données individuelles d'une méta-analyse d'essais cliniques randomisés. Ces méthodes définies dans le Chapitre 2 de ce manuscrit peuvent se classer en deux approches. Dans la première approche, le $rmstD(t^*)$ est directement estimé comme l'aire entre les deux courbes de survie agrégées tenant compte ou non de l'effet essai. Les méthodes Naïve Kaplan-Meier, Stewart-Parmar et Peto correspondaient à cette approche. Dans la deuxième approche, les méthodes à deux étapes Pooled Kaplan-Meier et Pooled Exponentiel estimaient le $rmstD(t^*)$ comme la moyenne pondérée des $rmstD_j(t^*)$ estimés dans chaque essai j par chacune de ces méthodes.

Nous avons dans un premier temps étudié l'impact du choix d'une méthode d'estimation dans le contexte d'une étude médico-économique (Chapitre 3). Ce premier travail empirique à partir des données de la *Meta-Analysis of Radiotherapy in Lung Cancer* a montré que les différentes méthodes étudiées conduisaient à des estimations différentes du bénéfice de survie restreint ($rmstD(t^*)$) et de son intervalle de confiance. Notre étude coût-efficacité à partir de la méta-analyse sur données individuelles (MADI) MAR-LC a montré que même de faibles différences dans l'estimation du $rmstD(t^*)$ pouvaient conduire à des conclusions économiques différentes selon le seuil de disposition à payer considéré (Lueza *et al.*, 2016a).

Dans un second temps, nous avons réalisé une étude de simulation afin de mieux comprendre les propriétés des méthodes d'estimation étudiées (Chapitre 4). Dans cette étude, nous avons comparé quatre méthodes d'estimation du $rmstD(t^*)$ à partir de MADI, dans différents scénarios faisant varier plusieurs paramètres d'intérêt (Lueza *et al.*, 2016b). Parmi ces quatre méthodes, il est apparu que la méthode Pooled Kaplan-Meier avec un modèle à effets aléatoires de DerSimonian-Laird permettait d'obtenir le meilleur compromis en termes de biais moyen et d'erreur-type pour estimer le $rmstD(t^*)$ à travers les scénarios investigués dans notre étude de simulation. C'est donc la méthode que nous avons choisi d'illustrer dans trois exemples de MADI.

L'objet de la troisième partie (Chapitre 5) était d'estimer le bénéfice de survie restreint pour trois MADI. Pour deux MADI dans le cancer du nasopharynx, l'estimation du $rmstD(t^*)$ correspondait à une analyse secondaire à la suite de l'analyse princeps réalisée avec le *hazard ratio* (HR) (Lueza *et al.*, 2016b). Pour la troisième MADI dans le cancer du poumon RTT-SCLC, l'analyse princeps a été réalisée à l'aide à la fois du HR (mesure principale) et du $rmstD(t^*)$ (mesure secondaire) estimé par la méthode Pooled Kaplan-Meier (De Ruyscher *et al.*, 2016).

A travers ces exemples, nous avons pu comparer les estimations du $rmstD(t^*)$ et du HR à la fois pour l'effet traitement global et dans chaque essai. Nous avons constaté que les conclusions relatives à l'effet traitement étaient similaires dans 40 des 42 essais (95 %) inclus dans trois des MADI que nous avons étudiées (MAR-LC, MAC-NPC 2, et RTT-SCLC ; Figure 5.12). De même, l'effet traitement global était similaire avec les deux mesures pour les quatre MADI (MAR-LC, MAC-NPC 1, MAC-NPC 2, et RTT-SCLC). Cette thèse a donc permis de confirmer les résultats des études précédentes situées dans le même contexte (Wei *et al.*, 2015; Trinquart *et al.*, 2016) et d'accumuler des données empiriques sur l'utilisation du $rmstD(t^*)$ pour estimer l'effet d'un traitement.

Les méta-analyses posent des questions méthodologiques propres liées à la structure hiérarchique des données. Les méthodes non-paramétriques d'estimation du bénéfice de survie à partir d'une MADI ont été peu étudiées, et ont nécessité d'être implémentées dans les logiciels SAS et R. Un des enseignements de nos travaux est que la méthode de Peto, utilisée classiquement en méta-analyse pour représenter les courbes de survie en tenant compte de l'effet essai, n'est pas la méthode la plus adaptée pour estimer le $rmstD(t^*)$. Ce travail de thèse est une des premières contributions sur cette thématique et demande que les recherches soient poursuivies.

Par ailleurs, notre étude de simulation (Chapitre 4) a porté sur de nombreux scénarios et a permis de mettre en évidence les paramètres qui influencent le plus l'estimation du bénéfice de survie en MADI : l'hétérogénéité de l'effet d'un traitement entre les essais et la proportionnalité des risques.

Enfin, dans cette thèse, à l'aide des différentes MADI étudiées, nous avons également souligné les avantages du bénéfice de survie restreint qui est cliniquement plus parlant que le *hazard ratio* pour quantifier l'effet traitement, et qui, de plus, est adapté même lorsque l'hypothèse de proportionnalité des risques n'est pas vérifiée.

6.2. Le bénéfice de survie restreint dans la littérature

Jusqu'à présent, dans les publications des essais cliniques ou des méta-analyses, les analyses de survie sont illustrées par les courbes de survie et le *hazard ratio*, mesure relative de l'effet traitement. Des mesure absolues de survie sont données en complément, telles que (la différence ou le rapport) des médianes de survie ou (la différence) des taux de survie à des temps donnés pour les deux bras de traitement. Plusieurs études ont montré que les résultats d'un essai présentés sur une échelle relative influencent plus la décision clinique que ceux présentés sur une échelle absolue, alors même que la mesure cliniquement pertinente est une amélioration absolue de la survie (Forrow *et al.*, 1992; Naylor *et al.*, 1992; Bobbio *et al.*, 1994).

Les recommandations PRISMA-IPD (*Preferred Reporting Items for a Systematic review and Meta-Analysis of Individual Participant Data*), qui sont l'équivalent des recommandations CONSORT mais adaptées pour les MADI, suggèrent de rapporter à la fois une mesure relative et une mesure absolue de l'effet traitement (Stewart *et al.*, 2015). En effet, Stewart *et al.* rappellent qu'un fort effet traitement relatif est peu pertinent si le risque de décès de base sous-jacent et donc l'effet du traitement absolu sont faibles. De plus, pour un même effet relatif, l'effet absolu peut être modifié en fonction du risque de base. C'est ce que nous illustrions dans la Figure 4.1 (variation du risque de base pour un effet relatif fixé) et la Figure 4.2 (variation de l'effet relatif pour un risque de base fixé). Le $rmstD(t^*)$ est donc une mesure complémentaire du HR, puisqu'il permet de quantifier de manière absolue l'effet du traitement en combinant à la fois le risque de décès de base de la population, et l'effet relatif du traitement. De plus, le bénéfice de survie restreint peut être utilisé en évaluation médico-économique où la mesure d'un effet absolu est nécessaire (nombre d'années de vie gagnées pondérées ou non par la qualité de vie). Le $rmstD(t^*)$ est une mesure résumée non-biaisée de la courbe de survie, contrairement au ratio des médianes de survie ou à un *odds ratio* (OR) à un temps donné, qui bien qu'encore souvent utilisés sont à éviter car ils peuvent être biaisés (Michiels *et al.*, 2005b; Cortés *et al.*, 2014). De plus, le $rmstD(t^*)$ peut être utilisé même si la médiane de survie n'est pas atteinte.

Le bénéfice de survie restreint a fait l'objet de travaux méthodologiques ayant montré son intérêt pour mesurer l'effet d'un traitement (Royston & Parmar, 2011, 2013, 2016, Zhao *et al.*, 2012, 2016; Zhang, 2013; Uno *et al.*, 2014; Wei *et al.*, 2015; Trinquart *et al.*, 2016). Royston et Parmar ont proposé d'utiliser le $rmstD(t^*)$ comme une mesure alternative au *hazard ratio* pour concevoir le design d'un essai avec le nombre de sujets nécessaire basé sur ce critère (Royston & Parmar, 2013). Ces auteurs ont également développé un test du log-rank fondé sur le $rmstD(t^*)$ pour tenir compte des risques non-proportionnels (Royston & Parmar, 2016). Dans une

publication d'un essai clinique de phase III dont Parmar était co-auteur, le $rmstD(t^*)$ était le critère principal d'efficacité en présence de risques non proportionnels (Oza *et al.*, 2015). Cette publication rapportait à la fois le HR, les courbes de survie de Kaplan-Meier, le test de Grambsch-Therneau de proportionnalité des risques basé sur les résidus de Schoenfeld (Grambsch & Therneau, 1994), les survies moyennes restreintes dans les deux bras, et le bénéfice de survie restreint à $t^* = 5$ ans. Dans la méta-analyse RIT-SCLC (Chapitre 5), nous n'avions pas spécifié l'utilisation du $rmstD(t^*)$ dans le protocole, mais nous avons rapporté les résultats de la survie globale avec le $rmstD(t^*)$ dans une analyse post-hoc. Dorénavant, dans les méta-analyses sur données individuelles réalisées à Gustave Roussy, nous rapporterons dans les publications la survie moyenne restreinte dans chaque bras de traitement, ainsi que le bénéfice de survie restreint à un/des horizons t^* pré-spécifiés dans le protocole, en sus du *hazard ratio* global, des courbes de survie, et des taux de survie à au moins deux temps donnés.

Afin de comparer sur la même échelle relative le *hazard ratio* et la survie moyenne, Uno *et al.* et Trinquart *et al.* ont utilisé le rapport des survies moyennes restreintes défini comme suit (Uno *et al.*, 2014; Trinquart *et al.*, 2016) :

$$rmstR(t^*) = \frac{RMST_{Exp}(t^*)}{RMST_{Cont}(t^*)} \quad (6.1)$$

où $RMST_{Cont}(t^*)$ et $RMST_{Exp}(t^*)$ sont respectivement les survies moyennes restreintes estimées dans le bras contrôle et dans le bras expérimental. Trinquart *et al.* insistent sur le fait que bien que le *hazard ratio* et le $rmstR(t^*)$ soient deux façons distinctes de quantifier l'effet traitement, ces mesures sont cependant comparables car définies sur la même échelle relative. Ainsi, le *hazard ratio* peut être recalculé en inversant le bras contrôle par le bras expérimental comme bras de référence, afin que ce nouvel HR et le $rmstR(t^*)$ reflètent tous deux la supériorité du traitement expérimental lorsque supérieurs à 1. Trinquart *et al.* ont par exemple montré que le HR indiquait un effet traitement plus grand comparé au $rmstR(t^*)$ pour 41 des 54 essais (76 %) qu'ils ont étudiés, et que la différence était significative dans 20 essais (37 %).

Nous avons souligné les avantages du bénéfice de survie restreint par rapport aux mesures usuelles de l'effet traitement. Cependant, le $rmstD(t^*)$ présente certains inconvénients. L'inconvénient majeur est que l'estimation du $rmstD(t^*)$ est intrinsèquement dépendante du choix de l'horizon temporel t^* . Plusieurs auteurs ont fait des recommandations pour le choix de t^* dans le cadre d'un essai clinique, que le critère soit statistique (Karrison, 1987, 1997) ou clinique (Royston & Parmar, 2011, 2013; Wei *et al.*, 2015). Dans le cas plus particulier d'une méta-

analyse sur données individuelles, les essais peuvent avoir différents suivis et certains essais peuvent ne plus présenter de patient à risque à l'horizon temporel choisi t^* . Afin de remédier à ce problème, nous avons utilisé dans cette thèse la méthode d'extrapolation de Brown et al. afin d'extrapoler la fonction de survie dans ces essais au-delà du dernier temps d'évènement observé t_{max} jusqu'à t^* pour les méthodes Naïve Kaplan-Meier et Pooled Kaplan-Meier (Brown *et al.*, 1974; Lamb *et al.*, 2015). Royston et Parmar, et Wei et al. ont également suggéré de présenter le graphique du $rmstD(t^*)$ en fonction de l'horizon t^* (Royston & Parmar, 2011, 2013; Wei *et al.*, 2015). Ce type de graphique permet de juger de la variation du $rmstD(t^*)$ et de sa significativité en variant t^* . Par exemple, sur la Figure 5.3-A pour la méta-analyse MAC-NPC 1, nous pouvions voir à l'aide de la région de confiance à 95% point par point que le $rmstD(t^*)$ n'était pas significativement différent de 0 pour $t^* \in [0; 10]$ ans. Enfin, certains auteurs ont proposé l'utilisation du bénéfice de survie restreint relatif ($rmstRD(t^*)$) afin de s'affranchir du fait que le $rmstD(t^*)$ s'interprète relativement à l'horizon t^* (Royston & Parmar, 2011; Zhao *et al.*, 2012; Wei *et al.*, 2015). Le $rmstRD(t^*)$ est alors présenté en termes relatifs comme un pourcentage de t^* .

Lorsque le $rmstD(t^*)$ est estimé comme l'aire entre deux courbes de Kaplan-Meier, un des inconvénients est qu'il n'est pas possible d'ajuster simplement sur des covariables, ce qui est possible pour l'estimation du *hazard ratio* avec un modèle de Cox multi-varié ou dans le cas d'un modèle de survie paramétrique. Certains auteurs ont proposé une estimation du $rmstD(t^*)$ ajusté sur des covariables en utilisant des méthodes paramétriques ou semi-paramétriques pour estimer des courbes de survie ajustées sur des covariables (Karrison, 1987, 1997; Zhang *et al.*, 2007; Zhang, 2013; Zucker, 2014). Cependant, ces méthodes sont encore peu utilisées en pratique.

Le bénéfice de survie, mesure absolue de l'effet d'un traitement, dépend à la fois de l'effet traitement relatif et du risque de base. Par conséquent, le test d'hétérogénéité lorsqu'on agrège les $rmstD_j(t^*)$ estimés dans chaque essai reflète à la fois l'hétérogénéité du risque de base et celle de l'effet traitement relatif. Deeks avait montré, en étudiant 551 méta-analyses avec un critère binaire pour mesurer l'effet traitement, que l'hétérogénéité était plus grande avec une mesure absolue plutôt qu'avec une mesure relative (Deeks, 2002). Nous pouvions donc nous attendre à une plus grande hétérogénéité de l'effet traitement, avec des données dépendantes du temps, lorsque celui-ci est mesuré avec le $rmstD(t^*)$ plutôt qu'avec le HR. Ceci n'a cependant pas été vérifié dans les quatre MADI que nous avons étudiées (MAR-LC, MAC-NPC 1, MAC-NPC 2, et RTT-SCLC) pour lesquelles l'hétérogénéité de l'effet traitement entre les essais était similaire avec l'utilisation du HR ou du $rmstD(t^*)$. Ce résultat était principalement dû au fait que,

dans ces méta-analyses, l'hypothèse des risques proportionnels était vérifiée à un seuil de 5 %, et que les risques de base étaient relativement homogènes entre les essais.

6.3. Limites de la thèse

Certains points limitants de cette thèse sont également à considérer, afin de mieux s'interroger sur les perspectives à envisager pour nos travaux. Le choix des méthodes d'estimation du bénéfice de survie étudiées dans cette thèse (Naïve Kaplan-Meier, Stewart-Parmar, Peto, Pooled Kaplan-Meier, et Pooled Exponentiel) peut être discuté. Concernant le choix de méthodes paramétriques, nous n'avons retenu que le modèle exponentiel. Ainsi, nous n'avons par exemple pas inclus le modèle de survie paramétrique flexible développé par Royston et Parmar qu'avait utilisé Wei et al. dans leur étude de simulation (Royston & Parmar, 2002; Wei *et al.*, 2015; modèle défini en Annexes). En effet, nous avons préféré nous focaliser sur des méthodes non-paramétriques (Naïve Kaplan-Meier, Stewart-Parmar et Peto) utilisées en méta-analyse (Stewart & Parmar, 1993; Durand-Zaleski *et al.*, 1997; Early Breast Cancer Trialists' Collaborative Group, 2012; Blanchard *et al.*, 2015). Cependant, les modèles paramétriques ont l'avantage de permettre l'extrapolation à la vie entière qui peut être nécessaire en évaluation médico-économique. Ainsi, plusieurs études se sont intéressées au choix d'un modèle paramétrique dans le contexte de l'évaluation médico-économique et ont montré que le choix d'un modèle d'extrapolation est décisif et que la sensibilité des résultats aux différentes méthodes d'extrapolation doit être testée (Jackson, 2010; Connock *et al.*, 2011; Gong & Fang, 2012; Latimer, 2013). Une récente revue de la littérature par Jackson et al. a décrit différentes méthodes utilisées pour extrapoler les courbes de survie en évaluation médico-économique à l'aide de données externes (Jackson *et al.*, 2016). Dans un article méthodologique auquel nous avons contribué, Guyot et al. proposaient une méthode d'extrapolation dans un cadre bayésien combinant des données d'un essai clinique pour le court-terme et des données de survie externes (tables de mortalité en population générale, le registre des cancers américains *SEER*, et une méta-analyse sur données individuelles) pour le long-terme (Guyot *et al.*, 2016). Dans leur exemple d'application, un modèle paramétrique basé sur des splines (cf. Annexes) était le modèle le plus flexible permettant la meilleure adéquation aux données à la fois à court-terme et à long-terme.

Ensuite, les résultats de notre étude de simulation étaient dépendants du choix des paramètres que nous avons fait varier. Ainsi, nous n'avons pas fait varier le taux de censures administratives et nous n'avons simulé que des essais avec un ratio d'allocation du traitement 1:1. Nous avons souligné dans le paragraphe 4.4 du Chapitre 4 qu'un biais aurait pu apparaître pour la

méthode Naïve Kaplan-Meier dans le cas d'essais avec un ratio d'allocation du bras de traitement variable et différent de 1:1.

Par ailleurs, dans les graphiques représentant le $rmstD(t^*)$ en fonction de l'horizon t^* , dans la Figure 5.3 par exemple, nous avons représenté l'incertitude associée au $rmstD(t^*)$ à l'aide d'une zone de confiance à 95 % point par point (*95% pointwise confidence interval* en anglais). Dans deux récents articles méthodologiques, Tian et al. et Zhao et al. ont proposé le développement d'un intervalle de confiance simultané (*simultaneous confidence interval*) du $rmstD(t^*)$ dans le contexte d'un essai clinique (Tian *et al.*, 2014; Zhao *et al.*, 2016).

Enfin, cette thèse s'est focalisée sur le contexte des méta-analyses sur données individuelles, qui représentent le meilleur niveau de preuve pour l'estimation de l'efficacité d'un traitement (Riley *et al.*, 2010). Cependant, les données individuelles ne sont pas toujours aisées à récupérer et les données de certains essais peuvent manquer. Des techniques récentes, comme l'algorithme proposé par Guyot et al. par exemple (Guyot *et al.*, 2012), permettent cependant de reconstruire les données individuelles de survie à l'aide des courbes de Kaplan-Meier publiées sans biais si des informations complémentaires sont disponibles dans l'article (nombre de patients à risque et nombre de décès). Trinquart et al. ont par exemple utilisé cet algorithme dans leur travail pour pouvoir estimer le HR et le $rmstD(t^*)$ pour des essais publiés (Trinquart *et al.*, 2016).

6.4. Conclusion et perspectives

Le bénéfice de survie restreint peut apporter une information complémentaire et cliniquement pertinente. De plus, le bénéfice de survie restreint est un paramètre essentiel pour l'évaluation médico-économique. Cette thèse plaide donc en faveur de l'utilisation du bénéfice de survie restreint en recherche clinique. Le $rmstD(t^*)$ devrait être un critère secondaire systématiquement rapporté dans l'analyse des essais et des méta-analyses.

Dans les publications des résultats d'un essai clinique ou d'une méta-analyse, l'hypothèse de proportionnalité des risques devrait être systématiquement testée (Grambsch & Therneau, 1994). En cas de risques proportionnels, il est alors licite de ne rapporter qu'un seul *hazard ratio* pour l'ensemble de la période de suivi. A l'inverse, en cas de risques non proportionnels, plusieurs HRs correspondant aux périodes avec un effet stable, voire un HR en fonction du temps devrait être rapporté (Uno *et al.*, 2014). Quelle que soit la proportionnalité des risques, les courbes de survie de Kaplan-Meier (ou de Peto en MADI), les survies moyennes restreintes dans

les deux bras, et le bénéfice de survie restreint à un ou plusieurs horizons t^* pré-spécifiés dans le protocole devraient être rapportés dans les publications cliniques, comme dans celle d'Oza et al (Oza *et al.*, 2015). Enfin, un graphique représentant le bénéfice de survie restreint, voire également le bénéfice relatif ($rmstRD(t^*)$), en fonction du temps permettrait de juger de la variation de l'effet traitement au cours du temps.

Néanmoins, des recherches sur cette thématique devraient être conduites. En réponse aux limites évoquées dans la section précédente, certaines perspectives peuvent être envisagées pour prolonger ce travail de thèse. Concernant le développement d'un intervalle de confiance simultané du $rmstD(t^*)$ proposé par Tian et al. et Zhao et al. (Tian *et al.*, 2014; Zhao *et al.*, 2016), une extension au contexte des méta-analyses ou des essais multicentriques n'a pas encore été proposée et pourrait faire l'objet de futures recherches. Dans notre étude de simulation, nous avons étudié quatre méthodes d'estimation $rmstD(t^*)$ principalement non-paramétriques. D'autres méthodes paramétriques existent dans la littérature et pourraient également être étudiées. Wei et al. avaient par exemple estimé le $rmstD(t^*)$ dans le contexte d'une MADI à l'aide d'une méthode de survie paramétrique flexible développée par Royston et Parmar (Royston & Parmar, 2002; Wei *et al.*, 2015; méthode définie en Annexes). Dans le cas où l'extrapolation à la vie entière serait nécessaire, dans le cas d'une évaluation médico-économique par exemple, un modèle plus flexible que nos méthodes Pooled Kaplan-Meier ou Pooled Exponentiel pourrait être d'estimer le $rmstD(t^*)$ dans chaque essai en utilisant le modèle paramétrique le plus adéquat à la distribution des temps de décès de chaque essai, puis de les agréger. Des modèles paramétriques permettant d'estimer des courbes de survies en prenant en compte l'effet essai ont également été développés pour la méta-analyse en réseau (Ouwens *et al.*, 2010; Jansen, 2011). Ces méthodes n'ont cependant pas encore été appliquées en MADI dans le but d'estimer un $rmstD(t^*)$. Par ailleurs, dans notre étude de simulation, nous avons sélectionné certains paramètres dans nos scénarios pour rester lisible. D'autres paramètres mériteraient d'être étudiés tels qu'un ratio d'allocation différent de 1:1, et une variation du taux de censures administratives. Enfin, cette thèse a permis de comparer empiriquement l'estimation de l'effet traitement à l'aide du bénéfice de survie restreint et du *hazard ratio*. Il serait par ailleurs intéressant d'étudier les propriétés statistiques et analytiques de trois mesures fondées sur la survie moyenne restreinte : la différence des survies moyennes restreintes ($rmstD(t^*)$), de la différence rapportée comme pourcentage de t^* ($rmstRD(t^*)$), et le ratio des survies moyennes restreintes ($rmstR(t^*)$).

Références

- Altman DG, Deeks JJ (2002) Meta-analysis, Simpson's paradox, and the number needed to treat. *BMC Medical Research Methodology*, **2**, 3.
- Andersen PK, Pohar Perme M (2010) Pseudo-observations in survival analysis. *Statistical Methods in Medical Research*, **19**, 71–99.
- Andersen PK, Hansen MG, Klein JP (2004) Regression analysis of restricted mean survival time based on pseudo-observations. *Lifetime Data Analysis*, **10**, 335–350.
- Anell A, Norinder A (2000) Health outcome measures used in cost-effectiveness studies: a review of original articles published between 1986 and 1996. *Health Policy*, **51**, 87–99.
- Baujat B, Audry H, Bourhis J et al. (2006) Chemotherapy in locally advanced nasopharyngeal carcinoma: an individual patient data meta-analysis of eight randomized trials and 1753 patients. *International Journal of Radiation Oncology, Biology, Physics*, **64**, 47–56.
- Blanchard P, Lee A, Marguet S et al. (2015) Chemotherapy and radiotherapy in nasopharyngeal carcinoma: an update of the MAC-NPC meta-analysis. *The Lancet Oncology*, **16**, 645–655.
- Bobbio M, Demichelis B, Giustetto G (1994) Completeness of reporting trial results: Effect on physicians' willingness to prescribe. *Lancet*, **343**, 1209–1211.
- Bowden J, Tierney JF, Simmonds M, Copas AJ, Higgins JP (2011) Individual patient data meta-analysis of time-to-event outcomes: one-stage versus two-stage approaches for estimating the hazard ratio under a random effects model. *Research Synthesis Methods*, **2**, 150–162.
- Braud A-C, Lévy-Piedbois C, Piedbois P et al. (2003) Direct treatment costs for patients with lung cancer from first recurrence to death in france. *Pharmacoeconomics*, **21**, 671–9.
- Brown J, Hollander M, Korwar R (1974) Nonparametric tests of independence for censored data with application to heart transplant studies. *Reliability and Biometry, Statistical Analysis of Lifelength*, 327–54.
- Cates CJ (2002) Simpson's paradox and calculation of number needed to treat from meta-analysis. *BMC Medical Research Methodology*, **2**, 1.
- Chen PY, Tsiatis A (2001) Causal inference on the difference of the restricted mean lifetime between two groups. *Biometrics*, **57**, 1030–1038.
- Connock M, Hyde C, Moore D (2011) Cautions Regarding the Fitting and Interpretation of Survival Curves Drugs for Cancer. *Pharmacoeconomics*, **29**, 827–837.
- Cortés J, González J a, Campbell MJ, Cobo E (2014) A hazard ratio was estimated by a ratio of median survival times, but with considerable uncertainty. *Journal of clinical epidemiology*, **67**, 1172–7.

- Cox D (1972) Regression Models and Life-Tables. *Journal of the Royal Statistical Society, Series B*, **34**, 187–220.
- Deeks JJ (2002) Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine*, **21**, 1575–1600.
- DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Controlled clinical trials*, **7**, 177–188.
- Durand-Zaleski I, Roche B, Buyse M et al. (1997) Economic implications of hepatic arterial infusion chemotherapy in treatment of nonresectable colorectal liver metastases. *Journal Of The National Cancer Institute*, **32**, 790–795.
- Dwan K, Gamble C, Williamson PR, Kirkham JJ (2013) Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias - An Updated Review. *PLoS ONE*, **8**.
- Early Breast Cancer Trialists' Collaborative Group (1990) *Treatment of early breast cancer vol.1: worldwide evidence 1985-1990* (ed Oxford University Press). Oxford.
- Early Breast Cancer Trialists' Collaborative Group (2012) Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *The Lancet*, **379**, 432–444.
- Fisher R (1932) *Statistical Methods for Research Workers 4th edition* (ed Boyd O and). London, 99-101 pp.
- Forrow L, Taylor WC, Arnold RM (1992) Absolutely relative: How research results are summarized can affect treatment decisions. *American Journal of Medicine*, **92**, 121–124.
- Fried DB, Morris DE, Poole C et al. (2004) Systematic review evaluating the timing of thoracic radiation therapy in combined modality therapy for limited-stage small-cell lung cancer. *Journal of Clinical Oncology*, **22**, 4837–4845.
- Glasziou PP, Simes RJ, Gelber RD (1990) Quality adjusted survival analysis. *Statistics in Medicine*, **9**, 1259–1276.
- Glidden D V., Vittinghoff E (2004) Modelling clustered survival data from multicentre clinical trials. *Statistics in Medicine*, **23**, 369–388.
- Goeree R, Villeneuve J, Goeree J, Penrod JR, Orsini L, Tahami Monfared AA (2016) Economic evaluation of nivolumab for the treatment of second-line advanced squamous NSCLC in Canada: a comparison of modeling approaches to estimate and extrapolate survival outcomes. *Journal of Medical Economics*, **19**, 630–644.
- Gong Q, Fang L (2012) Asymptotic properties of mean survival estimate based on the Kaplan-Meier curve with an extrapolated tail. *Pharmaceutical Statistics*, **11**, 135–140.
- Grambsch P, Therneau T (1994) Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*.
- Gray A, Clarke P, Wolstenholme J, Wordsworth S (2011) Applied methods of cost-effectiveness analysis in health care. 313.

- Guyot P, Welton NJ, Ouwens MJNM, Ades A (2011) Survival time outcomes in randomized, controlled trials and meta-analyses: the parallel universes of efficacy and cost-effectiveness. *Value in health*, **14**, 640–6.
- Guyot P, Ades A, Ouwens MJ, Welton NJ (2012) Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Medical Research Methodology*, **12**, 9.
- Guyot P, Ades AE, Beasley M, Lueza B, Pignon J-P, Welton NJ (2016) Extrapolation of Survival Curves from Cancer Trials Using External Information. *Medical Decision Making*, (in press).
- HAS (2011) *Choix méthodologiques pour l'évaluation économique à la HAS*. 1-84 pp.
- HAS (2013) *Niveau de preuve et gradation des recommandations de bonne pratique*. 1-92 pp.
- Higgins J, Green S (2011) *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley-Blackwell, 649 pp.
- Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. *BMJ*, **327**, 557–60.
- Hosmer DW, Lemeshow S, May S (2008a) The Delta Method. *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*, 355–358.
- Hosmer DW, Lemeshow S, May S (2008b) *Applied Survival Analysis: Regression Modeling of Time to Event Data, 2nd Edition*. John Wiley & Sons, Ltd, New-York.
- Irwin JO (1949) The standard error of an estimate of expectation of life, with special reference to expectation of tumourless life in experiments with mice. *The Journal of Hygiene*, **47**, 188.
- Jackson C (2010) Survival models in health economic evaluations: balancing fit and parsimony to improve prediction. *The International Journal of Biostatistics*, **6**.
- Jackson C, Stevens J, Ren S, Latimer N, Bojke L, Manca A, Sharples L (2016) Extrapolating Survival from Randomized Trials Using External Data: A Review of Methods. *Medical Decision Making*.
- Jansen JP (2011) Network meta-analysis of survival data with fractional polynomials. *BMC Medical Research Methodology*, **11**, 61.
- Kaplan EL, Meier P (1958) Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, **53**, 457–481.
- Karrison T (1987) Restricted mean life with adjustment for covariates. *Journal of the American Statistical Association*, **82**, 1169–1176.
- Karrison T (1997) Use of Irwin's restricted mean as an index for comparing survival in different treatment groups—interpretation and power considerations. *Controlled clinical trials*, **18**, 151–167.
- Klein JP, Moeschberger ML (2003) *Survival Analysis: techniques for censored and truncated data, 2nd ed.* Springer-Verlag, New York.
- Lamb KE, Williamson EJ, Coory M, Carlin JB (2015) Bias and precision of measures of survival gain from right-censored data. *Pharmaceutical Statistics*, **14**, 409–17.

- Latimer NR (2013) Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. *Medical Decision Making*, **33**, 743–54.
- Legrand C, Ducrocq V, Janssen P, Sylvester R, Duchateau L (2005) A Bayesian approach to jointly estimate centre and treatment by centre heterogeneity in a proportional hazards model. *Statistics in Medicine*, **24**, 3789–3804.
- Lueza B, Mauguén A, Pignon J-P, Rivero-Arias O, Bonastre J (2016a) Difference in Restricted Mean Survival Time for Cost-Effectiveness Analysis Using Individual Patient Data Meta-Analysis: Evidence from a Case Study. *PLoS ONE*, **11**, e0150032.
- Lueza B, Rotolo F, Bonastre J, Pignon J-P, Michiels S (2016b) Bias and precision of methods for estimating the difference in restricted mean survival time from an individual patient data meta-analysis. *BMC Medical Research Methodology*, **16**, 37.
- Mantel N (1966) Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*, **50**, 163–70.
- Martin P (2003) Le Coût de la radiothérapie. *Bulletin du Cancer*, **90**, 969–975.
- Mauguén A, Le Péchoux C, Saunders MI et al. (2012) Hyperfractionated or accelerated radiotherapy in lung cancer: an individual patient data meta-analysis. *Journal of Clinical Oncology*, **30**, 2788–97.
- Van Meerbeeck JP, Fennell DA, De Ruyscher DKM (2011) Small-cell lung cancer. *The Lancet*, **378**, 1741–1755.
- Michiels S, Baujat B, Mahé C, Sargent DJ, Pignon JP (2005a) Random effects survival models gave a better understanding of heterogeneity in individual patient data meta-analyses. *Journal of Clinical Epidemiology*, **58**, 238–45.
- Michiels S, Piedbois P, Stewart L, Pignon J (2005b) Meta-analysis when only the median survival times are known: a comparison with individual patient data results. *International Journal of Technology Assessment in Health Care*, **21**, 119–125.
- Minacori R, Bonastre J, Lueza B, Marguet S, Levy P (2015) How to Model Survival In Cost-Effectiveness Analysis? Differences Between Markov and Partitioned Survival Analysis Models. *Value in Health*, **18**, A704.
- Moher D, Hopewell S, Schulz KF et al. (2010) CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ*, **340**, c869.
- Munda M, Legrand C (2014) Adjusting for centre heterogeneity in multicentre clinical trials with a time-to-event outcome. *Pharmaceutical Statistics*, **13**, 145–152.
- Murray S, Tsiatis AA (1999) Sequential Methods for Comparing Years of Life Saved in the Two-Sample Censored Data Problem. *Biometrics*, **55**, 1085–1092.
- Naylor CD, Chen E, Strauss B (1992) Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Annals of Internal Medicine*, **117**, 916–21.
- Neymark N, Adriaenssen I, Gorlia T (2001) Cost-effectiveness of the addition of early hormonal therapy in locally advanced prostate cancer: results decisively determined by the cut-off time-point chosen for the. *European Journal of Cancer*, **37**, 1768–1774.

- Neymark N, Adriaenssen I, Gorlia T (2002) Estimating survival gain for economic evaluations with survival time as principal endpoint: A cost-effectiveness analysis of adding early hormonal therapy to radiotherapy in patients with locally advanced prostate cancer. *Health Economics*, **248**, 233–248.
- NICE (2013) Guide to the methods of technology appraisal. 1–93.
- Ouwens MJNM, Philips Z, Jansen JP (2010) Network meta-analysis of parametric survival curves. *Research Synthesis Methods*, **1**, 258–271.
- Oza AM, Cook AD, Pfisterer J et al. (2015) Standard chemotherapy with or without bevacizumab for women with newly diagnosed ovarian cancer (ICON7): Overall survival results of a phase 3 randomised trial. *The Lancet Oncology*, **16**, 928–936.
- Parmar M, Torri V, Stewart L (1998) Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Statistics in Medicine*, **2834**.
- Pignon J, Hill C (2001) Meta-analyses of randomised clinical trials in oncology. *The Lancet Oncology*, 475–482.
- Pignon J-P, Arriagada R, Ihde DC et al. (1992) A Meta-Analysis of Thoracic Radiotherapy for Small-Cell Lung Cancer. *New England Journal of Medicine*, **327**, 1618–1624.
- Pijls-Johannesma M, De Ruyscher D, Vansteenkiste J, Kester A, Rutten I, Lambin P (2007) Timing of chest radiotherapy in patients with limited stage small cell lung cancer: A systematic review and meta-analysis of randomised controlled trials. *Cancer Treatment Reviews*, **33**, 461–473.
- Ramaekers BLT, Joore M, Lueza B et al. (2013) Cost Effectiveness of Modified Fractionation Radiotherapy versus Conventional Radiotherapy for Unresected Non-Small-Cell Lung Cancer Patients. *Journal of Thoracic Oncology*, **8**, 1295–1307.
- Riley R, Lambert P, Abo-Zaid G (2010) Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ*, **340**, c221–c221.
- Rochau U, Jahn B, Qerimi V et al. (2015) Decision-analytic modeling studies: An overview for clinicians using multiple myeloma as an example. *Critical Reviews in Oncology/Hematology*, **94**, 164–178.
- Rondeau V, Michiels S, Liqueur B, Pignon J (2008) Investigating trial and treatment heterogeneity in an individual patient data meta-analysis of survival data by means of the penalized maximum likelihood approach. *Statistics in Medicine*, **27**, 1894–1910.
- Rotolo F, Michiels S (2014) Testing the treatment effect on competing causes of death in oncology clinical trials. *BMC Medical Research Methodology*, **14**.
- Royston P, Parmar MKB (2002) Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in Medicine*, **21**, 2175–97.
- Royston P, Parmar M (2011) The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Statistics in Medicine*, **30**, 2409–2421.

- Royston P, Parmar MK (2013) Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Medical Research Methodology*, **13**, 152.
- Royston P, Parmar MKB (2016) Augmenting the logrank test in the design of clinical trials in which non-proportional hazards of the treatment effect may be anticipated. *BMC Medical Research Methodology*, **16**, 16.
- Rücker G, Schumacher M (2008) Simpson's paradox visualized: the example of the rosiglitazone meta-analysis. *BMC Medical Research Methodology*, **8**, 34.
- De Ruyscher D, Lueza B, Le Péchoux C et al. (2016) Impact of thoracic radiotherapy timing in limited-stage small-cell lung cancer: usefulness of the individual patient data meta-analysis. *Annals of Oncology*, **27**, 10.
- Schaubel DE, Wei G (2011) Double Inverse-Weighted Estimation of Cumulative Treatment Effects Under Nonproportional Hazards and Dependent Censoring. *Biometrics*, **67**, 29–38.
- Simmonds MC, Higgins JPT, Stewart L a, Tierney JF, Clarke MJ, Thompson SG (2005) Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clinical Trials*, **2**, 209–17.
- Spiro SG, James LE, Rudd RM et al. (2006) Early compared with late radiotherapy in combined modality treatment for limited disease small-cell lung cancer: a London Lung Cancer Group multicenter randomized clinical trial and meta-analysis. *Journal of Clinical Oncology*, **24**, 3823–30.
- Stewart LA, Clarke MJ (1995) Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. *Statistics in Medicine*, **14**, 2057–79.
- Stewart LA, Parmar M (1993) Meta-analysis of the literature or of individual patient data: is there a difference? *The Lancet*, **341**, 25–28.
- Stewart LA, Tierney JF (2002) To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Evaluation & the health professions*, **25**, 76–97.
- Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, Stewart L (2012) Statistical analysis of individual participant data meta-analyses: a comparison of methods and recommendations for practice. *PLoS ONE*, **7**, e46042.
- Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF (2015) Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data. *JAMA*, **313**, 1657.
- Sun J-M, Ahn YC, Choi EK et al. (2013) Phase III trial of concurrent thoracic radiotherapy with either first- or third-cycle chemotherapy for limited-disease small-cell lung cancer. *Annals of Oncology*, **24**, 2088–92.
- Tian L, Zhao L, Wei LJ (2014) Predicting the restricted mean event time with the subject's baseline covariates in survival analysis. *Biostatistics*, **15**, 222–233.
- Trinquart L, Jacot J, Conner SC, Porcher R (2016) Comparison of Treatment Effects Measured by the Hazard Ratio and by the Ratio of Restricted Mean Survival Times in Oncology Randomized Controlled Trials. *Journal of Clinical Oncology*, **34**, 1813–1819.

- Tudur Smith C, Williamson PR, Marson AG (2005) Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Statistics in Medicine*, **24**, 1307–19.
- Tudur Smith C, Williamson PR (2007) A comparison of methods for fixed effects meta-analysis of individual patient data with time to event outcomes. *Clinical Trials*, **4**, 621–30.
- Uno H, Claggett B, Tian L et al. (2014) Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *Journal of Clinical Oncology*, **32**, 2380–5.
- Uno H, Wittes J, Fu H et al. (2015) Alternatives to hazard ratios for comparing the efficacy or safety of therapies in noninferiority studies. *Annals of Internal Medicine*, **163**, 127–134.
- Wei Y, Royston P, Tierney JF, Parmar MKB (2015) Meta-analysis of time-to-event outcomes from randomized trials using restricted mean survival time: application to individual participant data. *Statistics in Medicine*, **34**, 2881–98.
- Whitehead A (2002) *Meta-Analysis Of Controlled Clinical Trials*. John Wiley & Sons, Ltd, Chichester, UK.
- Yusuf S, Peto R, Lewis J, Collins R, Sleight P (1985) Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Progress in Cardiovascular Diseases*, **27**, 335–71.
- Zhang X, Loberiza FR, Klein JP, Zhang M-J (2007) A SAS macro for estimation of direct adjusted survival curves based on a stratified Cox regression model. *Computer Methods and Programs in Biomedicine*, **88**, 95–101.
- Zhang X (2013) Comparison of restricted mean survival times between treatments based on a stratified Cox model. *Bio-Algorithms and Med-Systems*, **9**, 183–189.
- Zhao H, Tsiatis AA (1997) A consistent estimator for the distribution of quality adjusted survival time. *Biometrika*, **84**, 339–348.
- Zhao H, Tsiatis A (1999) Efficient Estimation of the Distribution of Quality-Adjusted Survival Time. *Biometrics*, **55**, 1101–1107.
- Zhao L, Tian L, Uno H, Solomon SD, Pfeffer MA, Schindler JS, Wei LJ (2012) Utilizing the integrated difference of two survival functions to quantify the treatment contrast for designing, monitoring, and analyzing a comparative clinical study. *Clinical Trials*, **9**, 570–7.
- Zhao L, Claggett B, Tian L et al. (2016) On the restricted mean survival time curve in survival analysis. *Biometrics*, **72**, 215–221.
- Zucker DM (2014) Restricted Mean Life With Covariates: Modification and Extension of a Useful Survival Analysis Method. *Journal of the American Statistical Association*, **93**, 702–709.

Annexes

Modèle de survie paramétrique flexible de Royston et Parmar

Le modèle de survie paramétrique flexible développé par Royston et Parmar permet de définir une grande variété de distributions de façon flexible (Royston & Parmar, 2002). Ce modèle est défini sur l'échelle logarithmique du risque cumulé et est basé sur des splines cubiques restreintes. Le logarithme de la fonction de risque cumulée (fonction définie à l'équation (2.2)) est ainsi définie par :

$$\log[\Lambda(t; z)] = \log[-\log\{S(t; z)\}] = s(x; \gamma) + \beta z \quad (\text{A.1})$$

où $x = \log(t)$, z un vecteur de covariables, et $s(x; \gamma)$ est une fonction spline cubique restreinte.

Les fonctions splines cubiques restreintes (ou naturelles) sont définies comme des splines cubiques contraintes à être linéaires au-delà des nœuds extrêmes k_{min} et k_{max} . En plus de ces deux nœuds, m nœuds internes distincts $k_1 < \dots < k_m$ avec $k_1 > k_{min}$ et $k_m < k_{max}$ sont également spécifiés. Une fonction spline cubique restreinte peut s'écrire comme :

$$s(x; \gamma) = \gamma_0 + \gamma_1 x + \gamma_2 v_1(x) + \dots + \gamma_{m+1} v_m(x) \quad (\text{A.2})$$

avec la j -ième fonction de base $v_j(x)$ définie pour $j = 1, \dots, m$ par :

$$v_j(x) = (x - k_j)_+^3 - \lambda_j (x - k_{min})_+^3 - (1 - \lambda_j) (x - k_{max})_+^3 \quad (\text{A.3})$$

$$\text{où} \quad \lambda_j = \frac{k_{max} - k_j}{k_{max} - k_{min}} \quad (\text{A.4})$$

$$(x - a)_+ = \max(0; x - a) \quad (\text{A.5})$$

Il est intéressant de noter que ce modèle est à risques proportionnels et qu'il correspond au modèle de Weibull dans le cas où $m=0$. Cependant, en ajoutant une interaction entre les D covariables dépendantes du temps et les splines, ce modèle peut s'adapter au cas des risques non proportionnels.

$$\log[\Lambda(t; z)] = s(x; \gamma) + \beta z + \sum_{i=1}^D s(x; \gamma_i) z_i \quad (\text{A.6})$$

Tableau A.1 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint.
Scénario avec 20 essais et 100 patients par essai, et avec non-proportionnalité des risques

Scénario hétérogénéité	Méthodes	$\beta = 0$						$\beta = -0,2$						$\beta = -0,7$					
		$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$			$t^* = 5$			$t^* = 10$		
		Vrai $rmstD = 0$			Vrai $rmstD = 0$			Vrai $rmstD = -0,1$			Vrai $rmstD = 0,1$			Vrai $rmstD = -0,3$			Vrai $rmstD = 0,3$		
		Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM	Biais	ETE	ETM
$(\sigma^2, \tau^2) = (0,01;0,01)$	Naïve Kaplan-Meier	0,00	0,08	0,08	-0,01	0,18	0,17	0,00	0,08	0,08	0,01	0,19	0,17	0,00	0,08	0,08	0,01	0,18	0,17
	Pooled Kaplan-Meier	0,00	0,08	0,08	-0,01	0,20	0,20	0,00	0,08	0,08	-0,01	0,20	0,20	0,01	0,08	0,09	-0,05	0,20	0,20
	Pooled Exponentiel	0,00	0,07	0,07	-0,01	0,18	0,19	0,12	0,07	0,07	0,03	0,19	0,19	0,41	0,08	0,08	0,08	0,20	0,20
	Peto-quintile	0,00	0,08	0,07	-0,01	0,16	0,15	0,01	0,08	0,07	-0,05	0,17	0,15	0,02	0,08	0,07	-0,19	0,16	0,15
$(\sigma^2, \tau^2) = (0,01;0,10)$	Naïve Kaplan-Meier	0,00	0,11	0,08	0,00	0,26	0,17	0,00	0,11	0,08	0,01	0,27	0,17	0,00	0,11	0,08	0,00	0,26	0,17
	Pooled Kaplan-Meier	0,00	0,11	0,11	0,00	0,27	0,27	0,01	0,11	0,11	-0,01	0,28	0,27	0,00	0,11	0,11	-0,06	0,27	0,28
	Pooled Exponentiel	0,00	0,10	0,10	0,00	0,26	0,27	0,12	0,10	0,10	0,03	0,27	0,27	0,40	0,11	0,11	0,06	0,28	0,28
	Peto-quintile	0,00	0,10	0,07	0,00	0,23	0,15	0,01	0,10	0,07	-0,04	0,24	0,15	0,01	0,10	0,07	-0,18	0,23	0,15
$(\sigma^2, \tau^2) = (0,10;0,01)$	Naïve Kaplan-Meier	0,00	0,08	0,08	-0,01	0,18	0,17	0,00	0,08	0,08	0,00	0,19	0,17	0,00	0,08	0,08	-0,01	0,18	0,17
	Pooled Kaplan-Meier	0,00	0,08	0,08	-0,01	0,20	0,19	0,00	0,08	0,08	-0,02	0,19	0,19	0,00	0,08	0,08	-0,07	0,20	0,19
	Pooled Exponentiel	0,00	0,07	0,07	-0,01	0,18	0,18	0,12	0,07	0,07	0,02	0,18	0,18	0,39	0,08	0,07	0,05	0,20	0,19
	Peto-quintile	0,00	0,08	0,07	-0,01	0,17	0,15	0,00	0,08	0,07	-0,05	0,17	0,15	0,00	0,08	0,07	-0,20	0,17	0,15
$(\sigma^2, \tau^2) = (0,10;0,10)$	Naïve Kaplan-Meier	0,00	0,11	0,08	-0,01	0,26	0,17	0,00	0,11	0,08	0,01	0,27	0,17	0,00	0,12	0,08	-0,01	0,27	0,17
	Pooled Kaplan-Meier	0,00	0,11	0,11	-0,01	0,27	0,27	0,01	0,11	0,11	0,00	0,28	0,27	0,00	0,12	0,11	-0,07	0,27	0,27
	Pooled Exponentiel	0,00	0,10	0,10	-0,01	0,26	0,26	0,12	0,10	0,10	0,03	0,27	0,26	0,38	0,11	0,11	0,04	0,28	0,27
	Peto-quintile	0,00	0,10	0,07	-0,01	0,24	0,16	0,01	0,11	0,07	-0,04	0,24	0,16	0,00	0,11	0,07	-0,20	0,25	0,16

Un modèle à effets aléatoires est utilisé pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile.

β : Taille de l'effet traitement ($=\log(HR)$); σ^2 : hétérogénéité du risque de base; τ^2 : hétérogénéité de l'effet traitement; ETE: erreur-type empirique;

ETM: erreur-type moyenne; $rmstD$: bénéfice de survie restreint; t^* : horizon temporel.

Tableau A.2 - Résultats des simulations pour la comparaison de méthodes d'estimation du bénéfice de survie restreint.
Scénario avec 5 essais et un nombre de patients par essai aléatoire ($N = 1\,000$), avec proportionnalité des risques, et avec $\beta = -0,7$

Hétérogénéité scenario	Méthodes	$t^* = 5$			$t^* = 10$		
		Vrai $rmstD = 0,8$			Vrai $rmstD = 2,0$		
		Biais	ETE	ETM	Biais	ETE	ETM
$(\sigma^2, \tau^2) = (0,01;0,01)$	Naïve Kaplan-Meier	0,00	0,12	0,11	0,01	0,28	0,24
	Pooled Kaplan-Meier	0,01	0,12	0,13	0,02	0,29	0,29
	Pooled Exponential	0,01	0,11	0,11	0,02	0,27	0,28
	Peto-quintile	-0,04	0,12	0,11	-0,21	0,26	0,24
$(\sigma^2, \tau^2) = (0,01;0,10)$	Naïve Kaplan-Meier	-0,01	0,19	0,11	-0,02	0,47	0,24
	Pooled Kaplan-Meier	-0,01	0,19	0,19	-0,03	0,47	0,46
	Pooled Exponential	-0,01	0,18	0,18	-0,02	0,45	0,44
	Peto-quintile	-0,05	0,18	0,11	-0,22	0,45	0,24
$(\sigma^2, \tau^2) = (0,10;0,01)$	Naïve Kaplan-Meier	0,00	0,13	0,11	0,00	0,28	0,24
	Pooled Kaplan-Meier	0,00	0,13	0,13	0,01	0,28	0,29
	Pooled Exponential	-0,01	0,12	0,12	0,01	0,27	0,27
	Peto-quintile	-0,03	0,13	0,11	-0,17	0,27	0,24
$(\sigma^2, \tau^2) = (0,10;0,10)$	Naïve Kaplan-Meier	0,01	0,20	0,11	0,03	0,47	0,24
	Pooled Kaplan-Meier	0,01	0,20	0,19	0,04	0,46	0,44
	Pooled Exponential	0,01	0,19	0,18	0,04	0,45	0,43
	Peto-quintile	-0,02	0,19	0,11	-0,12	0,45	0,24

Un modèle à effets aléatoires est utilisé pour les méthodes Pooled Kaplan-Meier, Pooled Exponentiel et Peto-quintile.

β : Taille de l'effet traitement ($=\log(HR)$); σ^2 : hétérogénéité du risque de base; τ^2 : hétérogénéité de l'effet traitement; ETE: erreur-type empirique; ETM: erreur-type moyenne; $rmstD$: bénéfice de survie restreint; t^* : horizon temporel.

Figure A.1 - Publication de la méta-analyse RTT-SCLC acceptée à *Annals of Oncology*

La version finale de l'article au format *Annals of Oncology* est disponible via ce lien: [10.1093/annonc/mdw263](https://doi.org/10.1093/annonc/mdw263)

Impact of thoracic radiotherapy timing in limited-stage small-cell lung cancer: usefulness of the individual patient data meta-analysis

D. De Ruysscher^{1,*}, B. Lueza^{2,*}, C. Le Péchoux³, D. H. Johnson⁴, M. O'Brien⁵, N. Murray⁶, S. Spiro⁷, X. Wang⁸, M. Takada⁹, B. Lebeau¹⁰, W. Blackstock¹¹, D. Skarlos¹², P. Baas¹³, H. Choy¹⁴, A. Price¹⁵, L. Seymour¹⁶, R. Arriagada¹⁷, J-P. Pignon², on behalf of the RTT-SCLC Collaborative Group^{\$}.

Running head: Radiotherapy timing in limited-stage SCLC: an IPD meta-analysis

¹ Department of Radiation Oncology (MAASTRO clinic), GROW - School for oncology and developmental biology, Maastricht University Medical Center, Maastricht, the Netherlands; KU Leuven - Department of Oncology, Experimental Radiation Oncology, Leuven, Belgium

² Service de Biostatistique et Epidémiologie and Ligue Nationale Contre le Cancer meta-analysis platform, Gustave Roussy, Villejuif, France; CESP, INSERM U1018, Université Paris-Sud, Université Paris-Saclay, Villejuif, France;

³ Département d'Oncologie et de Radiothérapie, Gustave Roussy, Villejuif, France ; Université Paris-Sud, Université Paris-Saclay, Villejuif, France

⁴ UT Southwestern University School of Medicine, Dallas, Texas, USA

⁵ EORTC Data Center, Brussels, Belgium

⁶ British Columbia Cancer Agency, Vancouver, Canada

⁷ University College London Hospitals, London, UK

⁸ Alliance Data and Statistical Center, Duke University, Durham, USA

⁹ Osaka Prefectural Habikino Hospital, Osaka, Japan

¹⁰ Hôpital St Antoine, Paris, France

¹¹ Wake Forest University School of Medicine, Winston-Salem, USA

¹² Second Department of Medical Oncology, Metropolitan Hospital N. Faliro, Athens, Greece

¹³ The Netherlands Cancer Institute, Amsterdam, The Netherlands

¹⁴ Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, USA

¹⁵ NHS Lothian and University of Edinburgh, Edinburgh Cancer Centre, Western General Hospital, Edinburgh, UK

¹⁶ NCIC Clinical Trials Group and Queen's University, Kingston, Canada

¹⁷ Gustave Roussy, Villejuif, France; Karolinska Institutet, Stockholm, Sweden.

*: DDR and BL have equally contributed to this work as co-first authors

\$ Members list at the end of the manuscript

Corresponding author:

Dr Jean-Pierre Pignon

Meta-analysis Unit, Biostatistics and Epidemiology Department, Gustave Roussy, 114 rue Edouard Vaillant, 94805 Villejuif Cedex, France

Tel: 33 1 42 11 45 65 e-mail: jean-pierre.pignon@gustaveroussy.fr

This work was partly presented at the World Lung Cancer Conference 2011, Amsterdam, The Netherlands

ABSTRACT

Background Chemotherapy combined with radiotherapy is the standard treatment of “limited-stage” small-cell lung cancer. However, controversy persists over the optimal timing of thoracic radiotherapy and chemotherapy.

Material and methods We performed a meta-analysis of individual patient data in randomised trials comparing earlier versus later radiotherapy, or shorter vs. longer radiotherapy duration, as defined in each trial. We combined the results from trials using the stratified log-rank test to calculate pooled hazard ratios (HRs). The primary outcome was overall survival.

Results Twelve trials with 2,668 patients were eligible. Data from nine trials comprising 2,305 patients were available for analysis. The median follow-up was 10 years. When all trials were analysed together, “earlier or shorter” vs. “later or longer” thoracic radiotherapy did not affect overall survival. However, the HR for overall survival was significantly in favour of “earlier or shorter” radiotherapy among trials with a similar proportion of patients who were compliant with chemotherapy (defined as having received 100% or more of the planned chemotherapy cycles) in both arms (HR 0.79, 95% CI 0.69–0.91) and in favour of “later or longer” radiotherapy among trials with different chemotherapy compliance (HR 1.19, 1.05–1.34, interaction test $p < 0.0001$). The absolute gain between “earlier or shorter” vs. “later or longer” thoracic radiotherapy in 5-year overall survival for similar and for different chemotherapy compliance trials was 7.7% (95% CI 2.6–12.8 %) and -2.2% (-5.8–1.4 %), respectively. However, “earlier or shorter” thoracic radiotherapy was associated with a higher incidence of severe acute oesophagitis than “later or longer” radiotherapy.

Conclusion “Earlier or shorter” delivery of thoracic radiotherapy with planned chemotherapy significantly improves 5-year overall survival at the expense of more acute toxicity, especially oesophagitis.

Key words: individual participant data meta-analysis, randomised clinical trials, thoracic radiotherapy, radiotherapy timing, small-cell lung cancer, chemotherapy compliance

Key message:

The optimal timing and sequencing of thoracic radiotherapy and chemotherapy, which is the standard treatment of “limited-stage” small-cell lung cancer, has fuelled debate for many years. This individual patient data meta-analysis provides the best evidence of the beneficial effect of “earlier or shorter” radiotherapy when chemotherapy is administered with good compliance.

INTRODUCTION

Small-cell lung cancer (SCLC) is a rapidly disseminating cancer so that its primary treatment is chemotherapy, whatever the stage [1]. Approximately 25% of patients present with localised disease, formerly known as “limited-stage” disease, now called stage I-IIIb [2]. It is well known that optimal survival is achieved when chemotherapy can be administered at the total intended dose and at the required intervals [1,3]. Nevertheless, due to loco-regional failures after chemotherapy alone, the adjunction of thoracic radiotherapy was investigated. A worldwide meta-analysis showed that adding thoracic radiotherapy to chemotherapy improved long-term survival [4]. Concurrent chemotherapy comprising cisplatin and etoposide and thoracic radiotherapy has become the standard of care [1,5,6]. In non-progressing patients, this can be followed by prophylactic cranial irradiation, at the optimal dose of 25 Gy, as this treatment further prolongs survival [7,8].

However, the optimal timing and sequencing of thoracic radiotherapy with chemotherapy has fuelled debate for many years. When all trials were pooled together, no survival gain was detected whether thoracic radiotherapy was delivered early with chemotherapy or later [9-12]. However, in trials where patients were treated with cisplatin-based chemotherapy at full dose, early administration of thoracic radiotherapy seemed to confer a long-term survival advantage. There is considerable variation in the definition of early or late radiotherapy : early radiotherapy was defined as starting before 9 weeks following the beginning of chemotherapy and before the third cycle of chemotherapy in two previous literature-based meta-analyses [12,13], whilst a 30-day cut-off was used in other literature-based meta-analyses [9-11,14] (Table S1 for description of previous meta-analyses). One of these meta-analyses suggested that early delivery of thoracic radiotherapy yielded higher survival rates if all the intended cycles of chemotherapy could be administered [12], implying that the question of optimal radiotherapy timing and fractionation [15,16] could only be addressed with precise information on individual patient compliance with chemotherapy administration. Such information can only be provided by an individual patient data (IPD) meta-analysis. We therefore undertook such a study, aiming to define the best approach for combining thoracic radiotherapy with chemotherapy in stage I-IIIB SCLC.

METHODS-MATERIAL

The meta-analysis was performed according to a pre-specified protocol that is available on the Gustave Roussy website (<https://www.gustaveroussy.fr/sites/default/files/meta-analyses-protocol-rtt-sclc.pdf>).

Selection criteria and search strategy

To be eligible, trials had to compare two timing schedules of curative thoracic radiotherapy, i.e. earlier versus later within an individual trial in patients with limited-stage SCLC treated with chemo-radiotherapy. Our post-hoc criteria to define early radiotherapy was similar to the one used by Fried et al [13] and Spiro et al [12]: radiotherapy should have been initiated before 9 weeks after randomisation and before the third cycle of chemotherapy. Trials comparing two radiotherapy durations, i.e. a shorter vs. a longer course within an individual trial with at least a two-week treatment difference observed between the two arms, were also eligible. In this paper, we will use the term "earlier or shorter" for arms where earlier and/or shorter radiotherapy was used and the term "later or longer" for later and/or longer radiotherapy arms. Trials had to start after 1969 and to end before 2006, and be properly randomised. The planned chemotherapy schedule (drugs, doses, number of cycles) had to be the same in both arms, but radiotherapy modalities could be different. The total dose of radiotherapy had to be at least 30 Gy. Orthovoltage radiotherapy was an exclusion criterion. Eligible patients should have had a WHO (or equivalent) performance status 0-2 and should not have received previous treatment for this cancer. To limit publication bias, we searched both published and unpublished trials without language restriction (see Web-Appendix 1 for search strategy).

Statistical Analysis

We describe IPD collection and quality control in Web-Appendix 2. The main endpoint was overall survival and the secondary endpoints were progression-free survival and severe acute toxicities. Overall survival was defined as the time from randomisation until death from any cause or the last follow-up for surviving patients. Progression-free survival was defined as the time from randomisation until first progression or death from any cause, or the last follow-up for surviving patients without progression. We did not perform analyses on loco-regional

control, cancer deaths and late toxicities due to lack of data. The median follow-up was estimated using the reverse Kaplan-Meier method [17].

We carried out all analyses on an intention-to-treat basis. Survival analyses were stratified by trial, and the log-rank expected number of deaths and variance were used to calculate individual and overall pooled hazard ratios (HR) by the fixed effect model [15]. A similar model was used to estimate odds ratios (OR) for the comparison of toxicity between arms. χ^2 tests and the I^2 statistic were used to study heterogeneity between trials [18]. Hazard ratios were calculated using a DerSimonian-Laird random effects model if heterogeneity had a p-value <0.10 [19]. Stratified survival curves were estimated for control and experimental groups, using annual death rates and the pooled hazard ratio, and were used to estimate the absolute benefit at 3 and 5 years with their 95% confidence intervals [20]. Five-year mean survival times, parameters commonly used in economic evaluation, were also estimated (Web-Appendix 3) [21-23].

Subsets analyses according to trial characteristics were pre-planned. We investigated whether the treatment effect was dependent on any difference in the proportion of patients who were compliant with chemotherapy between the treatment arms within each trial. A patient was defined as compliant if he/she received 100% or more of the planned number of CT cycles, except for the CALGB8083 trial in which patients receiving 6 CT cycles or more were considered as compliant. A trial was considered as having different “between-arm” compliance if the difference was $\geq 10\%$ and as having similar “between-arm” compliance if it was $<10\%$ [12]. No other information on chemotherapy administration, such as the actual drug dose received or delays in chemotherapy administration, was available. χ^2 tests for interaction or trend were used to assess treatment effects across trial subsets. Overall heterogeneity was decomposed into the sum of between-subset and residual (within-subset) heterogeneity: the lower residual heterogeneity, the greater overall heterogeneity of the treatment effect between trials was explained by the trial characteristic [24]. χ^2 tests for interaction or trend were also used to test whether there was any evidence that a particular type of patient benefited more or less from “earlier or shorter” radiotherapy according to predefined subgroups. If there was substantial overall heterogeneity, then subgroup analyses were planned within treatment categories. All p-values were two-sided. Analyses were performed using SAS version 9.3.

Role of the funding source

The funding sources had no role in the study design, data collection, data analysis, data interpretation, or manuscript writing. BL and J-PP had full access to all the raw data. The corresponding author had the final responsibility for the decision to submit for publication.

RESULTS

Twelve randomised trials [12,16,25-34] including 2,668 patients were eligible. Data on nine trials and 2,305 patients (86% of potentially eligible patients) were available for this IPD meta-analysis (Figure S1). Data from one trial were lost [32] and we did not succeed in contacting the investigator of two other trials [33,34]. Table 1 depicts the nine trials included [12,16,25-31] and Table S2 summarises the trials with no available data. Four trials [16,27,30,31] had different radiotherapy modalities between the two arms, including three trials [16,30,31] comparing shorter vs. longer radiotherapy duration. Central randomisation was used in all trials, except one that used sealed envelopes [25]. In total, out of the 80 patients initially excluded from the individual trial analyses, data concerning 75 patients were recovered. The median follow-up was 10 years without any difference between the treatment arms. Patient characteristics were well balanced between the two arms of the analysis (Table

S3). Three trials [16,26,28] were categorised as having similar chemotherapy compliance in both arms, and they have a proportion of at least 79% of patients who were compliant with chemotherapy (i.e. receiving all their cycles) (Table S4). Five trials [12,25,27,29,31] had different chemotherapy (CT) compliance, with all of them exhibiting a lower compliance rate in the “earlier or shorter” arm. For the CCWFU62286 trial, we had no data available on individual CT compliance neither in the patient-level data provided by the investigator nor in the publication [30]: the CCWFU62286 trial was thus excluded from the trial subset analysis based on CT compliance. In the “later or longer” arm, 88% of patients started radiotherapy as compared to 93% in the “earlier or shorter” arm (Table S5). Among the five trials [12,25,26,27,29] comparing earlier and later radiotherapy with individual data on radiotherapy compliance, the observed difference in median times between the two arms from randomisation to the start of radiotherapy ranged from 63 to 93 days compared with 56 to 84 days for the planned difference (Table S6). There was also a significant association between individual RT compliance and CT compliance (Cochran-Mantel-Haenszel test stratified by trial: $p < 0.0001$). The more a patient was compliant with CT (i.e. receiving all their cycles), the more he/she was compliant with RT (i.e. receiving 90% of the total RT dose).

Overall survival and progression-free survival

In our main analysis, when all trials were pooled together, “earlier or shorter” radiotherapy did not have a significant impact on overall survival compared to “later or longer” radiotherapy (HR 0.99, 95% CI 0.91–1.08, $p=0.78$) (Figure S2). Treatment effect heterogeneity was observed ($p=0.006$, $I^2=63\%$). With a random effects model, the HR was not significant (0.99, 0.85–1.15, $p=0.90$).

Data on tumour progression were not available for two trials [27,31], thus the progression-free survival analysis concerned only seven trials comprising 1,764 patients and 1,596 events. There was no significant impact of radiotherapy timing on progression-free survival (HR 0.93, 95% CI 0.84–1.02, $p=0.13$) (Figure S3).

Trial subsets

Table 2 shows the HRs for overall survival according to the different pre-planned subsets analyses, described in Table S7, with overall between-trial heterogeneity decomposed into the sum of between-subset and residual (within-subset) heterogeneity. Trial subsets were in decreasing order of residual heterogeneity: the lower residual heterogeneity was for one trial subset, the greater studied characteristic (CT compliance, RT dose per fraction, etc.) explained overall heterogeneity. In Table 2, between-subset heterogeneity was associated with an interaction test between the treatment received (“earlier or shorter” RT vs “later or longer” RT) and the studied characteristic of the subset, and also with a trend test when the studied subset categories were ordinal (RT dose per fraction and RT overall treatment time). Five trial characteristics were found to be associated with an improvement in overall survival with “earlier or shorter” radiotherapy (Table 2): similar CT compliance in both arms, a dose per fraction lower than 1.8 Gy, hyperfractionated radiotherapy, overall treatment time of less than 30 days, and platin-based chemotherapy. It should be emphasised that trials using hyperfractionated radiotherapy delivered fractions of less than 1.8 Gy, and overall treatment time was less than 30 days.

The “between-arm” CT compliance (number of cycles actually given) is the factor that best explained between-trial heterogeneity, i.e. with the lowest residual heterogeneity (Table 2).

Compliance of chemotherapy and overall survival

The HR for overall survival was significantly in favour of “earlier or shorter” radiotherapy among trials in which the defined chemotherapy compliance was similar in both arms (Figure

1; HR 0.79, 95% CI 0.69–0.91) and in favour of “later or longer” radiotherapy among trials with different CT compliance: (1.19, 1.05–1.34). There was a significant interaction between chemotherapy compliance and the treatment effect (interaction test, $p < 0.0001$). In trials with similar CT compliance in both arms, “earlier or shorter” radiotherapy compared to “later or longer” radiotherapy increased the absolute 3-year and 5-year overall survival by 5.7% (from 24.4% to 30.1%) and by 7.7% (from 16.5% to 24.2%), respectively (Figure 2). In trials with different CT compliance, “earlier or shorter” radiotherapy decreased absolute 3-year and 5-year overall survival respectively by 3.8% (from 16.1% to 12.3%) and 2.2% (from 10.5% to 8.3%) (Figure 2). In other words, “earlier or shorter” radiotherapy extended the 5-year mean survival time by 4.2 months (95% CI 1.8–6.7) from 24.7 to 28.9 months in trials with similar CT compliance. In trials with different CT compliance, “earlier or shorter” radiotherapy shortened the 5-year mean survival time by 3.1 months (95% CI 1.3–4.9) from 20.6 to 17.5 months.

Compliance with chemotherapy and progression-free survival

The HR for progression-free survival favours trials in which “earlier or shorter” radiotherapy was delivered with similar CT compliance in both arms (HR for similar CT compliance: 0.81, 95% CI 0.71–0.92; for different CT compliance: 1.12, 0.95–1.31) (Figure 3). In trials in which CT compliance was similar, “earlier or shorter” radiotherapy increased the 3-year progression-free survival by 6.3% (95% CI 1.0–11.6%) and the 5-year progression-free survival by 5.6% (0.7–10.5%) (Figure S4).

Compliance with chemotherapy and landmark analysis

As the observed effect of CT compliance may be due to early treatment interruption because of progression or death, a post-hoc landmark analysis on the impact of individual CT compliance on overall survival and progression-free survival was performed among patients who survived (or had no disease progression) for at least 120 days. This landmark was chosen because most of the patients finished their chemo-radiation treatment at 120 days. Patients with good CT compliance, i.e. those receiving the planned total number of chemotherapy cycles had higher overall survival and progression-free survival than those with poor CT compliance (HR: 0.56, 95% CI 0.49–0.64 and 0.70, 0.59–0.83 respectively; Table S8).

Sub-group analyses

When the two subsets of trials with similar and different CT compliance were considered separately, no variation in the treatment effect was seen according to age, sex or the performance status (Figure S5).

Sensitivity analyses

Table S9 shows the results of pre-planned sensitivity analyses after excluding some trials. The results were similar to those of the main analysis, in particular to those related to chemotherapy compliance.

Toxicity

Three types of severe acute toxicities were significantly more frequent in patients receiving “earlier or shorter” thoracic radiotherapy: neutrophil, oesophageal and cardiac toxicity (Table 3) [35]. The toxicity odds ratios according to trial subsets based on CT compliance are shown in Table S10. We did not perform analyses on late toxicities as IPD were available only for two trials [26,27].

DISCUSSION

Based on this IPD meta-analysis of nine trials evaluating the optimal timing of thoracic radiotherapy in SCLC, overall there was no survival difference between “earlier or shorter” and “later or longer” thoracic radiotherapy (HR=0.99; $p=0.78$). As individual trials favoured either “earlier or shorter” or “later or longer” thoracic radiotherapy, it seemed relevant to further analyse these data and perform a subset analysis focusing on CT compliance. For trials with different CT compliance, in which lower compliance was always observed in the “earlier or shorter” arm, “earlier or shorter” delivery had a deleterious effect on survival compared to “later or longer” radiotherapy (HR 1.19, 95% CI 1.05–1.34). For trials that had similar (and good, i.e. at least 79% of compliant patient per arm) CT compliance, “earlier or shorter” delivery of thoracic radiotherapy improved overall survival (HR 0.79, 0.69–0.91). “Earlier or shorter” thoracic radiotherapy, when delivered with similar and good CT compliance, yielded an absolute survival gain of 5.7% at 3 years and 7.7% at 5 years compared with “later or longer” thoracic radiotherapy. Similar results were found for progression-free survival. We performed sensitivity analyses by only taking into account trials in which patients received concomitant chemoradiation and trials that exclusively addressed the timing of thoracic radiotherapy in their design. In these sensitivity analyses, the survival gain of delivering “earlier or shorter” thoracic radiotherapy with similar CT compliance remained significant (Table S9). Using a landmark analysis it was possible to confirm with IPD that good CT compliance was associated with longer survival. Of note, there was a significant association at patient-level between RT compliance and CT compliance which could explain our results.

Hyperfractionated accelerated radiotherapy also improved survival when delivered “earlier or shorter”, but this finding was driven by two large trials, JCOG9104 [28] and ECOG3588 [16], with good CT compliance. In the ECOG3588 trial [16], no dose adjustment was allowed for the first two cycles. Cisplatin-based chemotherapy seems to be more beneficial when combined with “earlier or shorter” thoracic radiotherapy. Issues such as the total radiotherapy dose and the dose per fraction are more difficult to interpret, because they are tightly correlated (Tables 1 and 2).

“Earlier or shorter” thoracic radiotherapy was associated with a higher incidence of acute severe esophagitis than “later or longer” radiotherapy (OR 1.93 [1.45–2.56]), but had no consequence on compliance with either chemotherapy or radiotherapy. Mauguen et al [15] also showed that hyperfractionated accelerated radiotherapy increased oesophageal toxicity. In this IPD meta-analysis, neutropenia was more frequent with “earlier or shorter” radiotherapy (OR 1.54, 95% CI: 1.19–2.00) and this effect was observed exclusively in trials with similar CT compliance (Table S10). Acute severe pulmonary toxicity was similar in “earlier or shorter” or “later or longer” thoracic radiotherapy groups, while acute severe cardiac toxicity was higher when “earlier or shorter” radiotherapy was delivered (OR 3.12, 1.46–6.68). The latter finding should be interpreted with caution, for it is based on only 26 cardiac events occurring in 1,648 patients among whom this toxicity was documented.

The results of this IPD meta-analysis primarily reinforce the evidence that chemotherapy should be delivered as intended whenever possible [1,36]. Cisplatin-based chemotherapy administered with good CT compliance appeared to be the best treatment when combined with “earlier or shorter” thoracic radiotherapy as all the three trials [16,26,28] with similar CT compliance used this regimen. This is in line with previous literature-based meta-analyses [9–14] in particular that reported by Spiro et al [12] which focused on CT compliance (Table S1). Interestingly, a recently published randomised trial [37], where all patients had early hyperfractionated radiotherapy given concomitantly with the first cycle of etoposide, showed a 5-year survival rate of 34.3% that the authors attributed to better patient selection and radiotherapy quality control. It will be interesting to observe the results of the on-going CALGB 30610 (NCT00632853) and the completed CONVERT (NCT00433563) randomised

trials comparing early hyperfractionated radiotherapy to early standard radiotherapy with a higher total dose and concomitant cisplatin plus etoposide in both arms.

The present IPD meta-analysis has some shortcomings. First, the trials were conducted at a time when imaging was not as advanced as it is today. However, the observed 5-year survival rate of about 25%, when “earlier or shorter” thoracic radiotherapy was combined with good chemotherapy compliance, remains among the best published results. These results continue to support their applicability today, as there has been no major change in the standard of care of SCLC (NCCN and ESMO guidelines) [6]. A recently published Korean phase III trial [38], which was not included in this meta-analysis as it was closed to accrual in 2010 (Table S2), showed a similar 5-year survival rate of approximately 24%. This trial did not show a significant difference in terms of overall survival between the two arms (HR 0.93, 0.67–1.29), but the study included only 222 patients. Secondly, data were not available for two other trials [32,34] (Table S2). However, when we included these three trials for which we have only published data (two in the similar CT compliance group [34,38] and one in different CT compliance group [32]) in a post-hoc analysis, we found similar effects on overall survival (HR 0.81, 95% CI 0.72–0.90 vs. 1.18, 1.06–1.32 for similar and different CT compliance subsets respectively). Third, only the number of chemotherapy cycles administered were available, but not doses or delays in treatment. However, consistency across endpoints and between the main analysis and sensitivity analyses underscore the robustness of our results. Another limitation is that data on long-term toxicity were not available, but less toxicity would be expected with the newer radiotherapy techniques. Lastly, the quality of radiotherapy could not be addressed in this meta-analysis as it was not explored in the studies included.

To improve the still dismal prognosis of patients with stage I-IIIB SCLC, we postulate that the optimal treatment should be full-dose but acceptable chemotherapy combined with “earlier or shorter” thoracic radiotherapy (i.e. before 9 weeks) preferably within a short overall treatment time. Our IPD meta-analysis provides the best evidence of the beneficial effect of “earlier or shorter” radiotherapy when chemotherapy is administered with good compliance.

Table 1. Description of trials

Trials	Inclusion period	Initiation of thoracic radiation (day)	RT dose (Gy)/ fraction/ duration (wks)	CT (mg/m²)	Nb CT cycles (before RT, during RT, after RT)	Nb of patients randomised^s	Median follow-up (years)
Earlier vs. Later Radiotherapy							
CALGB8083 [25]	1981-84	EoS: Day 1 LoL: Day 64	50 Gy / 24 fr / 5 wks	C: 1000 mg/m ² , every 3 wks V: 1.4 mg/m ² , every 3 wks E: 80 x 3 mg/m ² , every 3 wks Starting at cycle 7 for odd-numbered cycles: C: 1000 mg/m ² , every 3 wks V: 1.4 mg/m ² , every 3 wks A: 50 mg/m ² , every 3 wks	about 26 cycles “Earlier or shorter”: 2 cycles during RT, up to 24 cycles after RT “Later or longer”: 3 cycles before RT, 2 cycles during RT, up to 21 cycles after RT	292	17.2
BR.6 [26]	1985-88	EoS: Day 22 LoL: Day 106	40 Gy / 15 fr / 3 wks	“Earlier or shorter”: P: 25 mg/m ² x 3days, wks _{4,11,17} E: 100 mg/m ² x 3days, wks _{4,11,17} alternating with C: 1000 mg/m ² , wks _{1,8,14} A: 50mg/m ² , wks _{1,8,14} V: 2 mg, wks _{1,8,14} “Later or longer”: P: 25 mg/m ² x3days, wks _{4,10,16} E: 100 mg/m ² x3days, wks _{4,10,16} alternating with C: 1000 mg/m ² , wks _{1,7,13} A: 50mg/m ² , wks _{1,7,13} V: 2 mg, wks _{1,7,13}	“Earlier or shorter”: 6 cycles (1 before RT, 1 during RT, 4 after RT) “Later or longer”: 6 cycles (5 before RT, 1 during RT)	332	11.2
EORTC08877 [27]	1989-95	“Earlier or shorter”: Day 43 “Later or longer”: Day 99	“Earlier or shorter”: 12.5 Gy / 5 fr / 1 wk + break 3wks + 12.5 Gy / 5 fr / 1wk + break 3wks + 12.5 Gy / 5 fr / 1 wk + break 3wks + 12.5 Gy / 5 fr / 1 wk “Later or longer”: 50 Gy / 20 fr / 4 wks	“Earlier or shorter”: C: 1000 mg/m ² , wks _{1,5,9,13,17} A: 45 mg/m ² , wks _{1,5,9,13,17} E: 100x3 mg/m ² , wks _{1,5,9,13,17} “Later or longer”: C: 1000 mg/m ² , wks _{1,4,7,10,13} A: 45 mg/m ² , wks _{1,4,7,10,13} E: 100x3 mg/m ² , wks _{1,4,7,10,13}	“Earlier or shorter”: 5 cycles (1 before RT, 4 alternating with RT ^E) “Later or longer”: 5 cycles (5 before RT)	349	7.2
JCOG9104 [28]	1991-95	“Earlier or shorter”: Day 2 “Later or longer”: Day 85	45 Gy / 30 fr / 3 wks bid	“Earlier or shorter”: P: 80 mg/m ² , wks _{1,5,9,13} E: 100x3 mg/m ² , wks _{1,5,9,13} “Later or longer”: P: 80 mg/m ² , wks _{1,4,7,10} E: 100x3 mg/m ² , wks _{1,4,7,10}	“Earlier or shorter”: 4 cycles (1 during RT, 3 after RT) “Later or longer”: 4 cycles (4 before RT)	231	6.8
LLCG93 [12]	1993-99	“Earlier or shorter”: Day 22 “Later or longer”: Day 106	40 Gy / 15 fr / 3 wks	P: 25 x 3 mg/m ² , wks _{4,10,16} E: 100 x 3 mg/m ² , wks _{4,10,16} alternating with C: 1000 mg/m ² , wks _{1,7,13} A: 50 mg/m ² , wks _{1,7,13} V: 2 mg, wks _{1,7,13}	“Earlier or shorter”: 6 cycles (1 before RT, 1 during RT, 4 after RT) “Later or longer”: 6 cycles (5 before RT, 1 during RT)	325	5.3
HeCOG93 [29]	1993-99	“Earlier or shorter”: Day 1 “Later or longer”: Day 57	45 Gy / 30 fr / 3 wks bid	Cb: 6 AUC, wks _{1,4,7,10,13,16} E: 100 x 3 mg/m ² , wks _{1,4,7,10,13,16}	“Earlier or shorter”: 6 cycles (1 during RT, 5 after RT) “Later or longer”: 6 cycles (3 before RT, 1 during RT, 2 after RT)	81	11.8

Trials	Inclusion period	Start of thoracic radiation (day)	RT dose (Gy)/ fraction/ duration (wks)	CT (mg/m ²)	Nb CT cycles (before RT, during RT, after RT)	Nb of patients randomised ^s	Median follow-up (years)
Shorter vs. Longer Radiotherapy Duration							
CCCWFU62286 [30]	1987-92	“Earlier or shorter”: Day 1 “Later or longer”: Day 8	“Earlier or shorter”: 50 Gy / 25 fr / 5 wks “Later or longer”: 20 Gy / 8 fr / 2 wks + break 1wk + 20 Gy / 8 fr / 2 wks + break 1wk + 10 Gy / 4 fr / 1 wks	C: 750 mg/m ² , wks _{7,10,16} A: 60 mg/m ² , wks _{7,10,16} V: 2 mg, wks _{7,10,16} alternating with P: 60 mg/m ² , wks _{1,4,13} E: 120 x 3 mg/m ² , wks _{1,4,13}	“Earlier or shorter”: 6 cycles (2 during RT, 4 after RT) “Later or longer”: 6 cycles (3 alternating with RT, 3 after RT)	114	17.3
03PCL88 [31]	1988-94	“Earlier or shorter”: Day 30 “Later or longer”: Day 36	“Earlier or shorter”: 50 Gy / 20 fr / 5 wks “Later or longer”: 20 Gy / 8 fr / 2 wks + break 2wks + 20 Gy / 8 fr / 2 wks + break 2wks + 15 Gy / 6 fr / 1.5 wks	C: 1000 mg/m ² , wks _{1,13,17,21} A: 45 mg/m ² , wks _{1,13,17,21} E: 150 x 2 mg/m ² , wks _{1,13,17,21} Alternating with C: 1000 mg/m ² , wks _{5,9} Vd: 3 mg/m ² , wks _{5,9} E: 150 x 2 mg/m ² , wks _{5,9}	“Earlier or shorter”: 6 cycles (2 before RT, 1 during RT, 3 after RT) “Later or longer”: 6 cycles (2 before RT, 2 alternating with RT, 2 after RT)	164	6.5
ECOG3588 [16]	1989-92	Both arms: Day 1	“Earlier or shorter”: 45 Gy / 30 fr / 3 wks bid “Later or longer”: 45 Gy / 25 fr / 5 wks	P: 60 mg/m ² , wks _{1,4,7,10} E: 120 x 3 mg/m ² , wks _{1,4,7,10}	4 cycles (2 during RT, 2 after RT)	417	13.0

Trials are chronologically ordered within each category of trials (earlier vs. later RT, and shorter vs. longer RT).

^s Number of patients analysed equals the number of patients randomised, except for the HeCOG trial for which data on 81 patients were available out of the 86 randomised patients.

[£] Publication [27] stated that: “RT started [...] on the 14th day of the second and subsequent courses of chemotherapy in arm Earlier RT”

Abbreviations: **bid** = RT given twice a day; **CT** = chemotherapy; **EoS** = “Earlier or shorter” radiotherapy; **fr** = fraction; **Gy** = Gray; **LoL** = “Later or longer” radiotherapy; **RT** = Radiotherapy; **wks** = weeks;

A = Adriamycin; **C** = Cyclophosphamide; **Cb** = Carboplatin; **E** = Etoposide; **P** = Cisplatin; **V** = Vincristine; **Vd** = Vindesine;

BR = Bronchus; **CALGB** = Cancer and Leukaemia Group B; **CCCWFU** = Comprehensive Cancer Centre of Wake Forest University; **ECOG** = Eastern Cooperative Oncology Group; **EORTC** = European Organisation for Research and Treatment of Cancer; **HeCOG** = Hellenic Cooperative Oncology Group; **JCOG** = Japan Clinical Oncology Group; **LLCG** = London Lung Cancer Group; **PCL** = Petites Cellules Limitées

Table 2. Effect of “earlier or shorter” radiotherapy versus “later or longer” radiotherapy on overall survival according to different trial subsets

Trials Characteristics		HR [95% CI] [£]	Heterogeneity [§]	
			Between-subset	Residual (or within-subset)
CT compliance between arms	Similar	0.79 [0.69–0.91]	19.5***	1.9
	Different	1.19 [1.05–1.34]		
RT dose per fraction	< 1.8 Gy	0.82 [0.71–0.96]	7.5* ($p_{trend} = 0.02$) ^{\$\$}	14.1
	1.8- 2.4 Gy	1.11 [0.90–1.35]		
	> 2.4 Gy	1.06 [0.94–1.20]		
Type of RT	Hyperfractionated	0.82 [0.71–0.96]	7.4**	14.2
	Standard	1.07 [0.96–1.19]		
RT Overall Treatment Time	≤ 30 days in both arms	0.89 [0.78–1.02]	5.6 ($p_{trend} = 0.02$) ^{\$\$}	16.0
	One arm ≤ 30 days, one > 30 days	0.99 [0.85–1.15]		
	> 30 days in both arms	1.16 [0.98–1.38]		
Platin-based CT during RT in both arms	Yes	0.89 [0.79–1.01]	5.5**	16.1
	No	1.09 [0.97–1.24]		
Concurrent CT in both arms	Yes	0.95 [0.85–1.06]	1.5	20.1
	No	1.06 [0.92–1.22]		
Same RT in the 2 arms	Yes	0.96 [0.85–1.08]	0.5	21.1
	No	1.02 [0.90–1.16]		

* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$;

£ : Hazard ratio of death following “earlier or shorter” versus “later or longer” radiotherapy.

§ : Total heterogeneity is the sum of between-subset and residual (within-subset) heterogeneity, and is equal to 21.6 (analysis based on 9 trials) except for CT compliance 21.4 (8 trials). The test associated with between-subset heterogeneity corresponds to the interaction test. The lower residual heterogeneity, the greater overall heterogeneity of the treatment effect between trials was explained by the trial characteristic.

\$\$: test for trend;

Abbreviations: CI = Confidence Interval; CT = Chemotherapy; HR = Hazard Ratio; RT = Radiotherapy

Table 3. Acute toxicity according to radiotherapy arm

Severe toxicity (grade 3-5)	Availability	Toxicity rate		Results	p-value efficacy	I ²	p-value heterogeneity
	No. of trials (patients)	"Later or longer" RT	"Earlier or shorter" RT*	OR [95% CI]			
Neutrophil	6 (1,453)	59	69	1.54 [1.19–2.00]	0.001	79%	<0.001
Haemoglobin	6 (1,476)	21	24	1.17 [0.91–1.52]	0.22	31%	0.21
Platelets	7 (1,817)	18	21	1.22 [0.96–1.55]	0.11	45%	0.09
Oesophageal	8 (1,950)	8	14	1.93 [1.45–2.56]	<0.001	45%	0.08
Pulmonary	5 (1,207)	4	6	1.50 [0.86–2.62]	0.16	0%	0.68
Cardiac	6 (1,648)	1	3	3.12 [1.46–6.68]	0.003	0%	0.95

Toxicity was graded according to the National Cancer Institute Common Toxicity Criteria, World Health Organization criteria, or Eastern Cooperative Oncology Group Common Toxicity Criteria depending on the trials. Severe toxicity was defined as grade 3 to 5 toxicity. Grade 5 was present only for pulmonary toxicity ($n=4$) and cardiac toxicity ($n=1$).

*: The difference in the rate of toxicity between the two treatment arms was computed based on the rate in the “later or longer” radiotherapy arm and the odds ratio [35].

Abbreviations: CI = Confidence Interval; OR = Odds Ratio of the “earlier or shorter” RT arm compared with “later or longer” RT arm; RT = Radiotherapy

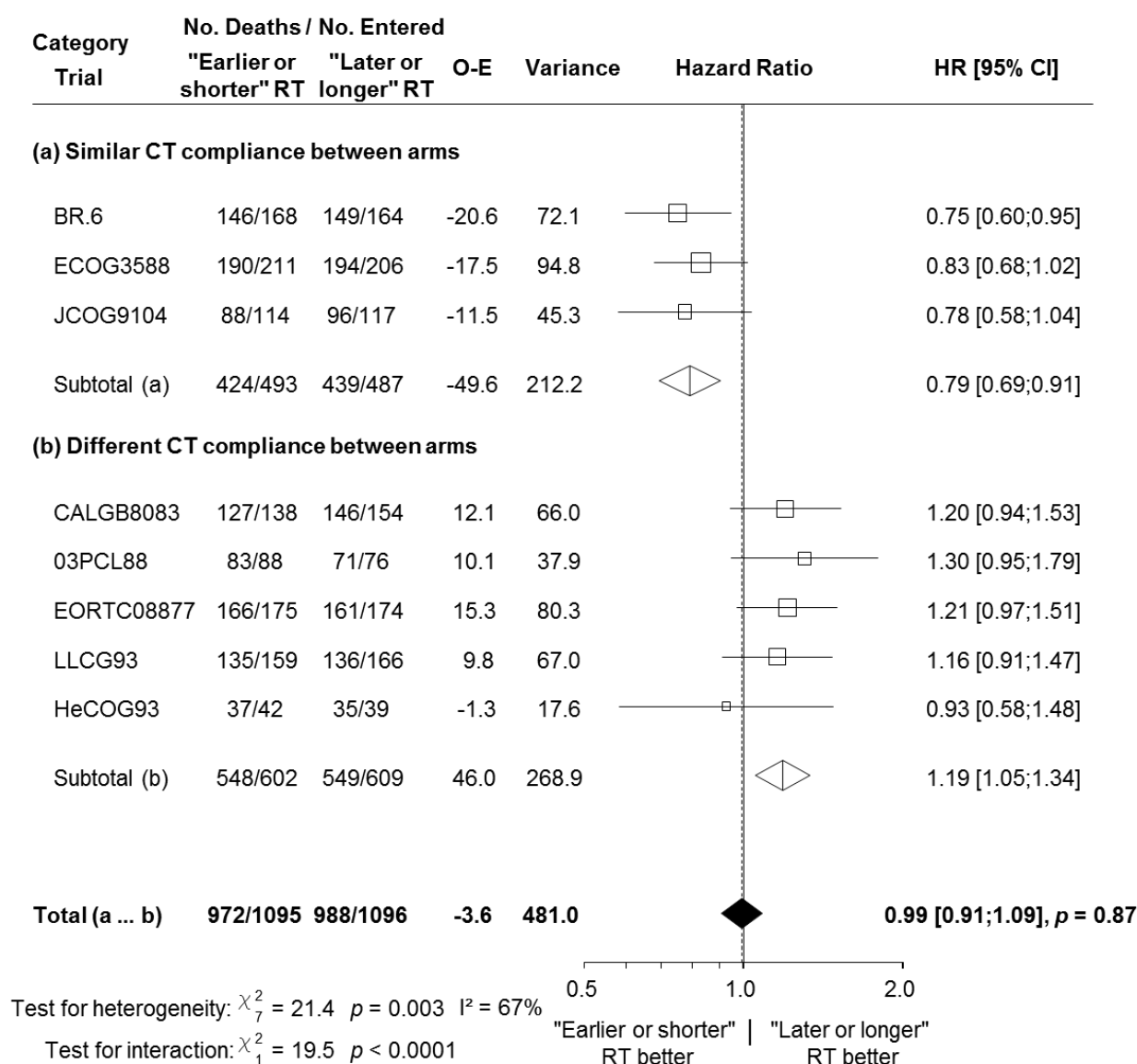
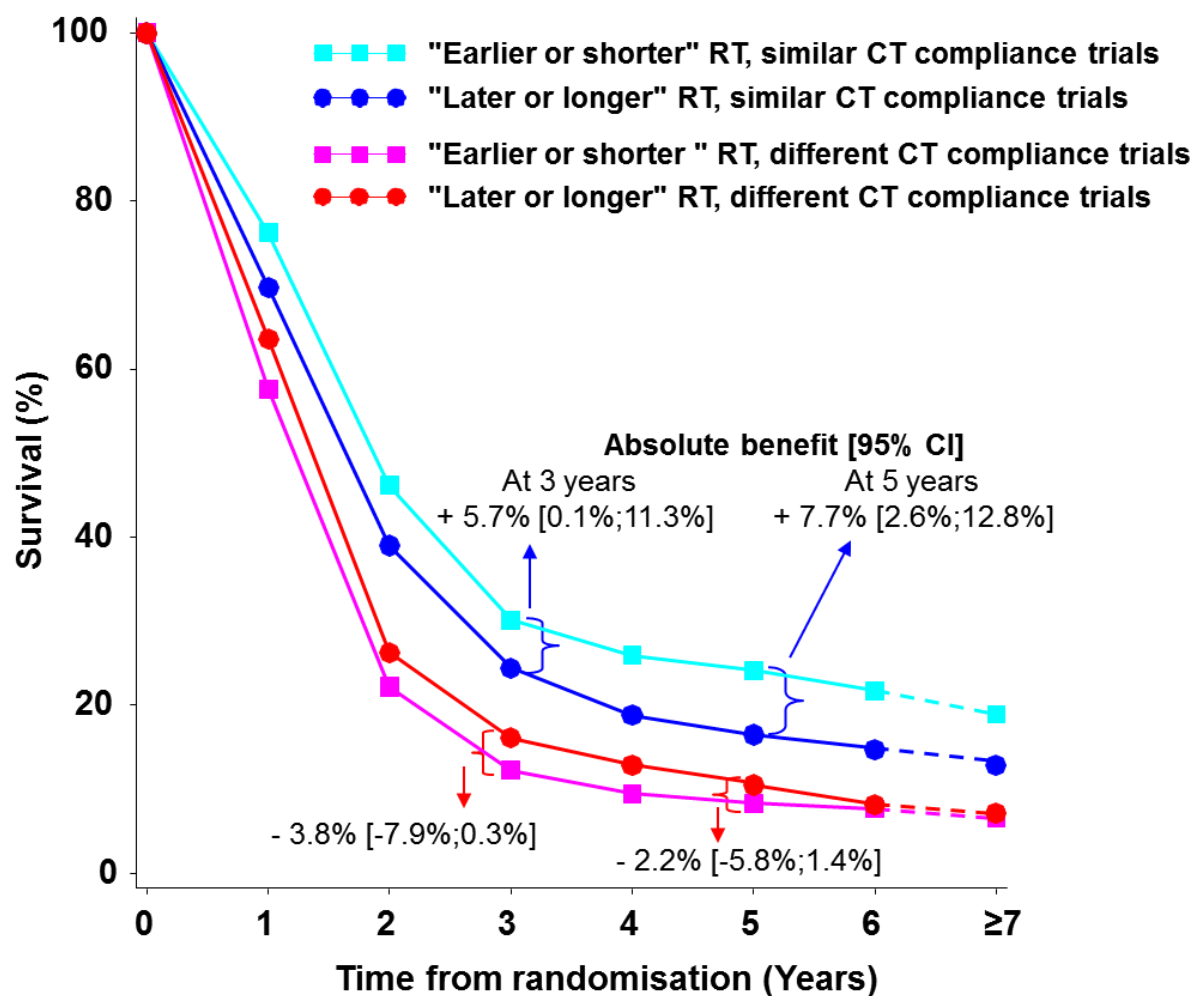


Figure 1. Effect of “earlier or shorter” radiotherapy versus “later or longer” radiotherapy on overall survival according to chemotherapy compliance

Each trial is represented by a square, the centre of which denotes the hazard ratio of death for that trial comparison with the horizontal lines showing the 95% confidence intervals (CIs). The size of the square is directly proportional to the amount of information contributed by the trial. The clear diamonds represent pooled hazard ratios for the trial groups and the black diamond the overall hazard ratio, with the centre denoting the hazard ratio and the extremities the 95% CI. The fixed effect model was used. Trials were chronologically ordered within each category of trials. Of note, data on CT compliance were not available for the CCCWFU62286 trial which is thus not included in this analysis.

Abbreviations: CI = Confidence Interval; CT = Chemotherapy; HR = Hazard ratio; O-E = Observed-Expected; RT = Radiotherapy



Number of deaths/ PY by period	Years 0-2	Years 3-5	Years ≥ 6
Similar CT compliance			
"Earlier or shorter" RT	262 / 735	107 / 437	55 / 425
"Later or longer" RT	302 / 575	104 / 319	33 / 263
Different CT compliance			
"Earlier or shorter" RT	462 / 675	69 / 175	17 / 133
"Later or longer" RT	441 / 760	82 / 239	26 / 152

Figure 2. Survival curves for overall survival according to chemotherapy compliance

Abbreviations: CI = Confidence Interval; CT = Chemotherapy; HR = Hazard ratio; PY = Person-Year; RT = Radiotherapy

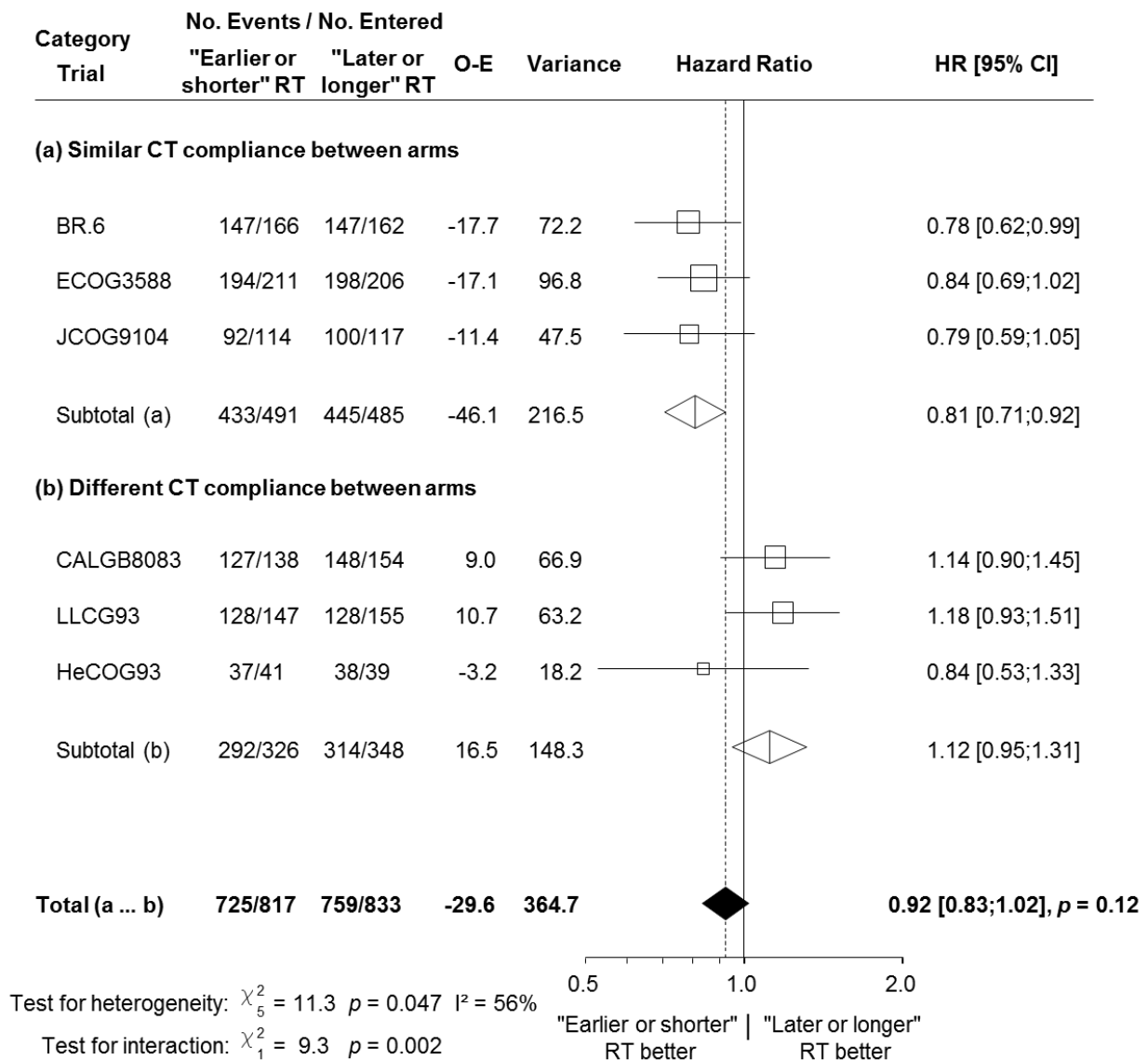


Figure 3. Effect of “earlier or shorter” radiotherapy versus “later or longer” radiotherapy on progression-free survival according to chemotherapy compliance

Each trial is represented by a square, the centre of which denotes the hazard ratio of death or tumour progression for that trial comparison with the horizontal lines showing the 95% confidence intervals (CIs). The size of the square is directly proportional to the amount of information contributed by the trial. The clear diamonds represent pooled hazard ratios for the trial groups and the black diamond the overall hazard ratios, with the centre denoting the hazard ratio and the extremities the 95% CI. The fixed effect model was used.

Abbreviations: CI = Confidence Interval; CT = Chemotherapy; HR = Hazard ratio; O-E = Observed-Expected; RT = Radiotherapy

ACKNOWLEDGEMENTS

We thank the patients and the clinical investigators who took part in the trials and contributed to this research. The meta-analysis would not have been possible without their participation or without the collaborating institutions or groups that provided their trial data: The Alliance for Clinical Trials in Oncology (formerly Cancer and Leukemia Group B), Comprehensive Cancer Centre of Wake Forest University, ECOG-ACRIN Cancer Research Group (formerly Eastern Cooperative Oncology Group), European Organization for Research and Treatment of Cancer, Hellenic Cooperative Oncology Group, Japan Clinical Oncology Group, London Lung Cancer Group, National Cancer Institute of Canada Clinical Trial Group, “Petites Cellules” Group. We are grateful to Lorna Saint-Ange for editorial assistance.

RTT-SCLC Collaborative Group

Secretariat

D De Ruyscher, C Le Pechoux, B Lueza, E Paris, JP Pignon, M Pijls-Johannesma, AS Veillard

Advisory board

Rodrigo Arriagada, Paul Baas, Hak Choy, Allan Price, Lesley Seymour.

Investigators

Rodrigo Arriagada (Gustave Roussy, Villejuif, France; Karolinska Institutet, Stockholm, Sweden), Paul Baas (Netherlands Cancer Institute – Antoni van Leeuwenhoek Ziekenhuis, Amsterdam, The Netherlands), William Blackstock (Wake Forest University School of Medicine, Winston-Salem, NC, USA), Sylvie Chevret (CRESS - UMR 1153, Inserm, Paris Diderot University, Paris, France), Hak Choy (University of Texas Southwestern, Dallas, Texas, USA), Jeffrey Crawford (Duke University School of Medicine, Durham, NC, USA), Urania Dafni (University of Athens, Greece), Suzanne Dahlberg (Dana-Farber Cancer Institute, Boston, MA, USA), Dirk De Ruyscher (Maastricht University Medical Center, Maastricht, the Netherlands; KU Leuven, Leuven, Belgium), Allan Hackshaw (University College London, UK), Baktiar Hasan (EORTC data center, Brussels, Belgium), David H. Johnson (UT Southwestern University School of Medicine, Dallas, USA), Cécile Le Pechoux (Gustave Roussy, Villejuif, France), Bernard Lebeau (Hôpital St Antoine, Paris, France), James Lovato (Wake Forest University Health Sciences, Winston-Salem, NC, USA), Béranger Lueza (Gustave Roussy, Villejuif, France), Nevin Murray (British Columbia Cancer Agency, Vancouver, Canada), Mary O'Brien (Royal Marsden Hospital, London, UK), Emmanuelle Paris (Gustave Roussy, Villejuif, France), Jean-Pierre Pignon (Gustave Roussy, Villejuif, France), Madelon Pijls-Johannesma (MAASTRO clinic, Maastricht, the Netherlands), Allan Price (University of Edinburgh, Edinburgh, UK), Lesley Seymour (NCIC-CTG, Kingston, Ontario, Canada), Taro Shibata (JCOG Data Center, National Cancer Center Coordinating, Tokyo, Japan), Dimosthenis Skarlos (Metropolitan Hospital N. Falirio, Athens, Greece), Stephen Spiro (University College London Hospital, London, UK), Minoru Takada (Osaka Prefectural Habikino Hospital, Osaka, Japan), Anne-Sophie Veillard (Gustave Roussy, Villejuif, France), Xiaofei Wang (Alliance Data and Statistical Center, NC, USA)

FUNDING

The meta-analysis was funded by the French National Cancer Institute (Programme Hospitalier de Recherche Clinique), the Ligue Nationale Contre le Cancer, and partly by Sanofi-Aventis (unrestricted grants). The investigators meeting was funded also by Gustave Roussy, Lilly and Astra-Zeneca (unrestricted grants). No grant number is applicable.

DISCLOSURE

Consultant or Advisory Role: David H Johnson, Peloton Therapeutics/miRNA Therapeutics; Paul Baas, Merck Sharp Dohme/Bristol-Myers Squibb; Lesley Seymour, Boehringer Ingelheim

Stock Ownership: Lesley Seymour, AstraZeneca

Honoraria: Paul Baas, AstraZeneca/Verastem; Lesley Seymour, Innate Pharma

Research Funding: Paul Baas, Merck Sharp Dohme/Bristol-Myers Squibb; Lesley Seymour, Pfizer, AstraZeneca, Astex Pharmaceuticals

Travel, Accommodations, Expenses: Paul Baas, Merck Sharp Dohme

All remaining authors have declared no conflicts of interest.

REFERENCES

1. van Meerbeeck JP, Fennell DA, De Ruysscher DK. Small-cell lung cancer. *Lancet* 378:1741–55, 2011
2. Shepherd FA, Crowley J, Van Houtte P, et al; International Association for the Study of Lung Cancer International Staging Committee and Participating Institutions. The International Association for the Study of Lung Cancer lung cancer staging project: proposals regarding the clinical staging of small cell lung cancer in the forthcoming (seventh) edition of the tumor, node, metastasis classification for lung cancer. *J Thorac Oncol* 2:1067–77, 2007
3. Arriagada R, Le Chevalier T, Pignon JP, Rivi  re A, Monnet I, Chomy P, Tuchsais C, Tarayre M, Ruffi   P. Initial chemotherapy doses and survival in limited small cell lung cancer. *N Engl J Med* 329:1848–52, 1993
4. Pignon JP, Arriagada R, Ihde DC, et al. A meta-analysis of thoracic radiotherapy for small-cell lung cancer. *N Engl J Med* 327:1618–24, 1992
5. Stahel R, Thatcher N, Fruh M, et al. 1st ESMO Consensus Conference in lung cancer; Lugano 2010: small-cell lung cancer. *Ann Oncol* 22:1973–80, 2011
6. NCCN Clinical Practice Guidelines in Oncology: Small Cell Lung Cancer, version 1.2015; http://www.nccn.org/professionals/physician_gls/pdf/sclc.pdf (accessed on October 29th, 2014).
7. Aup  rin A, Arriagada R, Pignon JP, et al. Prophylactic cranial irradiation for patients with small-cell lung cancer in complete remission. *N Engl J Med* 341:476–84, 1999
8. Le P  choux C, Dunant A, Senan S, et al. Standard-dose versus higher-dose prophylactic cranial irradiation (PCI) in patients with limited-stage small-cell lung cancer in complete remission after chemotherapy and thoracic radiotherapy (PCI 99-01, EORTC 22003-08004, RTOG 0212, and IFCT 99-01): a randomised clinical trial. *Lancet Oncol* 10:467–74, 2009
9. Pijls-Johannesma MC, De Ruysscher D, Lambin P, et al. Early versus late chest radiotherapy for limited stage small cell lung cancer. *Cochrane Database Syst Rev* 1:1–40, 2005
10. De Ruysscher D, Pijls-Johannesma M, Vansteenkiste J, et al. Systematic review and meta-analysis of randomised, controlled trials of the timing of chest radiotherapy in patients with limited-stage, small-cell lung cancer. *Ann Oncol* 17:543–52, 2006
11. Pijls-Johannesma M, De Ruysscher D, Vansteenkiste J, et al. Timing of chest radiotherapy in patients with limited stage small cell lung cancer: a systematic review and meta-analysis of randomised controlled trials. *Cancer Treat Rev* 33:461–73, 2007
12. Spiro SG, James LE, Rudd RM, et al. Early compared with late radiotherapy in combined modality treatment for limited disease small-cell lung cancer: a London Lung Cancer Group multicenter randomized clinical trial and meta-analysis. *J Clin Oncol* 24:3823–30, 2006
13. Fried DB, Morris DE, Poole C, et al. Systematic review evaluating the timing of thoracic radiation therapy in combined modality therapy for limited-stage small-cell lung cancer. *J Clin Oncol* 22:4837–45, 2004. Erratum in: *J Clin Oncol* 23:248, 2005

14. Huncharek M, McGarry R. A meta-analysis of the timing of chest irradiation in the combined modality treatment of limited-stage small cell lung cancer. *The Oncologist* 9:665–72, 2004
15. Mauguen A, Le Péchoux C, Saunders MI, et al. Hyperfractionated or accelerated radiotherapy in lung cancer: an individual patient data meta-analysis. *J Clin Oncol* 30:2788–97, 2012
16. Turrisi AT, Kim K, Blum R, et al. Twice-daily compared with once-daily thoracic radiotherapy in limited small-cell lung cancer treated concurrently with cisplatin and etoposide. *N Engl J Med* 340:265–71, 1999
17. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials* 17:343–346, 1996
18. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 21:1539–58, 2002
19. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 7: 177–88, 1986
20. Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic or immune therapy: 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women. *Lancet* 339:1–15, 1992
21. Wei Y, Royston P, Tierney JF, Parmar MKB. Meta-analysis of time-to-event outcomes from randomized trials using restricted mean survival time: application to individual participant data. *Stat Med* 34:2881–98, 2015
22. Lueza B, Mauguen A, Pignon JP, Rivero-Arias O, Bonastre J. Difference in Restricted Mean Survival Time for Cost-Effectiveness Analysis Using Individual Patient Data Meta-Analysis: Evidence from a Case Study. *PLoS One* 11: e0150032. doi:10.1186/1471-2288-14-72, 2016
23. Lueza B, Rotolo F, Bonastre J, Pignon JP, Michiels S. Bias and precision of methods for estimating the difference in restricted mean survival time from an individual patient data meta-analysis. *BMC Med Res Meth* 16:37. doi: 10.1186/s12874-016-0137-z, 2016
24. Non-small Cell Lung Cancer Collaborative Group. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. *BMJ* 311:899–909, 1995
25. Perry MC, Herndon JE, Eaton WL, et al. Thoracic radiation therapy added to chemotherapy for small-cell lung cancer: an update of Cancer and Leukemia Group B Study 8083. *J Clin Oncol* 16:2466–67, 1998
26. Murray N, Coy P, Pater JL, et al. Importance of timing for thoracic irradiation in the combined modality treatment of limited-stage small-cell lung cancer. The National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* 11:336–44, 1993
27. Gregor A, Drings P, Burghouts J, et al. Randomized trial of alternating versus sequential radiotherapy/chemotherapy in limited-disease patients with small-cell lung cancer: a European Organization for Research and Treatment of Cancer Lung Cancer Cooperative Group Study. *J Clin Oncol* 15:2840–9, 1997
28. Takada M, Fukuoka M, Kawahara M, et al. Phase III study of concurrent versus sequential thoracic radiotherapy in combination with cisplatin and etoposide for limited-stage small-cell lung cancer: results of the Japan Clinical Oncology Group Study 9104. *J Clin Oncol* 20:3054–60, 2002
29. Skarlos DV, Samantas E, Briassoulis E, et al. Randomized comparison of early versus late hyperfractionated thoracic irradiation concurrently with chemotherapy in limited disease small-cell lung cancer: a randomized phase II study of the Hellenic Cooperative Oncology Group (HeCOG). *Ann Oncol* 12:1231–8, 2001

30. Blackstock AW, Bogart JA, C. Matthews C, et al. Split-course versus continuous thoracic radiation therapy for limited-stage small-cell lung cancer: final report of a randomized phase III trial. *Clin Lung Cancer* 6:287–92, 2005
31. Lebeau B, Urban T, Brechot JM, et al. A randomized clinical trial comparing concurrent and alternating thoracic irradiation for patients with limited small cell lung carcinoma. “Petites Cellules Group”. *Cancer* 86:1480–7, 1999
32. Work E, Nielsen OS, Bentzen SM, et al. Randomized study of initial versus late chest irradiation combined with chemotherapy in limited-stage small-cell lung cancer. Aarhus Lung Cancer Group. *J Clin Oncol* 15: 3030–7, 1997
33. Park SK, Kim GH, Jeong SS, et al. The effects according to the timing of thoracic radiotherapy in limited stage small cell lung cancer. *Tuberc Respir Dis* 43:903–15, 1996
34. Jeremic B, Shibamoto Y, Acimovic L, et al. Initial versus delayed accelerated hyperfractionated radiation therapy and concurrent chemotherapy in limited small-cell lung cancer: a randomized study. *J Clin Oncol* 5:893–900, 1997
35. Stewart L, Parmar M. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 341:25–8, 1993
36. Pelayo AM, Gallego RÓ, Bonfill CX, Agra VY. Chemotherapy versus best supportive care for extensive small cell lung cancer. *Cochrane Database Syst Rev* 4:1–44, 2009
37. Kubota K, Hida T, Ishikura S, et al for the Japan Clinical Oncology Group. Etoposide and cisplatin versus irinotecan and cisplatin in patients with limited-stage small-cell lung cancer treated with etoposide and cisplatin plus concurrent accelerated hyperfractionated thoracic radiotherapy (JCOG0202): a randomised phase 3 study. *Lancet Oncol* 15:106–13, 2014
38. Sun JM, Ahn YC, Choi EK, et al. Phase III trial of concurrent thoracic radiotherapy with either first- or third-cycle chemotherapy for limited-disease small-cell lung cancer. *Ann Oncol* 24:2088–92, 2013

Supplementary Material – Online Only

Web-Appendix 1: Search strategy

An initial search was done in 2007, and then updated. The search in PubMed MEDLINE was conducted on January 11th 2007 and yielded 905 results (+138 references from other sources). The search conducted on February 21st 2014, yielded 1,285 results in Pubmed. Only one new trial (Sun et al. 2013) was considered appropriate to be included in the meta-analysis and therefore a sensitivity analysis including this trial was performed. The initial search is described below.

Electronic Databases

A modified search strategy for retrieving randomised controlled trials (RCTs) from Medline, developed by the Cochrane Collaboration^{A1} was used.

In addition, the following electronic bibliographic databases were searched: The Cochrane Central Register of Controlled Trials (CENTRAL) and the Proceedings of ASCO 1995 – 2007.

Trial Registers

Trial registers used to supplement searches of electronic databases with trials that may not (yet) be published or still recruiting patients were: UKCCCR Trials Register, ClinicalTrials.gov, Physicians Data Query Protocols (open and closed) and the Current Controlled Trials 'metaRegister' of controlled trials

Hand Searches

The following hand searches were carried out to identify trials that may have only been reported as abstracts or that might have been missed in the searches described above:

- Proceedings of the American Society for Clinical Oncology (ASCO) 1993-2007
- Proceedings of the World Lung Cancer Conference 1997-2007
- Proceedings of the European Society of Medical Oncology (ESMO) 1996 - 2007
- Proceedings of the European Cancer Conference Organisation (ECCO) 1995 - 2007
- Bibliographies of all identified trials and review articles were hand searched

Experts in the field

We asked all participating trialists to review and supplement a provisional list of trials.

Search Strategy for Medline

("Lung Neoplasms/radiotherapy"[MeSH] AND Randomized Controlled Trial[ptyp])
 OR ("Lung Neoplasms/radiotherapy"[MeSH] AND "Randomized Controlled Trials"[MeSH Terms])
 OR ("Lung Neoplasms/radiotherapy"[MeSH] AND random*[Title/Abstract])

OR ("Lung Neoplasms/radiotherapy"[MAJR] AND Randomized Controlled Trial[ptyp])
 OR ("Lung Neoplasms/radiotherapy"[MAJR] AND "Randomized Controlled Trials"[MeSH Terms])
 OR ("Lung Neoplasms/radiotherapy"[MAJR] AND random*[Title/Abstract])

OR ("Lung Neoplasms"[MAJR] AND (radiother*[Title] OR radiat*[Title]) AND Randomized Controlled Trial[ptyp])
 OR ("Lung Neoplasms"[MAJR] AND (radiother*[Title] OR radiat*[Title]) AND "Randomized Controlled Trials"[MeSH Terms])
 OR ("Lung Neoplasms"[MAJR] AND (radiother*[Title] OR radiat*[Title]) AND random*[Title/Abstract])

OR (lung[Title] AND (radiother*[Title] OR radiat*[Title]) AND Randomized Controlled Trial[ptyp])
 OR (lung[Title] AND (radiother*[Title] OR radiat*[Title])] AND "Randomized Controlled Trials"[MeSH Terms])
 OR (lung[Title] AND (radiother*[Title] OR radiat*[Title]) AND random*[Title/Abstract])

AND ("1980"[PDAT] : "3000"[PDAT])

Web-Appendix 2: Individual patient data, data collection and quality control

For all trials, survival, compliance and toxicity data, as well as baseline characteristics were sought for all patients randomised in each trial. An updated follow-up was requested in order to report on both short and longer-term outcomes.

The following items were collected: patient identifier (preferably not patient name), date of birth or age at randomisation, sex, performance status, tumour TNM (or stage (extensive versus limited stage) if TNM was not available), involvement of lymph nodes: no versus supra-clavicular/mediastinal lymph nodes, date of randomisation, treatment allocated (specify each arm of radiotherapy), date of start of chemotherapy, number of chemotherapy cycles received, compliance with chemotherapy: radiotherapy started / not started, date first day chest radiotherapy, date last day chest radiotherapy, total administered dose of radiotherapy, number of fractions of radiotherapy, once-daily (QD) vs. twice-daily (BID) schedule, if BID schedule: Time between the two daily fractions, chest radiotherapy concurrent with chemotherapy (yes/ no), Prophylactic Cranial Irradiation (PCI: yes/ no), survival status, date of last follow-up, vital status, cause of death, local recurrence status, date of local recurrence, distant recurrence status, date of distant recurrence, second malignancy status, date of second malignancy, acute toxicity (neutrophil, haemoglobin, platelets, cardiac, oesophageal and pulmonary) and specification of toxicity grading system used, late toxicity (oesophageal and pulmonary) and specification of toxicity grading system used, whether excluded from the trial analysis and reason for exclusion. Toxicity was graded according to the National Cancer Institute Common Toxicity Criteria, World Health Organisation criteria, or Eastern Cooperative Oncology Group Common Toxicity Criteria depending on the trials. Severe toxicity was defined as grade 3 to 5 toxicity.

Data were checked^{A2} for missing values and for data validity and consistency across variables, and were also compared with published results if there were any. To assess randomisation integrity, we looked for unusual patterns in the sequencing of allocation or imbalances in baseline characteristics between treatment arms. Patient follow-up was also assessed to ensure that it was well balanced between treatment arms, and as updated as possible. Any queries were resolved with the investigator responsible for the trial or statisticians.

We carried out all analyses on an intention-to-treat basis, i.e. all randomised patients were included and analysed according to the treatment allocated, irrespective of whether they actually received that treatment.

Web-Appendix 3: Restricted mean survival time

The mean survival time restricted up to a specific time point $t^* = 5$ years was estimated in each arm. It is defined as:

$$\mu_{arm}(t^*) = \int_0^{t^*} S(t) dt$$

where the survival function $S(t)$ can be substituted by the Kaplan-Meier estimate $\hat{S}(t)$.

Graphically, $\mu_{arm}(t^*)$ is equivalent to the area under the Kaplan-Meier survival curve restricted at t^* . The 5-year mean survival time can thus take a maximum value of 5 years, if no patient died during the first 5 years. It is noteworthy that this parameter represents an average survival time (i.e. life expectancy) and is different from the median survival time.

The difference in the restricted mean survival time (rmstD) between two arms is often used in economic evaluation to quantify a treatment effect as this measure is still valid with the presence of non-proportional hazards, unlike the common hazard ratio. The difference in 5-year mean survival times can range from a 5-year survival gain to a 5-year survival loss, and can be equal to 0 which indicates the absence of a treatment effect.

To take into account the trial effect in this individual patient data meta-analysis, we estimated the difference in 5-year mean survival times as follows. First, we estimated rmstD in each trial as the area between trial-specific Kaplan-Meier curves. Second, we pooled the rmstDs using a fixed-effect meta-analysis model. This method has already been used by Wei and colleagues (as “Integrated difference of survival functions”) [21] and Lueza and colleagues (as “Pooled Kaplan-Meier”) [22,23].

References

- A1. Lefebvre C, Clarke MJ. Identifying Randomised Trials. In: Egger M, Smith GD, Altman DG, Eds. *Systematic Reviews in Healthcare*. 2nd edition. London: BMJ Publishing Group, 2002: 69–87.
- A2. Stewart LA, Clarke MJ on behalf of the Cochrane Working Group on meta-analyses using individual patient data. Practical methodology of meta-analyses (overviews) using updated individual patients data. *Stat Med* 1995; **14**: 2057–79.

Supplementary Tables and Figures – Online Only

Table S1: Previous literature-based meta-analyses on the timing of radiotherapy in small-cell lung cancer

Meta-Analysis	Number of trials	Trials	Number of patients	Endpoints	Results for Overall survival [95%CI]	Heterogeneity	Subset analysis
Fried – 2004 [13]	7	Work; CALGB8083; BR6; Jeremic; EORTC08877; JCOG9104; HeCOG93	1,524	Overall Survival (Risk Ratios and risk difference)	RR _{2-year} [□] = 0.85 [0.74–0.98] RR _{3-year} [□] = 0.88 [0.72–1.09]	<i>p</i> =0.17 at 2 yrs <i>p</i> =0.18 at 3 yrs	Hyperfract RT Platin CT Concurrent CT-RT
Huncharek – 2004 [14]	8	Work; CALGB8083; BR6; Jeremic; 03PCL88; JCO9104*; HeCOG93	1,574	Overall Survival (Odds ratio)	OR _{2-year} [□] = 0.63 [0.50–0.78] OR _{3-year} [□] = 0.67 [0.52–0.87]	<i>p</i> =0.24 at 2 yrs <i>p</i> =0.81 at 3 yrs	No
Pijls – 2005 [9-11] [§]	7	Work; CALGB8083; BR6; Jeremic; JCOG9104; LLCG93 [£] ; HeCOG93	1,514	Overall Survival (Odds ratio/Hazard ratio) Toxicity (Odds ratio)	OR _{2-year} = 0.84 [0.56–1.28] OR _{5-year} = 0.80 [0.47–1.38] HR = 0.91 [0.75–1.10]	<i>p</i> =0.006 at 2 yrs <i>p</i> =0.05 at 5 yrs <i>p</i> =0.01	Platin CT Overall treatment time of RT
Spiro – 2006 [12]	8	Work; CALGB8083; BR6; Jeremic; EORTC08877; JCOG9104; LLCG93; HeCOG93	1,849	Overall Survival (Hazard ratio)	HR = 0.96 [0.84–1.10]	<i>p</i> =0.002	CT compliance Hyperfract RT Platin CT

*: JCOG was included twice as “Goto” and “Takada” trials corresponding to the same Japan Clinical Oncology Group study [28]

§: references 9, 10 and 11 are publications with similar methodology and really close results

□: The risk ratio for the Fried meta-analysis [13] and the Odds ratio for the Huncharek meta-analysis [14] were recalculated in order to have the same reference treatment as that used in the other meta-analyses

£: The LLCG93 trial was called the “James 2003” study in the De Ruysscher et al. meta-analysis [10]

Abbreviations: CI = Confidence interval; CT = Chemotherapy; HR = Hazard ratio; OR = Odds ratio; RR = Risk ratio; RT = Radiotherapy;

BR = Bronchus; CALGB = Cancer and Leukaemia Group B; CCCWFU = Comprehensive Cancer Centre of Wake Forest University; ECOG = Eastern Cooperative Oncology Group; EORTC = European Organisation for Research and Treatment of Cancer; HeCOG = Hellenic Cooperative Oncology Group; JCOG = Japan Clinical Oncology Group; LLCG = London Lung Cancer Group; PCL = Petites Cellules Limitées

Table S2: Trials not included

Trials	Inclusion period	Included in previous meta-analysis	Included in our IPD meta-analysis	Start of chest radiation (day)	RT dose (Gy)/ fraction/ duration (wks)	Nb CT cycles (before RT, during RT, after RT)	Nb of patients randomised
Work [32]	1981-89	Fried 2004 Huncharek 2004 De Ruysscher 2006 Spiro 2006 Pijls 2007	In a sensitivity analysis	“Earlier or shorter”: Day 1 “Later or longer”: Day 127	20 Gy in 11 fractions (1.8 Gy) / 2 wks 3 wks break 22.5 Gy in 11 fractions (2 Gy) / 2 wks	“Earlier or shorter”: 9 cycles (1 alternating with RT, 8 after RT) “Later or longer”: 9 cycles (6 before RT, 1 alternating with RT, 2 after RT)	199
Jeremic [34]	1988-92	Fried 2004 Huncharek 2004 De Ruysscher 2006 Spiro 2006 Pijls 2007	In a sensitivity analysis	“Earlier or shorter”: Day 1 “Later or longer”: Day 36	54 Gy in 36 fractions (1.5 Gy bid) / 3.6 wks	“Earlier or shorter”: 4 cycles (1 injection x 5 per week during RT, 4 after RT) “Later or longer”: 4 cycles (2 before RT, 1 injection x 5 per week during RT, 2 after RT)	107
Park [33]*	Not mentioned in abstract	No	No	“Earlier or shorter”: Day 1 “Later or longer”: after 6 th CT cycle	“Earlier or shorter”: 45 Gy in 30 fractions (1.5 Gy bid) / 3 wks Late: 40-50 Gy in ? fractions / 5-6 wks	“Earlier or shorter”: 6 cycles (1 during RT, 5 after RT) “Later or longer”: 6 cycles (6 before RT)	51
Sun [38]	2010-13	No	In a sensitivity analysis	“Earlier or shorter”: Day 1 “Later or longer”: Day 43 (7 th wk)	52.5 Gy in 25 fractions (2.1 Gy) / 5 wks	“Earlier or shorter”: 4 cycles (4 during RT) “Later or longer”: 4 cycles (2 before RT, 2 during RT)	222

*: published in Korean with the abstract in English, publication not found

Abbreviations: **bid** = RT given twice a day; **CT** = chemotherapy; **fr** = fraction; **Gy** = Gray; **RT** = Radiotherapy; **wks** = weeks;

Table S3. Description of patient characteristics by arm

	<i>“Later or longer” RT</i>		<i>“Earlier or shorter” RT</i>	
	<i>N=1,153</i>		<i>N=1,152</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
Sex				
<i>Male</i>	775	67	745	65
<i>Female</i>	377	33	407	35
<i>Unknown</i>	1	0	0	0
Age Category				
<i>< 60</i>	474	41	474	41
<i>60-64</i>	256	22	294	26
<i>≥ 65</i>	419	36	381	33
<i>Unknown</i>	4	0	3	0
Performance Status (WHO/ECOG)				
<i>0</i>	449	39	417	36
<i>1</i>	602	52	622	54
<i>2/3</i>	91	7	104	9
<i>Unknown</i>	11	1	9	1
Disease extend^{\$}				
<i>Limited</i>	762	66	796	69
<i>Extensive</i>	17	2	4	0
<i>Unknown*</i>	374	32	352	31
Mediastinal nodes involved				
<i>No</i>	207	18	182	16
<i>Yes</i>	275	24	290	25
<i>Unknown[#]</i>	671	58	680	59
Supra-clavicular involvement				
<i>No</i>	477	41	437	38
<i>Yes</i>	100	9	126	11
<i>Unknown[□]</i>	576	50	589	51

N: number of patients

^{\$}: limited-stage disease was an inclusion criteria for all trials

^{*}: Data are missing for LLCG93, CCWFU62286 and CALG8083 trials.

[#]: Data are missing for 03PCL88, ECOG3588, CCCWFU62286, EORTC08877 and JCOG9104 trials.

[□]: Data are missing for ECOG3588, CCCWFU62286, EORTC08877 and JCOG9104 trials.

Table S4. Description of chemotherapy compliance by trial

Trial	Allocated Radiotherapy	N*	%**	Compliance category
CALGB8083	“Later or longer”	149	73	Different
	“Earlier or shorter”	130	52	
BR6	“Later or longer”	157	80	Similar
	“Earlier or shorter”	162	79	
03PCL88 ^{\$}	“Later or longer”	76	85	Different
	“Earlier or shorter”	88	75	
ECOG3588	“Later or longer”	202	91	Similar
	“Earlier or shorter”	206	85	
EORTC08877	“Later or longer”	172	90	Different
	“Earlier or shorter”	173	80	
JCOG9104	“Later or longer”	114	85	Similar
	“Earlier or shorter”	113	87	
LLCG93	“Later or longer”	166	80	Different
	“Earlier or shorter”	159	69	
HeCOG93	“Later or longer”	39	95	Different
	“Earlier or shorter”	41	83	

Note: It should be underlined that two trials (BR6 and LLCG93) with the same design and using the same chemoradiation modalities, differed in terms of chemotherapy compliance.

* N is the total number of patients with information on chemotherapy compliance in each treatment arm. Data on compliance were available for between 96% and 100% of patients depending on the trial (median = 98%).

** % is the percentage of patients compliant to CT, defined as the patients receiving 100% or more of the planned number of CT cycles, except for the CALGB8083 trial in which patients receiving 6 CT cycles or more were considered as compliant. If the difference between arms was $\geq 10\%$, the trial was considered as having different CT compliance between arms; if it was $< 10\%$, the trial was considered as having similar CT compliance in both arms.

\$ based on published data

Table S5. Description of patients who started radiotherapy according to arm and chemotherapy compliance subset

Trial*	N [§]	Nb patients who started RT (%)	
		“Earlier or shorter” RT	“Later or longer” RT
<i>Similar CT compliance</i>			
BR.6	318	156 (96)	136 (87)
JCOG9104	231	112 (98)	105 (90)
<i>Different CT compliance</i>			
CALGB8083	292	128 (93)	132 (86)
03PCL88	164	79 (90)	73 (96)
EORTC08877	349	155 (89)	156 (90)
LLCG93	324	147 (92)	136 (82)
HeCOG93	81	41 (98)	39 (100)
All trials	1,759	818 (93)	777 (88)

[§]: N is the number of patients with information about the initiation of RT

* For two trials (CCCWFU62281 and ECOG 3588) the information about the initiation of RT is unknown for all patients. For two trials, the information about the initiation of RT is unknown for some patients (14 for BR6 and 1 for LLCG93).

Abbreviations: CT = chemotherapy; RT = Radiotherapy;

Table S6. Description of radiotherapy compliance

Trials*	Arm	Initiation of chest radiation			Total RT dose (Gy)			Total number of fractions			Total RT duration (wks)		
		N	Theoretical	Observed [§]	N	Theoretical	Observed [§]	N	Theoretical	Observed [§]	N	Theoretical	Observed [§]
CCCWFU62286 [30]	Shorter Longer	0 0	Day 1 Day 8	MD	0 0	50 50	MD	0 0	25 20	MD	0 0	5 7	MD
<u>Similar CT compliance</u>													
BR.6 [26]	Earlier	156	Day 22	20 [0–41]	156	40	40 [5–43]	156	15	15 [2–17]	156	3	21 [1–33]
	Later	136	Day 106	113 [26–182]	136	40	40 [21–45]	136	15	15 [8–20]	136	3	21 [12–54]
ECOG3588 [16]	Shorter	0	Day 1	MD	0	45	MD	0	30	MD	0	3 bid	MD
	Longer	0	Day 1	MD	0	45	MD	0	25	MD	0	5	MD
JCOG9104 [28]	Earlier	0	Day 2	MD	112	45	45 [24–48]	0	30	MD	0	3 bid	MD
	Later	0	Day 85	MD	105	45	45 [20–60]	0	30	MD	0	3 bid	MD
<u>Different CT compliance</u>													
CALGB8083 [25] [£]	Earlier	128	Day 1	4 [0–61]	97	40 + 10	40 [10–52]	0	24	MD	113	5	5.4 [0.6–13]
	Later	132	Day 64	67 [0–106]	97	40 + 10	40 [4–50]	0	24	MD	115	5	5.9 [0–12.6]
03PCL88 [31]	Shorter	79	Day 30	32 [6–68]	67	50	50 [30–55]	65	20	20 [15–25]	79	5	5.6 [3.1–19.3]
	Longer	73	Day 36	37 [8–49]	72	55	55 [31–60]	66	22	22 [12–28]	73	9.5	9.6 [3.6–18.1]
EORTC08877 [27]	Earlier	155	Day 43	41 [10–56]	155	50	50 [13–58]	155	20	20 [5–25]	155	13	12.6 [0.6–21.7]
	Later	156	Day 99	105 [67–238]	156	50	50 [15–60]	155	20	20 [2–25]	156	4	4 [2–7.7]
LLCG93 [12]	Earlier	147	Day 22	24 [0–110]	146	40	40 [13–50]	146	15	15 [5–25]	146	3	2.9 [0.6–5.7]
	Later	135	Day 106	116 [68–171]	136	40	40 [30–40]	136	15	15 [9–15]	135	3	2.9 [1.9–7.1]
HeCOG93 [29]	Earlier	37	Day 1	3 [0–47]	30	45	45 [30–58]	0	30	MD	33	3 bid	4 [1.3–29.3]
	Later	34	Day 57	68 [41–125]	29	45	45 [28–60]	0	30	MD	31	3 bid	3 [1.3–29.3]

[§] Medians [Range] are given for observed values

* Trials are chronologically ordered within each category of trials (Similar CT compliance and Different CT compliance). CCCWFU62286 trial is not part of the CT compliance subset analysis.

[£] Patients were supposed to receive a total dose of 40 Gy plus a boost of 10 Gy. The exact boost dose that the patients received was not available, but the corresponding publication stated that the therapy of approximately 80% of the patients was deemed appropriate or appropriate with minor violations after the review by the Quality Assurance Review Centre.

Abbreviations: **bid** = RT given twice a day; **d** = day; **CT** = chemotherapy; **fr** = fraction; **Gy** = Gray; **MD** = Missing Data; **RT** = Radiotherapy; **wks** = weeks;

BR = Bronchus; **CALGB** = Cancer and Leukaemia Group B; **CCCWFU** = Comprehensive Cancer Centre of Wake Forest University; **ECOG** = Eastern Cooperative Oncology Group; **EORTC** = European Organisation for Research and Treatment of Cancer; **HeCOG** = Hellenic Cooperative Oncology Group; **JCOG** = Japan Clinical Oncology Group; **LLCG** = London Lung Cancer Group; **PCL** = Petites Cellules Limitées

Table S7: Trial subset description

Trial	CT compliance	RT Dose per fraction	Type of RT	Overall time of RT	Platin-based CT during RT in both arms	Concurrent CT in both arms	Same RT in the 2 arms
CALGB8083	Different	1.8 – 2.4 Gy	Standard RT	> 30 days in both arms	No	Yes	Yes
BR6	Similar	> 2.4 Gy	Standard RT	≤ 30 days in both arms	Yes	Yes	Yes
CCWFU62286	NA	1.8 – 2.4 Gy	Standard RT	> 30 days in both arms	No	No	No*
03PCL88	Different	> 2.4 Gy	Standard RT	> 30 days in both arms	No	No	No*
ECOG3588	Similar	< 1.8 Gy	Hyperfr RT	One arm ≤ 30 days, one > 30 days	Yes	Yes	No*
EORTC08877	Different	> 2.4 Gy	Standard RT	One arm ≤ 30 days, one > 30 days	No	No	No
JCOG9104	Similar	< 1.8 Gy	Hyperfr RT	≤ 30 days in both arms	No	No	Yes
LLCG93	Different	> 2.4 Gy	Standard RT	≤ 30 days in both arms	Yes	Yes	Yes
HeCOG93	Different	< 1.8 Gy	Hyperfr RT	≤ 30 days in both arms	Yes	Yes	Yes

Abbreviations: CT = Chemotherapy; Gy = Gray; **Hyperfr** = Hyperfractionated; **NA** = Not Available; **RT** = Radiotherapy

* Trials comparing a short versus a long duration of radiotherapy with a difference of less than one week in the initiation of radiotherapy

Table S8. Chemotherapy compliance and overall and progression-free survival using a landmark analysis.

Variable	Category	Overall Survival		Progression-Free Survival	
		HR [95% CI]	<i>p</i> -value (test for trend)	HR [95% CI]	<i>p</i> -value (test for trend)
CT compliance	Poor	1 (ref)		1 (ref)	
	Good	0.56 [0.49–0.64]	<0.0001	0.70 [0.59–0.83]	<0.0001
Sex	Female	1 (ref)		1 (ref)	
	Male	1.27 [1.14–1.40]	<0.0001	1.29 [1.15–1.45]	<0.0001
Performance status	Good	1 (ref)		1 (ref)	
	Mild	1.20 [1.08–1.33]	0.0001	1.11 [0.98–1.25]	0.02
	Poor	1.40 [1.16–1.69]	(<0.0001)	1.36 [1.01–1.69]	(0.006)
Age	< 60	1 (ref)		1 (ref)	
	60–64	1.02 [0.91–1.16]	0.07 (0.03)	1.01 [0.87–1.16]	0.03 (0.02)
	≥65	1.14 [1.01–1.27]		1.17 [1.03–1.33]	

We performed a post-hoc landmark analysis on the impact of individual CT compliance on overall and progression-free survival among patients who survived (or had no progression) at least 120 days. This landmark was chosen because most of patients finished their chemoradiation treatment at 120 days. We used a multivariate Cox model stratified by trial and adjusted on the radiotherapy arm, sex, age, and performance status.

Abbreviations: CI = Confidence Interval; CT = Chemotherapy; HR = Hazard Ratio, ref=reference

Table S9. Sensitivity analyses on overall survival.

		Availability	Results	I ²	Total heterogeneity	CT compliance subset			
		No. of trials (patients)	HR [95% CI]			Between-subset heterogeneity	Within-subset heterogeneity (residual)	Similar CT Compliance	Different CT Compliance
Exclusion of trials with:								HR [95% CI]	
-	no concurrent chemoradiation in one arm (JCOG9104)	8 (2,074)	1.01 [0.92–1.11]	63%	18.7**	16.4*** ^{\$}	1.9	0.80 [0.68–0.93]	1.19 [1.05–1.34]
-	shorter vs. longer RT comparison (03PCL88, ECOG3588, CCWFU62286) [□]	6 (1,610)	1.01 [0.91–1.12]	68%	15.5**	14.3***	1.2	0.76 [0.63–0.91]	1.17 [1.03–1.33]
-	different RT modalities (03PCL88, ECOG3588, CCWFU62286, EORTC08777) ^{□□}	5 (1,261)	0.96 [0.85–1.08]	67%	12.1*	11.1***	1.0	0.76 [0.63–0.91]	1.15 [0.98–1.35]

□ trials comparing shorter versus longer radiotherapy, but starting at the same time

□□ : this group of trials corresponds to the previous group plus one trial (EORTC08777) with the longer RT arm starting earlier than the shorter arm

* : $p < 0.05$ ** : $p < 0.01$ *** : $p < 0.001$

^{\$}: to be compared with total heterogeneity of 18.3 when both JCOG9104 and CCCWFU62286 are removed;

The test associated with between-subset heterogeneity corresponds to the interaction test.

Abbreviations: CI = Confidence Interval; HR = Hazard Ratio; RT = Radiotherapy

Table S10. Acute toxicity according to trial subsets based on chemotherapy compliance

Severe toxicity (grade 3-5)	Availability No. of trials (patients)	Different CT compliance	Similar CT compliance	<i>p</i> -value interaction	Total heterogeneity	Residual heterogeneity (within-subset)
		OR [95% CI]	OR [95% CI]			
Neutrophil	5 (1,342)	0.93 [0.56–1.54]	1.92 [1.40–2.65]	0.02	24.1***	18.3***
Haemoglobin	5 (1,365)	1.17 [0.52–2.61]	1.33 [0.99–1.77]	0.76	1.5	1.4
Platelets	6 (1,706)	1.53 [1.04–2.27]	1.07 [0.78–1.48]	0.17	10.7	8.8
Oesophageal	7 (1,839)	1.42 [0.87–2.31]	2.25 [1.57–3.23]	0.14	12.7*	10.5
Pulmonary	4 (996)	1.42 [0.63–3.21]	1.50 [0.67–2.62]	0.92	2.1	2.1
Cardiac	5 (1,537)	3.48 [0.78–15.5]	3.04 [1.20–7.72]	0.89	1.1	1.1

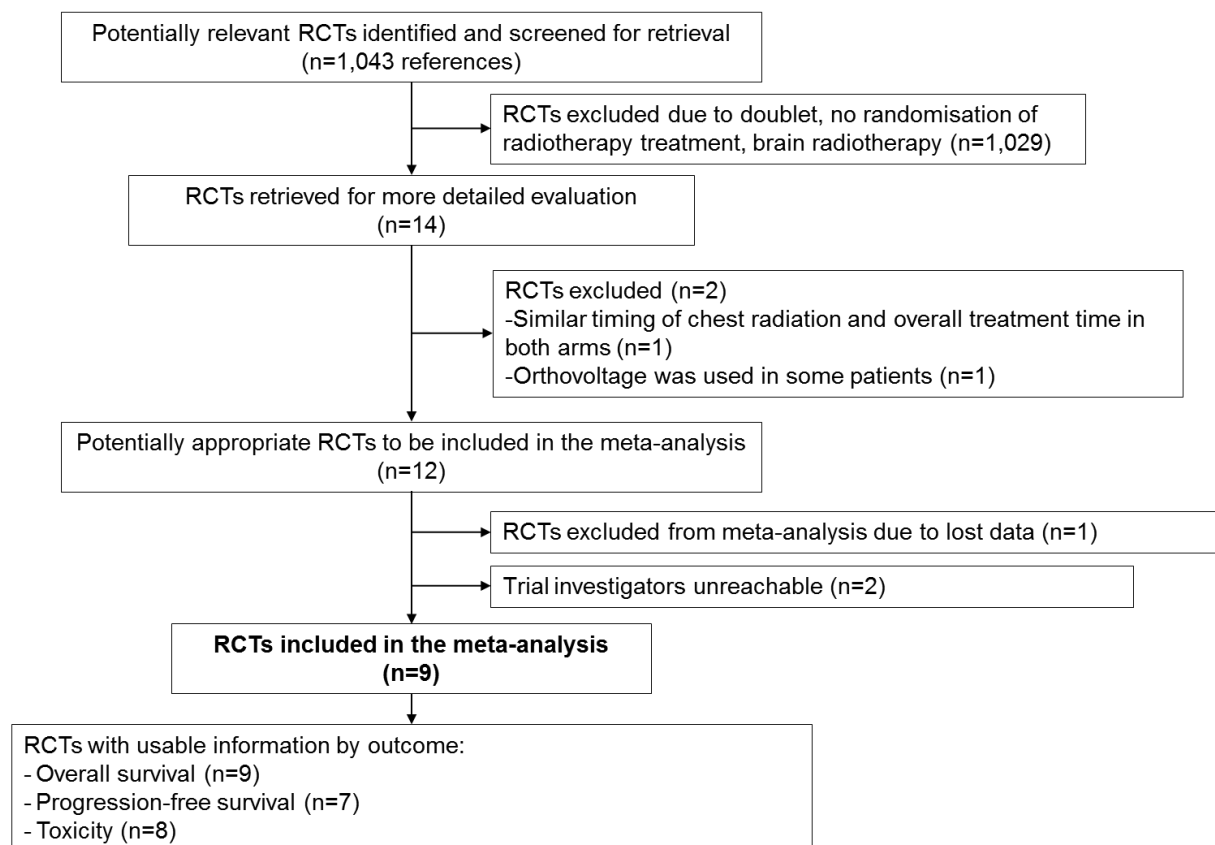
* : $p < 0.05$ ** : $p < 0.01$ *** : $p < 0.001$

Toxicity was graded according to the National Cancer Institute Common Toxicity Criteria, World Health Organisation criteria, or Eastern Cooperative Oncology Group Common Toxicity Criteria depending on the trials. Severe toxicity was defined as grade 3 to 5 toxicity. Grade 5 was present only for pulmonary toxicity ($n=4$) and cardiac toxicity ($n=1$).

Total heterogeneity is the sum of between-subset and within-subset heterogeneity. The test associated with between-subset heterogeneity corresponds to the interaction test. The lower residual heterogeneity, the greater overall heterogeneity of the treatment effect between trials is explained by between-subset heterogeneity.

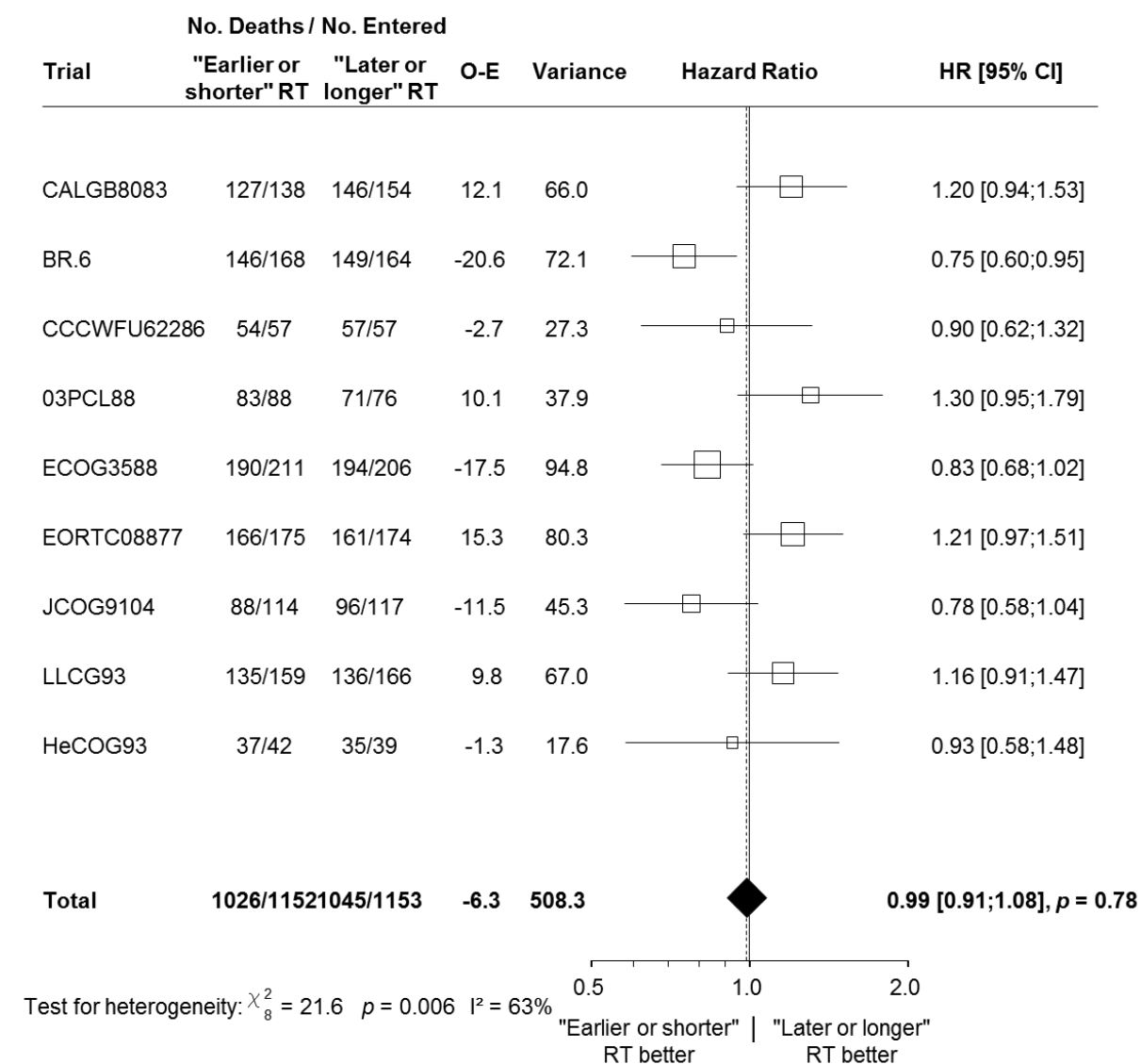
The only significant interaction was for neutrophil toxicity: the excess toxicity in the “earlier or shorter” arm was mainly driven by trials in the similar compliance group.

Abbreviations: **CI** = Confidence Interval; **OR** = Odds Ratio; **CT** = Chemotherapy

Figure S1: Flow chart

Abbreviations: RCT = Randomised Controlled Trial

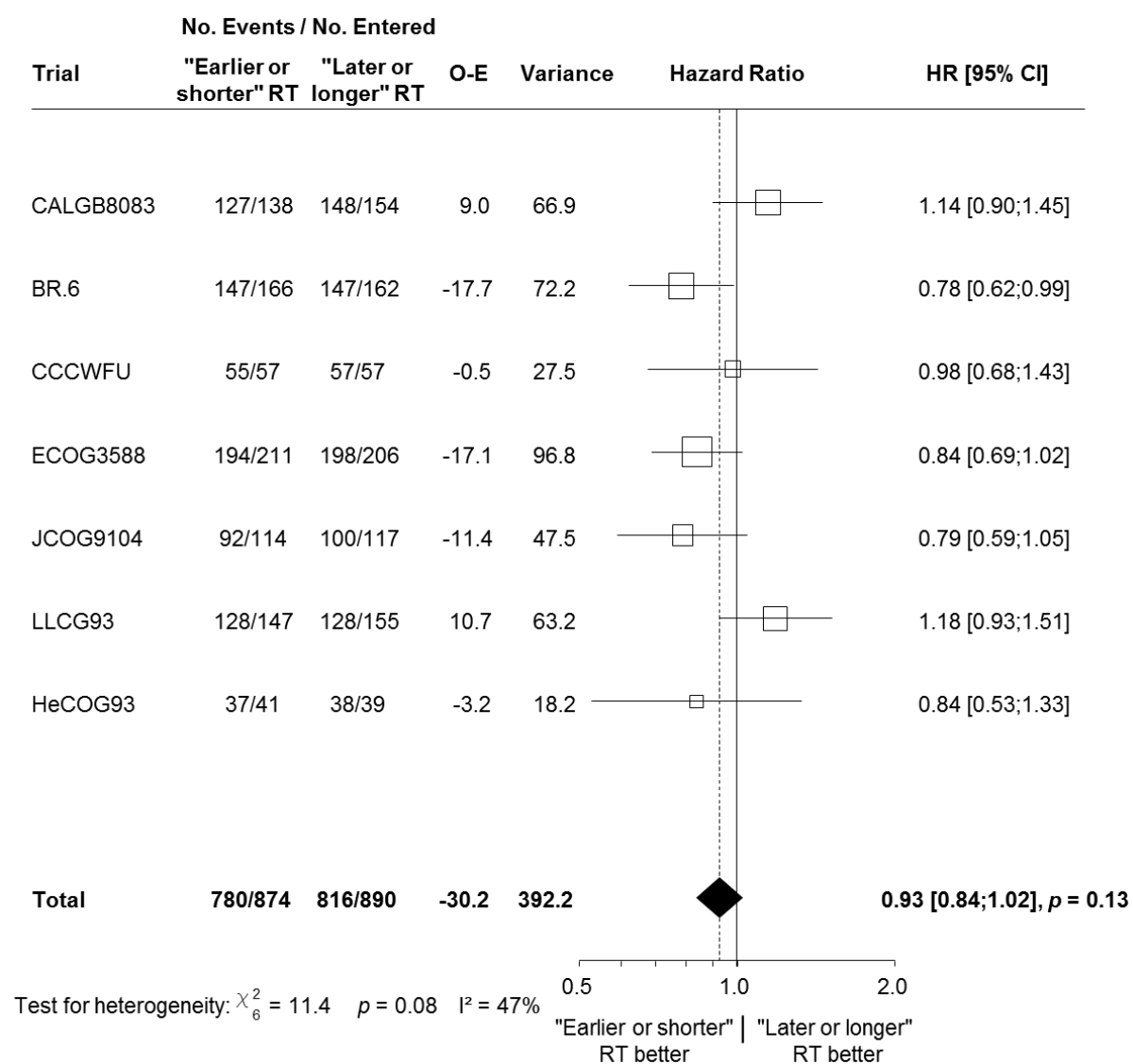
Figure S2: Effect of “earlier or shorter” radiotherapy versus “later or longer” radiotherapy on overall survival



Each trial is represented by a square, the centre of which denotes the hazard ratio of death for that trial comparison with the horizontal lines showing the 95% confidence intervals (CIs). The size of the square is directly proportional to the amount of information contributed by the trial. The black diamond represents the overall hazard ratios, with the centre denoting the hazard ratio and the extremities the 95% CI. The fixed effect model was used.

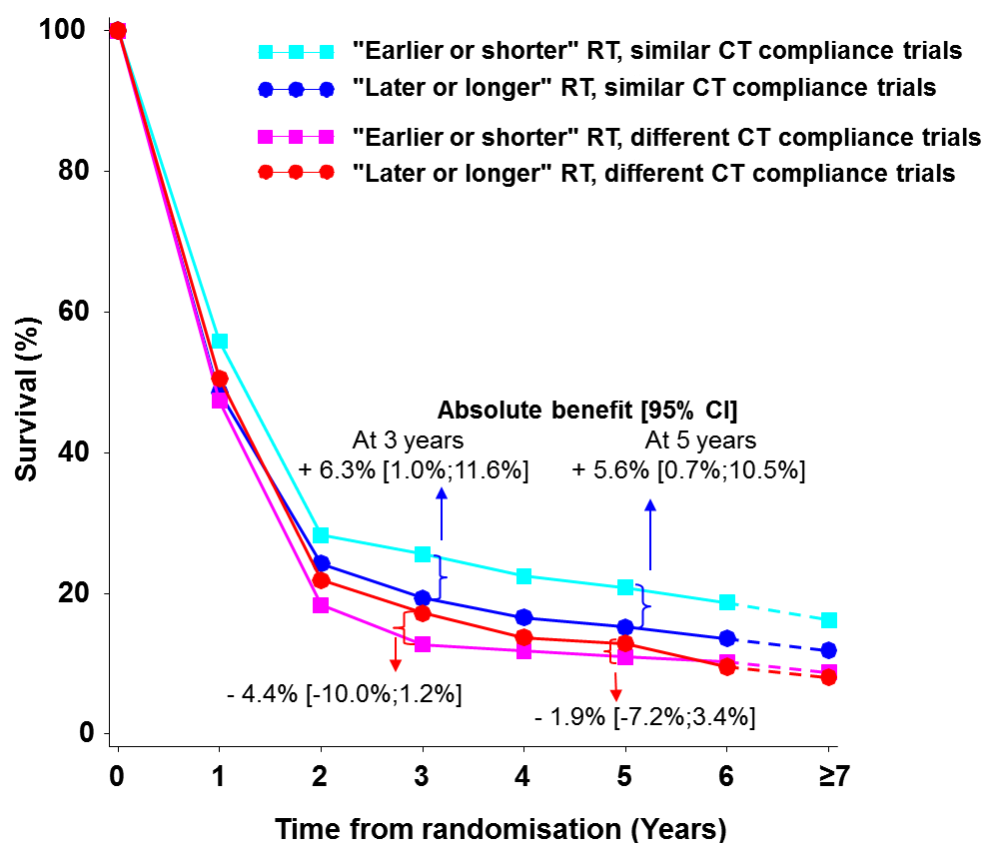
Abbreviations: CI = Confidence Interval; HR = Hazard ratio; O-E = Observed-Expected; RT = Radiotherapy

Figure S3: Effect of “earlier or shorter” radiotherapy versus “later or longer” radiotherapy on progression-free survival



Abbreviations: CI = Confidence Interval; HR = Hazard ratio; O-E = Observed-Expected; RT = Radiotherapy

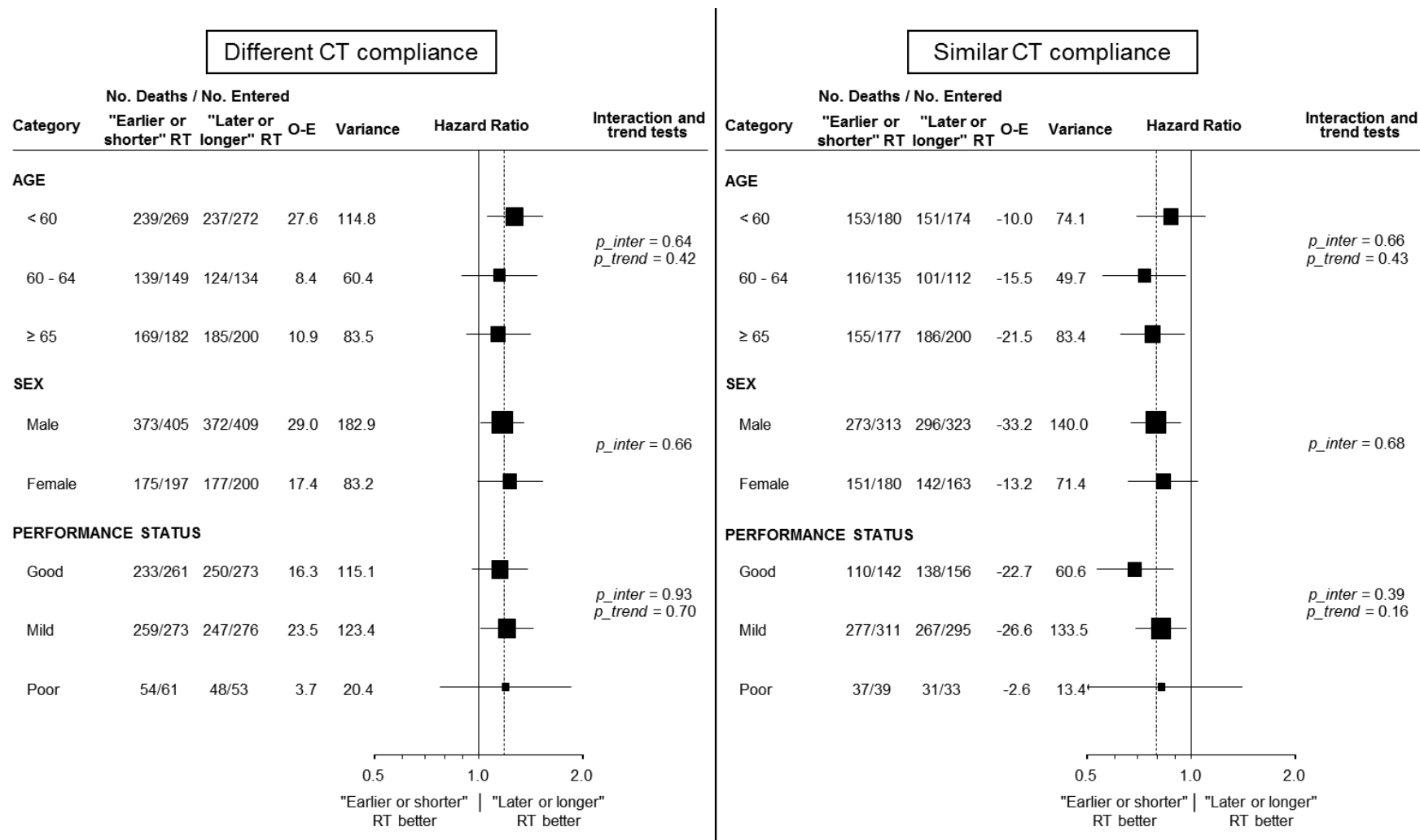
Figure S4: Survival curves for progression-free survival according to chemotherapy compliance



Number of deaths/ PY by period	Years 0-2	Years 3-5	Years ≥ 6
Similar CT compliance			
"Earlier or shorter" RT	349 / 580	36 / 344	48 / 356
"Later or longer" RT	374 / 514	41 / 238	30 / 228
Different CT compliance			
"Earlier or shorter" RT	261 / 320	19 / 103	12 / 103
"Later or longer" RT	268 / 374	26 / 130	20 / 87

Abbreviations: CI = Confidence Interval; CT = Chemotherapy; PY = Person-Year; RT = Radiotherapy

Figure S5: Effect of “earlier or shorter” radiotherapy versus “later or longer” radiotherapy on overall-survival according to chemotherapy compliance and patient characteristics



Abbreviations: CT= Chemotherapy; O-E= Observed-Expected; RT= Radiotherapy.

Titre : Estimation du bénéfice de survie à partir de méta-analyse sur données individuelles et évaluation économique.

Mots clés : analyse de survie, survie moyenne restreinte, méta-analyse sur données individuelles, hétérogénéité, essais cliniques randomisés, étude coût-efficacité, évaluation médico-économique.

Résumé : Le bénéfice de survie restreint à un horizon temporel donné a été proposé comme mesure alternative aux mesures relatives de l'effet d'un traitement notamment dans le cas de non proportionnalité des risques de décès. Le bénéfice de survie restreint correspond à la différence des survies moyennes entre deux bras de traitement et s'exprime en nombre d'années de vie gagnées. Dans la littérature, cette mesure est présentée comme plus intuitive que le hazard ratio et plusieurs auteurs ont proposé son utilisation pour le design et l'analyse d'un essai clinique. Toutefois, ce n'est pas actuellement la mesure qui est utilisée de manière courante dans les essais randomisés. Cette mesure s'applique quelle que soit la distribution des temps de survie et est adaptée si l'hypothèse des risques proportionnels n'est pas respectée. De plus, le bénéfice de survie restreint peut être utilisé en évaluation médico-économique où la mesure d'un effet absolu est nécessaire (nombre d'années de vie gagnées pondérées ou non par la qualité de vie). Si l'on souhaite estimer le bénéfice de survie restreint à partir d'une méta-analyse sur données individuelles, se pose alors la question de prendre en compte l'effet essai dû à la structure hiérarchique des données. L'objectif de cette thèse était de comparer des méthodes statistiques d'estimation du bénéfice de survie restreint à partir de données individuelles d'une méta-analyse d'essais cliniques randomisés. Le point de départ a été une étude de cas (étude coût-efficacité) réalisée à partir des données de la Meta-Analysis of Radiotherapy in Lung Cancer. Cette étude a montré que les cinq méthodes d'estimation étudiées conduisaient à des estimations différentes du bénéfice de survie et de son intervalle de confiance. Le choix de la méthode d'estimation avait également un impact sur les résultats de l'analyse coût-efficacité. Un second travail a consisté à mener une étude de simulation pour mieux comprendre les propriétés des méthodes d'estimation considérées en termes de biais moyen et d'erreur-type. Enfin, la dernière partie de la thèse a mis en application les enseignements de cette étude de simulation au travers de trois méta-analyses sur données individuelles dans le cancer du nasopharynx et dans le cancer du poumon à petites cellules.

Title : Estimation of the survival benefit from individual participant data meta-analysis and economic evaluation.

Keywords : survival analysis, restricted mean survival time, individual participant data meta-analysis, heterogeneity, randomized clinical trials, cost-effectiveness analysis, economic evaluation.

Abstract : The survival benefit restricted up to a certain time horizon has been suggested as an alternative measure to the common relative measures used to estimate the treatment effect, especially in case of non-proportional hazards of death. The restricted survival benefit corresponds to the difference of the two restricted mean survival times estimated for each treatment arm, and is expressed in terms of life years gained. In the literature, this measure is considered as more intuitive than the hazard ratio and many authors have suggested its use for the design and the analysis of clinical trials. However, it is not currently the most used measure in randomized trials. This measure is valid under any distribution of the survival times and is adapted if the proportional hazards assumption does not hold. In addition, the restricted survival benefit can be used in medico-economic evaluation where an absolute measure of the treatment effect is needed (number of [quality adjusted] life years gained). If one wants to estimate the restricted survival benefit from an individual participant data meta-analysis, there is a need to take into account the trial effect due to the hierarchical structure of the data. The aim of this thesis was to compare statistical methods to estimate the restricted survival benefit from an individual participant data meta-analysis of randomized trials. The starting point was a case study (cost-effectiveness analysis) using data from the Meta-Analysis of Radiotherapy in Lung Cancer. This study showed that the five investigated methods yielded different estimates for the restricted survival benefit and its confidence interval. The choice of a method to estimate the survival benefit also impacted on cost-effectiveness results. Our second project consisted in a simulation study to have a better understanding of the properties of the investigated methods in terms of bias and standard error. Finally, the last part of the thesis illustrated the lessons learned from the simulation study through three examples of individual participant data meta-analysis in nasopharynx cancer and in small cell lung cancer.