



**HAL**  
open science

# Meta-analyse des traitements néoadjuvants des cancers de l'œsophage et de la jonction oesogastrique

Matthieu Faron

► **To cite this version:**

Matthieu Faron. Meta-analyse des traitements néoadjuvants des cancers de l'œsophage et de la jonction oesogastrique. Cancer. Université Paris-Saclay, 2021. Français. NNT : 2021UPASR033 . tel-03692702

**HAL Id: tel-03692702**

**<https://theses.hal.science/tel-03692702>**

Submitted on 10 Jun 2022

**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.

Méta-analyse des traitements néoadjuvants  
des cancers de l'œsophage et de la jonction  
œsogastrique

*Meta-analysis of neoadjuvant treatment of esophageal and  
gastroesophageal junction carcinoma*

**Thèse de doctorat de l'université Paris-Saclay**

École doctorale n°570 : Santé publique (EDSP)

Spécialité de doctorat : Santé Publique – Recherche clinique

Unité de recherche : Université Paris-Saclay, UVSQ, Inserm, CESP, 94807, Villejuif, France

Réfèrent : Faculté de médecine

**Thèse présentée et soutenue à Paris-Saclay,  
le 14 décembre 2021, par**

**Matthieu FARON**

**Composition du Jury**

<b>Frédérique PESCHAUD</b> PUPH, Université Paris Sorbonne	Présidente
<b>Isabelle BOUTRON</b> PUPH, Université Paris Sorbonne	Rapporteur & Examinatrice
<b>Denis COLLET</b> PUPH, Université Bordeaux	Rapporteur & Examinateur
<b>Silvy LAPORTE</b> MCU/HDR, Université Saint Etienne	Examinatrice
<b>Caroline MOLLEVI</b> HDR ; ICM, Montpellier	Examinatrice
<b>Sarah WATSON</b> PhD/HDR, Université Paris Sorbonne	Examinatrice

**Direction de la thèse**

<b>Pierre BLANCHARD</b> PUPH, Université Paris SACLAY	Directeur de thèse
<b>Gwenael LE TEUFF</b> PhD, HDR, Université Paris SACLAY	Co-Directeur de thèse



A Joséphine et Valentine

# Valorisation de la thèse

## Communications orales et affichées

**ISCB 2020 (oral) :** Indirect Treatment Effect Estimation In Presence Of Modifier In Network Meta Analyses Of Time-to-event Data

**EPICLIN 2020 (oral) :** Estimation de l'effet du traitement en présence d'un modificateur de son effet dans les méta-analyses en réseau pour données censurées : étude de simulation d'un réseau simple et application à des données réelles

**ASCO annual meeting (2021) :** Abstract 4067: Individual patient data meta-analysis of neoadjuvant chemotherapy followed by surgery versus upfront surgery in esophageal or gastro-esophageal carcinoma

## Publications

**Individual patient data meta-analysis of neoadjuvant chemotherapy followed by surgery versus upfront surgery for carcinoma of the oesophagus or the gastrooesophageal junction**

Matthieu Faron, Arnel Maurice Cheugoua-Zanetsie, Pierre Thirion, Matthew Nankivell, athryn Winter, David Cunningham, Ate Van der Gaast, Simon Law, Ruth Langley, Florent de Vathaire, Michele Valmasoni, Muriel Mauer, Jack Roth, Val GebSKI, Bryan H. Burmeister, Xavier Paoletti, Johanna van Sandick, Jianhua Fu, Michel Ducreux, Pierre Blanchard, Jayne Tierney, Jean-Pierre Pignon, Stefan Michiels on behalf of the MANATEC-02 collaborative group

In : *European Journal of Cancer*, 2021

**A frequentist one-step model for a simple network meta-analysis of time-to-event data in presence of an effect modifier**

Matthieu Faron, Pierre Blanchard, Laureen Ribassin-Majed, Jean-Pierre Pignon, Stefan Michiels, Gwénaél Le Teuff

In *PlosOne November 2021 (vol 16- n 11)*

# Remerciements

A Jean-Pierre Pignon, mon directeur de M2 et directeur de thèse initial. Merci pour votre encadrement rigoureux tout en étant bienveillant au cours de toutes ces années. Je suis très fier de pouvoir dire que j'aurais été votre étudiant en thèse. Si je ne vous avais pas rencontré en 2010 lors de ma recherche d'un stage de M2, je ne serais probablement jamais venu à Gustave Roussy et je n'en serais pas là où j'en suis aujourd'hui....

A Pierre Blanchard et Gwenael Le Teuff, merci d'avoir accepté de prendre la suite de Jean-Pierre pour mon encadrement. Merci aussi pour tous vos conseils et soutient lors de la thèse.

A Stefan Michiels, merci de m'avoir accueilli dans l'unité Oncostat. Merci également pour les innombrables aides : statistiques, logistique (le serveur Oxygène 2) diplomatiques (pour la récupération des données), etc. Merci également de supporter toutes les facéties des cliniciens qui veulent faire de la recherche.

A la direction (générale et médicale) de Gustave Roussy, merci d'avoir accepté de me libérer un jour par semaine pour mener mes travaux de recherche en parallèle de mes activités cliniques. Ceci représente bien à mes yeux l'engagement de notre centre comme Hôpital Universitaire et tout l'intérêt de travailler dans cette structure.

Au Pr. Ducreux, vous étiez déjà impliqué dans mon M2, impliqué dans ma thèse, président de mon jury de praticien spécialiste : merci de tout ce que vous faites pour propulser vos jeunes collaborateurs.

A mes chefs de service Diane Goere puis Charles Honoré et mes co-praticiens Léonor Benhaïm, Maximilliano Gelli et Isabelle Sourouille merci d'avoir pu créer les conditions propices à mener mes activités de recherche en parallèle de mes activités cliniques. Par ailleurs un véritable plaisir de travailler avec vous au jour le jour.

Impossible de remercier chacun individuellement, mais merci à tous les membres du SBE et d'Oncostat, les praticiens des comités 040, 070 et 100 votre aide n'a parfois été qu'un petit conseil, une réponse à une question ou un message de soutien mais ça m'a bien aidé.

Merci également à tous nos collaborateurs internationaux et aux sociétés savantes. Sans votre aide une méta-analyse sur données individuelles n'aurait jamais été possible.

Merci à l'INCA d'avoir soutenu ce projet en lui attribuant un PHRC.

Enfin merci à tous les patients inclus dans ces essais et à leur famille, sans leur aide la lutte contre le cancer serait tout simplement impossible.



## Table des matières

Valorisation de la thèse.....	4
Communications orales et affichées .....	4
Publications .....	4
Remerciements .....	5
Abréviations .....	10
Première partie : Introduction générale et modèles utilisés.....	11
Chapitre 1. Le cancer de l'œsophage.....	13
Anatomie.....	13
Embryologie et histologie.....	14
Epidémiologie .....	15
Histologie et facteurs de risques.....	15
Classification histo-pronostique et bilan d'extension.....	15
Traitements disponibles.....	17
Stratégie thérapeutique.....	17
Objectif clinique de la thèse .....	19
Chapitre 2. Meta-analyses.....	21
Méta-analyse par paire.....	21
Méta-analyse en réseau.....	22
Données individuelles et données agrégées .....	24
Interaction dans les méta-analyses .....	25
Objectif statistique de la thèse .....	26
Chapitre 3. Modèles mathématiques utilisés dans les MAR.....	27
Estimation de l'effet traitement pour des données censurées.....	27
Modèles basés sur les données individuelles.....	29
Modèles basés sur les données agrégées.....	32
Cas particulier des modèles additifs .....	35
Approches fréquentiste et Bayésienne.....	37
Seconde partie : Evaluation d'un modèle IPD en une étape.....	39
Chapitre 1. But de l'étude.....	41
Chapitre 2. Méthodes.....	42
Etude de simulation.....	42
Présentation des résultats des simulations .....	46
Données réelles : MARCH/MACH-NC .....	46



Modèles utilisés .....	47
Logiciels .....	48
Chapitre 3. Résultats .....	49
Résultats généraux .....	49
Configuration I : estimation de A-B indirecte uniquement.....	49
Configuration II : estimation mixte directe-indirecte de A-B .....	50
Application sur données réelles .....	53
Chapitre 4. Discussion et perspectives.....	57
Chapitre 5. Conclusions .....	61
Troisième partie : MANATEC-02 .....	63
Chapitre 1. Introduction .....	65
MANATEC-01 .....	65
But de l'étude.....	65
Chapitre 2. Méthodes.....	69
Advisory board et collaboration internationale.....	69
Essais incluables et stratégie d'identification .....	69
Analyse de la qualité des données (Checking) .....	70
Critères de jugement étudiés .....	71
Méthodes statistiques .....	73
Chapitre 3. Résultats .....	79
Caractéristiques des essais identifiés .....	79
Caractéristiques des patients inclus dans l'analyse .....	87
MA par paire 1 : chirurgie seule contre chimiothérapie suivie de chirurgie.....	93
MA par paire 2 : chirurgie seule contre chimioradiothérapie suivie de chirurgie.	104
MA par paire 3 : chimiothérapie suivie de chirurgie contre chimioradiothérapie suivie de chirurgie.....	115
Méta-Analyse en réseau .....	123
Chapitre 4. Discussion et perspectives.....	134
Résultats principaux .....	134
Forces de l'étude .....	138
Limites de l'étude.....	138
Perspectives .....	139
Chapitre 5. Conclusions .....	141
Références.....	142
Annexes.....	151

Partie simulation.....	151
Partie méta-analyse.....	156
Advisory board.....	156
Equations de recherche .....	157
Repartitions des critères secondaires dans les essais .....	159
Figures supplémentaires des méta-analyses deux à deux .....	164

# Abréviations

MA : Méta-analyse

MAR : Méta analyse en réseau

ET : Effet traitement

IPD : données individuelles des patients

AD : Données agrégées

HR : Hazard Ratio (rapport des risques instantanées de décès)

IC : intervalle de confiance

RR : Rapport des Risques

JOG : Jonction Oeso-Gastrique

OT : œsophage thoracique

ADK : adénocarcinome

CE : carcinome épidermoïde

RCT : Essai randomisé contrôlé

TNM : Tumor Node Metastasis

UICC : Union Internationale Contre le Cancer

ESMO : European Society of Medical Oncology

NCNN : National Comprehensive Cancer Network

GLMM : modèle linéaire généralisé mixte (ou hiérarchique)

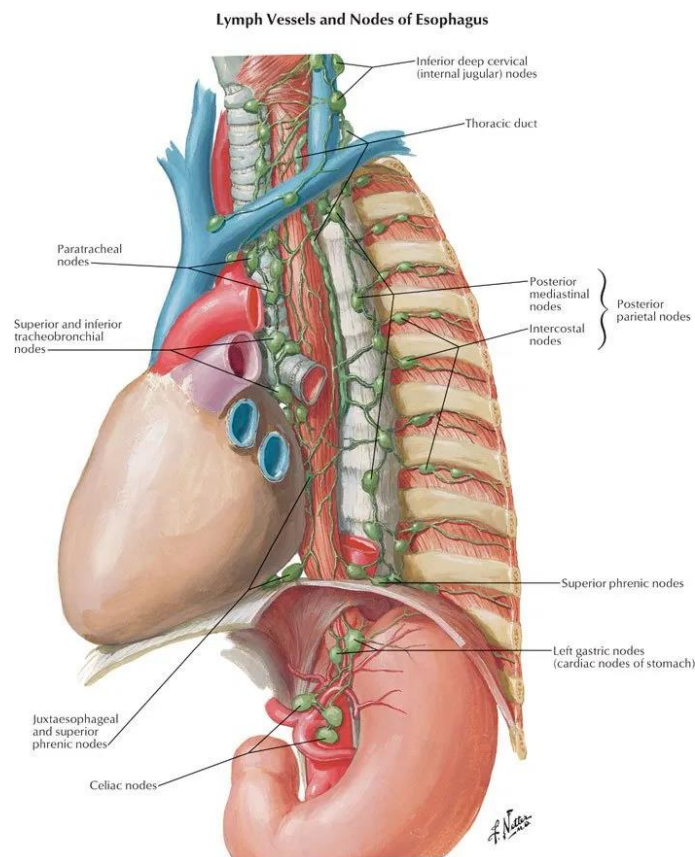
Première partie :  
Introduction générale et  
modèles utilisés



# Chapitre 1. Le cancer de l'œsophage

## Anatomie

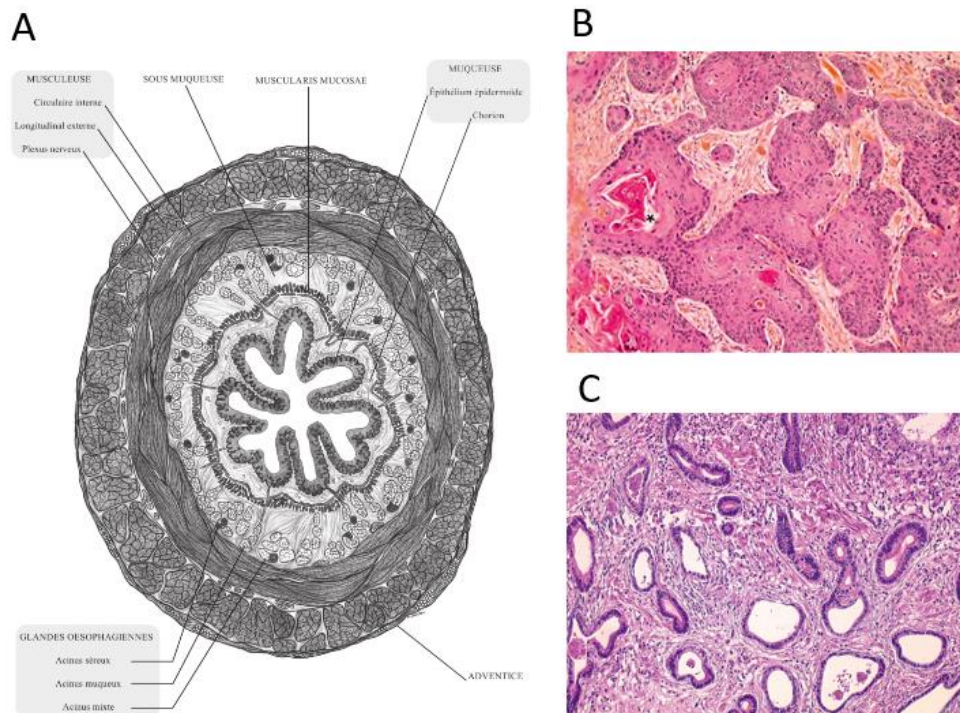
L'œsophage est le premier segment du tube digestif. Il mesure en moyenne 25 cm de longueur pour deux à trois cm de diamètre.(1) Il débute à l'extrémité inférieure du pharynx, au niveau de la 6<sup>ème</sup> vertèbre cervicale et se termine dans l'estomac au niveau de la 11<sup>ème</sup> vertèbre thoracique (**Figure 1**). La zone de terminaison de l'œsophage dans l'estomac s'appelle le cardia ou jonction œsogastrique (JOG). L'œsophage se trouve donc interposé entre deux territoires fonctionnels différents : la sphère ORL vers le haut et le tube digestif vers le bas. Il se situe à cheval sur trois territoires anatomiques : le cou, le thorax et l'abdomen car la JOG se situe sous le diaphragme. L'œsophage cervical étant très court il est rarement le siège de lésions. En pratique il existe donc deux localisation anatomiques distinctes : l'*œsophage thoracique* (OT) et la *jonction œsogastrique*. L'OT anatomique s'étend depuis l'entrée de l'œsophage dans le thorax jusqu'au hiatus œsophagien qui se situe environ 5 cm au-dessus de l'estomac. La JOG est définie comme une zone comprenant les cinq derniers cm d'œsophage et les cinq premiers cm d'estomac. (2) En l'absence d'hernie hiatale, cette zone est totalement intra-abdominale.



**Figure 1:** Anatomie de l'œsophage (d'après Netter)

## Embryologie et histologie

Chez l'embryon, l'œsophage se développe à partir de l'intestin primitif antérieur duquel dérive également l'estomac. D'un point de vue histologique, il est composé de quatre couches, qui sont, de l'intérieur vers l'extérieur : la muqueuse, la sous-muqueuse, la musculuse et l'adventice (**Figure 2 A**). Les cancers les plus fréquents de l'œsophage sont les carcinomes et ceux-ci naissent à partir de la muqueuse. Celle-ci est un épithélium malpighien non kératinisé du même type que celui de la sphère ORL. Ce type d'épithélium donne naissance aux *carcinomes épidermoïdes* (CE) (**Figure 2 B**). Plus bas, au niveau de l'estomac, l'épithélium est différent : il s'agit d'un épithélium glandulaire. Celui-ci donne naissance à un autre type de carcinome : l'*adénocarcinome* (ADK) (**Figure 2 C**). En condition normale, il ne devrait donc pas se développer d'adénocarcinome au niveau de l'œsophage. Toutefois, dans certaines conditions (notamment de reflux gastro-œsophagien), la muqueuse de l'œsophage se transforme en une muqueuse de type gastrique, phénomène qu'on appelle *métaplasie*. Cette métaplasie est un état pré-cancéreux qui pourra donner naissance à un adénocarcinome.



**Figure 2:** Coupes histologiques de l'œsophage normal et pathologique

(A) section transversale de l'œsophage normal avec ses quatre couches : muqueuses, sous muqueuse, musculuse et adventice (Source Université de Montpellier<sup>i</sup>). (B) Vue de la muqueuse présentant un carcinome épidermoïde. L'étoile montre la kératinisation (C) Vue de la muqueuse siège d'un adénocarcinome (Source Cours du Collège Français des Pathologistes, UMVF<sup>ii</sup>)

<sup>i</sup> <https://doc-pedagogie.umontpellier.fr/medecine/histologieLV/index.php?module=detail&subaction=desc&vue=4&itm=190&g=1&d=1>

<sup>ii</sup> [http://campus.cerimes.fr/anatomie-pathologique/enseignement/anopath\\_20/site/html/cours.pdf](http://campus.cerimes.fr/anatomie-pathologique/enseignement/anopath_20/site/html/cours.pdf)

## Epidémiologie

D'après les données du réseau FRANCIM, en 2018 en France, 5 445 patients ont eu un diagnostic de cancer de l'œsophage dont une majorité d'homme (4 251 vs 1 194). **(3)** Ceci le positionne en 11<sup>ème</sup> position chez l'homme et 18<sup>ème</sup> position chez la femme des cancers les plus fréquents. **(3)** Au niveau mondial, ce sont 604 100 nouveaux cas chaque année ce qui le place en 8<sup>ème</sup> position. **(4)** Le pronostic reste sombre avec **en France**, un taux de survie nette à 5 ans, de 14% chez les hommes et 16% chez les femmes, ce qui le positionne dans les deux sexes parmi les cinq cancers avec la survie globale à 5 ans la plus faible. **(3)** En termes de mortalité, ce sont 2 851 hommes (8<sup>ème</sup> position) et 874 femmes (14<sup>ème</sup> position) qui en décèdent chaque année et 544 076 au niveau mondial (6<sup>ème</sup> position). Dans l'œsophage, les néoplasies non carcinomateuses sont exceptionnelles (léiomyosarcomes, GIST, TNE, ...) et ne seront pas abordées.

## Histologie et facteurs de risques

Les deux types de carcinomes ont des facteurs de risque différents. Les *carcinomes épidermoïdes* ressemblent aux tumeurs de la sphère ORL. Survenant préférentiellement dans l'œsophage cervical et thoracique, ils sont fortement liés à *l'intoxication alcoolo-tabagique*. Les adénocarcinomes, quant à eux, surviennent préférentiellement dans le bas-œsophage ou la JOG et ressemblent aux adénocarcinomes de l'estomac. Leur principal facteur de risque est le *reflux gastro-œsophagien* car celui-ci engendre la métaplasie du bas œsophage qui constitue l'état précancéreux. Le principal facteur de risque de reflux gastro-œsophagien étant le surpoids, il existe une association statistique entre le surpoids et l'adénocarcinome. D'autres facteurs de risques ou habitudes alimentaires ont été identifiés mais sont, au moins en France métropolitaine, négligeables en fréquence comparativement à l'intoxication alcoolo-tabagique et à l'obésité.

## Classification histo-pronostique et bilan d'extension

La classification la plus utilisée est la classification *Tumor Node Metastasis* (TNM) de l'Union Internationale Contre le Cancer (UICC). La version la plus récente est la 8<sup>ème</sup> édition datant de 2017 (**Tableau 1**). Un bilan pré-thérapeutique complet permet d'estimer cette classification et de déterminer le stade (**Tableau 2**). Le bilan classique comprend une fibroscopie qui va permettre de localiser la tumeur et d'assurer le prélèvement histologique, une écho-endoscopie qui va permettre d'estimer la stade T et le stade N, un scanner cervico-thoraco-abdo-pelvien qui va permettre de confirmer en partie le stade T, le N et de rechercher des métastases à distance. Ce bilan peut être complété par un PET-Scan au <sup>18</sup>F<sup>18</sup>FDG afin d'éliminer une localisation à distance.



*Tableau 1: Classification TNM (UICC 2017)*

Stade	Description
<b>Tis</b>	Carcinome in situ ou dysplasie de haut grade
<b>T1</b>	Tumeur envahissant la lamina propria, la muscularis mucosae ou la sous-muqueuse
T1a	Tumeur envahit la lamina propria ou la muscularis
T1b	La tumeur envahit la sous-muqueuse
<b>T2</b>	Tumeur envahissant la musculature
<b>T3</b>	Tumeur envahissant l'adventice
<b>T4</b>	Tumeur envahissant les structures adjacentes
T4a	La tumeur envahit la plèvre, le péricarde, la veine azygos, le diaphragme ou le péritoine
T4b	La tumeur envahit d'autres structures adjacentes comme l'aorte, un corps vertébral ou la trachée
<b>N0</b>	Pas de signe d'atteinte des ganglions lymphatiques régionaux
<b>N1</b>	Métastases dans 1 à 2 ganglions lymphatiques régionaux
<b>N2</b>	Métastases dans 3 à 6 ganglions lymphatiques régionaux
<b>N3</b>	Métastases dans 7 ganglions régionaux ou plus
<b>M0</b>	Pas de métastase à distance
<b>M1</b>	Présence de métastase(s) à distance

*Tableau 2: Stades de la classification TNM (UICC 2017)*

Stade	T	N	M
<b>Stade 0</b>	Tis	N0	M0
<b>Stade IA</b>	T1a	N0	M0
<b>Stade IB</b>	T1b	N0	M0
	T2	N0	M0
<b>Stade II</b>	T1	N1	M0
	T3	N0	M0
<b>Stade IIIA</b>	T1	N2	M0
	T2	N1	M0
<b>Stade IIIB</b>	T2	N2	M0
	T3	N1, N2	M0
	T4a	N0, N1	M0
<b>Stade IVA</b>	T4a	N2	M0
	T4b	Tous N	M0
	Tous T	N3	M0
<b>Stade IVB</b>	Tous T	Tous N	M1

Les stades surlignés en bleu sont ceux qui feront l'objet de ce travail et correspondent aux tumeurs localement avancées mais non métastatiques

## Traitements disponibles

### Chirurgie

La chirurgie est le seul traitement potentiellement curatif des cancers de l'œsophage. Elle fait donc partie de toutes les stratégies à visée curatrice. Le geste habituel résèque en partie l'œsophage avec une marge de sécurité, la partie supérieure de l'estomac et les ganglions des premiers relais ganglionnaires. Compte tenu de la localisation à cheval sur plusieurs territoires anatomiques, cette intervention nécessite le plus souvent la combinaison de deux, voire trois voies d'abord chirurgicales : abdominale, thoracique et parfois cervicale. La continuité digestive est reconstruite le plus souvent à partir de l'estomac qui est tubulé puis réanastomosé à l'œsophage restant. La technique chirurgicale la plus courante est une œso-gastrectomie polaire supérieure avec curage ganglionnaire deux champs et anastomose œsogastrique intra-thoracique par voie d'abord abdominale et thoracique (technique de Lewis-Santy). Plusieurs variantes existent permettant de répondre à des situations différentes. Bien que de très gros progrès aient été faits ces dernières années avec la préhabilitation et la chirurgie mini-invasive, ceci reste une intervention grevée d'une morbidité et d'une mortalité postopératoire significatives. Elle ne peut donc être proposée qu'aux malades dont l'état général le permet.

### Chimioradiothérapie

La radiothérapie est le second traitement loco-régional des cancers de l'œsophage. Elle est le plus souvent chimio-sensibilisée par un sel de platine associé à une fluoropyrimidine ou un taxane. Les doses délivrées sont en général de 40 à 45 Grays en 20 à 25 fractions étalées sur 5 semaines. La dose délivrée peut être plus importante en cas d'irradiation exclusive. Cette technique permet de traiter à la fois la tumeur œsophagienne mais également les ganglions de proximité.

### Chimiothérapie

La chimiothérapie est un traitement systémique actif localement mais également sur les métastases à distance (potentielles ou avérées). Elle fait le plus souvent appel à un sel de platine et une fluoropyrimidine. Quand elle est donnée en préopératoire, deux à quatre cycles sont habituellement délivrés. En situation métastatique, elle est délivrée jusqu'à progression.

## Stratégie thérapeutique

Dans les stades I (tumeur ne dépassant pas la musculature, sans ganglions atteints) la chirurgie (voire l'endoscopie interventionnelle) est le principal traitement et est potentiellement curatrice à elle seule. Aucun traitement complémentaire n'ayant fait la preuve de son efficacité, la chirurgie est le plus souvent exclusive. A l'inverse dans les stades IV, la priorité étant au contrôle de la maladie systémique, la chimiothérapie est donc le plus souvent exclusive.

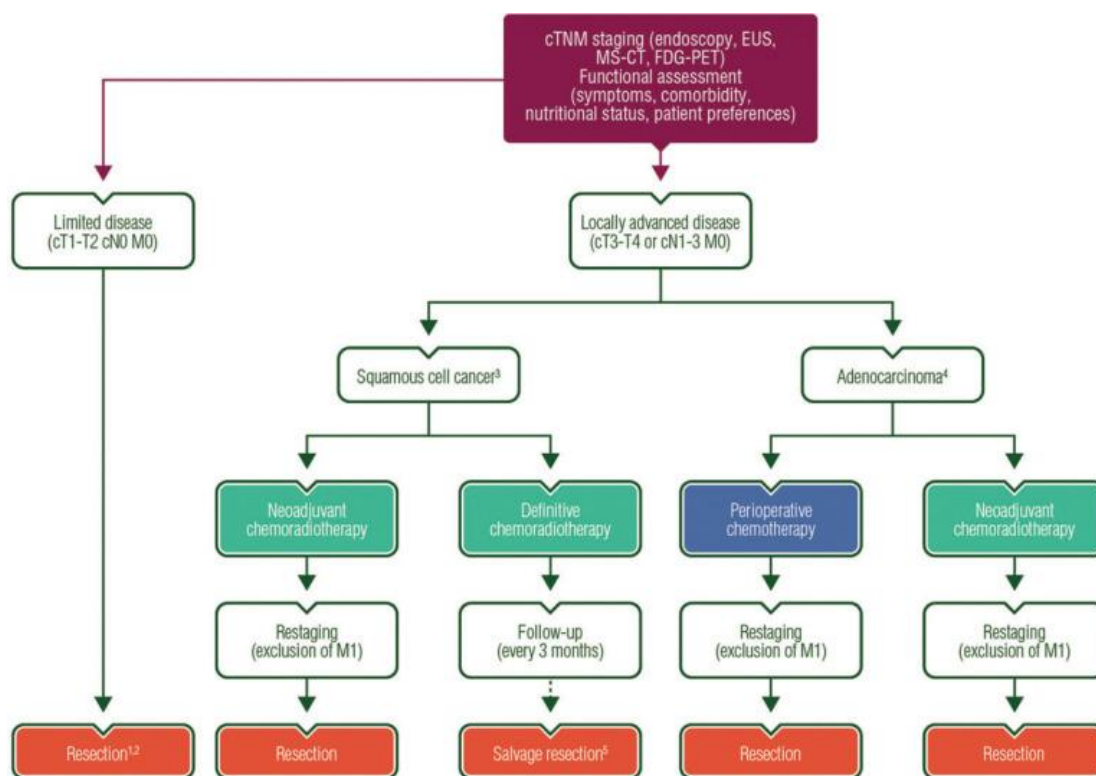
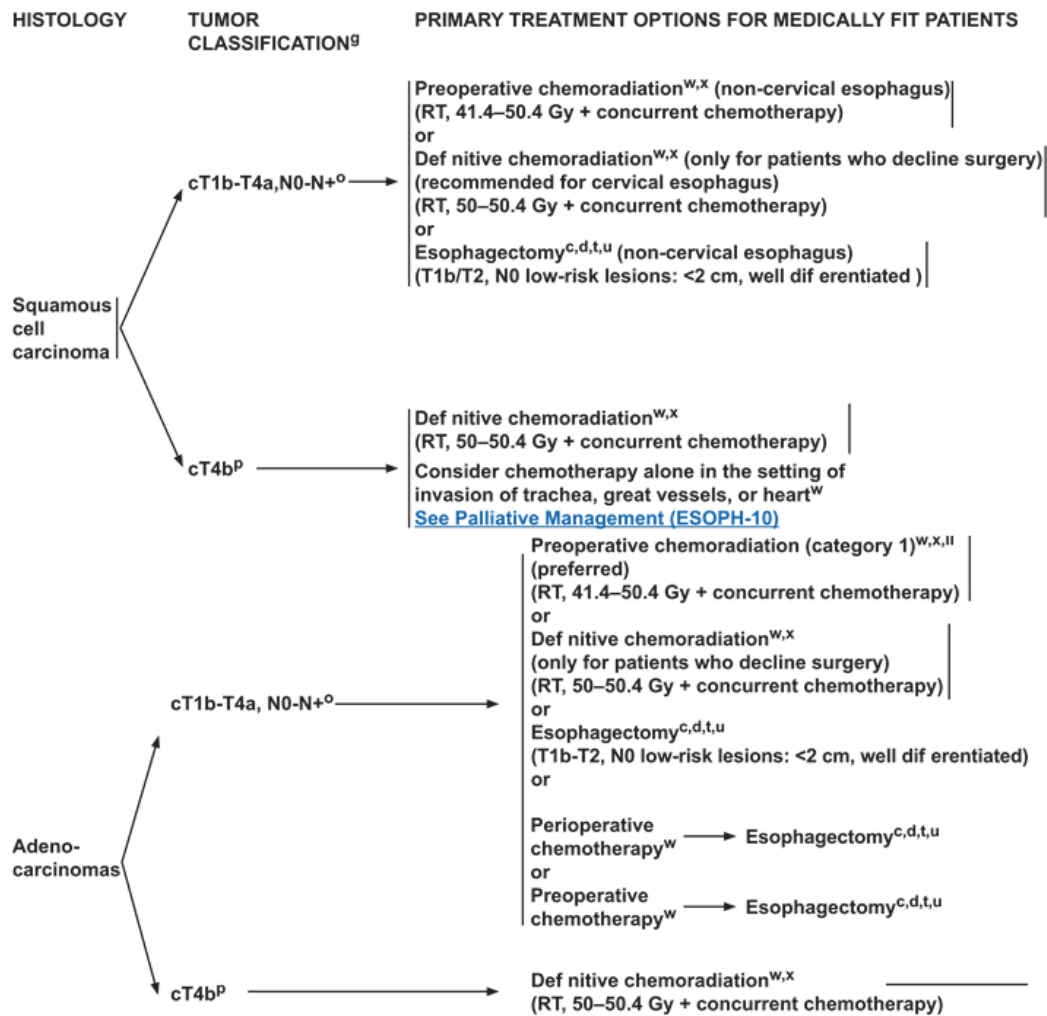


Figure 3 : Recommendations de European Society of Medical Oncology (ESMO) 2016

Les stades II et III sont également des situations dans lesquelles le traitement est potentiellement curatif. Toutefois la chirurgie seule (stratégie « S ») donnait des résultats décevants et n'est plus recommandée par aucune société savante. Deux principales stratégies multimodales ont été développées : chimiothérapie suivie de chirurgie (stratégie « CS ») ou chimioradiothérapie suivie de chirurgie (stratégie « CRS »). Toutefois ces deux stratégies néoadjuvantes (CS et CRS) sont considérées comme une alternative l'une de l'autre. Cette situation est liée au fait que si de nombreux essais ont comparé ces stratégies à la chirurgie d'emblée, peu d'essais sont disponibles pour leur comparaison directe. Plusieurs essais sont en cours : Re-Evaluation (NCT02442440), TOPGEAR(5) (ACTRN12609000035224), CMISG1701(6) (NCT03001596), ESOPEC(7) (NCT02509286), Neo-AEGIS(8) (NCT01726452), NExtT (JCOG 1109, UMIN000009482) mais les résultats définitifs de ceux-ci ne seront pas connus avant plusieurs années.

Les essais ont parfois inclus à la fois des CE et des ADK ou exclusivement l'un des sous-types. Les méta-analyses (standard ou en réseau) disponibles sont toutes basées sur des données agrégées (AD).(9–11) De ce fait, ces études ne pouvaient rechercher une interaction entre l'effet traitement et le sous-type histologique ou la localisation tumorale rendant difficile l'identification de critères prédictifs. A l'heure de la médecine « de précision » (i.e. « personnalisée »), il serait pourtant logique que la stratégie thérapeutique puisse être guidée par ces facteurs. Les recommandations européennes de l'European Society of Medical Oncology (ESMO) (Figure 3) et les Américaines du National Comprehensive Cancer Network (NCNN) (Figure 4) stratifient bien la prise en charge



*Figure 4 : Recommandations américaines du National Comprehensive Cancer Network (NCCN) version 1.2017*

thérapeutique sur le type histologique mais, faute de données, considèrent les deux stratégies comme des alternatives

## Objectif clinique de la thèse

D'un point de vue clinique, l'objectif principal de cette thèse est d'utiliser les méthodes de méta-analyse en réseau sur données individuelles afin de déterminer la meilleure séquence thérapeutique. Les données individuelles des participants (IPD) seront utilisées afin de rechercher une interaction avec la localisation tumorale (OT vs. JOG) et le type histologique (CE vs ADK) et de déterminer la meilleure stratégie dans les différents sous-groupes à titre d'objectif secondaire.



## Chapitre 2. Meta-analyses

### Méta-analyse par paire

Comme toute variable aléatoire, l'effet d'un traitement (ET) est soumis à la fluctuation d'échantillonnage et varie donc, par simple hasard d'un essai à l'autre. En présence d'effet traitement faible et/ou d'essai de petite taille, cette fluctuation peut amener à une situation d'apparence paradoxale où les résultats des essais semblent incohérents. En effet, la domination dans les méthodes statistiques actuelles, du test de significativité de l'hypothèse nulle avec un seuil dichotomique arbitrairement choisi à 0,05 peut conduire à une situation où certains essais sont dits « positifs », d'autres indéterminés (le plus souvent appelés « négatifs ») voire dont les résultats semblent opposés. Cette situation est pourtant parfois parfaitement normale car il ne s'agit que de fluctuation d'échantillonnage. Les techniques de méta-analyse (MA) permettent de combiner les estimations obtenues dans chaque essai afin de donner une estimation globale, plus précise, de l'effet traitement. Il ne s'agit pas d'une simple moyenne, mais d'une moyenne pondérée. La pondération est faite de telle sorte que les essais ayant estimé le plus précisément l'effet traitement contribuent le plus à l'estimation. La méthode de pondération la plus utilisée consiste à pondérer l'estimateur de l'ET dans un essai par *l'inverse de sa variance*. La variance dépend entre autres de la taille de l'échantillon, les essais avec le plus grand nombre de patients ayant habituellement un poids plus important.

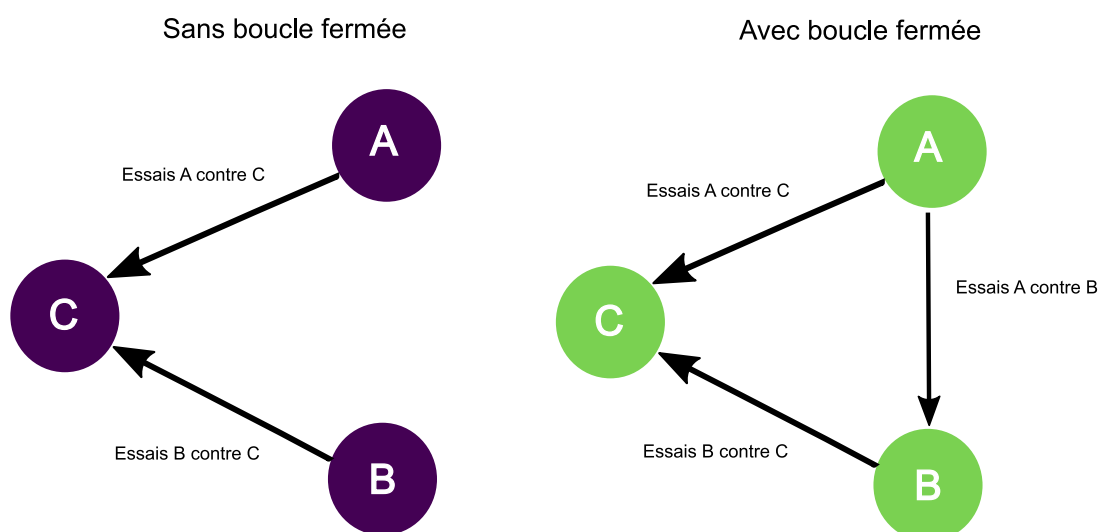
L'*hétérogénéité* désigne à quel point les ET varient d'un essai à l'autre. S'il existe un ET unique que tous les essais ont essayé d'estimer, on devrait observer une hétérogénéité faible, compatible avec de simples fluctuations aléatoires autour de cet effet unique. Toutefois, dans certaines situations, on observe une variation de l'ET plus importante que celle à laquelle on s'attendrait par simple fluctuation d'échantillonnage. Cette situation peut alors remettre en cause l'idée que tous les essais ont estimé le même ET unique. En effet cet ET peut également varier en raison de caractéristiques propres aux patients inclus dans l'essai, à la manière dont le traitement a été réalisé ou à la période temporelle. Cette situation ne correspond plus à un ET unique mais à une *distribution* d'ET autour d'un ET moyen. Les modèles statistiques utilisés permettent alors de prendre en compte cette situation. Quand l'hétérogénéité est faible, on utilise un *modèle à effets fixes*. Cette modélisation suppose que la seule cause de variabilité d'un essai à l'autre est la fluctuation d'échantillonnage. Quand l'hétérogénéité est forte, on utilise un *modèle à effets aléatoires*. Cette modélisation permet d'ajouter au modèle à effet fixe, une seconde source de variabilité dépendante de l'essai.

Les MA sont actuellement considérées comme la méthode donnant le plus haut niveau de preuve scientifique lors d'une comparaison entre deux traitements. Toutefois la qualité de la méta-analyse est très dépendante de la qualité des études la composant. Réaliser une MA ne permet pas de s'affranchir de la plupart des biais présents dans les études et ces biais limiteront le niveau de preuve final. Une évaluation des biais des études incluses doit donc être réalisé. Nous ne parlerons ici que de MA d'essai randomisé, car ces études fournissent le plus haut niveau de preuve. Toutefois les concepts sont applicables à de nombreux autres designs d'études et ne portent pas que sur l'étude d'un traitement.

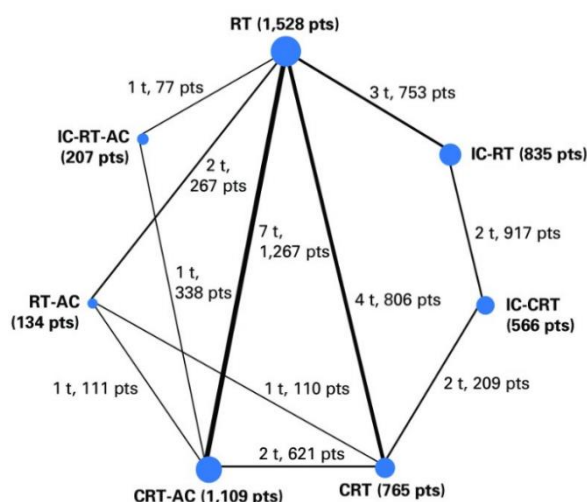
Une dernière condition est nécessaire pour que les conclusions soient valides : il faut impérativement avoir identifié la *totalité des essais* réalisés. En effet, ne combiner que les essais ayant mis en évidence un effet traitement favorable conduit à une estimation biaisée de l'ET (biais de sélection). Cette situation est malheureusement fréquente car, pour diverses raisons, les études n'ont pas la même probabilité d'être publiées (ou publiées dans un journal indexé) selon leurs résultats (*biais de publication*). Il faut donc s'acharner à identifier tous les essais, publiés ou non, en croisant les méthodes d'identification. Certaines équipes testent cette hypothèse en regardant la distribution des ET en fonction de la taille de l'échantillon.

## Méta-analyse en réseau

Les méthodes de MA standard sont applicables dans la comparaison de deux traitements l'un face à l'autre. Pourtant, les situations cliniques où il n'existe que deux traitements possibles sont de plus en plus rares. Les méta-analyses en réseau (MAR), étendent le principe des MA à une situation où plus de deux traitements existent. Décrites depuis le début des années 2000,(12) elle ne se sont couramment utilisées que depuis moins de 10 ans. Dans cette approche, un « *réseau de traitement* » est établi. Chaque *nœud* du réseau représente une option thérapeutique (*un traitement*) et les *arêtes* reliant ces nœuds représentent les comparaisons (ici *des essais*) disponibles entre ces deux traitements. La **Figure 5** montre les deux plus simples réseaux, constitués uniquement de trois traitements A, B et C. Dans le réseau de gauche, il n'existe que deux types d'essais : A contre C et B contre C. Dans le réseau de droite il existe un troisième type d'essais A contre B réalisant une *boucle fermée*.



**Figure 5 :** Un réseau simple à trois traitements A, B et C sans (gauche) ou avec (droite) boucle fermée



**Figure 6 :** Exemple de réseau de traitement dans une méta-analyse en réseau d'essais d'ORL (d'après Ribassin-Majed et al, J Clin Oncol, 2017)

Chaque point représente un traitement avec le nombre de patients l'ayant reçu. Chaque trait représente un ou plusieurs essais et l'épaisseur du trait est proportionnel au nombre de patients. RT : radiothérapie, CRT : chimioradiothérapie, AC : chimiothérapie adjuvante, IC : chimiothérapie d'induction

Chaque comparaison entre deux traitements peut ne jamais avoir été réalisé (la branche n'est alors pas dessinée) ou avoir été réalisé une ou plusieurs fois (l'épaisseur de la branche est alors proportionnelle à la quantité d'informations disponibles). Il n'y a pas de limite au nombre de traitement pouvant être inclus et les réseaux obtenus peuvent prendre des géométries plus ou moins complexes. La **Figure 6** montre un exemple de réseau plus complexe (7 traitements, 11 types d'essais) récemment publié par notre équipe en cancérologie ORL.

D'un point de vue clinique, ce type d'étude présente deux avantages : l'utilisation des données indirectes pour la comparaison et le classement des traitements. Ce premier avantage désigne le fait que, dans ce type d'étude, l'évaluation d'une comparaison ne repose pas uniquement sur les études les ayant comparés directement. L'utilisation du réseau permet de combiner ces informations que l'on appelle *directes*, aux informations passant via le réseau par des comparateurs communs que l'on appelle *indirectes*. Dans l'exemple de la **Figure 6**, la comparaison entre CRT-AC et CRT n'est donc pas estimée uniquement par les deux essais les ayant comparés directement mais également indirectement via le traitement RT qui a été comparé dans sept essais à CRT-AC et dans quatre essais à CRT. Ceci est particulièrement intéressant quand peu, voire aucun essai n'ont estimé la comparaison d'intérêt. Toujours dans le réseau de la **Figure 6**, la comparaison RT contre IC-CRT est donc également estimable via le réseau, bien qu'aucun essai n'ait comparé directement ces deux traitements. Cette méthode a toutefois ses limites : il faut quatre fois plus de patients dans une comparaison indirecte pour obtenir la même puissance que dans une comparaison directe. Le second avantage est de pouvoir, à partir du réseau, identifier la meilleure stratégie thérapeutique. Il est en effet possible de classer les traitements même si toutes les comparaisons ne sont pas disponibles, ce qui constitue un intérêt majeur.



Toutefois, pour que les conclusions d'une MAR soient valides, deux hypothèses statistiques *supplémentaires* par rapport aux MA sont nécessaires : la *transitivité* et la *cohérence*. L'hypothèse de *transitivité* désigne le fait qu'un patient inclus dans l'un des essais du réseau aurait également pu être inclus dans n'importe lequel des autres essais. Cette hypothèse ne peut pas être testée statistiquement. L'évaluation du respect de cette hypothèse repose donc sur l'analyse des critères d'inclusions des essais et des indications des traitements que l'on a comparés. L'hypothèse de *cohérence* désigne le fait que l'estimation de l'ET apportée par les comparaisons directes doit être la même que celle apportée par les comparaisons indirectes. Il est uniquement possible de vérifier cette hypothèse dans les comparaisons disposant d'une *boucle fermée* (i.e. comportant une comparaison directe). Dans l'exemple de la **Figure 6** il est donc possible de tester cette hypothèse pour la comparaison CRT-AC contre CRT car on dispose d'informations directes et indirectes. Il est en revanche impossible de la tester dans la comparaison RT contre IC-CRT. Si de *l'incohérence* est détectée, il faut essayer d'expliquer celle-ci (par exemple en ajustant sur des covariables) avant de pouvoir combiner les estimations.

## Données individuelles et données agrégées

Pour réaliser une MA ou une MAR il faut disposer de l'ET et de sa variabilité, le plus souvent représentée par son écart type. Deux méthodes sont possibles : les *données agrégées* (AD) réalisées en pratique dans 99% des cas (13) et les *données individuelles des participants* (IPD).

Dans la méthode basée sur les *données agrégées*, l'estimateur de l'effet traitement et son écart type sont récupérés agrégés par étude à partir des résultats publiés ou donnés par les investigateurs. Cette méthode a pour principal avantage d'être plus simple et rapide car elle simplifie les contacts avec les équipes ayant réalisé les essais et les mesures réglementaires. Le principal désavantage est que le contrôle des données est très limité. Un autre désavantage est que la forme sous laquelle est présenté l'effet traitement n'est pas forcément homogène d'une étude à l'autre. Par exemple pour une étude de type temps avant événements, celui-ci peut être présenté sous la forme d'un hazard-ratio, d'une différence de médiane, de comparaison de pourcentage de survie à un temps donné, de courbe de survie uniquement ou simplement sous la forme d'une p-value de test du log-rank. Une double extraction par deux investigateurs différents est souvent réalisée afin de limiter les erreurs. Les résultats sont par la suite combinés en pondérant les effets traitements par l'inverse de leur variance. L'analyse de l'effet des covariables est très limitée, restreinte à la répartition de celle-ci et parfois à la variabilité de la survie dans les bras.

Dans la méthode basée sur les *données individuelles des participants*, les critères de jugement ainsi que toutes les covariables de chaque patient sont récupérés. L'effet traitement est alors recalculé avec son écart type. Le principal frein à l'utilisation de cette méthode est qu'elle prend plus de temps car elle nécessite de pouvoir effectivement contacter et obtenir la coopération des autres équipes. De plus les contraintes réglementaires liées à la protection des données, qui varient d'un pays à l'autre, ajoute de la complexité. Une autre difficulté est le traitement d'un volume de données qui peut devenir important et l'utilisation de méthodes statistiques plus complexes. En revanche les avantages sont nombreux :

- Meilleur contrôle des données
- Données de survie mise à jour par rapport à la publication
- Possibilité de recalculer l'effet traitement sous la forme désirée
- Explication de l'hétérogénéité
- Contrôle des exclusions après randomisation
- Etude de l'interaction (cf. infra)

Ces différents avantages font que cette méthode est souvent considérée comme le « gold-standard » de la méta-analyse.

## Interaction dans les méta-analyses

On appelle variable en interaction avec l'ET ou modificateur de l'ET, toute variable pour laquelle l'ET varie en fonction de la valeur prise par la variable. Ce modificateur peut être qualitatif ou quantitatif, et constitue à l'ère de la médecine de précision, une voie de recherche majeure. En effet, identifier une interaction peut permettre d'identifier un sous-groupe de patients pour lequel l'effet traitement est nul et qui ne devrait donc pas recevoir le traitement, ou au contraire un sous-groupe de patients pour lequel l'effet traitement est encore plus grand.

Cette interaction pourrait être recherchée dans chaque essai mais les essais sont le plus souvent dimensionnés afin de démontrer un bénéfice dans la population globale et non dans un sous-groupe. Il est donc souvent difficile de mettre en évidence celle-ci à l'intérieur d'un essai unique, d'autant que les tests la recherchant sont peu puissants. En combinant différents essais, et donc en augmentant la puissance, les MA permettent d'étudier plus finement ces interactions. Toutefois, cette étude n'est complètement possible que dans le cadre d'une MA-IPD car elle nécessite de pouvoir calculer des effets ajustés sur le modificateur.

Outre l'aspect médecine de précision, un autre risque est celui de biais dans l'effet traitement. En l'absence d'interaction il est bien démontré que les MA basées sur les IPD et les AD aboutissent aux mêmes résultats. En revanche en présence d'une interaction, l'estimation faite dans le cadre d'une AD-MA va dépendre de la répartition/distribution moyenne de la variable dans les essais. En effet, si les essais ont inclus des patients plus favorables, on observera un effet plus favorable qu'il ne l'est dans la population générale sans possibilités de correction. Cette situation est connue et très bien étudiée dans le cadre des MA standard.

Le problème se révèle plus complexe dans le cadre d'une MAR. En effet, l'interaction peut exister dans certaines comparaisons mais pas dans d'autres. De même, la distribution du modificateur peut être différente d'une comparaison à l'autre. Dans une telle situation, il existe un biais potentiel dans l'estimation de l'ET si l'on se base uniquement sur les AD. Si le problème commence à être décrit dans le cadre d'un critère de jugement binaire, celui-ci est très peu étudié dans le cadre de données censurées. Une des difficultés provient des modèles utilisés en survie, plus complexes pour un critère binaire,

notamment quand on utilise les IPD. Les rares équipes travaillant sur la thématique utilisent le plus souvent une approche Bayésienne.

## Objectif statistique de la thèse

L'objectif statistique de la thèse était de proposer un modèle statistique permettant d'étudier l'interaction dans une MAR faites sur les IPD pour un critère de jugement de type survie avec un modèle estimé sous paradigme fréquentiste. Cette thématique complète bien la problématique clinique puisque celle-ci a pour critère principal un critère de survie et que l'interaction entre l'ET et les caractéristiques du patient constitue le cœur de la problématique clinique. Le nouveau modèle développé sera donc utilisé pour répondre à la problématique clinique.

## Chapitre 3. Modèles mathématiques utilisés dans les MAR

### Estimation de l'effet traitement pour des données censurées

Lorsque le critère principal de jugement est une survie, on souhaiterait pouvoir mesurer le *temps avant qu'un événement* ne se produise. La différence de temps entre les deux groupes estimerait alors l'effet du traitement. Toutefois, dans la plupart des études cliniques, de nombreux patients ne présentent pas l'événement d'intérêt au cours de l'étude. Leur délai de survie est donc inconnu et l'on dit de ces observations qu'elles sont *censurées*. Si le temps avant de présenter un événement est inconnu pour certains patients, le *risque* de présenter un événement chaque jour qui passe est en revanche toujours calculable. L'estimation est basée, chaque jour, sur le rapport entre le nombre d'événements et le nombre de personnes à risque le jour dit. La plupart des méthodes de survie sont donc basées sur l'estimation du *risque de décès* dans chaque groupe. L'effet traitement est alors exprimé sous la forme du *rapport de ces risques instantanés* (appelé *Hazard Ratio* (HR)) entre les groupes à comparer. Trois méthodes d'estimation du HR ont été utilisées au cours de cette thèse : le log-rank, le modèle de Cox et le modèle de Poisson. Au cours des applications statistiques le  $\log(\text{HR})$  était souvent présenté au lieu du HR lui-même car il correspond au coefficient estimé par le modèle. Pour les applications cliniques, le HR lui-même était présenté.

#### Méthode du Log-rank

Peto(14,15) a proposé d'utiliser le rapport entre le nombre d'événements observés et le nombre d'événements attendus (O/A) comme indicateur du HR. Cet estimateur peut être calculé avec son écart type (SE) dans chaque essai de la façon suivante :

$$\log(\text{HR}) = \frac{(O - A)}{V}$$

$$SE = \frac{1}{\sqrt{V}}$$

Avec O : la somme du nombre d'événements observés et E : la somme du nombre d'événement attendus dans le groupe expérimental à chaque pas de temps et V la variance de la statistique du log-rank. Cette méthode est assez simple et peut même être calculée avec une simple calculatrice. Elle est disponible dans tous les logiciels. Toutefois cette statistique est de plus en plus rarement rapportée dans les publications d'essais randomisés au profit du HR provenant du modèle de Cox.

## Modèle de Cox

Le modèle de Cox s'écrit pour une seule variable X :

$$h(t) = \lambda_0(t)\exp(X\beta)$$

Avec  $h(t)$  : le risque au temps  $t$ ,  $\lambda_0(t)$  la fonction de risque de base au temps  $t$  (non paramétrisée) et  $\beta$  le log(HR) associée à la variable X.

Même dans cette situation simple, il n'existe pas de formule pour calculer  $\beta$  et son écart type. Ceux-ci sont donc estimés (et non calculés) par *maximum de vraisemblance*. Il s'agit toutefois d'une méthode courante disponible dans la plupart des logiciels.

## Modèle de Poisson

Le HR peut également être estimé par un modèle linéaire généralisé de Poisson au prix de transformations minimales du jeu de données. La principale modification est la nécessité de découper le temps en intervalles (ou périodes). Chaque patient apparaît alors sur une ou plusieurs lignes du nouveau jeu de données en fonction de sa survie. Un patient présentant l'évènement ou une censure au cours d'une période a un délai dans cette période égal au délai entre le début de la période et l'évènement. Il ne sera pas reconduit sur la période suivante. En revanche, l'observation des patients n'ayant pas présenté l'évènement au cours d'une période est notée comme censurée avec un temps de suivi égal à la totalité de la période, puis est reconduit dans la période suivante. A l'issue de la dernière période, tous les patients sans évènements sont censurés. On obtient un jeu de données de « pseudo-patients » plus long que le jeu de données originales. Le modèle de Poisson s'écrit alors pour une seule variable X :

$$\log(y) = \beta_b + \beta_p * \text{periode} + \beta_1 X + \log(D)$$

Avec  $y$  l'indicateur d'évènement,  $\beta_b$  la fonction de risque de base,  $\beta_p$  une fonction de la période permettant de faire varier le risque au cours du temps,  $\beta_1$  le log(HR) associé à la variable X, et  $\log(D)$  le logarithme du temps de participation dans la période. Ce dernier est utilisé comme un « offset » dans le modèle et permet de transformer la variable de compte caractéristique du modèle de Poisson en un taux. Il est possible de modéliser plus ou moins finement la variation du risque de base en augmentant le nombre d'intervalles (même si celui-ci joue peu sur la qualité de l'estimation) et en modélisant différemment la fonction de la période dans le modèle. La période a été représentée par une variable continue modélisée par une spline naturelle qui présente l'avantage d'être une fonction continue et continuellement dérivable. La modélisation comme un facteur (décomposition en variable indicatrice) est également possible. Ce modèle peut alors être estimé par maximum de vraisemblance comme un *modèle linéaire généralisé* à fonction de lien logarithmique et à distribution des résidus selon une loi de Poisson. Parmi les avantages de ce type de modèle sur le modèle de Cox, on peut noter la possibilité d'estimer la fonction de risque de base et une plus grande facilité à modéliser une non-proportionnalité des risques (même si celle-ci est également possible dans le Cox). Si la fonction permettant de transformer le jeu de données initiales en pseudo-patient n'est pas toujours disponible, l'estimation d'un modèle linéaire généralisé est disponible dans la plupart des logiciels. Ce type de modèle bénéficie par ailleurs d'implémentation plus rapide et plus variée dans les modèles multi-niveaux complexes que le modèle de Cox.

## Modèles basés sur les données individuelles

### Modèle en une ou deux étapes

Deux grandes méthodes d'exploitation des données individuelles sont possibles : en une ou en deux étapes. Dans la méthode en *deux étapes*, les essais sont analysés séparément afin de recalculer l'effet traitement et son écart type, en ajustant si nécessaire sur d'autres variables. Ces estimateurs sont alors combinés par exemple selon la méthode habituelle de l'inverse variance.

Dans la méthode en *une seule étape*, les données sont utilisées de façon simultanée dans un *modèle unique*. Il s'agit alors d'un *modèle hiérarchique* (on dit aussi mixte ou multi-niveaux). Ce type de modèle prend en compte la structure sous-jacente des données et le fait que deux patients inclus dans le même essai se ressemblent plus que deux patients pris au hasard. Dans le contexte d'une NMA les observations des patients sont nichées dans leurs essais et les essais sont éventuellement nichés dans les comparaisons. Ce type de modèle permet également l'utilisation d'*effet aléatoire*. Ceci signifie que l'on peut laisser varier la valeur d'un paramètre d'un essai à l'autre tout en contraignant leur distribution globale pour éviter les effets extrêmes. On peut ainsi laisser varier le *risque de base* on parle alors d'*intercept aléatoire* ou l'effet d'une variable sur le critère de jugement on parle alors de *pente aléatoire*. Selon la situation, l'utilisation de ces effets aléatoires en alternative à la stratification permet de ne pas perdre le bénéfice de la randomisation. Ce type de modèle est plus séduisant dans le principe puisqu'il permet notamment d'inclure plus facilement les interactions. Il peut toutefois être la source de difficultés d'estimation, notamment dans les modèles complexes.

Il est actuellement considéré que, dans la plupart des situations, les approches en une ou deux étapes aboutissent à des estimations très proches à condition que des modèles partageant les mêmes suppositions statistiques soient utilisés.(16,17) Un avantage au modèle en une étape pourrait être observé dans le cas d'essais de petite taille,(18) alors que le modèle en deux étapes permettrait quant à lui de combiner plus facilement données individuelles et données agrégées si l'on pas réussi à récupérer toutes les IPD.

### Codage des traitements

#### *Paramètres principaux et paramètres fonctionnels*

Dans une IPD-MAR les variables indicatrices du traitement reçu doivent être choisies de façon minutieuse. En effet ces modèles doivent respecter l'hypothèse de cohérence. Si l'on compare  $n$  traitements, seulement  $n-1$  *paramètres principaux* sont nécessaires pour coder les  $n$  effets traitements. Le dernier, appelé *paramètre fonctionnel*, est déduit par contraste à partir des autres. Ce choix est arbitraire et n'a pas de conséquence sur les résultats. Par exemple, dans un réseau à trois traitements A, B et C, si l'on décide de mettre le paramètre fonctionnel sur la comparaison AB, celle-ci est alors estimée à partir des paramètres AC et BC selon

$$\log(HR_{AB}) = \beta_{AB} = \beta_{AC} - \beta_{BC}$$

Le même principe est appliqué aux interactions. Le paramètre  $\gamma_{AB}$  associé à l'interaction entre A-B et une variable X peut être estimé à partir des paramètres associés aux interactions entre X et AC ( $\gamma_{AC}$ ) et X et BC ( $\gamma_{BC}$ ) selon :

$$\gamma_{AB} = \gamma_{AC} - \gamma_{BC}$$

Ainsi, pour une valeur x de la variable X, le  $\log(\text{HR})$  de l'ET A-B est donné par :

$$\beta_{AB}^{X=x} = (\beta_{AC} - \beta_{BC}) + (\gamma_{AC} - \gamma_{BC}) \times x$$

### Signe des effets traitements

La valeur donnée aux variables doit également refléter cette contrainte. Dès lors qu'il existe une boucle fermée et/ou s'il y a plus de trois nœuds, il faut alors également définir arbitrairement un sens pour chaque comparaison possible. Le sens de la comparaison est représenté par convention par une flèche partant du traitement « évalué » vers le traitement « de référence » :  $A \rightarrow B$  signifiant A versus B. L'indicatrice de traitement associé à cette comparaison peut alors prendre la valeur 1 si le patient reçoit A et 0 si le patient reçoit B, toutes les autres indicatrices de traitement recevant la valeur 0. Toutefois, dans un modèle avec intercept aléatoire, il est préférable de respectivement les coder +0.5/-0.5 afin que la variance soit correctement estimée dans tous les groupes. (19) Le codage du traitement reçu par les patients inclus dans les essais dont la comparaison est codée par un paramètre fonctionnel se fait alors par les autres indicatrices. Les signes étant choisis pour respecter l'hypothèse de cohérence. Un exemple de codage est donné dans le **Tableau 3**.

*Tableau 3: Codage des 3 effets traitements par 2 indicatrices dans un réseau à 3 nœuds*

Type d'essai	Bras de randomisation	Indicatrice de traitement 1	Indicatrice de traitement 2
<b>A contre C</b>	A	+0.5	0
	C	-0.5	0
<b>B contre C</b>	B	0	+0.5
	C	0	-0.5
<b>A contre B</b>	A	+0.5	-0.5
	B	-0.5	+0.5

Le paramètre fonctionnel est ici celui de la comparaison A contre B

### Le modèle de Cox hiérarchique

Le modèle de Cox, peut s'écrire sous la forme d'un modèle mixte avec des effets aléatoires sur « l'intercept » et les covariables. (20,21) Pour  $X^{(1)}$  une variable à effet fixe et  $X^{(2)}$  une variable à effet aléatoire, le modèle s'écrit pour le patient i inclus dans l'essai j :

$$h(t) = \lambda_0(t) \exp\left(b_j^{(0)} + X_{ij}^{(1)}\beta^{(1)} + X_{ij}^{(2)}(\beta^{(2)} + b_j^2)\right)$$

$$\begin{bmatrix} b_j^{(0)} \\ b_j^{(2)} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & \tau^2 \end{bmatrix}\right)$$

Avec  $h(t)$  la fonction de risque au temps  $t$ ,  $\lambda_0(t)$  la fonction de risque de base au temps  $t$  commune à tous les essais,  $b_j^{(0)}$  la déviation du risque de base propre à l'essai  $j$ ,  $\beta^{(1)}$  le  $\log(\text{HR})$  associé à  $X^{(1)}$ , commun à tous les essais,  $\beta^{(2)}$  le  $\log(\text{HR})$  associé à  $X^{(2)}$  commun à tous les essais,  $b_j^2$  une déviation autour du  $\log(\text{HR})$  de  $X^{(2)}$ , propre à chaque essai. Les deux effets aléatoires sont distribués normalement avec  $\sigma^2$  la variance de l'effet aléatoire associé au risque de base et  $\tau^2$  la variance de l'effet aléatoire associé à  $X^{(2)}$ .

## Le modèle de Poisson hiérarchique

Le modèle de Poisson peut lui aussi s'écrire sous la forme d'un modèle hiérarchique appartenant à la famille des modèles linéaires généralisés mixtes (GLMM). (22,23) Celui-ci s'écrit alors pour  $X^{(1)}$  une variable à effet fixe et  $X^{(2)}$  une variable à effet aléatoire pour le patient  $i$  inclus dans l'essai  $j$  :

$$\log(y) = (\beta_0 + b_{0j}) + \beta_p * \text{periode} + X_{ij}^{(1)}\beta^{(1)} + X_{ij}^{(2)}(\beta^{(2)} + b_j^2) + \log(D)$$

$$\begin{bmatrix} b_{0j} \\ b_j^2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & \tau^2 \end{bmatrix}\right)$$

Avec  $\log(y)$  le logarithme de l'indicateur d'évènement,  $\beta_0$  l'intercept du modèle représentant le risque de base, commun à tous les essais,  $b_{0j}$  une variation autour du risque de base spécifique à chaque essai,  $\beta_p * \text{periode}$  la variation du risque au cours des intervalles généralement modélisés par une spline pour permettre une relation non linéaire tout en maintenant la continuité (22) ,  $\beta^{(1)}$  le  $\log(\text{HR})$  associé à  $X^{(1)}$ , commun à tous les essais,  $\beta^{(2)}$  le  $\log(\text{HR})$  associé à  $X^{(2)}$  commun à tous les essais,  $b_j^2$  une déviation autour du  $\log(\text{HR})$  de  $X^{(2)}$ , propre à chaque essai et  $\log(D)$  le logarithme du temps de participation dans l'intervalle utilisé comme un offset du modèle.

## Décomposition de l'interaction

Les deux modèles hiérarchiques précédents permettent d'introduire des interactions entre des variables (continues ou catégorielles). L'interaction entre une variable catégorielle  $X^{(1)}$  et une numérique  $X^{(2)}$  s'écrit alors :

$$X_{ij}^{(1)}\beta^{(1)} + X_{ij}^{(2)}\beta^{(2)} + X_{ij}^{(1)}X_{ij}^{(2)}\gamma^{(1)}$$

Avec  $\gamma^{(1)}$  la variation dans le  $\log(\text{HR})$  de  $X^{(1)}$  pour chaque augmentation d'une unité de  $X^{(2)}$  (ou de façon similaire la variation du  $\log(\text{HR})$  associée à  $X^{(2)}$  pour la catégorie alternative de  $X^{(1)}$ ). Dans cette paramétrisation  $\beta^{(1)}$  est le  $\log(\text{HR})$  associé à  $X^{(1)}$  pour un individu ayant  $X^{(2)}$  égal à zéro. Dans certaines situations, ce  $\log(\text{HR})$  peut alors être celui d'un patient ne pouvant exister (e.g. patient de taille ou de poids nuls) et donc ne pas être directement interprétable. On peut alors décider de centrer  $X^{(2)}$ , ce qui permettra au HR de s'interpréter comme celui d'un patient ayant pour valeur de  $X^{(2)}$  sa moyenne.



Dans l'équation ci-dessus, le paramètre  $\gamma^{(1)}$  estime ici l'interaction globale dans le jeu de données. Dans le cadre des méta-analyses, cette formulation a pour risque de ne pas séparer l'interaction intra-essais et l'interaction inter-essais, ce qui peut conduire à des résultats biaisés. (24) Hua (25) a récemment proposé de séparer ces deux composantes en remplaçant le terme  $X_{ij}^{(1)} X_{ij}^{(2)} \gamma^{(1)}$  par la formulation suivante :

$$\gamma_{intra}^{(1)} * X_{ij}^{(1)} * (X_{ij}^{(2)} - \overline{X_j^{(2)}}) + \gamma_{inter}^{(1)} * X_{ij}^{(1)} * \overline{X_j^{(2)}}$$

Avec  $\gamma_{intra}^{(1)}$  le terme interaction intra-essais,  $\gamma_{inter}^{(1)}$  le terme d'interaction inter-essais,  $\overline{X_j^{(2)}}$  la valeur moyenne de  $X^{(2)}$  dans l'essai j. Seuls les résultats intra-essais devraient être utilisés pour estimer la présence d'une interaction. Par ailleurs, l'existence d'une différence entre  $\gamma_{intra}^{(1)}$  et  $\gamma_{inter}^{(1)}$  signale l'existence d'un biais écologique dans les données.

## Modèles basés sur les données agrégées

Les modèles basés sur les données agrégées nécessitent d'obtenir au préalable pour chaque essai, l'ET ainsi que sa variance.

### Modèle basé sur les contrastes

L'extension naturelle de la méta-analyse deux à deux pondérée par l'inverse variance est la *méthode basée sur les contrastes*. (26) Un contraste est la différence d'efficacité entre deux traitements. Ceci signifie qu'un essai A contre B n'est pas utilisé pour estimer l'effet d'A ou l'effet de B, mais plutôt la différence d'efficacité entre les deux (i.e. le contraste). Un essai à deux bras évalue donc un contraste, mais un essai multi-bras en évalue plusieurs. L'approche opposée est appelée : « *basée sur les bras* ». Elle mesure l'effet du traitement dans chaque bras. Elle peut être intéressante dans certaines situations : nombres d'évènements, effet du traitement sur le poids, etc... Elle ne sera pas abordée ici.

### Matrice d'incidence ou de design

Les traitements reçus n'étant plus codés au niveau patient par des indicatrices, il faut pouvoir décrire la géométrie du réseau. On utilise pour ceci une *matrice d'incidence* qui décrit comment sont reliés les nœuds (ou traitement) par des arêtes (comparaison). Cette notation va, comme dans les modèles IPD, nécessiter de définir arbitrairement un ordre pour les traitements et un sens pour les comparaisons. La matrice d'incidence a autant de lignes que de comparaisons et autant de colonnes que de traitements possibles. Pour chaque comparaison (chaque ligne), la valeur 1 est attribuée à la colonne correspondant à la colonne depuis lequel le traitement part et la valeur -1 à celle dans laquelle elle arrive. La matrice d'incidence d'un réseau de trois essais 1 : AC, 2 : BC et 3 : AB s'écrirait donc, avec les essais 1, 2 et 3 en lignes et les traitements A, B et C en colonne :

$$I = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

Cette représentation permet de décrire n'importe quel réseau aussi complexe soit-il. Elle est particulièrement utile si l'information que l'on a est l'effet du traitement dans chaque

bras (« basée sur les bras »). Si l'on souhaite plutôt représenter le réseau sous la forme de contraste, on peut utiliser une matrice différente appelée *matrice de design*. Chaque ligne de la matrice de design représente également une comparaison d'un essai. En revanche, chaque colonne représente maintenant un des  $k$  contrastes. S'il y a autant de lignes que d'essais, seules  $k-1$  colonnes sont nécessaires dans la pratique, la dernière étant déduite à partir des autres. Pour déterminer les valeurs de la matrice, il faut à nouveau définir arbitrairement des sens au contraste et choisir quel contraste sera représenté par un paramètre fonctionnel. Si on reprend la notation avec les flèches de la **Figure 5** réseau de droite (page 22), un 1 est mis dans la colonne quand la comparaison passe dans le sens de la flèche, un -1 si elle prend la flèche à l'envers. La matrice D s'écrit donc avec les trois essais AC, BC et AB en ligne et les contrastes AC et BC en colonne :

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{bmatrix}$$

### *Écriture du modèle*

Pour une série de  $j$  essais à deux bras estimant  $k$  contrastes, le modèle s'écrit :

$$y = X\mu + b + \varepsilon$$

Où  $y$  représente un vecteur de  $j$   $\log(\text{HR})$ ,  $X$  une matrice de design de dimension  $(j, k)$ ,  $\mu$  un vecteur des  $(k)$  contrastes sous-jacents,  $b$  un vecteur des  $j$  effets aléatoires propre à chaque essai et  $\varepsilon$  un vecteur des  $j$  erreurs aléatoires avec :

$$b \sim \mathcal{N}(0, \tau^2)$$

$$\varepsilon \sim \mathcal{N}(0, V); V = \text{diag}\{v_j\}$$

Avec  $V$  une matrice diagonale dont la  $j^{\text{ème}}$  entrée correspond à  $v_j$  la variance du  $\log(\text{HR})$  dans l'essai  $j$ .

Par exemple la méta-analyse en réseau de la **Figure 5** **Erreur ! Source du renvoi introuvable.** réseau de droite (page 22) qui comprend trois traitements et trois contrastes (AC, BC et AB) s'écrirait :

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \mu_{AC} \\ \mu_{BC} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{AC}^2 & 0 & 0 \\ 0 & \tau_{BC}^2 & 0 \\ 0 & 0 & \tau_{AB}^2 \end{bmatrix} \right)$$

Le 3<sup>ème</sup> contraste (A contre B) s'obtiendrait alors par  $\mu_{AB} = \mu_{AC} - \mu_{BC}$ .

On peut alors estimer les paramètres du modèle. Toutefois cette estimation n'est pas toujours facile, notamment quand le réseau devient complexe ou comprend des essais multi-bras. Pour cette raison, de nombreuses d'équipes utilisent une approche Bayésienne dans l'estimation des paramètres. Une autre possibilité est de simplifier le modèle. Par exemple, considérer que les variances de tous les effets aléatoires sont identiques  $\tau_{AC}^2 =$

$\tau_{AC}^2 = \tau_{AC}^2 = \tau^2$  ce qui permet de retomber sur un modèle identique à une méta-régression.

### Approche par théorie des graphes

En raison de ces difficultés calculatoires, une approche des méta-analyses en réseau basée sur la *théorie des graphes* a récemment été proposée par Rücker(27). Cette approche fréquentiste est basée sur l'analogie quasi-parfaite qu'il existe entre une méta-analyse en réseau et un réseau électrique (**Tableau 4**). De façon simplifiée, en utilisant la loi d'Ohm ( $U = R \times I$ ), cette analogie suppose que le courant qui circule entre deux points du réseau (le contraste pondéré entre deux traitements) est égal à la différence de tension entre les points (le contraste brut) divisée par la résistance (la variance). Les avancées mathématiques disponibles en *théorie des graphes* et en *calcul matriciel* permettent alors d'estimer les paramètres.(27)

Cette approche a été montrée comme équivalente aux autres approches (notamment Bayésienne), tout en étant moins intensive en temps de calcul. Elle permet de traiter des situations variées, dont des effets aléatoires et l'inclusion d'essais multi-bras. Elle ne permet en revanche pas, dans sa version actuelle, l'étude de l'interaction.

**Tableau 4 :** Correspondance entre un réseau électrique et une méta analyse en réseau

Réseau électrique	Méta-analyse en réseau
Noeuds du réseau (i)	Traitements
Connexion (k)	Comparaisons existantes
Résistance $R_k$	Variance $V_k$
Conductivité $\frac{1}{R_k}$	Précision $w_k = \frac{1}{V_k}$
Potentiel sur le noeud i	Résultats du traitement i
Tension entre $i_1$ et $i_2$	Effet traitement $i_1$ contre $i_2$
Courant entre $i_1$ et $i_2$	Effet traitement pondéré

Adapté d'après Rücker, Res Synth Meth, 2012

### Méta-régression

Sous l'hypothèse de cohérence, une AD-MAR peut également être vue comme un cas particulier de méta-régression.(26) Cette méthode, plus simple d'un point de vue calculatoire a l'avantage de pouvoir introduire facilement des interactions entre l'effet traitement et une variable mesurée au niveau de l'essai (interaction inter-essais). Cette dernière pouvant être une caractéristique de l'essai ou une variable agrégeant une information mesurée au niveau patient. En revanche, les essais multi bras sont très difficiles à prendre en compte.

Comme dans les méthodes précédentes, seules  $n-1$  variables indicatrices sont nécessaires pour coder les n effets traitements et il faut choisir arbitrairement une comparaison

comme étant la référence. Le **Tableau 5** résume le codage d'une NMA à trois traitements :

*Tableau 5 : Codage des variables indicatrices dans une méta-régression de trois traitements*

Type d'essai	Variable indicatrice 1 (« essai AC »)	Variable indicatrice 2 (« essai BC »)
<b>A contre C</b>	1	0
<b>B contre C</b>	0	1
<b>A contre B</b>	1	-1

Si la comparaison AB n'est pas présente, la deuxième variable indicatrice n'est pas nécessaire en raison de la présence d'un intercept dans le modèle.

Pour une NMA à 3 traitements A, B, C comportant des essais AC et BC, le modèle s'écrit pour l'essai j :

$$y_j = (\beta_0 + b_{0j}) + \beta_1 * \text{essai}_j^{AC} + \beta_2 * \text{essai}_j^{BC} + e_j$$

$$b_{0j} \sim \mathcal{N}(0, \tau^2)$$

$$e_j \sim \mathcal{N}(0, v_j)$$

Où  $y_j$  représente le log(HR) observé dans l'essai j,  $\beta_0$  l'effet traitement moyen du réseau,  $b_{0j}$  une variation autour de l'effet moyen propre à chaque essai,  $\text{essai}_j^{AC}$  et  $\text{essai}_j^{BC}$  deux variables indicatrices du type d'essai (**Tableau 5**)  $\beta_1$  et  $\beta_2$  la variation de l'ET par rapport à la moyenne s'il s'agit respectivement d'un essai de type AC ou BC.  $v_j$  représente la variance du log(HR) dans l'essai j, afin de pondérer chaque essai par son inverse-variance et  $\tau^2$  la variabilité du log(HR) moyen d'un essai à l'autre.

On peut alors estimer les effets traitements principaux par :

$$\beta_{AC} = \beta_0 + \beta_1$$

$$\beta_{BC} = \beta_0 + \beta_2$$

Et le paramètre fonctionnel  $\beta_{AB}$  par contraste :

$$\beta_{AB} = \beta_{AC} - \beta_{BC}$$

## Cas particulier des modèles additifs

Dans les modèles précédents, nous avons abordé une situation où les trois traitements A, B et C étaient considérés comme totalement différents. Dans certaines situations (fréquentes en oncologie), ces traitements peuvent en réalité être des séquences thérapeutiques ou des associations de composantes dont certaines peuvent être commune à plusieurs protocoles. (28) Soit quatre protocoles thérapeutiques A, B, C et D combinant différemment trois composantes : e, f et g selon :

- A = e

- $B = e + f$
- $C = e + g$
- $D = e + f + g$

Et cinq essais comparant ces protocoles et estimant donc des composantes différentes :

1. Essai A-B : estime  $f$
2. Essai A-C : estime  $g$
3. Essai A-D : estime  $f + g$
4. Essai B-D : estime  $g$
5. Essai C-D : estime  $f$

Une première approche pourrait être de considérer les quatre protocoles comme autant de traitements différents et utiliser les modèles classiques. Une approche alternative, reflétant mieux la structure sous-jacente, est de décomposer les protocoles pour estimer leurs composantes respectives. Cette approche est appelée : *modèle additif*.

## Approche par théorie des graphes

Pour estimer de façon adéquate ce réseau, la matrice de design doit refléter la structure sous-jacente des traitements et donc être modifiée.

Si le réseau de l'étude est composé des cinq essais ci-dessus, la *matrice de structure* du réseau, avec en ligne les cinq essais dans l'ordre et en colonne les quatre protocoles de traitements (A, B, C, D), serait :

$$S = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Ou 1 signifie le premier bras de l'étude, -1 le second bras et 0 que cette étude n'étudiait pas ce protocole.

La *matrice de combinaison* décrit comment les protocoles sont composés. Il faut choisir une composante de référence qui sera ici  $e$  et n'apparaîtra pas. Cette matrice de combinaison avec en ligne les protocoles (A, B, C et D) et en colonne les composantes ( $f$  et  $g$ ) serait :

$$C = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Ou 1 signifie que le protocole comprend cette composante et zéro qu'il ne la comprend pas. Compte tenu du choix de  $e$  comme référence, le protocole A ne comprend aucune composante et est donc une ligne de zéro et la colonne  $e$  n'apparaît pas.

La *matrice de design* est alors le produit matriciel de ces deux matrices :

$$D = S \times C = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ -1 & -1 \\ 0 & -1 \\ -1 & 0 \end{pmatrix}$$

Le modèle peut alors être estimé. Le package R proposé par l'équipe de Rücker dispose déjà d'une fonction [ `netcomb()` ] capable d'estimer un tel réseau.(28)

## Modèle IPD en une étape

Si on désire réaliser la même représentation dans un modèle IPD, il faut modifier le codage des indicatrices de traitement. On choisit à nouveau e comme catégorie de référence.

*Tableau 6 : Codage des indicatrices de traitement dans un modèle IPD additif*

Essai	Traitement	Indicatrice f	Indicatrice g
<b>A vs B</b>	e	-0,5	0
	e + f	+0,5	0
<b>A vs C</b>	e	0	-0,5
	e + g	0	+0,5
<b>A vs D</b>	e	-0,5	-0,5
	e + f + g	+0,5	+0,5
<b>B vs D</b>	e + f	0	-0,5
	e + f + g	0	+0,5
<b>C vs D</b>	e + g	-0,5	0
	e + f + g	+0,5	0

Le modèle est alors estimé de la même façon que dans l'approche classique. Toutefois, l'interprétation des  $\log(\text{HR})$  change. Le  $\log(\text{HR})$  associé à l'indicatrice de f s'interprète alors comme l'effet d'ajouter f à un protocole comprenant e, ajusté sur la présence de g. Celui de g, comme l'effet d'ajouter g à un protocole comprenant e ajusté sur la présence de f. Si suffisamment de données sont disponibles, on peut même envisager de rechercher un effet synergique (ou antagoniste) en ajoutant un terme d'interaction entre les deux indicatrices de traitement. Un coefficient du terme d'interaction positif montrerait alors une synergie et un coefficient négatif un effet antagoniste.

## Approches fréquentiste et Bayésienne

Les approches fréquentiste et Bayésienne divergent autant dans leur « philosophie » sous-jacente que dans la façon d'estimer les paramètres. Si l'approche fréquentiste domine très largement la littérature actuelle, la plupart des modèles peuvent être approchés par les deux méthodes avec des résultats très proches. Aucune des deux approches n'est considérée comme supérieure à l'autre, leurs avantages respectifs dépassant largement le cadre des MAR. (29)

Dans la philosophie fréquentiste, les paramètres à estimer sont ponctuels. Il n'y a donc pas de probabilité que le paramètre soit égal à une certaine valeur, soit il l'est soit il ne l'est pas. Les probabilités sont vues comme des *fréquences* d'observation de chaque valeur si on répétait un nombre infini de fois l'expérience sous certaines conditions. Une probabilité de 5% signifie donc que si on répétait l'expérience une infinité de fois, 5% des échantillons aurait cette valeur. Dans la philosophie Bayésienne, les paramètres sont des distributions (plus ou moins serrées) de valeurs. Les probabilités sont vues comme la *crédibilité* que l'on peut donner à chacune de ces valeurs. Une probabilité de 5% signifie, selon nos connaissances, une faible crédibilité pour cette valeur. Cette approche nécessite d'estimer la connaissance que l'on a avant l'expérience (le « prior ») pour la modifier par les valeurs observées en cours d'expérience pour former une nouvelle connaissance (le « posterior »). Cette nouvelle connaissance prend la forme d'une nouvelle distribution. L'approche fréquentiste exprime donc la probabilité que l'on avait d'observer nos données si notre hypothèse est vraie alors que l'approche Bayésienne estime la probabilité de l'hypothèse connaissant les données observées.

L'approche fréquentiste estime le plus souvent les paramètres par *maximum de vraisemblance*. La vraisemblance est la probabilité qu'on avait d'observer les données sous certaines conditions. L'algorithme cherche le jeu de paramètres qui maximisent cette valeur en se rapprochant progressivement par itération successive. Cette approche est assez rapide, mais dans les modèles complexes, il se peut qu'elle échoue. L'algorithme ne « trouve pas dans quelle direction aller », on dit qu'il *ne converge pas*. Dans l'approche Bayésienne, différents jeux de paramètres sont tirés aléatoirement à chaque itération puis évalués (Monte Carlo Markov Chain). Même dans un modèle complexe, ce type d'exploration finira donc toujours par trouver la meilleure solution, mais explorer suffisamment de combinaison peut être très long. Cette approche est souvent donc favorisée dans les MAR à modèle complexe au prix de temps de calculs de plusieurs heures voire plusieurs jours.

Seconde partie :

Evaluation d'un modèle IPD  
en une étape





## Chapitre 1. But de l'étude

Compte tenu de ses nombreux avantages, notre équipe recommande depuis longtemps de réaliser des IPD-MA plutôt que des AD-MA.(30) Outre le conséquent travail supplémentaire que cette démarche représente, l'approche IPD en une étape a un autre désavantage : la complexité du modèle nécessaire pour estimer les paramètres. Celle-ci découle de la structure hiérarchique des données dans lesquelles les patients sont inclus dans des essais. On souhaite à la fois prendre en compte la corrélation intra-essais et ne pas perdre le bénéfice de la randomisation en ne comparant à un patient donné qu'un patient du même essai. Ce problème s'est majoré avec le développement des NMA puisque les modèles sont encore plus complexes. L'approche traditionnelle consistait alors en une approche en deux étapes, ce qui permettait de bénéficier des avantages de l'IPD tout en se retrouvant dans une situation de calcul proche de l'approche sur AD.

La disponibilité d'ordinateurs de plus en plus puissants et d'algorithmes de plus en plus performants permet depuis plusieurs années d'estimer les paramètres de modèles multi-niveaux complexes. Les premières IPD-NMA ont alors été approchées sous paradigme Bayésien, probablement en raison de sa capacité à estimer des modèles complexes.

Ces problématiques sont encore moins étudiées pour l'analyse de données censurées. Si plusieurs articles ont analysé les performances pour un critère binaire ou continu (31), peu se sont intéressés aux données de survie. En effet, ces modèles un peu plus complexes et moins ubiquitaires bénéficient en règle générale des développements plus tardivement. Nous avons donc souhaité évaluer les performances d'un modèle IPD en une étape sous paradigme fréquentiste pour données censurées, situation jusqu'à présent très peu étudiée. L'évaluation des performances se fera en présence d'interactions puisque c'est notamment dans ce contexte que les modèles IPD sont le plus intéressants.

Pour évaluer ce modèle, une étude de simulation et une application sur données réelles ont été réalisées. L'étude de simulation est inspirée d'une étude précédente où nous avons observé une interaction dans une seule branche du réseau, avec une distribution du modificateur différente selon les branches. Il s'agit d'une situation à risque de biais, notamment pour les estimations indirectes ou mixtes directes/indirectes qui représentaient le paramètre principal d'intérêt. Les approches sur AD ont été utilisées comme référence de la magnitude du biais que l'on peut observer dans cette situation. Les mêmes modèles ont ensuite été appliqués sur les données réelles d'une NMA d'ORL.

## Chapitre 2. Méthodes

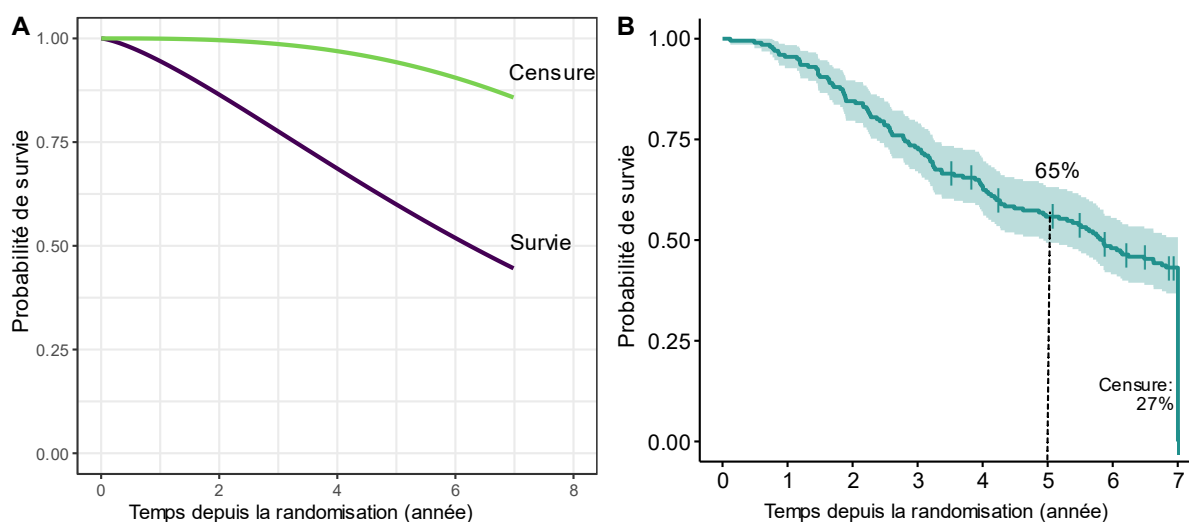
### Etude de simulation

#### Paramètres de simulations

Les simulations étaient basées sur une situation où il existe trois traitements : A, B et C. Ceux-ci étaient comparés dans des essais à deux bras avec une randomisation équilibrée. L'âge était considéré comme le modificateur de l'ET.

Les temps de survie étaient tirés selon une distribution de Weibull (intercept = 8, scale = 0,73) qui donnait une survie de 65% à 5 ans et 41 % à 8 ans. Les délais de censure étaient tirés selon une seconde distribution de Weibull (intercept = 8,5 et scale = 0,35) (**Figure 7**). Les patients pour qui le délai de censure était inférieur au délai de survie étaient considérés comme censurés. Toutes les observations étaient arbitrairement censurées après 7 ans de suivi pour imiter un essai ayant inclus des patients pendant 2 ans et les ayant suivis un minimum de 5 ans. Dans cette configuration le taux de censure dans chaque essai est voisin de 27%. (**Figure 7**). Pour le modèle de Poisson, l'étude était divisée en 7 intervalles d'un an. Ces paramètres ont été choisis pour mimer des délais de survie typiquement observés dans des essais de tumeurs solides localisées ou localement avancées.

Dans toutes les simulations, l'effet traitement était identique dans les comparaisons A-C et B-C. Sous l'hypothèse de cohérence, l'effet traitement de la comparaison A-B devait donc être nul. Des coefficients de  $\log(0.82) = -0.2$  et  $\log(0.61) = -0.5$  (HR = 0.61) ont été simulés avec un signe arbitrairement choisi de façon qu'A et B soit supérieur à C.



**Figure 7:** Simulation de temps de survies par la loi de Weibull

A : Lois de Weibull utilisées pour simuler la survie et la censure, B : simulation de 200 patients selon ces lois

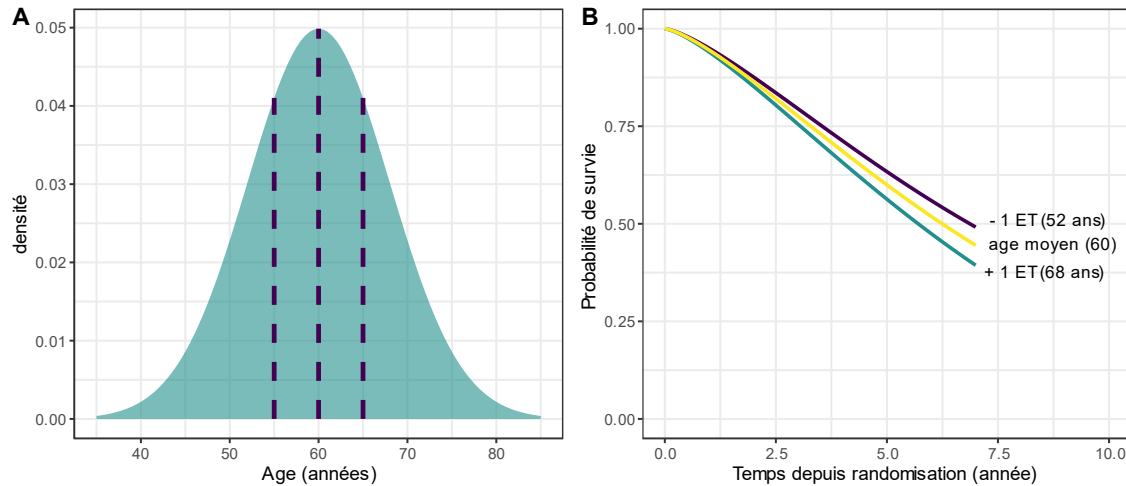


Figure 8 : Distribution de l'âge (A) et effet de l'âge sur la survie globale (B)

(A) les lignes pointillés sont à la moyenne  $\pm$  un écart type, (B) ET : écart type

Tableau 7: Paramètres variables dans les simulations

Paramètres	Valeurs possibles
Effet traitement	<ul style="list-style-type: none"> <li>-0.2 (hazard ratio=0.82)</li> <li>-0.5 (hazard ratio=0.61)</li> </ul>
Variance de l'effet aléatoire pour le risque de base ( $\delta_i^2$ , $i=1,2$ )	<ul style="list-style-type: none"> <li>0.1<sup>2</sup></li> <li>0.01<sup>2</sup></li> </ul>
Variance de l'effet aléatoire pour l'effet traitement ( $\tau_i^2$ , $i=1,2$ )	<ul style="list-style-type: none"> <li>0.1<sup>2</sup></li> <li>0.01<sup>2</sup></li> </ul>
Configurations	<ul style="list-style-type: none"> <li>I*: Pas d'essais A-B</li> <li>II : 3 essais de 200 patients dans la comparaison directe A-B</li> </ul>
Scenarios	<ul style="list-style-type: none"> <li>Rien (pas d'interaction avec l'âge ; même distribution de l'âge)</li> <li>Interaction (interaction âge/traitement dans A-C ; même distribution de l'âge)</li> <li>Combiné (interaction âge/traitement dans A-C ; âge moyen 52 ans dans les essais A-C et 68 ans dans les essais B-C)</li> </ul>
Nombre total de possibilités	$2 \times 2 \times 2 \times 2 \times 3 = 48$
Nombre de répliquons de chaque possibilités	1 000
<b>Nombre total de simulations :</b>	<b><math>48 \times 1\,000 = 48\,000</math></b>

Une variation du risque de base dans chaque essai était simulée suivant une distribution  $N(0, \delta^2)$ . Une variation inter-essais des effets traitements était également simulée

suivant une distribution  $N(0, \tau^2)$ . Une variance faible ( $0,01^2$ ) ou importante ( $0,1^2$ ) de ces deux variations étaient considérée.

Une variable continue, représentant l'âge, était tirée selon une loi normale (moyenne = 60, écart-type = 8). Celle-ci était par la suite considérée soit comme une variable continue (modèle IPD-Poisson et Metareg), soit catégorisée (modèles IPD-Poisson) en quatre classes définies a priori pour obtenir une répartition équilibrée (moins de 55, 55 à 60, 60 à 65 et plus de 65). L'âge avait un effet pronostique sur la survie de  $\log(\text{HR}) = 0.011$  pour chaque augmentation d'une unité d'âge (**Figure 8**).

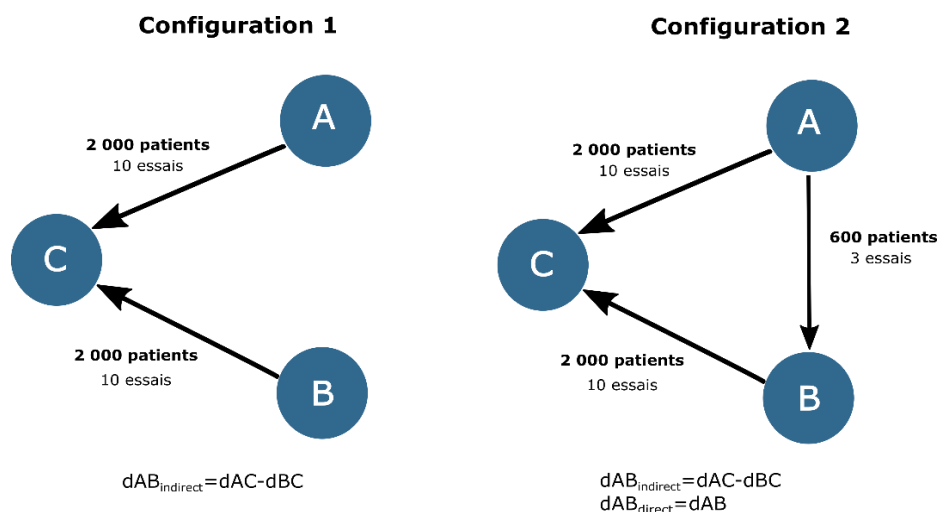
Chacune des combinaisons était répétée 1 000 fois pour deux configurations de réseau et trois scénarios (cf. infra). Les paramètres variables au cours des simulations sont résumés dans le

**Tableau 7.**

### Configuration du réseau

Deux configurations de réseaux évaluant les trois traitements A, B et C ont été simulées. Dans la *configuration I*, les deux traitements A et B ont séparément été comparés à un comparateur commun C, mais aucun essai comparant directement A et B n'était disponible. Pour chaque comparaison, 10 essais de 200 patients chacun étaient simulés (**Figure 9**). La comparaison A contre B (A-B) ne reposait donc, dans cette configuration, que sur les *comparaisons indirectes*.

Dans la *configuration II*, la boucle était fermée par trois essais A-B de 200 patients chacun. La comparaison A-B y était donc *mixte*, reposant à la fois sur des comparaisons *directes* et *indirectes*. Un nombre inférieur d'essais a été choisi dans la comparaison A-B afin de



*Figure 9 : Réseau de traitement simulé*

Le sens de la flèche indique le sens de la comparaison. Les chiffres au-dessus des flèches indiquent le nombre total de patients/d'essais disponibles dans chaque comparaison.

ne pas sur-représenter les informations directes dans la comparaison mixte. En effet une trop grande quantité d'informations aurait pu masquer le biais induit par les comparaisons indirectes. (Figure 9)

## Scénarios

Pour chacune de ces deux configurations, trois scénarios ont été simulés, variant selon la présence d'une interaction et la distribution de l'âge dans les comparaisons.

- *Scénario 1 – « Rien »* : Les âges des patients dans les essais A-C et B-C étaient tirés de la même distribution et aucune interaction n'était simulée. Sous l'hypothèse de cohérence, ce scénario aboutit à un effet traitement nul sans interaction dans la comparaison A-B
- *Scénario 2 – « Interaction »* : Les âges des patients dans les essais A-C et B-C étaient tirés de la même distribution, une interaction était présente dans A-C mais pas dans B-C. La valeur de cette interaction était fixée de sorte que l'effet traitement augmente de 25% pour chaque augmentation d'un écart type de l'âge. Sous l'hypothèse de cohérence, ce scénario aboutit pour la comparaison A-B à un effet traitement nul pour un patient d'âge égal à la moyenne de la distribution (60 ans) et à la même interaction que dans A-C
- *Scénario 3 « Interaction + distribution »* : les âges dans les comparaisons A-C et B-C étaient tirés de distributions différentes, un écart type plus jeune dans A-C et un écart type plus vieux dans B-C. La même interaction que dans le scénario 2 était présente dans A-C et il n'y avait pas d'interaction dans B-C. Sous l'hypothèse de cohérence, ce scénario aboutit pour la comparaison A-B à un effet traitement

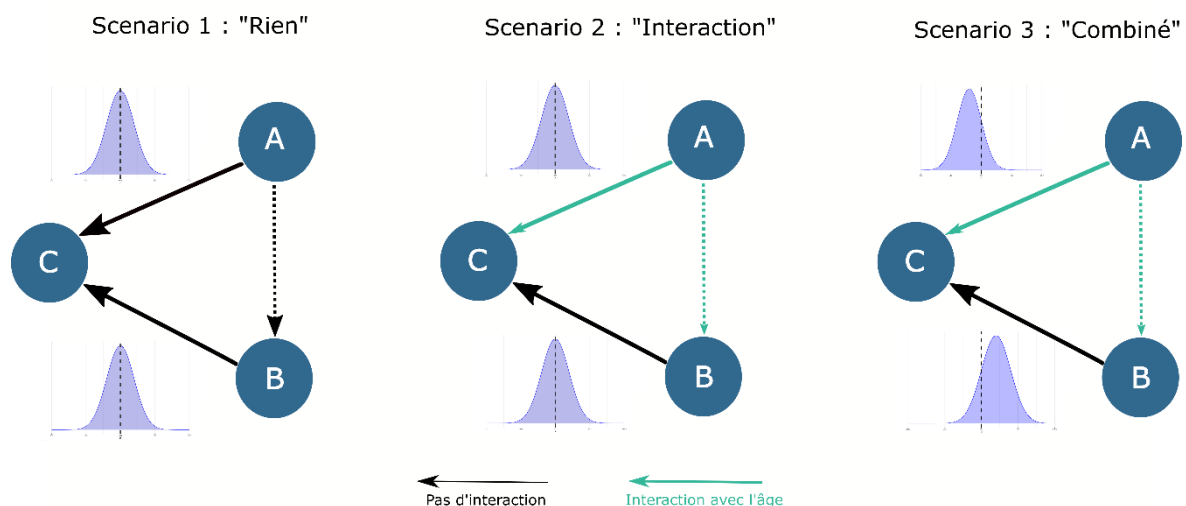


Figure 10 : Les 3 scénarios utilisés dans la simulation

Le sens de la flèche indique le sens de la comparaison, sa couleur la présence d'une interaction ou non, les flèches pleines indiquent que des essais sont disponibles dans les deux configurations, une flèche en pointillés que des essais peuvent être absent (configuration I) ou présent (configuration II). La courbe de densité indique la distribution de l'âge dans les différentes comparaisons avec la barre verticale positionnée sur 60 ans.

nul pour un âge égal à la moyenne de la distribution (60 ans) et à la même interaction que dans A-C

Dans la configuration II, l'âge des patients des essais A-B étaient tirés dans tous les scénarios selon la même distribution que dans le scénario 1.

## Présentation des résultats des simulations

Le paramètre d'intérêt principal était le *contraste* AB car il s'agissait de la comparaison estimée de façon *indirecte* uniquement ou *mixte* et donc la plus à risque de biais. Pour chacune des combinaisons de paramètres, de configurations et de scénarios, la moyenne, sur les 1 000 réplifications, des paramètres fonctionnels (*logHR*) associés à la comparaison A-B était calculée. Quand l'âge était considéré comme une variable quantitative, ces paramètres étaient : (i) le paramètre associé à l'effet traitement pour un patient d'âge moyen, noté  $ET_{\overline{age}}$  (correspondant ici à  $ET_{age=60}$ ) et appelé *effet conditionnel* du traitement et (ii) le paramètre associé à l'interaction (âge x ET) représentant la *variation d'ET* (*vET*) pour une augmentation d'une unité de l'âge, noté  $vET_{age+1}$ . Quand l'âge était considéré comme une variable qualitative, les paramètres d'intérêt étaient *l'effet traitement dans chaque classe d'âge*, notés  $ET_{<55}$ ,  $ET_{55-60}$ ,  $ET_{60-65}$  et  $ET_{>65}$ . Les résultats étaient également présentés sous la forme de biais moyen, défini comme la moyenne des différences entre la vraie valeur et la valeur estimée.

Les erreurs étaient présentées sous la forme d'Erreurs Empiriques Standards (EES) et d'Erreurs Standards Moyennes (ESM). L'EES était défini comme l'écart type de la distribution du paramètre estimé sur les 1 000 réplifications. Il montrait comment le paramètre estimé variait d'une simulation à l'autre. L'ESM était défini comme la moyenne des écart-types des coefficients prédits par le modèle. Il représentait donc l'incertitude moyenne autour du coefficient estimé par chaque modèle.

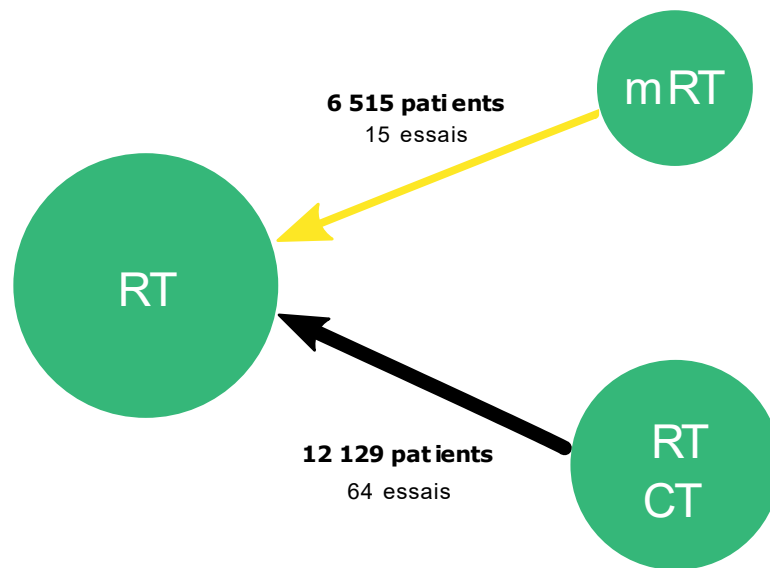
## Données réelles : MARCH/MACH-NC

La méta-analyse des chimiothérapies pour les cancers de la tête et du cou (MACH-NC) (32) a comparé la chimioradiothérapie (CTRTR) à la radiothérapie seule (RT) et comportait dans sa première mise à jour 64 essais et 12 129 patients. La méta-analyse de la radiothérapie dans les carcinomes de la tête et du cou (MARCH) (33) a comparé la radiothérapie modifiée (mRT) à la radiothérapie (RT) et comportait initialement 15 essais et 6 515 patients). Ces deux MA avaient comme critère de jugement principal la survie globale (Figure 11).

En les combinant,<sup>iii</sup> on forme une MAR comportant trois traitements (CTRTR, mRT et RT) mais uniquement deux comparaisons disponibles (pas d'essais CTRTR contre mRT). (34) Il s'agit donc d'une configuration équivalente à notre configuration I (Figure 11).

---

<sup>iii</sup> Des simplifications ont été faites pour ce travail méthodologique, aussi nous invitons un lecteur intéressé par les résultats à se rapporter aux publications originales.



*Figure 11 : Réseau MARCH/MACH-NC*

RT : radiothérapie, mRT : radiothérapie modifiée, RTCT : chimioradiothérapie. La taille des cercles est proportionnelle au nombre de patients ayant reçu ce traitement. L'épaisseur des flèches est proportionnelle au nombre de patients dans la comparaison. Leur sens indique le sens arbitrairement choisi pour la comparaison. La flèche jaune indique la comparaison où il y avait de l'interaction

Par ailleurs, les résultats publiés à partir des données individuelles dans MARCH montraient une interaction entre l'âge et la comparaison mRT-RT, et les patients inclus dans MACH-NC étaient en moyenne plus jeunes que ceux inclus dans MARCH générant une situation équivalente à celle de notre scénario 3.

## Modèles utilisés

Le principal modèle d'intérêt était le modèle IPD en une étape qui pouvait correspondre soit au modèle de Cox, soit au modèle de Poisson. Les modèles basés sur les données agrégées étaient utilisés à titre de référence du biais attendu.

## Etude de simulation

Dans l'étude de simulation, le modèle de Cox hiérarchique aurait pu être utilisé, mais les temps de calcul étant longs, seul le modèle de Poisson a été utilisé. Les interactions ayant été simulées par essai sans variation de l'âge moyen d'un essai à l'autre, la décomposition de l'interaction n'a pas été utilisée.

- Modèles basés sur les IPD :
  - **IPD-Poisson (âge qualitatif)** : modèle de Poisson hiérarchique avec interaction entre les effets traitement et l'âge en continu, effet aléatoire sur le risque de base et les effets traitement
  - **IPD Poisson (âge quantitatif)** : le même modèle de Poisson hiérarchique était utilisé, mais l'âge était introduit comme une variable qualitative à quatre classes
- Modèles basés sur les AD :



- **AD-Netmeta** : modèle basé sur la théorie des graphes ne permettant aucune interaction. L'âge n'était donc pas considéré
- **AD-Metareg** : metarégression introduisant une interaction entre l'âge moyen (en continu) dans l'essai et la (configuration I) ou les (configuration II) indicatrices. L'âge n'était pas considéré comme une variable qualitative.

## Etude sur données réelles

L'âge a uniquement été considéré comme une variable continue comme dans les publications originales. Si l'on avait souhaité prendre en compte l'âge avec une approche de type Netmeta, la catégorisation aurait alors été impérative. On aurait calculé un effet traitement ajusté par classe d'âge (première étape), puis on aurait combiné séparément les résultats obtenus dans chaque classe avec Netmeta (deuxième étape).

Les deux modèles supplémentaires suivants, basés sur les IPD, ont été considérés par rapport à l'étude de simulation :

- **IPD-Cox** : modèle de Cox hiérarchique avec interaction entre les effets traitement et l'âge en continu, effet aléatoire sur le risque de base et les effets traitement
- **IPD-Poisson2** : variante du modèle de IPD-Poisson1 décomposant l'interaction totale en interaction inter- et intra essais et recherchant ici un biais écologique

## Logiciels

Toutes les simulations ont été réalisées dans le logiciel R 3.5.2 (The R Core Team, Vienne, Autriche). Une fonction permettant de simuler des données de survie avec censure et permettant de faire varier le risque de base et les effets traitement a été écrite pour ce travail. Le package `furrr` 1.14.0 (Vaughan, 2018) a été utilisé de façon à paralléliser les calculs. Un générateur de nombre aléatoire de type L'Ecuyer-CMRG a été utilisé pour permettre à chaque instance du programme de partir d'une graine différente de façon reproductible. Les packages `tibble`, `tidyr` et `purrr` (Tidyverse, Wickham, 2019) étaient utilisés pour la manipulation des données, le package `ggplot2` (Wickham, 2019) pour les graphiques.

Le modèle de Cox multiniveau a été régressé à l'aide du package `Coxme` 2.2-16 (Therneau 2020), le modèle de Poisson multiniveau grâce à `lme4` 1.1.21 (Bates, 2020), le modèle basé sur les contrastes par `netmeta` 1.1.0 (Grücker, 2019) et la méta-régression par le package `metafor` 3.0.2 (Viechtbauer, 2019).

Le code des simulations est disponible sur le compte Github de l'unité.<sup>iv</sup>

---

<sup>iv</sup> [https://github.com/Oncostat/IPD\\_vs\\_AD\\_NMA](https://github.com/Oncostat/IPD_vs_AD_NMA)

## Chapitre 3. Résultats

### Résultats généraux

Aucun des modèles IPD en une étape n'a présenté de problème de convergence. En revanche, le modèle de méta-régression a présenté des problèmes de convergence dans 47 (0,2%) simulations de la configuration I et dans 3 (0,01%) de la configuration II. Les deux effets traitement directs (A-C et B-C) étaient estimés sans biais (Tableau annexe 1). De la même façon, l'effet âge, l'interaction et les deux interactions avec l'effet traitement étaient estimés sans biais.

### Configuration I : estimation de A-B indirecte uniquement

Dans cette première configuration, l'estimation de A-B reposait uniquement sur les comparaisons indirectes puisqu'il n'existait pas d'essai A-B.

#### Scenario 1 : rien

L'effet conditionnel A-B ( $ET_{\overline{age}}$ ) était estimé avec un biais très faible que l'âge soit considéré comme une variable continue ou catégorielle dans les trois modèles (AD-Netmeta, AD-Metareg et IPD-Poisson). A la différence des autres, le modèle AD-Metareg montrait une variabilité importante de son estimation au cours des réplifications avec une ESE pour  $ET_{\overline{age}}$  de 0,159.

#### Scenario 2 : interaction

Quand l'âge était utilisé comme une variable continue, l'effet conditionnel  $ET_{\overline{age}}$  était estimé sans biais par les trois modèles (AD-Netmeta : 0,006 ; Ad-Metareg : 0,005 ; IPD-Poisson 0,004). Les deux modèles qui permettent une interaction (IPD-Poisson et AD-Metareg) estimaient également sans biais le  $\nu TE_{age+1}$  (Biais = 0,000 et 0,004 respectivement). Toutefois, le modèle AD-Metareg avait une variabilité d'estimation beaucoup plus importante que celle du modèle IPD-Poisson (ESE=0,170, ASE=0,032, vs ASE=ESE=0,010). Le dernier modèle (AD-Netmeta), qui ne permet pas d'interaction, estimait le même effet traitement pour tous les âges et était, de fait, biaisé.

Quand l'âge était utilisé comme une variable catégorielle, ne pas prendre en compte l'interaction (AD-Netmeta) conduisait à une sous-estimation de l'effet traitement pour les classes d'âge les plus jeunes (biais = 0,160 pour les âges < 55 ans) et une surestimation pour les classes les plus âgées (biais = -0,148 pour la classe > 65 ans). Le modèle IPD-Poisson estimait quant à lui toutes les classes avec un biais < 0,01.

#### Scenario 3 : interaction + distribution

Quand l'âge était utilisé comme une variable continue, seul le modèle IPD-Poisson estimait correctement l'effet conditionnel  $ET_{\overline{age}}$  (biais = -0,003) et sa variation avec l'âge  $\nu TE_{age+1}$  (biais = 0,000). Le modèle AD-Netmeta obtenait des estimations biaisées à la

fois pour  $ET_{\overline{age}}$  (biais = -0,120) et pour  $\nu TE_{age+1}$  (biais = -0,016). Bien que le modèle AD-Metareg, ait présenté, *en moyenne* un biais faible pour l' $ET_{\overline{age}}$  (biais = 0,041) et le  $\nu TE_{age+1}$  (biais = -0,002), ses estimations étaient très instables d'une simulation à l'autre, aussi bien pour  $ET_{\overline{age}}$  (ESE=1,381, ASE = 2,121) que pour  $\nu TE_{age+1}$  (ESE=0,179 ASE=0,033).

Quand l'âge était utilisé comme une variable catégorielle, le modèle IPD-Poisson avait des estimations non biaisées (biais maximum = 0,034 pour la classe > 65 ans). Comme dans cette simulation les patients inclus dans les essais A-C étaient plus jeunes, le modèle AD-Netmeta avait un biais faible dans la classe < 55 ans (biais = -0,034) mais un biais important dans la classe > 65 ans (biais = -0,275). Ce modèle présentait une variabilité plus faible que le modèle IPD-Poisson sur l'ESE (0,084 contre 0,259 à 0,264) et sur l'ASE (0,084 contre 0,259 à 0,264).

## Configuration II : estimation mixte directe-indirecte de A-B

Dans cette seconde configuration, quelques essais A-B étaient simulés. La comparaison faisait donc appel à une combinaison entre les informations directes et indirectes.

### Scenario 1 : rien

Ajouter de l'information directe ne modifiait pas de façon substantielle les faibles biais déjà observés dans la configuration I. La variabilité des résultats (ASE, ESE) était toutefois plus faible.

### Scenario 2 : interaction

Ajouter de l'information directe ne modifiait pas de façon substantielle les résultats observés dans la configuration I. La variabilité des résultats (ASE, ESE) était toutefois plus faible.

### Scenario 3 : combiné

Quand l'âge était utilisé comme une variable continue, ajouter un peu d'information directe permettait de réduire, sans toutefois l'annuler, le biais pour l'effet conditionnel  $ET_{\overline{age}}$  qui était observé avec le modèle AD-Netmeta dans la configuration I (biais = -0,075 contre -0,12). De même, le modèle AD-Metareg avait une variabilité qui était plus faible, mais toujours présente ( $ET_{\overline{age}}$  : ESE=0,110 et ASE=1,276;  $\nu TE_{age+1}$  : ESE=0,135 et ASE=0,000). Les résultats du modèle IPD-Poisson étaient globalement inchangés.

Quand l'âge était utilisé comme une variable catégorielle, le modèle AD-Netmeta présentait un biais plus faible que dans la configuration I, mais toujours important (biais = 0,229 pour la catégorie > 65 ans versus -0,275). Les résultats du modèle IPD-Poisson étaient globalement inchangés.

**Tableau 8 :** Résultats de l'estimation de l'effet traitement dans la comparaison A-B dans les deux configurations et les trois scénarios, avec l'âge comme une variable catégorielle, un  $\log(\text{HR})$  simulé à -0.5,  $\sigma = 0.01$  et  $\tau = 0.01$

Scénario	Âge	Vrai log (HR)	Moy.	AD-Netmeta			Moy.	IPD-Poisson		
				Biais	ESE	ASE		Biais	ESE	ASE
<b>Configuration I : comparaison directe uniquement</b>										
<b>1 : Rien</b>	<55	0,000	0,000	0,000	0,076	0,006	0,004	0,004	0,143	0,020
(pas d'interaction, âges distribués identiquement)	55-60	0,000	0,000	0,000	0,076	0,006	-0,008	-0,008	0,154	0,024
	60-65	0,000	0,000	0,000	0,076	0,006	0,000	0,000	0,154	0,024
	>65	0,000	0,000	0,000	0,076	0,006	0,001	0,001	0,143	0,020
<b>2 : Interaction</b>	<55	-0,154	0,006	0,160	0,081	0,032	-0,147	0,008	0,158	0,025
(interaction, âge distribués identiquement)	55-60	-0,038	0,006	0,044	0,081	0,008	-0,037	0,000	0,162	0,026
	60-65	0,038	0,006	-0,032	0,081	0,008	0,044	0,006	0,159	0,025
	>65	0,154	0,006	-0,148	0,081	0,028	0,156	0,002	0,155	0,024
<b>3 : Combiné</b>	<55	-0,154	-0,120	0,034	0,084	0,008	-0,186	-0,031	0,259	0,068
(interaction, âge distribués différemment)	55-60	-0,038	-0,120	-0,083	0,084	0,014	-0,044	-0,006	0,214	0,046
	60-65	0,038	-0,120	-0,158	0,084	0,032	0,029	-0,009	0,213	0,045
	>65	0,154	-0,120	-0,275	0,084	0,082	0,121	-0,034	0,264	0,071
<b>Configuration II : comparaison mixte (directe / indirecte)</b>										
<b>1 : Rien</b>	<55	0,000	-0,003	-0,003	0,057	0,003	-0,005	-0,005	0,112	0,013
(pas d'interaction, âges distribués identiquement)	55-60	0,000	-0,003	-0,003	0,057	0,003	-0,003	-0,003	0,125	0,016
	60-65	0,000	-0,003	-0,003	0,057	0,003	-0,002	-0,002	0,115	0,013
	>65	0,000	-0,003	-0,003	0,057	0,003	-0,002	-0,002	0,105	0,011
<b>2 : Interaction</b>	<55	-0,154	0,003	0,157	0,064	0,029	-0,153	0,002	0,129	0,017
(interaction, âge distribués identiquement)	55-60	-0,038	0,003	0,040	0,064	0,006	-0,033	0,005	0,135	0,018
	60-65	0,038	0,003	-0,035	0,064	0,005	0,036	-0,002	0,131	0,017
	>65	0,154	0,003	-0,152	0,064	0,027	0,157	0,003	0,113	0,013
<b>3 : Combiné</b>	<55	-0,154	-0,075	0,079	0,065	0,011	-0,164	-0,010	0,156	0,024
(interaction, âge distribués différemment)	55-60	-0,038	-0,075	-0,037	0,065	0,006	-0,049	-0,011	0,153	0,024
	60-65	0,038	-0,075	-0,113	0,065	0,017	0,033	-0,004	0,152	0,023
	>65	0,154	-0,075	-0,229	0,065	0,057	0,144	-0,010	0,160	0,026

HR : Hazard Ratio, moy. : moyenne, ESE : Empirical Standard Error, ASE : Average Standard Error

Seconde partie : Evaluation d'un modèle IPD en une étape

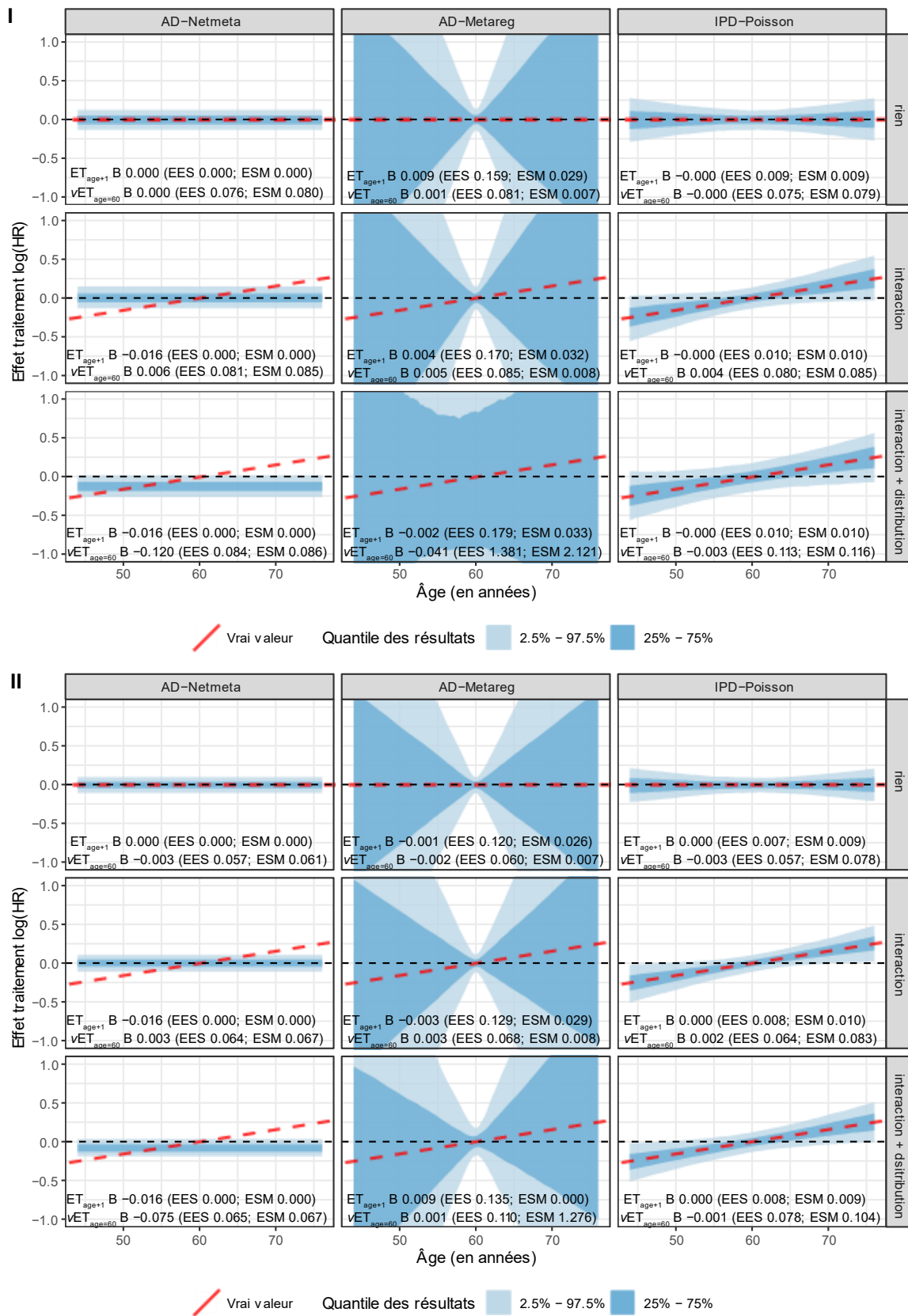


Figure 12: Résultat de la simulation dans les deux configurations et les trois scénarios quand l'âge était utilisé comme une variable quantitative,  $\log(HR) = -0.5$ ,  $\sigma = \tau = 0.01$

## Application sur données réelles

La méta-analyse MARCH/MACH-NC avait une géométrie identique à notre configuration I puisqu'il n'y avait pas de boucle fermée en l'absence d'essai comparant la mRT à la CTRT. De plus, les âges moyens dans les deux comparaisons étaient différents (59,9 +/- 10,1 ans vs 56,7 +/- 9,84 ans), ce qui représente 0,3 écart-type. Une interaction avec l'âge avait été observée dans la comparaison mRT contre RT. Cette configuration est donc proche de notre scénario 3, mais avec une différence d'âge moins importante.

Les deux effets directs (mRT contre RT et CTRT contre RT) étaient estimés à des valeurs proches par tous les modèles avec les mêmes écarts-types. Les trois modèles IPD (IPD-Poisson1, IPD-Poisson2 et IPD-Cox) estimaient une interaction significative (au seuil 5%) entre l'effet mRT contre RT et l'âge, et non significative entre CTRT contre RT et l'âge. La méta-régression estimait ces interactions à une plus grande valeur avec une incertitude plus forte. Ceci conduisant à des résultats discordants avec une interaction avec l'âge non significative pour mRT contre RT, mais significative pour CTRT contre RT (**Tableau 9, Figure 13**).

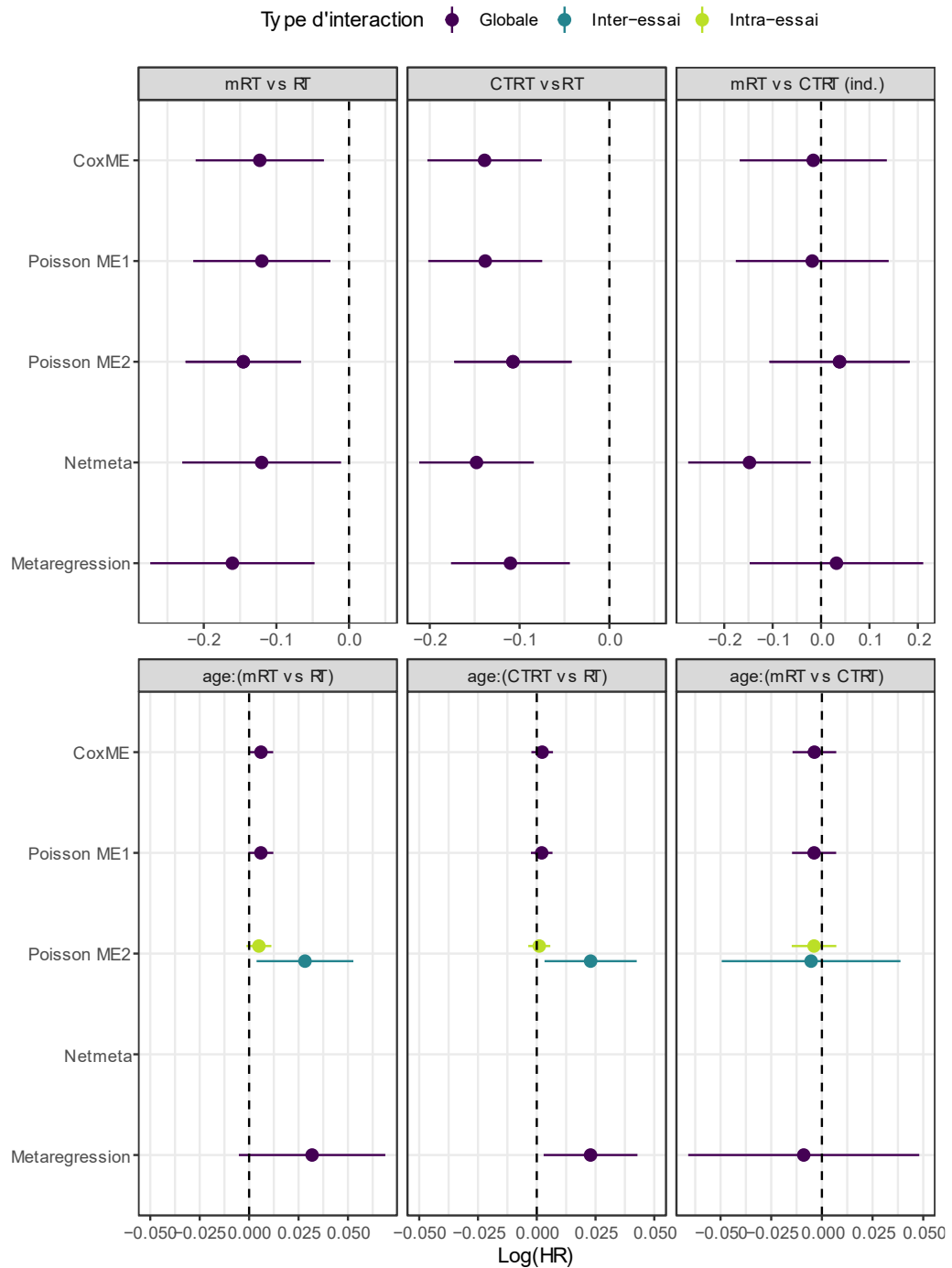
Le modèle de Poisson avec décomposition de l'interaction (IPD-Poisson2) explique en partie cette discordance. Ce modèle estimait : (i) une interaction *intra-essai* proche des modèles IPD, (ii) une interaction *inter-essai* proche de celle de la méta-régression. Ce résultat indique donc un *biais écologique* dans l'estimation des termes d'interactions par le modèle basé sur les AD. Bien que plus simple, la méta-régression n'était donc pas aussi performante que les modèles IPD pour estimer la structure sous-jacente des données.

Pour la comparaison indirecte (mRT contre CTRT), aucun effet significatif n'était observé dans les trois modèles IPD et dans la méta-régression puisque tous les intervalles de confiance des  $\log(\text{HR})$  comprenaient la valeur zéro. Toutefois, le modèle AD-netmeta concluait en une différence significative en faveur de la mRT. Les écarts types de cette comparaison étaient similaires dans tous les modèles.

*Tableau 9 : Estimation des paramètres par les différents modèles appliqués sur les données réelles de la méta-analyse en réseau MARCH/MACH-NC*

Paramètre	Modèles	Log(HR)	Ecart-type	Intervalle de confiance à 95%
<b>mRT vs RT (comparaison directe)</b>	CoxME	-0,123	0,045	-0,211 ; -0,035
	Poisson ME1	-0,120	0,048	-0,215 ; -0,026
	Poisson ME2	-0,146	0,041	-0,225 ; -0,066
	Netmeta	-0,120	0,056	-0,225 ; -0,066
	Meta régression	-0,161	0,058	-0,274 ; -0,047
<b>CTRT vs RT (comparaison directe)</b>	CoxME	-0,139	0,032	-0,203 ; -0,075
	Poisson ME1	-0,138	0,032	-0,202 ; -0,075
	Poisson ME2	-0,107	0,033	-0,173 ; -0,042
	Netmeta	-0,148	0,033	-0,212 ; -0,084
	Meta régression	-0,110	0,034	-0,176 ; -0,044
<b>Interaction entre : -Age (continu) - (mRT vs RT)</b>	CoxME	0,006	0,003	0,000 ; 0,012
	Poisson ME1	0,006	0,003	0,000 ; 0,012
	Poisson ME2b	0,028	0,012	0,004 ; 0,053
	Poisson ME2w	0,005	0,003	-0,001 ; 0,011
	Netmeta	NA	NA	
	Meta régression	0,032	0,019	-0,005 ; 0,069
<b>Interaction entre : -Age (continu) - (CTRT vs RT)</b>	CoxME	0,002	0,002	-0,002 ; 0,007
	Poisson ME1	0,002	0,002	-0,002 ; 0,007
	Poisson ME2b	0,023	0,010	0,003 ; 0,043
	Poisson ME2w	0,001	0,002	-0,004 ; 0,006
	Netmeta	NA	NA	
	Meta régression	0,023	0,010	0,003 ; 0,043
<b>mRT vs CTRT (comparaison indirecte)</b>	CoxME	-0,016	0,078	-0,168 ; 0,136
	Poisson ME1	-0,018	0,081	-0,176 ; 0,140
	Poisson ME2	0,038	0,074	-0,107 ; 0,183
	Netmeta	-0,148	0,065	-0,275 ; -0,021
	Meta régression	0,032	0,091	-0,147 ; 0,211
<b>Interaction entre : -Age (continu) - mRT vs CTRT (indirecte)</b>	CoxME	-0,004	0,006	-0,014 0,007
	Poisson ME1	-0,004	0,006	-0,015 0,007
	Poisson ME2b	-0,005	0,022	-0,049 0,039
	Poisson ME2bw	-0,004	0,006	-0,015 0,007
	Netmeta	NA		
	Meta régression	-0,009	0,029	-0,066 0,048

Dans ce réseau ouvert, les effets mRT contre RT et CTRT contre RT reposent sur des comparaisons directes alors que mRT contre CTRT repose sur des comparaisons indirectes. RT : Radiothérapie, mRT : Radiothérapie Modifiée, CTRT : chimioradiothérapie, CoxME : modèle mixte de Cox, Poisson ME1 : modèle mixte de Poisson classique, Poisson ME2 : modèle mixte de Poisson décomposant l'interaction (b : inter-essai ; w : intra-essai), Méta-régression : modèle basé sur les données agrégées utilisant les contrastes et ajusté sur l'âge.



**Figure 13 :** Estimation des  $\log(HR)$  des deux effets traitement directs, de l'effet traitement indirect (ligne supérieure) et des interactions (ligne inférieure) sur les données réelles de MARCH/MACH-NC

Dans ce réseau ouvert, les effets mRT contre RT et CTRT contre RT reposaient sur des comparaisons directes alors que mRT contre CTRT reposait sur des comparaisons indirectes. RT : Radiothérapie, mRT : Radiothérapie Modifiée, CTRT : chimioradiothérapie, CoxME : modèle mixte de Cox, Poisson ME1 : modèle mixte de Poisson classique, Poisson ME2 : modèle mixte de Poisson décomposant l'interaction (b : inter-essai ; w : intra-essai), Méta régression : modèle basé sur les données agrégées utilisant les contrastes et ajusté sur l'âge.





## Chapitre 4. Discussion et perspectives

Les NMA sont particulièrement intéressantes quand il n'y a pas (configuration I) ou peu (configuration II) de données disponibles pour une comparaison spécifique d'intérêt. Toutefois, cette estimation, principalement indirecte, dépend alors beaucoup de la qualité de l'estimation de l'effet traitement dans les autres comparaisons. Ces NMA sont le plus souvent réalisées, soit par une approche basée sur les données agrégées, soit en deux étapes, ou encore sous paradigme Bayésien. Dans cette étude, nous avons proposé un modèle différent, basé sur les IPD en une étape faisant appel à des modèles hiérarchiques de Poisson ou de Cox. Dans un réseau simulé à trois nœuds sans (configuration I) ou avec (configuration II) boucle fermée, ce modèle avait de bonnes performances dans tous les scénarios. Notamment, même dans la situation où il existait une interaction avec un modificateur distribué différemment selon les comparaisons, le modèle estimait de façon correcte l'effet traitement conditionnel à l'âge moyen et sa variation avec l'âge. Au cours des 1 000 réplifications de chaque configuration, ce modèle montrait également une grande stabilité avec une erreur empirique standard qui restait faible.

Dans cette étude, les modèles basés sur les AD servaient de référence de l'amplitude du biais dans les approches classiques. Comme on pouvait s'y attendre, ces modèles donnaient des résultats identiques à ceux basés sur les IPD quand il n'y avait pas d'interaction. Ceci n'était plus le cas en présence d'interaction. Le type d'erreur commise dépendait alors de si l'on s'intéresse à évaluer uniquement l'effet traitement moyen (i.e. à l'âge moyen) ou l'effet traitement à tous les âges. Si l'on s'intéresse à l'effet traitement à tous les âges, les modèles qui ne permettent pas de prendre en compte l'interaction (Netmeta) sont logiquement biaisés dès que celle-ci est présente (Scénarios 2 et 3). Le modèle AD permettant l'introduction d'une interaction (Metareg) était beaucoup moins biaisé. Toutefois, si en moyenne l'estimation était correcte, ce modèle montrait un ASE et un ESE très importants au cours des 1 000 réplifications, suggérant que pour une situation unique on pourrait obtenir des estimateurs très éloignés de la réalité. En revanche, si l'on s'intéresse uniquement à l'estimation de l'effet « moyen », les modèles basés sur les AD n'étaient pas systématiquement biaisés en présence d'interaction. La présence d'un biais dépendait alors de la distribution dans le réseau du modificateur. Dans la situation où le modificateur était distribué de façon équivalente (Scénario 2), l'effet traitement conditionnel à l'âge moyen n'était pas biaisé. En revanche, quand le modificateur était distribué de façon différente entre les comparaisons (Scénario 3), les modèles basés sur les AD étaient alors tous biaisés.

Des résultats similaires ont été observés sur nos données réelles. Les deux modèles IPD n'estimaient pas de différence entre les deux traitements qui étaient comparés uniquement via les données indirectes. En revanche, un chercheur qui se baserait uniquement sur les AD en ignorant l'âge des patients aurait conclu que la mRT était supérieure à la chimioradiothérapie au seuil 5%. Dans ce jeu de données disposant d'une grande puissance statistique, la méta-régression estimait correctement ce contraste avec une différence non significative. Toutefois, ce jeu de données illustre également une faiblesse de la méta-régression dont ne souffre pas l'approche IPD : le biais écologique. En effet, la méta-régression utilise l'âge moyen des patients dans les essais (qui est lui

aussi agrégé), et non l'âge individuel des patients. Si les essais qui ont estimé le plus grand effet ont également inclus des patients plus jeunes (ou plus âgés), on peut surestimer l'importance de l'interaction. La décomposition de l'interaction en interaction intra- et inter-essais dans le modèle IPD tel que proposé par Hua(25) montrait ici que l'interaction intra-essai (celle estimée par les modèles IPD) était en réalité bien plus faible que l'interaction inter-essai (celle estimée par la méta-régression).

Un autre résultat intéressant de notre étude était l'illustration de la possibilité d'utiliser le modèle de Poisson comme alternative au modèle de Cox en paradigme fréquentiste dans des données hiérarchiques. Ces deux modèles ont été utilisés avec succès sur ce jeu de données comprenant 79 essais et 18 000 patients, sans problème de convergence. Les temps de calcul étaient raisonnables, prenant, sur un ordinateur standard de bureau, 100 secondes pour le modèle de Poisson et 400 secondes pour le modèle de Cox. Dans notre expérience, cette différence se creuse progressivement quand la complexité du modèle augmente. Dans une application clinique où l'on n'a qu'un (ou peu de modèles) à régressé, de tel tant de calcul sont tout à fait raisonnables, surtout mis en perspective avec ceux requis dans les approches Bayésiennes. Toutefois, dans le contexte de simulations répétées, le modèle de Cox présentait un temps de calcul rédhibitoire. Notre simulation basée sur le modèle de Poisson a pris plusieurs jours de calcul malgré la programmation des calculs en parallèle et l'utilisation d'un serveur de calcul dédié de 80 cœurs. Une telle simulation avec le modèle de Cox aurait pris plus d'une semaine. Pour cette raison, nous nous sommes limités au modèle de Poisson. D'autres équipes ont fait le choix d'utiliser un autre modèle comme alternative au modèle de Cox, comme le modèle de Royston-Parmar. (19)

Estimer de façon correcte les paramètres d'une MAR peut être difficile. Pour les AD-NMA, des implémentations logicielles sont disponibles à la fois sous paradigme fréquentiste et Bayésien et délivre des résultats similaires. Pour les IPD-NMA, la plupart des équipes utilisent soit une approche en deux étapes sous paradigme fréquentiste, soit une approche Bayésienne. L'approche en deux étapes utilise les données individuelles pour calculer des effets traitement ajustés par essais avant de les combiner. L'approche Bayésienne en une étape permet l'estimation de modèles multiniveaux complexes, parfois au prix de temps de calcul très long. En pratique, très peu d'études utilisant un modèle fréquentiste IPD en une étape sous paradigme fréquentiste pour données de survie ont été publiées. En effet, spécifier correctement le modèle peut devenir compliqué au fur et à mesure que la complexité augmente et celui-ci peut ne jamais converger. Nous avons ici fait le choix (pour des raisons de temps de calcul) de remplacer le modèle de Cox par un modèle linéaire généralisé multiniveau de Poisson. En effet, les modèles de type GLMM, largement plus ubiquitaires que les modèles de Cox, bénéficient d'implémentations logicielles beaucoup plus performantes. Récemment, Ollier et al ont proposé d'étendre le modèle de Poisson par une pénalisation de type LASSO. Celle-ci est utilisable dans des NMA plus complexes comprenant des effets principaux, des termes d'incohérence, des interactions traitements\*covariables et des effets non proportionnels.(35) Une autre approche, proposée par Jackson et al pour les modèles binomiaux, pourrait être utilisée dans ce contexte.(36) Dans cette approche, le modèle complet est progressivement simplifié en retirant un à un les effets jusqu'à ce que le modèle converge.

Un autre problème potentiel créé par les interactions dans une NMA est l'incohérence. L'utilisation de modèles basés uniquement sur les AD peut donner des problèmes d'incohérence si des covariables sont distribuées de façon différente dans les comparaisons. Dans notre modèle, tous les effets traitement étaient simulés sous l'hypothèse de cohérence. Toutefois, les modèles IPD peuvent être reformulés par l'ajout d'un nouveau terme, afin de détecter une éventuelle incohérence. (37,38) Si de l'incohérence est détectée, ajuster sur un facteur de confusion ou introduire une interaction pourrait corriger en partie ce problème. (31,39,40) Cette possibilité n'est en général pas possible si on ne dispose que des AD. Un travail de Donegan(41) a toutefois montré que ce problème était potentiellement plus subtil, car il implique à la fois l'intercept et la pente de la covariable en rapport avec l'effet traitement. Les modèles IPD en une étape permettent, par leurs intercept et pentes aléatoires (nichés dans la comparaison) de prendre en compte ces configurations.

Notre étude avait plusieurs forces qui doivent être mises en avant. Les simulations ont été réalisées selon 48 combinaisons différentes représentant un large panel de situations possibles. Deux codages différents pour l'âge ont été utilisés (continu et catégoriel) afin d'évaluer plusieurs représentations d'une variable continue. Plusieurs modèles ont été utilisés dont deux modèles IPD. Enfin, l'application sur données réelles a montré que nos résultats n'étaient pas que théoriques, mais pouvaient entraîner des conséquences sur l'interprétation d'études dans le monde réel.

Nous devons également reconnaître plusieurs limites à cette étude. La plus importante est que nous avons limité notre travail à une situation simple pas forcément courante en pratique clinique. Plusieurs paramètres pourraient faire varier les performances de notre modèle et nécessiteront des travaux futurs. (i) La taille du réseau : notre modèle n'avait que trois nœuds or plus le nombre de nœuds augmente, plus l'estimation des paramètres devient difficile. Si en oncologie le nombre d'options pour une situation est souvent limitée, d'autres secteurs peuvent avoir des réseaux très larges. (ii) Le nombre de patients et d'études par comparaison : chaque comparaison principale de notre étude comprenait 2 000 patients et 10 études. La stabilité des estimateurs dans une situation où il y aurait peu de patients pourrait être moins bonne. De même, d'éventuels problèmes de convergence pour des nombres très élevés ou très faibles de patients pourrait survenir. Nous avons toutefois pu estimer sans difficulté le réseau de MARCH/MACH-NC qui comprenait 79 essais et 18 000 patients. (iii) La géométrie du réseau (42): notre réseau avait une géométrie simple et bien équilibrée. Ces performances pourraient être moins bonnes dans un réseau très déséquilibré, ce qui est une situation relativement fréquente en présence d'une comparaison ou d'un traitement « de référence ». D'autre part, la répartition entre la quantité d'informations directes et indirectes joue directement sur la magnitude du biais. (iv) La complexité des interactions : dans notre réseau il n'y avait qu'un modificateur. Dans un réseau plus grand, plusieurs interactions et plusieurs modificateurs peuvent être présents. (v) La cohérence : notre modèle était cohérent par design, ce qui permettait de se concentrer sur l'interaction. Un certain degré d'incohérence est probable dans la plupart des situations cliniques habituelles. Cette incohérence pourrait perturber l'estimation de l'interaction. Même si ce travail ne permet pas de répondre aux performances du modèle dans ces situations plus complexes, il s'agit du premier travail sur le sujet. Celui-ci pourra servir de base à de futurs travaux explorant

ces situations. Une autre limite, inhérente à toute simulation, est la nécessité de réaliser des choix arbitraires. Ces choix portent par exemple sur le taux de censure, la magnitude de l'effet traitement, l'importance de l'interaction ou du déséquilibre des modificateurs. Nous avons fait des choix représentatifs des situations que nous voyons habituellement en oncologie. Des chercheurs dans d'autres domaines d'activité auraient probablement fait d'autres choix et auraient pu aboutir à des résultats légèrement différents.

## Chapitre 5. Conclusions

Ce travail a évalué un modèle de type IPD en une étape afin d'estimer l'effet d'un traitement sur un critère de survie sous paradigme fréquentiste. Ce modèle gardait un biais faible même dans des situations où de l'interaction était présente avec une distribution déséquilibrée des modificateurs. Les estimations fournies étaient par ailleurs stables au cours des réplifications. Dans la situation d'un modèle simple, cette modélisation peut donc être utilisée comme alternative aux approches Bayésiennes.

Le modèle mixte de Poisson donnait des estimations très proches de celles du modèle de Cox et peut donc être utilisé, après des modifications minimales du jeu de données, en cas de difficulté de convergence ou quand la vitesse de calcul est importante (comme dans le cas de répétitions de simulations).

Bien que ceci ne soit pas une nouveauté, une approche basée uniquement sur les AD pouvait induire des conclusions erronées en présence d'interactions, notamment quand le modificateur était distribué de façon différente dans les comparaisons. Une approche par méta-régression pourrait limiter ce biais avec comme limitation une exposition au biais écologique et une instabilité sur des petits jeux de données.

Les modèles IPD devraient être considérés comme le gold standard d'une NMA. Si une approche par AD était initialement choisie, la constatation d'un déséquilibre de distribution d'un modificateur potentiel entre les comparaisons, ou la présence d'une incohérence non expliquée par les variables mesurées au niveau de l'essai, ou un faible nombre de participants, ou la suspicion d'un biais écologique devraient faire discuter de compléter par une approche IPD. Il en va de même si on cherche à prédire précisément le risque des patients comme dans la médecine personnalisée.



# Troisième partie :

## MANATEC-02





# Chapitre 1. Introduction

L'étude MANATEC-02 (Méta Analyse des traitements Néoadjuvants des cancers de l'œsophage ou de la jonction œsogastrique) est une méta-analyse en réseau basée sur les données individuelles. Elle a pour but de comparer les traitements néoadjuvants potentiels utilisés dans les cancers de l'œsophage localement avancés. Initiée par l'institut Gustave Roussy (Villejuif, France), elle est menée par un groupe collaboratif international multi-spécialités.

## MANATEC-01

En 2007, notre équipe avait déjà réalisé deux méta-analyses classiques séparées sur les traitements néoadjuvants de l'œsophage. Ces études avaient été rapportées sous forme d'abstract en congrès international, mais jamais publiées (ASCO 2007 : abstract 4512 et ASTRO 2008 : Abstract 158). Il s'agissait de deux méta-analyses sur données individuelles dont le critère de jugement principal était la survie.

La méta-analyse sur la chimiothérapie néoadjuvante (ASCO 2007 : abstract 4512) avait obtenu les IPD de neuf essais sur les 12 éligibles, ce qui représentait 2 102 patients. Cette étude démontrait la supériorité de la chimiothérapie néoadjuvante sur la chirurgie seule avec un HR = 0,87 [0,79-0,95]. Il existait une hétérogénéité importante ( $p = 0,03$ ,  $I^2=52\%$ ). Le bénéfice absolu de survie à 5 ans était de 4,3%. Un bénéfice en survie sans récurrence était également observé HR = 0,82 [0,74 ; 0,91].

La méta-analyse sur la chimioradiothérapie néoadjuvante (ASTRO 2008 : Abstract 158) avait obtenu les IPD de neuf essais sur les neuf identifiés, ce qui représentait 1 210 patients. L'étude démontrait un bénéfice en survie globale avec HR = 0,82 [0,72-0,93]. Il existait une hétérogénéité importante ( $p=0,04$ ,  $I^2=51\%$ ). Le bénéfice absolu à 5 ans était de 6,5%. Un bénéfice en survie sans récurrence était également observé HR = 0,85 [0,72 ; 0,99].

De nombreuses autres méta-analyses, standards ou en réseaux, mais reposant sur données agrégées ont également été publiées depuis. Toutefois aucune d'entre elles ne disposait des IPD, ne permettant pas de réaliser d'analyses en sous-groupes et de répondre à la principale question clinique actuelle : le meilleur traitement est-il différent selon le type histologique et/ou la localisation de la tumeur ?

## But de l'étude

Depuis nos premiers résultats, la chirurgie seule n'est plus considérée comme un standard et le traitement néoadjuvant est systématique pour les tumeurs localement avancées. Toutefois peu d'essais ont comparé les deux traitements principaux de façon directe. Plusieurs essais sont en cours partout dans le monde (**Tableau 10**), mais leurs résultats ne seront pas disponibles avant plusieurs années. De ce fait la meilleure stratégie thérapeutique globale est inconnue et on ne sait pas si elle est la même dans les différents sous-groupes.

Parallèlement, depuis notre première étude, de nombreux nouveaux essais ont été rapportés et certains essais inclus dans nos méta-analyses précédentes ont été mis à jour. De plus, l'essor des méta-analyses en réseau rend possible d'estimer la comparaison CS contre CRS en combinant les informations apportées par les quelques comparaisons directes avec les nombreuses comparaisons indirectes.

MANATEC-02 avait donc trois buts : (i) mettre à jours les deux méta-analyses S contre CS et S contre CRS, (ii) estimer la comparaison CS contre CRS dans une méta-analyse deux à deux, (iii) effectuer une méta-analyse en réseau combinant toute l'information disponible. Au cours de toutes ces méta-analyses, des analyses en sous-groupes pré-spécifiées notamment sur le type histologique (CE/ADK) et la localisation anatomique (OT/JOG) ont été effectuées. Le protocole est disponible sur le site de Gustave Roussy.<sup>5</sup>

L'étude avait été enregistrée sur le registre PROSPERO sous le numéro : CRD42018107158

---

<sup>5</sup> <https://www.gustaveroussy.fr/sites/default/files/meta-analysis-protocole-ma-f.pdf>

Tableau 10: Essais thérapeutiques en cours

Essai	Pays	Localisation Histologie	Bras de traitement	Critère principal	N	Statu
<b>Re-Evaluation (NCT02442440)</b>	Chine	OT CE	<b>A</b> : Chirurgie seule <b>B</b> : 3 cycles Cisplatine-Paclitaxel puis chirurgie	OS 5 ans	528	Recrute depuis juin 2015 2.5 ans de recrutement 5 ans de suivi
<b>TOPGEAR (ACTRN12609000035224) (5)</b>	International	JOG ADK	<b>A</b> : ECF Périopératoire (3+3) <b>B</b> : ECF Périopératoire (3+3plus chimioradiothérapie (45 Gy / 25 fr)	OS 5 ans	720	318 patients à recruter à partir de novembre 2016
<b>CMISG1701 (NCT03001596)(6)</b>	Chine	OT CE	<b>A</b> : 2 cycles de Cisplatine / Paclitaxel <b>B</b> : 4 semaines de chimioradiothérapie (40 Gy/ 20 fr) Cisplatine / Paclitaxel	OS 3 ans	364	40 patients inclus en mars 2017
<b>ESOPEC (NCT02509286)(7)</b>	Allemagne	OT/JOG ADK	<b>A</b> : Chimiothérapie péri-opératoire (protocole FLOT : 5-FU/leucovorine/oxaliplatine/docetaxel) <b>B</b> : Chimioradiothérapie néoadjuvante (protocole CROSS : 41.4Gy plus carboplatine/paclitaxel) suivie de chirurgie	OS 3 ans	550	En cours de recrutement
<b>Neo-AEGIS (NCT01726452)(8)</b>	Europe	OT/JOG ADK	<b>A</b> : MAGIC péri-opératoire modifié (Epirubicine Cisplatine/Oxaliplatine 5-FU/Capecitabine) <b>B</b> : CROSS (Carboplatine and Paclitaxel avec radiothérapie concomitante, 41.4Gy/23 fr, sur 5 semaines) suivie de chirurgie	OS 3 ans	594	En cours de recrutement
<b>NExT (JCOG 1109) (UMIN000009482)</b>	Japon	OT CE	<b>A</b> : Deux cycles de Cisplatine / 5FU <b>B</b> : Trois cycles de Docetaxel /Cisplatine/5FU <b>C</b> : Chimioradiothérapie (41.4 Gy/23 fr) (Cisplatine / 5FU)	OS 5 ans	483	Recrute depuis 2013

OT : œsophage thoracique, JOG : jonction œsogastrique, CE : carcinome épidermoïde, ADK : adénocarcinome, fr : fractions, Gy : Grays, OS : Survie globale



## Chapitre 2. Méthodes

### Advisory board et collaboration internationale

#### Advisory board

Un « advisory board » a été constitué regroupant différentes spécialités et expertises : statistiques, méta-analyse, oncologie médicale, radiothérapie et chirurgie. Ces spécialistes ont été choisis partout dans le monde, notamment parce que leur équipe avait participé aux essais randomisés sur le sujet. Ils venaient principalement d'Europe, mais aussi d'Australie ou de Chine.

#### Contact avec les équipes

Les investigateurs des essais présents dans la méta-analyse précédente ont été contactés afin de leur proposer de participer à la mise à jour, en fournissant des données de survie actualisées ou de données de toxicité qui n'avaient pas été utilisées avant. Les investigateurs principaux des nouveaux essais ont été contactés par courriel. En cas de non-réponse, une seconde proposition de collaboration était ou a été envoyée. En cas de non-réponse, tous les autres investigateurs de l'essai étaient contactés. Enfin, en l'absence de réponse, les membres de « l'advisory board » étaient sollicités à la recherche de contacts pour joindre les investigateurs.

### Essais incluables et stratégie d'identification

#### Critères d'inclusion/d'exclusion des essais

Les essais publiés ou non étaient incluables sans restriction de langage.

Les essais devaient :

- Être randomisés d'une façon ne permettant pas de connaître à l'avance le bras de traitement
- Être clos aux inclusions avant le 31 décembre 2015 inclus
- Avoir randomisé des patients :
  - Avec un carcinome de l'œsophage ou de la JOG (quel que soit l'histologie)
  - Avec une maladie localement avancée mais sans métastases à distance (définie par un Stade II ou III)
  - En première ligne thérapeutique
- Avoir comparé des stratégies thérapeutiques de :
  - Chirurgie d'emblée contre chimiothérapie néoadjuvante suivie de chirurgie
  - Chirurgie d'emblée contre chimio radiothérapie néoadjuvante suivie de chirurgie
  - Chimiothérapie néoadjuvante contre chimio radiothérapie néoadjuvante toutes les deux suivies de chirurgie

Les critères d'exclusions étaient :

- Nouveaux essais (non inclus dans la méta-analyse précédente) ayant inclus moins de 60 patients (30 patients par bras)
- Essais ayant comparé différents protocoles de chimiothérapie ou de radiothérapie

## Stratégie d'identification des essais

Afin d'éviter un biais de publications, tous les essais publiés ou non étaient éligibles. Afin d'identifier le plus d'essais possibles, une recherche systématique de plusieurs sources différentes a été réalisée. Une veille bibliographique a été réalisée pendant le projet à la recherche de nouveaux articles parus au cours du projet. En cas d'incertitude sur le caractère éligible d'un essai, une discussion au sein des investigateurs était menée jusqu'à ce qu'un consensus soit atteint.

### *Bases de données électroniques*

Les essais ont été recherchés, sans restriction de langage, en utilisant la méthode optimale de la Cochrane Collaboration, ainsi que des mots-clés MeSH et du texte libre en rapport avec les cancers de l'œsophage et de la jonction œsogastrique ainsi que leurs traitements.

Les trois bases suivantes ont été interrogées en fin d'année 2017 :

- Pubmed
- Web of Science
- Scopus

Les équations de recherches sont données dans l'appendice page 157.

### *Registres d'essais*

Les registres d'essais suivants ont été interrogés afin d'identifier des essais non (ou non encore) publiés ou encore en phase de recrutements :

- ClinicalTrials.gov
- Cochrane Central Register of Controlled Trials (CENTRAL)

### *Résumés de conférences*

Les résumés de conférences suivantes ont été interrogées afin d'identifier des essais uniquement rapportés sous forme de résumé à ce stade :

- Proceedings of the American Society of Clinical Oncology (ASCO)
- Proceedings of the European and American Society of Therapeutic Radiation Oncology (ESTRO, ASTRO)
- Proceedings of the European Society of Medical Oncology (ESMO)
- Proceedings of the European Cancer Conference Organization (ECCO)

### *Recherches additionnelles*

Les bibliographies des essais inclus, des revues de la littérature et des méta-analyses ont été analysées à la recherche d'essais non encore identifiés. De même, tous les participants à la méta-analyse (investigateurs et membres du « board ») ont été interrogés sur l'existence d'autres essais.

## Analyse de la qualité des données (Checking)

Les données des essais disponibles étaient contrôlées et les essais réanalysés pour identifier des erreurs ou des incohérences avec les données publiées. Ce contrôle était

réalisé selon un protocole standardisé suivant les recommandations du groupe méthodologique de la Cochrane pour les méta-analyses sur données individuelles et les recommandations PRISMA-IPD. (43) En cas de problèmes identifiés, les équipes étaient contactées afin de résoudre ceux-ci.

## Données manquantes

Dans la partie MA deux à deux, en cas de données manquantes ponctuelles, les patients concernés étaient exclus de l'analyse utilisant ces données. Quand un essai n'avait pas rapporté un critère de jugement, ou l'avait rapporté pour moins de 20% des patients, celui-ci était exclu de l'analyse correspondante.

Dans la partie MAR, afin de pouvoir ajuster les modèles sur les critères pronostiques répartis de façon différente entre les comparaisons, une imputation multiple des données a été réalisée. Vingt jeux de données complets ont été imputés selon la méthode d'imputations multiples par équations chaînées (MICE). Chaque chaîne avait 20 itérations et la convergence était estimée par l'examen graphique de celles-ci. Pour toutes les variables, un modèle hiérarchique prenant en compte l'essai était utilisé. Les variables continues étaient imputées selon la méthode du « predictive mean matching » et les variables catégorielles par un modèle linéaire généralisé mixte binomial.

## Critères de jugement étudiés

### Survie globale

La survie globale était le critère de jugement principal. Elle était définie comme le temps entre la randomisation et le décès, quelle que soit sa cause. Les patients vivants aux dernières nouvelles étaient censurés à cette date.

### Survie sans récurrence

Le décès ou une récurrence, qu'elle soit locale ou à distance, constituait un événement de survie sans récurrence. Les autres patients étaient censurés à la date des dernières nouvelles. L'intervention chirurgicale ne survenait pas au même moment dans les différents bras d'études. Or, une intervention chirurgicale est aussi une exploration capable de diagnostiquer des lésions non visibles en radiologie. Une méthode de « landmark » a donc été utilisée. (44) Sans celle-ci, les événements de récurrences pourraient être diagnostiqués plus précocement dans le bras chirurgie d'emblée, uniquement à cause de la chronologie des traitements. Le landmark a été fixé à 6 mois pour tenir compte de la longueur des différentes stratégies thérapeutiques. Dans cette méthode, tous les événements survenant avant 6 mois sont pris en compte au même moment dans les deux groupes.(44)

### Risques compétitifs de récurrences

Dans cette étude, trois risques compétitifs étaient considérés : la récurrence locale ou régionale (désigné par la suite par récurrence locale), la récurrence à distance et le décès sans récurrence. En cas de récurrence locale et à distance simultanée, le patient était classé comme récurrence à distance. Le « landmark » de six mois était appliqué.



## Survie spécifique

Les décès étaient considérés comme liés ou non au cancer selon la méthode de Peto. (45) Dans cette méthode, un décès est considéré comme non lié au cancer si sa cause est connue et il survient chez un patient n'ayant jamais récidivé. Dans toutes les autres situations, notamment si la cause est inconnue ou si le patient a récidivé, le décès est considéré comme lié au cancer. Les essais n'ayant pas rapporté d'information de récurrence étaient exclus de cette analyse.

## Qualité de la chirurgie

La qualité de la résection chirurgicale était évaluée par le score R. Il s'agit d'un score mixte chirurgical / anatomopathologique en trois classes : R0 : pas de résidu, R1 : résidu microscopique (i.e. : tumeur à moins d'un mm de la marge chirurgicale en anatomopathologie), R2 ; résidu macroscopique (signalé par le chirurgien). L'évènement d'intérêt était une chirurgie R0 : un RR > 1 indiquait donc une plus grande probabilité d'être R0.

## Morbidité et mortalité postopératoires

La mortalité postopératoire était définie comme tout décès survenant dans les 30 jours après l'intervention. La morbidité postopératoire était considérée comme sévère quand elle nécessitait une ré-opération ou une hospitalisation en réanimation, ce qui correspond aux stades  $\geq 3$  de l'actuelle classification de Clavien-Dindo.(46). Pour cette analyse un RR>1 indiquait une augmentation du risque de morbi-mortalité.

## Toxicités

Les informations de toxicité étaient rarement disponibles. Elles ne pouvaient être comparatives que dans la comparaison CS contre CRS, la toxicité de ces traitements dans un bras chirurgie seule étant par définition nulle. Dans les comparaisons S-CS et S-CRS, elles ne sont rapportées que de façon descriptive. Les toxicités étaient définies comme sévères selon la classification CTCAE dont la version a pu varier au cours du temps.(47)

## Analyses en sous-groupes définis a priori

Plusieurs analyses en sous-groupes ont été définies a priori sous réserve que les données étaient suffisantes. Les deux principales étaient le type histologique, et la localisation anatomique. Le type histologique était défini en deux classes (ADK et CE). Les carcinomes indifférenciés non classés étaient considérés comme des valeurs manquantes pour les analyses impliquant le type histologique, mais conservés dans le reste des analyses. La localisation anatomique était considérée en deux classes, l'œsophage thoracique quelle que soit la localisation et la jonction œsogastrique (quel que soit le type selon Siewert). Dans les essais qui avaient inclus aussi des cancers de l'estomac, ces derniers étaient exclus de toutes les analyses.

Quatre autres analyses en sous-groupes au niveau patient étaient prévues par le protocole : le sexe, l'âge, le T du TNM et le N du TNM. L'âge était considéré à la fois comme une variable quantitative et comme une variable qualitative à quatre classes : <60, 60-64, 65-69 et 70 ans et plus. Le T et le N du TNM étaient dichotomisés : T1T2 contre T3T4 et N0 contre N+.

Le protocole prévoyait également trois analyses au niveau essai : (i) la dose prévue de radiothérapie (<40 Gy, ≥40 Gy), (ii) le type de chimiothérapie (cisplatine-5FU contre autres) et (iii) la similarité de la chimiothérapie dans les bras CS et CRS.

## Méthodes statistiques

Toutes les analyses étaient menées sous le principe de l'intention de traiter (i.e. tous les patients randomisés sont analysés dans leur bras de randomisation initial). Tous les tests étaient bilatéraux et un seuil de 0,05 pour la valeur du p a été retenu comme indicatif de la significativité statistique.

### Données censurées

Les probabilités de survie au cours du temps étaient calculées selon la méthode de Kaplan-Meier. Le bénéfice absolu au temps t était défini comme la différence entre les deux pourcentages à ce temps. La survie était également exprimée sous forme de *temps de survie moyen restreint* (Restricted Mean Survival Time : RMST) qui correspond au temps de survie moyen entre la randomisation et un horizon temporel donné t\* (RMST<sub>t\*</sub>). Elle était exprimée en mois. La différence de RMST (dRMST) entre deux groupes correspond au gain moyen (en mois) que l'on peut attendre grâce au traitement jusqu'à l'horizon t\*. Ceci correspond également à l'aire entre les deux courbes. Cette méthode a pour avantage de ne pas être soumise à la proportionnalité des risques.

Pour les données censurées soumises à un landmark, la méthode du test du log-rank partiellement groupé a été appliquée.(44) Dans cette méthode, les événements survenant avant le landmark sont tous regroupés, une statistique du  $\chi^2$  est calculée et utilisée dans le test du log-rank.

Le suivi médian était calculé selon la méthode du Kaplan-Meier inversé (i.e. méthode de Schemper).

### Critères binaires

Les critères binaires étaient décrits sous la forme N (%). La différence entre les deux groupes était présentée sous forme de rapport des risques (RR). En univariée, ce rapport était calculé directement par le rapport des probabilités d'événements dans les deux groupes. En cas de nombre nul d'événements dans un groupe, un effectif de 0,5 était ajouté à chaque case du tableau. En cas d'ajustement, ce rapport était calculé par un modèle linéaire généralisé avec une fonction de lien logarithmique et une distribution binomiale des résidus.

## Méta-analyse deux à deux

### Test de l'hétérogénéité

Un test du  $\chi^2$  d'hétérogénéité a été utilisé pour tester l'hétérogénéité de l'effet traitement entre les essais. Une mesure de l'index I<sup>2</sup>(48) était utilisée en combinaison avec celui-ci et une valeur inférieure à 25% était considérée comme témoignant d'une faible hétérogénéité. En cas d'hétérogénéité faible, un modèle à effet fixe était utilisé ; en cas d'hétérogénéité importante, un modèle à effet aléatoire

### *Variables censurées*

Le modèle à effet fixe utilisait un test du log-rank stratifié sur l'essai. Les différences entre les événements observés et attendus, ainsi que leurs variances étaient calculées. Ces valeurs étaient rapportées sur les forest-plot quel que soit le modèle final utilisé. Le modèle à effet fixe combinait les  $\log(\text{HR})$  en les pondérant par l'inverse de leurs variances. Le modèle à effet aléatoire était de type IPD en un temps. Il reposait sur un modèle mixte de Cox avec un risque de base spécifique à chaque essai (« intercept aléatoire ») et un effet traitement spécifique à chaque essai (« pente aléatoire »). (24)

La dRMST était calculée selon la méthode du Kaplan-Meier « poolé ». (49)

### *Critères binaires*

Le modèle à effet fixe combinait les  $\log(\text{RR})$  de chaque essai en les pondérant par l'inverse de leur variance. Le modèle à effet aléatoire partait des IPD avec un modèle linéaire généralisé mixte à fonction de lien logarithmique et distribution des résidus selon une loi binomiale. Ces modèles utilisaient un risque de base propre à chaque essai (« intercept aléatoire ») et un risque associé au traitement spécifique à chaque essai (« pente aléatoire »). Dans les cas rares où ce modèle ne convergait pas, dû le plus souvent à des probabilités estimées inférieures à zéro ou supérieures à un, un modèle basé sur le rapport des cotes (OR : odds ratio) était utilisé (fonction de lien logistique ; distribution binomiale des résidus).

### *Risques compétitifs*

Les risques compétitifs étaient évalués par un modèle stratifié de Fine et Gray. (50) Pour chacun des trois événements (récidive locale, récidive à distance et décès sans récidive), l'événement étudié était considéré comme l'événement principal et les autres comme des risques compétitifs. Les HR de sous-distribution (subHR) étaient estimés dans chaque essai par le paquet R « cmprsk » et le subHR global par le paquet « crrSC ». Des courbes d'incidences cumulées des événements étaient également tracées.

### *Mortalité spécifique*

L'effet du traitement sur la mortalité liée ou non liée au cancer était calculé selon la méthode de Peto. (50) Dans cette méthode de type actuarielle, stratifié sur l'essai, les taux de mortalité annuels et les HR étaient utilisés pour calculer la différence absolue.

### *Interaction et sous-groupes*

L'interaction était testée en introduisant dans le modèle un terme d'interaction covariable\*traitement en plus des effets principaux du traitement et de la covariable (modèle IPD à une étape).

## **Méta-analyse en réseau**

La méta-analyse en réseau était réalisée selon un modèle fréquentiste basé sur les IPD en une étape. Ce modèle utilisait un modèle de Cox à effets aléatoires comme décrit dans la première partie de la thèse.

### Codages des indicatrices de traitement

Le réseau n'avait que trois comparaisons de traitement possible : il devait donc être codé avec deux indicatrices de traitement. La troisième comparaison (le paramètre fonctionnel) étant déduit par contraste. Le codage suivant a été utilisé :

**Tableau 11 :** Codage des indicatrices de traitement dans MANATEC-02 considéré comme un réseau classique

Type d'essai	Bras de randomisation	Indicatrice de traitement 1	Indicatrice de traitement 2
<b>S contre CS</b>	CS	+0.5	0
	S	-0.5	0
<b>S contre CRS</b>	CRS	0	+0.5
	S	0	-0.5
<b>CS contre CRS</b>	CRS	-0.5	+0.5
	CS	+0.5	-0.5

S : chirurgie d'emblée, CS : chimiothérapie suivie de chirurgie, CRS : chimioradiothérapie suivie de chirurgie

Un tel codage implique donc que les choix arbitraires suivants ont été faits (le sens des comparaisons suit celui des flèches de la **Figure 45**) :

- Indicatrice 1 : CS contre S une valeur  $<1$  indiquant un plus grand effet de CS
- Indicatrice 2 : CRS contre S une valeur  $<1$  indiquant un plus grand effet de CRS
- Paramètre fonctionnel (combinaison des deux précédentes) sur la comparaison CRS contre CS une valeur  $<1$  indiquant une plus grande efficacité de CRS.

Le choix de coder la comparaison d'intérêt principal par un paramètre fonctionnel et non par une indicatrice était ici dictée par la présence de beaucoup plus d'informations dans les deux autres comparaisons.

Les traitements étaient classés par le P-Score.(51)

Alternativement, ce réseau pouvait également être vu comme un exemple de modèle additif puisque les trois protocoles de traitement pouvaient être considérés comme des compositions des trois composantes principales : Chirurgie (S), Chimiothérapie (C) et Radiothérapie (R). En prenant la chirurgie seule comme catégorie de référence, on peut donc coder le modèle avec deux indicatrices (une pour C, une pour R) selon le schéma suivant :

**Tableau 12:** Codage des indicatrices de traitement dans MANATEC-02 considéré comme un réseau additif

Type d'essai	Bras de randomisation	Indicatrice de Chimio	Indicatrice de Radio
<b>S contre CS</b>	CS	+0.5	0
	S	-0.5	0
<b>S contre CRS</b>	CRS	+0.5	+0.5
	S	-0.5	-0.5
<b>CS contre CRS</b>	CRS	0	+0.5
	CS	0	-0.5

S : chirurgie d'emblée, CS : chimiothérapie suivie de chirurgie, CRS : chimioradiothérapie suivie de chirurgie

### Essais multi bras

Les modèles IPD en une étape permettent en théorie de prendre en compte les essais multi bras en introduisant une corrélation entre les effets aléatoires des comparaisons provenant d'un même essai. En pratique, MANATEC-02 ne contenait qu'un seul essai multi bras de petite taille. Cette corrélation a donc été négligée et les effets aléatoires des traitements des patients inclus dans cet essai ont été considérés comme non corrélés. La méthode netmeta qui permet de prendre en compte les essais multi-bras a été utilisée comme analyse de sensibilité (cf. infra).

### Hétérogénéité et cohérence

L'hétérogénéité était quantifiée par le  $I^2$ .(48)

Ce réseau possédant une boucle fermée, il était possible de tester l'hypothèse de cohérence. Celle-ci était quantifiée par la statistique  $Q$ .(27) Dans les modèles IPD en une étape, il est également possible de tester l'hypothèse de cohérence par l'introduction d'une nouvelle indicatrice de traitement sans modifier le reste du codage. Celle-ci était introduite pour la comparaison CS contre CRS (-0,5 pour CS ; +0,5 pour CRS ; 0 dans tous les autres essais). Si l'intervalle de confiance du paramètre (i.e. le  $\log(HR)$ ) associé à cette indicatrice ne contenait pas la valeur 0, ceci indiquait un problème d'incohérence.

### Analyses de sensibilité

Deux analyses de sensibilité étaient prévues au protocole sur la sélection des essais : (i) exclusion des essais de moins de 60 patients, (ii) exclusion des essais clairement aberrant (« outlier »).

Des analyses de sensibilité ont également été réalisées afin de vérifier l'impact de la méthodologie statistique et de la représentation du réseau de traitement. Afin de tester la robustesse de notre approche, un modèle de type IPD en deux étapes était également utilisé. Dans ce modèle, les HR de chaque essai étaient calculés comme dans les méta-analyses standards par un modèle de Cox avant d'être combiné par l'approche basée sur les contrastes de type Netmeta. Cette méthode permettait de prendre en compte facilement l'essai à quatre bras. Ce modèle permettait également d'estimer l'effet

traitement dans chacun des niveaux de la covariable en interaction potentielle, mais pas de tester formellement que celle-ci était bien différente de zéro.

## Logiciels

Le « checking » et le recodage des essais a été réalisé par des macro spécialisées dans le logiciel SAS 9.4. Toutes les autres analyses ont été réalisées dans le logiciel R 4.1.0 (R foundation for Statistical Computing, Vienne, Autriche, 2021) en utilisant les paquets : Coxme 2.2-16 (Therneau, 2020), lme4 1.1-27 (Bates, 2015), metafor 3.0-2 (Viechtbauer, 2010), netmeta 1.4-0 (Rücker, 2021) et MICE 3.13-0 (van Buuren, 2011) pour l'imputation multiple.



## Chapitre 3. Résultats

### Caractéristiques des essais identifiés

Dix nouveaux essais ont été identifiés. Parmi les 25 essais déjà identifiés, 12 équipes ont été capables de nous envoyer des données mises à jour. En revanche, les équipes de 9 essais n'ont jamais répondu à aucune de nos tentatives de contact. Le diagramme des flux des essais est présenté en **Figure 14**. Un essai (Scandinavia) était un essai en plan factoriel 2x2 randomisant séparément l'utilisation de la radiothérapie et la chimiothérapie. Tous les bras n'ont pas été utilisés dans toutes les comparaisons.

### Essais comparant chirurgie seule à chimiothérapie puis chirurgie (S contre CS)

Les essais sont détaillés dans le **Tableau 13**. On notait dans cette comparaison la présence de deux de quatre bras de l'essai Scandinavia. Ces essais étaient plutôt anciens avec des périodes d'inclusion pour la plupart avant les années 2000.

Seize essais ont été identifiés, dont huit inclus dans la précédente méta-analyse. Les IPD étaient disponibles pour 12 essais représentant 2 478 patients, soit 90,4% des patients théoriques.

Tous les essais utilisaient un sel de platine (cisplatine) associé à la bléomycine, le fluoro-uracile ou l'épirubicine. Deux à quatre cycles étaient administrés.

Les suivis médians étaient variables selon les essais, allant de 3,3 ans à 26 ans. Aucun des essais n'avait de différence de suivi entre les bras de traitements. La médiane de suivi globale était de 5,6 [5,2 – 6,1] ans.

### Essais comparant chirurgie seule à chimioradiothérapie suivie de chirurgie (S contre CRS)

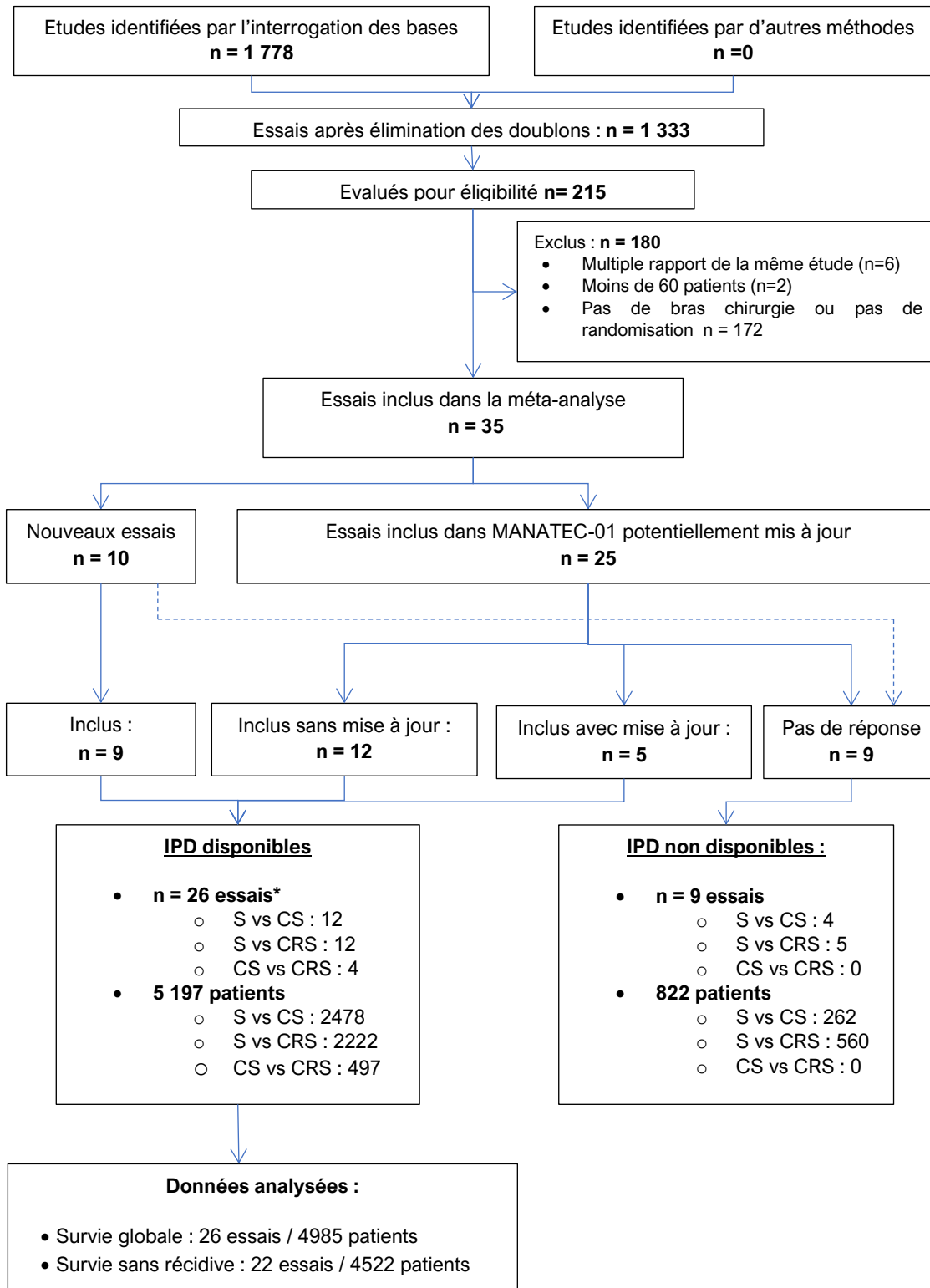
Les essais sont détaillés dans le **Tableau 14**. On notait la présence des deux essais Irlandais, présenté comme deux essais séparés réalisés en parallèle l'un de l'autre. L'un incluait des adénocarcinomes (Ireland 1), l'autre des carcinomes épidermoïdes (Ireland 2). Ces essais sont un peu plus récents que ceux de la comparaison précédente, plusieurs ayant inclus des patients au cours des années 2000.

Dix-sept essais ont été identifiées, dont huit inclus dans la précédente méta-analyse. Les IPD étaient disponibles pour 12 essais représentant 2 222 patients, soit 79,9% du total théorique.

La radiothérapie délivrait entre 35 et 50,4 Gy en 10 à 23 fractions. La chimiothérapie était le plus souvent concomitante basée sur le cisplatine et le fluoro-uracile principalement.

Les suivis médians étaient variables selon les essais, allant de 4,4 ans à 17 ans. Aucun des essais n'avait de différence de suivi entre les bras de traitements. La médiane de suivi globale était de 6,4 [6,1 – 6,9] ans.





*Figure 14: Diagramme des flux des essais*

\* un essai avec 4 bras inclus dans toutes les comparaisons, IPD : données individuelles des patients S : chirurgie, CS : chimiothérapie suivie de chirurgie, CRS : chimio radiothérapie suivie de chirurgie

## Essais comparant chimiothérapie contre chimioradiothérapie suivie de chirurgie (CS contre CRS)

Les essais sont détaillés dans le **Tableau 15**. Hormis la quatrième comparaison de l'essai Scandinavia, ces essais ont tous recruté des patients dans les années 2000.

Quatre essais étaient identifiés représentant 497 patients. Les IPD de chacun d'entre eux étaient disponibles (100%).

Les protocoles de chimiothérapie étaient basés sur le cisplatine principalement associé au fluoro-uracile. La radiothérapie délivrait 30 à 40 Gy en 15 à 20 fractions. La chimiothérapie utilisée dans les groupes chimioradiothérapie était la même que dans les groupes chimiothérapie en dehors de quelques adaptations de doses.

Les suivis médians étaient variables selon les essais, allant de 4,5 ans à 18 ans. Aucun des essais n'avait de différence de suivi entre les bras de traitement. La médiane de suivi globale était de 8,0 [6,2 – 9,6] ans.

*Tableau 13: Essais comparant chirurgie seule contre chimiothérapie puis chirurgie*

Essai Nom abrégé (Premier auteur, année)	Période d'inclusion	N Cycles	Protocole de chimiothérapie	Suivi médian année [IC 95%] <sup>§</sup>	N*
<b>IPD disponible</b>					
<b>MD Anderson</b> (Roth 1988) <sup>†</sup> (52)	1982-1986	2	Cisplatine 120 mg/m <sup>2</sup> /j ; J1 Bléomycine 10 UI/m <sup>2</sup> /j ; J3 à 6 Vindesine 3 mg/m <sup>2</sup> /j ; J1, 8, 15, 22	2,6[1,4-NR]	36
<b>Scandinavia</b> (Nygaard 1992) <sup>†§</sup> (53)	1983-1988	2	Cisplatine 20 mg/m <sup>2</sup> /j ; J1 à 5 Bléomycine 5 mg/m <sup>2</sup> /j ; J1 à 5	11[7,4-NR]	106
- <i>Scandinavia</i> : CS vs S				11[7,9-NR]	111
- <i>Scandinavia</i> : CRS vs RS			+ 35 Gy / 20 fractions		
<b>Oeso 2</b> (Giuli) <sup>†</sup> (Non publié)	1985-1989	2	Cisplatine 120 mg/m <sup>2</sup> /j, J1. Bléomycine 10 mg/m <sup>2</sup> /j, J3 à 6 Vinblastine 3 mg/m <sup>2</sup> /j, J1, 8, 15 & 22.	4,2[4,2-5,0]	122
<b>Songkla</b> (Maipang 1994) <sup>†</sup> (54)	1988-1990	2	Cisplatine 100 mg/m <sup>2</sup> ; j1 Bléomycine 10 mg/m <sup>2</sup> puis 10mg/m <sup>2</sup> /j ; J4 à 7 Vinblastine 3 mg/m <sup>2</sup> J1, 18, 15, 22	4,5[2,9-NR]	46
<b>Queen Mary</b> (Law 1997) <sup>†</sup> (55)	1989-1995	2	Cisplatine 100 mg/m <sup>2</sup> ; j1 Fluorouracile 500 mg/m <sup>2</sup> /j ; J1 à 5	26 [23-NR]	147
<b>Rotterdam</b> (Boonstra 2011) <sup>†</sup> (56)	1989-1996	2-4	Cisplatine 80 mg/m <sup>2</sup> ; j1 Etoposide (IV) 100 mg/m <sup>2</sup> ; j1-2 Etoposide (PO) 200 mg/m <sup>2</sup> ; j3-5	4,6[4,1-6,5]	169
<b>RTOG 8911</b> (Kelsen 1998) <sup>†</sup> (57,58)	1990-1995	3	Cisplatine 100 mg/m <sup>2</sup> ; j1 Fluorouracile 1000 mg/m <sup>2</sup> /j ; J1 à 5	9,2[8,8-10]	467
<b>Italy</b> (Ancona 2001) <sup>†</sup> (59)	1992-1997	2	Cisplatine 100 mg/m <sup>2</sup> ; j1 Fluorouracile 1000 mg/m <sup>2</sup> /j ; J1 à 5	6,9[6,6-7,8]	96
<b>MRC EO-02</b> (Medical Research Council 2002) <sup>†*</sup> (60,61)	1992-1998	2	Cisplatine 80 mg/m <sup>2</sup> ; j1 Fluorouracile 1000 mg/m <sup>2</sup> /j ; J1 à 4	4,0[3,8-4,5]	802
<b>MAGIC</b> (Cunningham 2006) <sup>μ</sup> (62)	1994-2002	3	Cisplatine 60 mg/m <sup>2</sup> ; j1 Fluorouracile 200 mg/m <sup>2</sup> /j ; J1 à J21 Epirubicine 50 mg/m <sup>2</sup> ; j1	3,3[2,9-4,0]	131
<b>ACCORD7</b> (Ychou 2011) <sup>μ</sup> (63)	1995-2003	2-3	Cisplatine 100 mg/m <sup>2</sup> ; j1 Fluorouracile 800 mg/m <sup>2</sup> /j ; J1 à 5	6,2[5,2-8,7]	169

<b>EORTC 40954</b> (Schuhmacher 2010) <sup>μ</sup> (64)	1999-2004	2	Cisplatine 50 mg/m <sup>2</sup> j ; J : 1-15-29 Fluorouracile 2000 mg/m <sup>2</sup> /j ; J : 1-8-15-22-36	4,4[3,9-5,4]	76
------------------------------------------------------------	-----------	---	---------------------------------------------------------------------------------------------------------------	--------------	----

**IPD non disponible**

<b>CAO</b> (Schlag 1992) (65)	NA	3	Cisplatine 20 mg/m <sup>2</sup> /j ; J1 à 5 Fluorouracile 1000 mg/m <sup>2</sup> /j ; J1 à 5		46
<b>China 01</b> (Wang 2000) (66)	1987-1988	1	FPLC <sup>¶</sup> 2x20 ml/j x 12.5 J		60
<b>China 02</b> (Wang 2001) (67)	1991-1994	NA	Cisplatine 30 mg J1 à J5		100
<b>Kagoshima</b> (Baba 2000) (68)	1993-1995	2	Cisplatine 70 mg/m <sup>2</sup> ; j1 Fluorouracile 700 mg/m <sup>2</sup> /j ; J1 à J5		56

<b>Total : 16 essais</b>	<b>2 740 (100%)</b>
<b>Disponibles : 12 essais</b>	<b>2 478 (90,4%)</b>
<b>Non disponibles : 4 essais</b>	<b>262 (9,6%)</b>

IV : intraveineux, PO : per-os, NA : Non disponible, NR : non atteinte (« not reached »)

\* Les nombres sont le nombre de patients reçus pour les essais disponibles et le nombre de patients dans la publication pour les autres. Les effectifs des sous-groupes peuvent ne pas correspondre au total en présence de données manquantes.

† Inclus dans la précédente méta-analyse dans mise à jour disponible

‡ Inclus dans la précédente méta-analyse avec mise à jour disponible

§ Essai avec quatre bras : Chirurgie seule (S), Chimiothérapie puis chirurgie (CS), Radiothérapie puis chirurgie (RS), Chimio radiothérapie puis chirurgie (CRS) ; comme dans la méta-analyse précédente, les deux comparaisons CS vs. S (n=106) et CRS vs RS (n=111) ont été incluses dans partie.

μ Ces essais incluait aussi des cancers de l'estomac. Le nombre total de patients randomisés était : 503 (MAGIC), 224 (ACCORD7 2011) et 144 (EORTC 40954).

¶ fluorouracili polyphase liposome composita pro orale (~ 2000 mg fluoro-uracile)

\$ Aucune étude n'avait de différence significative de suivi entre les bras

Tableau 14: Essais comparant chirurgie seule à chimioradiothérapie suivie de chirurgie

Essai Nom abrégé (Premier auteur, année)	Période d'inclusion	Protocole radiothérapie	de Protocole de chimiothérapie	Suivi médian	N*
<b>IPD disponible</b>					
<b>Scandinavia S_CRS</b> (Nygaard, 1992) <sup>§</sup> † (53)	1983-1988	35 Gy 20 fractions	Cisplatine 20 mg/m <sup>2</sup> /j ; J1 à 5 Bléomycine 5 mg/m <sup>2</sup> /j ; J1 à 5 RT 3 semaines après la chimiothérapie	9,7 (7,9-NR)	103
<b>Rennes</b> (Le Prise, 1994) † (69)	1988 -1991	20 Gy 10 fractions	Cisplatine 100 mg/m <sup>2</sup> ; J1 et J21 Fluorouracile 600 mg/m <sup>2</sup> ; J2 à J5 et J22 à J25 RT concomitante	17 (16-NR)	91
<b>Michigan</b> (Urba, 2001) † (70)	1989-1994	45 Gy	Cisplatine 20 mg/m <sup>2</sup> ; J1 à 5 and 17-21 Fluorouracile 300 mg/m <sup>2</sup> /j ; J1 à 21 Vinblastine 1 mg/m <sup>2</sup> /j ; J1 à 4 et 17 à 21 RT concomitante	9,9 (8,4-NR)	100
<b>EORTC88-05</b> (Bosset, 1997) † (71)	1989-1995	37 Gy	Cisplatine 85 mg/m <sup>2</sup> RT concomitante	4,6 (4,4-5,1)	294
<b>Ireland</b> (Bass, 2014) ‡ (72)(73)(74)	1990-1997	40 Gy 15 fractions	Cisplatine 75 mg/m <sup>2</sup> ; J7 et J49 Fluorouracile 15 mg/kg ; J1 à 5 et J42 à J47 RT concomitante		
- Ireland 1				14 (13-NR)	113
- Ireland 2				14 (11-NR)	97
<b>Australia</b> (Burmeister, 2005) ‡ (75)	1994-2000	35 Gy 15 fractions	Cisplatine 80 mg/m <sup>2</sup> ; j1 Fluorouracile 800 mg/m <sup>2</sup> ; J1 à 4 RT concomitante	5,4 (5,1-6,1)	257
<b>CALGB97-81</b> (Tepper, 2008) † (76)	1997-2000	50,4 Gy 28 fractions	Cisplatine 100 mg/m <sup>2</sup> ; J1 and 29 Fluorouracile 1000 mg/m <sup>2</sup> ; J1 à 4 et J29 à 32 RT concomitante	6,1 (5,8-7,0)	56
<b>Korea</b> (Lee, 2004) † (77)	1999-2002	45,6 Gy 38 fractions	Cisplatine 60 mg/m <sup>2</sup> ; J1 et 21 Fluorouracile 1000 mg/m <sup>2</sup> ; J2 à 5 and 22 à 26 RT concomitante	6,1 (5,5-6,4)	99
<b>FFCD99-01</b> (Mariette, 2014) (78)	2000-2009	45 Gy 25 fractions	Cisplatine 75 mg/m <sup>2</sup> ; J1 et 29 Fluorouracile 800 mg/m <sup>2</sup> ; J1 à 4 et J29 à J32 RT concomitante	11 (9,5-12)	195
<b>CROSS</b> (van Hagen, 2012) (79)(80)	2004-2008	41,4 Gy 23 fractions	Paclitaxel 50 mg/m <sup>2</sup> Carboplatine (AUC 2 mg/ml/min)	12 (12-12)	366

<b>Guangzhou</b> (Yang, 2012) <b>(81)</b>	2007-2011	40 Gy 20 fractions	RT concomitante Vinorelbine 25mg/m <sup>2</sup> ; J1-8-22-29 Cisplatine 75 mg/m <sup>2</sup> ; J1-22 RT concomitante	4,4 (4,2-4,7)	451
<b>IPD non disponible</b>					
<b>Songkla 2</b> (Apinop, 1994) <b>(82)</b>	1986-1992	40 Gy 20 fractions	Cisplatine 100 mg/m <sup>2</sup> ; J1 et 29 Fluorouracile 600 mg/m <sup>2</sup> ; J1 à 4 et 29 à 32 RT concomitante		69
<b>China 03</b> (An, 2003) <b>(83)</b>	1996-1997	36 Gy 12 fractions	Cisplatine Fluorouracile RT concomitante		97
<b>Nanjing</b> (Lv, 2010) <b>(84)</b>	1997-2004	40 Gy 20 fractions	Paclitaxel 135 mg/m <sup>2</sup> ; J1 et 22 Cisplatine 20 mg/m <sup>2</sup> /j ; J1 à 3 et 22 à 24 RT concomitante		238
<b>China04</b> (Peng, 2008) <b>(85)</b>	2000-2002	40 Gy NA	Cisplatine 70 mg/m <sup>2</sup> Fluorouracile 500 mg/m <sup>2</sup> RT concomitante		80
<b>Hebei</b> (Zhao, 2015) <b>(86)</b>	2012-2013	45 Gy 25 fractions	Capecitabine 2000 mg/m <sup>2</sup> /j ; J1 à J14 Oxaliplatine 130 mg/m <sup>2</sup> ; J1 RT concomitante		76
<b>Total 17 essais</b>					<b>2 782(100%)</b>
Disponibles : 12 essais					2 222(79,9%)
Non disponibles : 5 essais					560(20,1%)

IV : intraveineux, PO : per-os, NA : Non disponible, AUC : aire sous la courbe, NR : non atteinte (« not reached »)

\* Les nombres sont le nombre de patients reçue pour les essais disponibles et le nombre de patients dans la publication pour les autres. Les effectifs des sous-groupes peuvent ne pas correspondre au total en présence de données manquantes.

† Inclus dans la précédente méta-analyse dans mise à jour disponible

‡ Inclus dans la précédente méta-analyse avec mise à jour disponible

§ Essai avec quatre bras : Chirurgie seule (S), Chimiothérapie puis chirurgie (CS), Radiothérapie puis chirurgie (RS), Chimio radiothérapie puis chirurgie (CRS) ; seule la comparaison S vs. CRS (n=103) a été incluse dans cette partie.

*Tableau 15 : Essais comparant chimiothérapie à chimioradiothérapie suivie de chirurgie*

Essai Nom abrégé (Premier auteur, année)	Période d'inclusion	Protocole de chimiothérapie	Protocole de chimioradiothérapie	Suivi médian	N*
<b>IPD disponible</b>					
<b>Scandinavia (Nygaard, 1992) § † (53)</b>	1983-1988	2 cycles de : Cisplatine 20 mg/m <sup>2</sup> /j ; J1 à 5 Bléomycine 5 mg/m <sup>2</sup> /j ; J1 à 5	35 Gy 20 fractions 2 cycles de la même chimio que le groupe CT Début de la chimio 3 semaines après	9,7 (7,8-NR)	109
<b>POET (Stahl, 2009) (87) (88)</b>	2000-2005	2.5 cycles de : Cisplatine 50 mg/m <sup>2</sup> /j : bihebdomadaire Fluorouracile 2000 mg/m <sup>2</sup> ; J1-8-15-22-29-36	30 Gy 15 fractions 2 cycles de la même chimio que le groupe CT Suivie de CRT avec Cisplatine 50 mg/m <sup>2</sup> /j : J1 et J8 Etoposide 80 mg/m <sup>2</sup> /j ; J3 à 5	11 (9,7-12)	119
<b>IG9401 (Burmeister, 2011) (89)</b>	2000-2006	2 cycles de : Cisplatine 80 mg/m <sup>2</sup> ; j1 Fluorouracile 1000 mg/m <sup>2</sup> /j ; J1 à J4	35 Gy 15 fractions 2 cycles de la même chimio que le groupe CT avec réduction du 5-FU à 800 mg/m <sup>2</sup> Débutant à J22	18 (15-18)	88
<b>NeoRes (Klevebro, 2016) (90)</b>	2006-2013	3 cycles de : Cisplatine 100 mg/m <sup>2</sup> ; j1 Fluorouracile 750 mg/m <sup>2</sup> /j ; J1 à 5	40 Gy 20 fractions 3 cycles de la même chimio que le groupe CT Concomitante aux cycles 2 et 3	4,5 (4,2-5,3)	181
<b>Total 4 essais</b>					<b>497(100%)</b>
<b>Disponibles 4 essais</b>					<b>497(100%)</b>
<b>Non disponible 0 essais</b>					<b>0(0%)</b>

IV : intraveineux, PO : per-os, NA : Non disponible

\* Les nombres sont le nombre de patients reçus pour les essais disponibles et le nombre de patients dans la publication pour les autres. Les effectifs des sous-groupes peuvent ne pas correspondre au total en présence de données manquantes.

† Inclus dans la précédente méta-analyse dans mise à jour disponible

§ Essai avec quatre bras : Chirurgie seule (S), Chimiothérapie puis chirurgie (CS), Radiothérapie puis chirurgie (RS), Chimio radiothérapie puis chirurgie (CRS) ; seule la comparaison CS vs. CRS (n=109) a été incluse dans cette partie.

## Caractéristiques des patients inclus dans l'analyse

### Comparaison chirurgie seule à chimiothérapie puis chirurgie (S contre CS)

La plupart des participants étaient des hommes (80%) âgés de 62 ans en médiane. Les tumeurs étaient plutôt localisées dans l'OT que dans la JOG (1 806/511). La répartition était plus équilibrée entre CE et ADK (1 131/1 296). Les informations sur le stade TNM pré thérapeutique de la tumeur étaient manquantes chez plus de la moitié des patients, rendant délicate toute interprétation. Les caractéristiques générales et détaillées par essai sont décrites dans le **Tableau 16**.

### Comparaison chirurgie seule à chimioradiothérapie suivie de chirurgie (S contre CRS)

La plupart des participants étaient des hommes (82%) âgés de 60 ans en médiane. Les tumeurs étaient en très large majorité localisées dans l'OT plutôt que dans la JOG (1 806/511) et correspondaient majoritairement à des CE plutôt que des ADK (1 131/1 296). Les stades T et N du TNM étaient manquants chez près de la moitié des patients, rendant délicate toute interprétation. Les caractéristiques générales et détaillées par essai sont décrites dans le **Tableau 17**.

### Comparaison chimiothérapie contre chimioradiothérapie suivie de chirurgie (CS contre CRS)

Ces essais avaient également inclus plutôt des hommes (84%) âgés de 62 ans en médiane. Les tumeurs étaient plus localisées dans l'OT que dans la JOG (437/55) et correspondaient plus à des ADK que des CE (327/170). La répartition entre les tumeurs T1T2 et T3T4 était équilibrée (234/261). Plus de la moitié des patients avaient une tumeur considérée comme N+ en pré thérapeutique (64%). Les caractéristiques générales et détaillées par essai sont décrites dans le **Tableau 18**.



**Tableau 16:** Caractéristiques des patients inclus dans les essais de chirurgie seule contre chimiothérapie suivie de chirurgie

	MD Anderson N = 36	Scandinavia : S_CS N = 106	Scandinavia : RS_CRS N = 111	Oeso 2 N = 122	Songkla N = 46	Queen Mary N = 147	Rotterdam N = 169
<b>Age (années) médiane (IQR)</b>		62 (56, 67)	62 (57, 68)	54 (48, 60)	66 (60, 70)	65 (57, 70)	60 (53, 68)
Manquant	36	0	3	4	0	0	0
<b>Sexe</b>							
Homme	30 (83%)	84 (79%)	73 (66%)	106 (87%)	40 (87%)	125 (85%)	126 (75%)
Femme	6 (17%)	22 (21%)	37 (34%)	16 (13%)	6 (13%)	22 (15%)	43 (25%)
Manquant	0	0	1	0	0	0	0
<b>Histologie</b>							
ADK	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
CE	36 (100%)	106 (100%)	111 (100%)	122 (100%)	46 (100%)	147 (100%)	169 (100%)
Manquant	0	0	0	0	0	0	0
<b>Localisation</b>							
ET	36 (100%)	104 (100%)	104 (100%)	0 (ND%)	46 (100%)	144 (100%)	150 (100%)
GEJ	0 (0%)	0 (0%)	0 (0%)	0 (ND%)	0 (0%)	0 (0%)	0 (0%)
Manquant	0	2	7	122	0	3	19
<b>ECOG</b>							
0	0 (0%)	0 (ND%)	0 (ND%)	47 (46%)	16 (35%)	0 (ND%)	0 (ND%)
1	36 (100%)	0 (ND%)	0 (ND%)	53 (51%)	26 (57%)	0 (ND%)	0 (ND%)
2-3-4	0 (0%)	0 (ND%)	0 (ND%)	3 (2.9%)	4 (8.7%)	0 (ND%)	0 (ND%)
Manquant	0	106	111	19	0	147	169
<b>T du TNM</b>							
T1-T2	36 (100%)	101 (97%)	106 (98%)	0 (ND%)	3 (6.5%)	0 (ND%)	0 (ND%)
T3-T4	0 (0%)	3 (2,9%)	2 (1,9%)	0 (ND%)	43 (93%)	0 (ND%)	0 (ND%)
Manquant	0	2	3	122	0	147	169
<b>N du TNM</b>							
N0	0 (ND%)	94 (99%)	94 (100%)	0 (ND%)	46 (100%)	0 (ND%)	0 (ND%)
N+	0 (ND%)	1 (1,1%)	0 (0%)	0 (ND%)	0 (0%)	0 (ND%)	0 (ND%)
Manquant	36	11	17	122	0	147	169

ADK : adénocarcinome, CE : carcinome épidermoïde, ET : Œsophage thoracique, JOG : jonction œsogastrique, ECOG : Score de performance de l'Eastern Cooperative Oncology Group, TNM : classification Tumor Node Metastasis pré-thérapeutique. ND : non disponible

*Suite du tableau : Caractéristiques des patients inclus dans les essais de chirurgie seule contre chimiothérapie suivie de chirurgie*

	<b>RTOG 8911 N = 467</b>	<b>Italy N = 96</b>	<b>MRC OE02 N = 802</b>	<b>MRC MAGIC N = 131</b>	<b>ACCORD 07 N = 169</b>	<b>EORTC 40954 N = 76</b>	<b>Global N = 2 478</b>
<b>Age (années) médiane (IQR)</b>	62 (53, 67)	58 (53, 64)	62 (55, 68)	60 (53, 68)	64 (56, 70)	57 (49, 62)	<b>62 (54, 68)</b>
Manquant	0	0	0	0	0	0	<b>43</b>
<b>Sexe</b>							
Homme	388 (83%)	77 (80%)	603 (75%)	114 (87%)	152 (90%)	63 (83%)	<b>1 981 (80%)</b>
Femme	79 (17%)	19 (20%)	199 (25%)	17 (13%)	17 (10%)	13 (17%)	<b>496 (20%)</b>
Manquant	0	0	0	0	0	0	<b>1</b>
<b>Histologie</b>							
ADK	244 (53%)	0 (0%)	533 (68%)	131 (100%)	167 (100%)	56 (100%)	<b>1 131 (47%)</b>
CE	216 (47%)	96 (100%)	247 (32%)	0 (0%)	0 (0%)	0 (0%)	<b>1 296 (53%)</b>
Manquant	7	0	22	0	2	20	<b>51</b>
<b>Localisation</b>							
ET	308 (67%)	96 (100%)	720 (90%)	73 (56%)	25 (15%)	0 (0%)	<b>1 806 (78%)</b>
GEJ	151 (33%)	0 (0%)	82 (10%)	58 (44%)	144 (85%)	76 (100%)	<b>511 (22%)</b>
Manquant	8	0	0	0	0	0	<b>161</b>
<b>ECOG</b>							
0	294 (66%)	76 (79%)	532 (66%)	89 (68%)	127 (75%)	51 (67%)	<b>1 232 (65%)</b>
1	145 (32%)	19 (20%)	247 (31%)	42 (32%)	42 (25%)	25 (33%)	<b>635 (33%)</b>
2-3-4	9 (2,0%)	1 (1,0%)	23 (2,9%)	0 (0%)	0 (0%)	0 (0%)	<b>40 (2,1%)</b>
Manquant	19	0	0	0	0	0	<b>571</b>
<b>T du TNM</b>							
T1-T2	285 (63%)	36 (38%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (0%)	<b>567 (62%)</b>
T3-T4	170 (37%)	60 (62%)	0 (ND%)	0 (ND%)	0 (ND%)	75 (100%)	<b>353 (38%)</b>
Manquant	12	0	802	131	169	1	<b>1,558</b>
<b>N du TNM</b>							
N0	160 (64%)	63 (66%)	0 (ND%)	0 (ND%)	0 (ND%)	2 (2.7%)	<b>459 (70%)</b>
N+	90 (36%)	33 (34%)	0 (ND%)	0 (ND%)	0 (ND%)	72 (97%)	<b>196 (30%)</b>
Manquant	217	0	802	131	169	2	<b>1 823</b>

ADK : adénocarcinome, CE : carcinome épidermoïde, ET : Œsophage thoracique, JOG : jonction œsogastrique, ECOG : Score de performance de l'Eastern Cooperative Oncology Group, TNM : classification Tumor Node Metastasis préthérapeutique. ND : non disponible

*Tableau 17 : Caractéristiques des patients inclus dans les essais de chirurgie seule contre chimioradiothérapie suivie de chirurgie*

	Scandinavia : S_CRS N = 103	Rennes N = 91	Michigan N = 100	EORTC 88- 05 N = 294	Ireland 1 N = 113	Ireland 2 N = 97	Australia N = 257
<b>Age (années) médiane (IQR)</b>	62 (56, 67)	57 (52, 64)	62 (55, 67)	58 (51, 63)	65 (57, 70)	66 (58, 70)	62 (55, 68)
<b>Manquant</b>	1	10	0	12	0	1	1
<b>Sexe</b>							
<b>Homme</b>	76 (74%)	77 (95%)	84 (84%)	263 (93%)	83 (73%)	50 (52%)	206 (80%)
<b>Femme</b>	27 (26%)	4 (4,9%)	16 (16%)	19 (6,7%)	30 (27%)	47 (48%)	51 (20%)
<b>Manquant</b>	0	10	0	12	0	0	0
<b>Histologie</b>							
<b>ADK</b>	0 (0%)	0 (0%)	76 (76%)	0 (0%)	113 (100%)	0 (0%)	158 (62%)
<b>CE</b>	103 (100%)	91 (100%)	24 (24%)	294 (100%)	0 (0%)	97 (100%)	95 (38%)
<b>Manquant</b>	0	0	0	0	0	0	4
<b>Localisation</b>							
<b>ET</b>	98 (100%)	80 (99%)	100 (100%)	275 (100%)	74 (65%)	89 (97%)	0 (ND%)
<b>GEJ</b>	0 (0%)	1 (1,2%)	0 (0%)	0 (0%)	39 (35%)	3 (3,3%)	0 (ND%)
<b>Manquant</b>	5	10	0	19	0	5	257
<b>ECOG</b>							
<b>0</b>	0 (ND%)	58 (72%)	0 (ND%)	154 (66%)	0 (ND%)	0 (ND%)	0 (ND%)
<b>1</b>	0 (ND%)	23 (28%)	0 (ND%)	73 (31%)	0 (ND%)	0 (ND%)	0 (ND%)
<b>2-3-4</b>	0 (ND%)	0 (0%)	0 (ND%)	7 (3,0%)	0 (ND%)	0 (ND%)	0 (ND%)
<b>Manquant</b>	103	10	100	60	113	97	257
<b>T du TNM</b>							
<b>T1-T2</b>	96 (96%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)
<b>T3-T4</b>	4 (4,0%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)
<b>Manquant</b>	3	91	100	294	113	97	257
<b>N du TNM</b>							
<b>N0</b>	88 (100%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)
<b>N+</b>	0 (0%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)	0 (ND%)
<b>Manquant</b>	15	91	100	294	113	97	257

ADK : adénocarcinome, CE : carcinome épidermoïde, ET : Œsophage thoracique, JOG : jonction œsogastrique, ECOG : Score de performance de l'Eastern Cooperative Oncology Group, TNM : classification Tumor Node Metastasis préthérapeutique. ND : non disponible

*Suite du tableau : Caractéristiques des patients inclus dans les essais de chirurgie seule contre chimioradiothérapie suivie de chirurgie*

	<b>CALGB 97-81 N = 56</b>	<b>Korea N = 99</b>	<b>FFCD 99-01 N = 195</b>	<b>CROSS N = 366</b>	<b>Guangzhou N = 451</b>	<b>Global N = 2 222</b>
<b>Age (années) médiane (IQR)</b>	60 (56, 68)	63 (57, 67)	58 (52, 64)	60 (54, 66)	57 (53, 61)	<b>60 (54, 65)</b>
<b>Manquant</b>	0	0	0	1	0	<b>26</b>
<b>Sexe</b>						
<b>Homme</b>	51 (91%)	91 (92%)	167 (86%)	286 (78%)	367 (81%)	<b>1 801 (82%)</b>
<b>Femme</b>	5 (8,9%)	8 (8,1%)	28 (14%)	80 (22%)	84 (19%)	<b>399 (18%)</b>
<b>Manquant</b>	0	0	0	0	0	<b>22</b>
<b>Histologie</b>						
<b>ADK</b>	42 (75%)	0 (0%)	57 (29%)	275 (77%)	0 (0%)	<b>721 (33%)</b>
<b>CE</b>	14 (25%)	99 (100%)	137 (71%)	84 (23%)	451 (100%)	<b>1 489 (67%)</b>
<b>Manquant</b>	0	0	1	7	0	<b>12</b>
<b>Localisation</b>						
<b>ET</b>	0 (ND%)	67 (68%)	195 (100%)	318 (87%)	451 (100%)	<b>1 747 (93%)</b>
<b>GEJ</b>	0 (ND%)	32 (32%)	0 (0%)	47 (13%)	0 (0%)	<b>122 (6.5%)</b>
<b>Manquant</b>	56	0	0	1	0	<b>353</b>
<b>ECOG</b>						
<b>0</b>	37 (66%)	5 (5,1%)	147 (77%)	307 (84%)	450 (100%)	<b>1 158 (78%)</b>
<b>1</b>	16 (29%)	94 (95%)	44 (23%)	59 (16%)	1 (0,2%)	<b>310 (21%)</b>
<b>2-3-4</b>	3 (5,4%)	0 (0%)	1 (0,5%)	0 (0%)	0 (0%)	<b>11 (0.7%)</b>
<b>Manquant</b>	0	0	3	0	0	<b>743</b>
<b>T du TNM</b>						
<b>T1-T2</b>	0 (ND%)	38 (38%)	157 (81%)	63 (17%)	72 (16%)	<b>426 (35%)</b>
<b>T3-T4</b>	0 (ND%)	61 (62%)	36 (19%)	298 (83%)	379 (84%)	<b>778 (65%)</b>
<b>Manquant</b>	56	0	2	5	0	<b>1,018</b>
<b>N du TNM</b>						
<b>N0</b>	50 (91%)	41 (41%)	141 (72%)	117 (33%)	61 (14%)	<b>498 (40%)</b>
<b>N+</b>	5 (9,1%)	58 (59%)	54 (28%)	236 (67%)	390 (86%)	<b>743 (60%)</b>
<b>Manquant</b>	1	0	0	13	0	<b>981</b>

ADK : adénocarcinome, CE : carcinome épidermoïde, ET : Œsophage thoracique, JOG : jonction œsogastrique, ECOG : Score de performance de l'Eastern Cooperative Oncology Group, TNM : classification Tumor Node Metastasis préthérapeutique, ND : non disponible

**Tableau 18 : Caractéristiques des patients inclus dans les essais de chimiothérapie suivie de chirurgie contre chimioradiothérapie suivie de chirurgie**

	Scandinavia : CS_CRS N = 109	POET N = 119	IG9401 N = 88	NeoRes N = 181	Global N = 497
<b>Age</b> (années) médiane (IQR)	62 (56, 67)	ND (ND, ND)	63 (55, 68)	63 (57, 67)	<b>63 (56, 67)</b>
<b>Manquant</b>	1	119	0	0	<b>120</b>
<b>Sexe</b>					
Homme	82 (75%)	108 (91%)	76 (86%)	149 (82%)	<b>415 (84%)</b>
Femme	27 (25%)	11 (9,2%)	12 (14%)	32 (18%)	<b>82 (16%)</b>
Manquant					
<b>Histologie</b>					
ADK	0 (0%)	119 (100%)	77 (88%)	131 (72%)	<b>327 (66%)</b>
CE	109 (100%)	0 (0%)	11 (12%)	50 (28%)	<b>170 (34%)</b>
Manquant					
<b>Localisation</b>					
ET	104 (100%)	119 (100%)	64 (73%)	150 (83%)	<b>437 (89%)</b>
JOG	0 (0%)	0 (0%)	24 (27%)	31 (17%)	<b>55 (11%)</b>
Manquant	5	0	0	0	<b>5</b>
<b>ECOG</b>					
0	0 (ND%)	71 (63%)	81 (92%)	152 (84%)	<b>304 (80%)</b>
1	0 (ND%)	41 (37%)	7 (8,0%)	29 (16%)	<b>77 (20%)</b>
2-3-4					
Manquant	109	7	0	0	<b>116</b>
<b>T du TNM</b>					
T1-T2	107 (99%)	0 (0%)	63 (72%)	64 (35%)	<b>234 (47%)</b>
T3-T4	1 (0,9%)	118 (100%)	25 (28%)	117 (65%)	<b>261 (53%)</b>
Manquant	1	1	0	0	<b>2</b>
<b>N du TNM</b>					
N0	94 (99%)	0 (ND%)	71 (81%)	67 (37%)	<b>232 (64%)</b>
N+	1 (1,1%)	0 (ND%)	17 (19%)	114 (63%)	<b>132 (36%)</b>
Manquant	14	119	0	0	<b>133</b>

ADK : adénocarcinome, CE : carcinome épidermoïde, ET : Œsophage thoracique, JOG : jonction œsogastrique, ECOG : Score de performance de l'Eastern Cooperative Oncology Group, TNM : classification Tumor Node Metastasis préthérapeutique. ND : non disponible

## MA par paire 1 : chirurgie seule contre chimiothérapie suivie de chirurgie

### Effet sur la survie globale

#### *Effet en population générale*

Deux mille quatre cent soixante-dix-huit patients étaient évaluable. Pendant le suivi, 2 010 décès ont été observés. La chimiothérapie préopératoire était associée à une meilleure survie globale que la chirurgie d'emblée  $HR_{aléatoire} = 0,83[0,72-0,96]$ ,  $p = 0,009$ . Il existait de l'hétérogénéité ( $I^2 = 50\%$ ,  $\tau^2 = 0.03$ ,  $p = 0.02$ ).

La CS conférait un bénéfice absolu à 3 ans de +5,5% ( $\pm 1,8\%$ ), passant de 23,9% à 29,4% et de +6,1% ( $\pm 1,7\%$ ) à 5 ans passant de 14,8% à 20,9%. La dRMST à 5 ans était de +3,3 ( $\pm 1,3$ ) mois passant de 21,7 mois à 25,1 mois.

Le suivi médian était de 5,6 [5,2-6,1] ans.

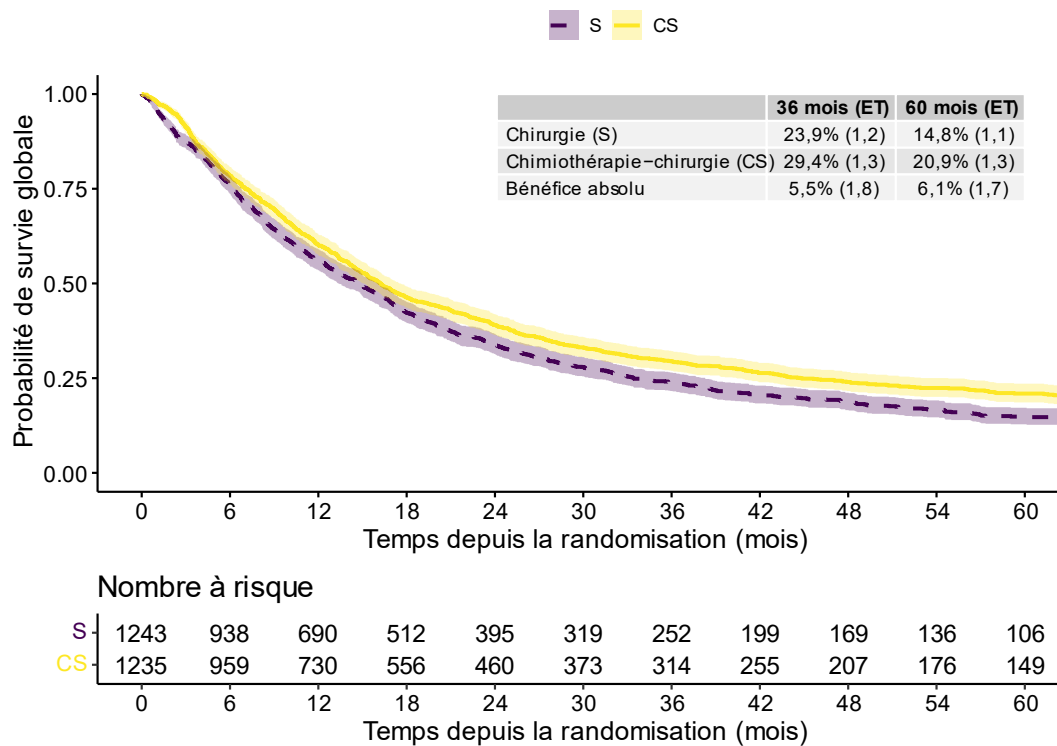


Figure 15 : Courbes de survie globale S contre CS

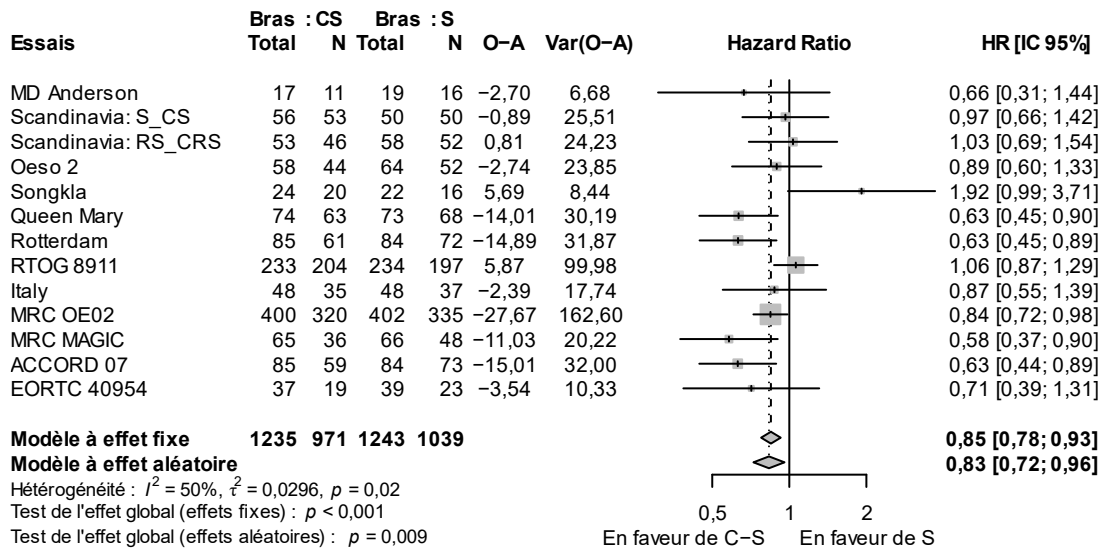


Figure 16 : Forest plot de l'effet de CS contre S sur la survie globale dans la population globale

### Effet en fonction des sous-groupes de participants

L'analyse selon la localisation suggérait un effet traitement potentiellement différent sans atteindre la significativité ( $p_{\text{interaction}} = 0,09$ ). L'effet de la CS était de  $HR_{\text{aléatoire}} = 0,88[0,76-1,03]$  pour les tumeurs situées dans l'œsophage thoracique et de  $HR_{\text{aléatoire}} = 0,69[0,53-0,90]$  pour celles situées à la jonction œsogastrique. Cette différence relative correspondait à une dRMST de  $+2,0[IC\ 95\% -0,7 ; +4,8]$  mois pour les lésions localisées dans l'ET, contre  $+6,9[IC95\% +2,4 ; +11,3]$  mois pour celles à la JOG.

On n'observait pas d'argument pour une variation selon le type histologique ( $p_{\text{interaction}} = 0,26$ ). Le  $HR_{\text{aléatoire}}$  était de  $0,91[0,76-1,08]$  pour les carcinomes épidermoïdes et de  $HR_{\text{aléatoire}} = 0,73[0,62-0,87]$  pour les adénocarcinomes.

Aucune interaction n'était mise en évidence avec l'âge, qu'il soit considéré comme qualitatif à quatre classes ( $p_{\text{interaction}} = 0,21$ ) ou continu ( $p_{\text{interaction}} = 0,98$ ) ou le sexe ( $p_{\text{interaction}} = 0,43$ ). Les données disponibles étaient insuffisantes pour estimer l'interaction avec les stades cT ou cN.

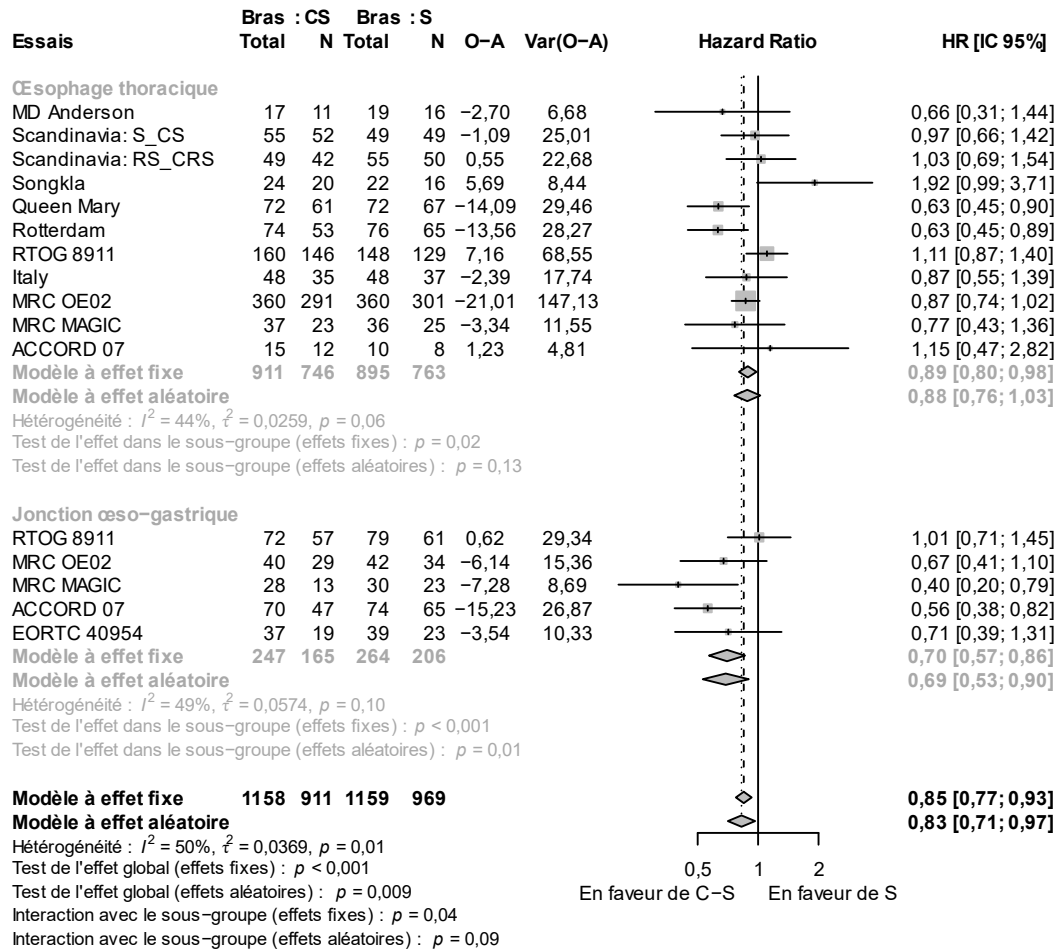


Figure 17 : Forest plot de l'effet de CS contre S sur la survie globale selon la localisation tumorale



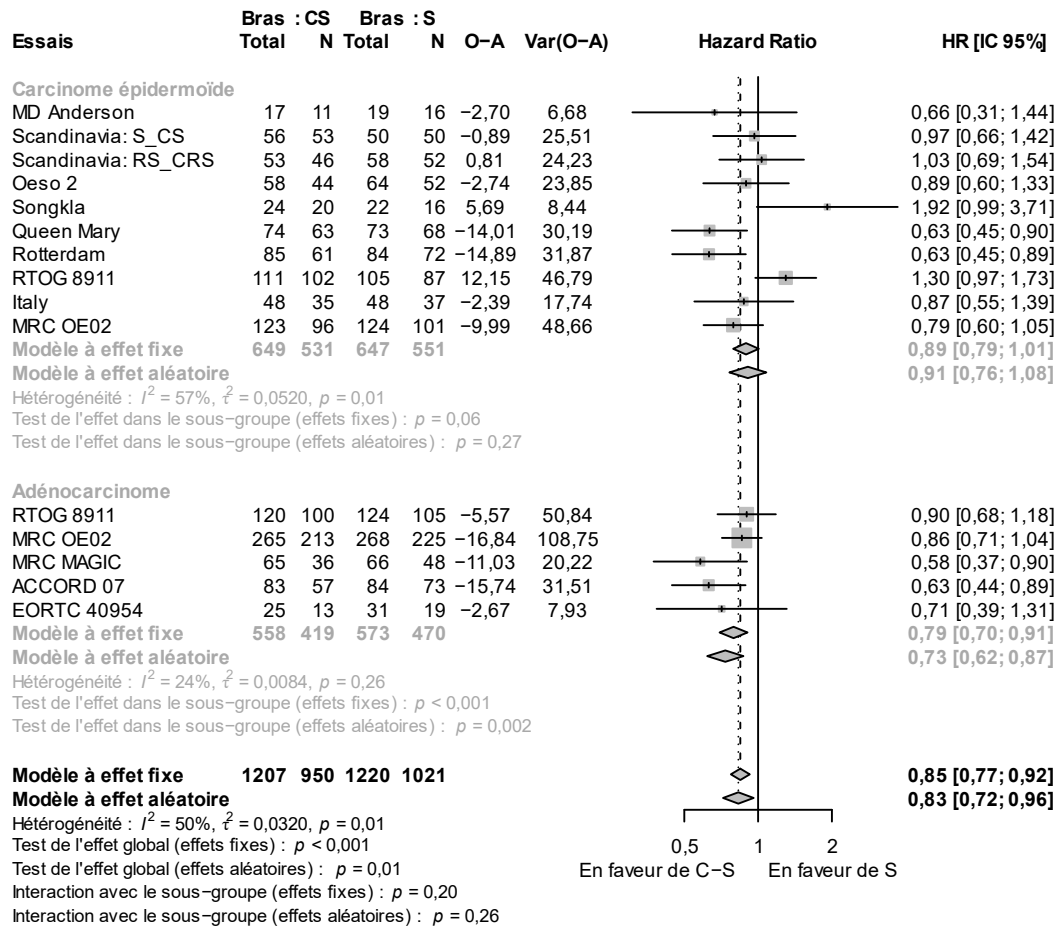


Figure 18: Forest plot de l'effet de CS contre S sur la survie globale en fonction du type histologique

## Effet sur les autres critères

### Survie sans récurrence

Des données de SSR étaient disponibles dans 10 essais (2 225 patients). Pendant le suivi 1 783 évènements de survie sans récurrence ont été observés, dont 705 survenant avant le « landmark » à 6 mois. La chimiothérapie préopératoire conférait une meilleure SSP que la chirurgie d'emblée :  $HR_{aléatoire} = 0,74[0,64-0,85]$ ,  $p < 0,0001$ . Il y avait de l'hétérogénéité ( $I^2 = 56\%$ ,  $\tau^2 = 0,03$ ,  $p = 0,01$ ). Les bénéfices absolus de SSP étaient de +7,6% ( $\pm 1,8\%$ ) à 3 ans (de 18,4% à 26,0%) et +6,8% ( $\pm 1,6$ ) à cinq ans (de 14,9% à 21,7%). La différence de RMST à 5 ans était de +5,5( $\pm 1,4$ ) mois passant de 19,6 à 25,0 mois.

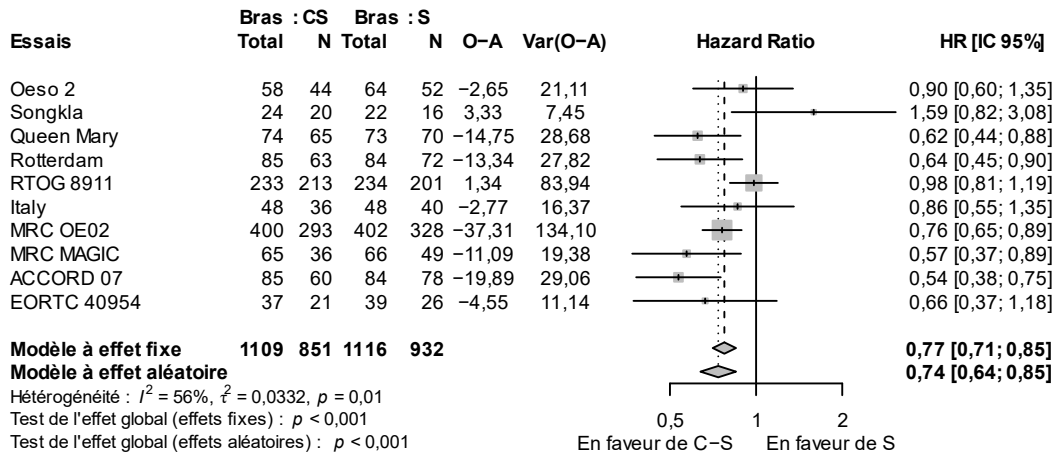


Figure 19 : Forest plot de l'effet sur la survie sans récurrence de CS contre S dans la population globale

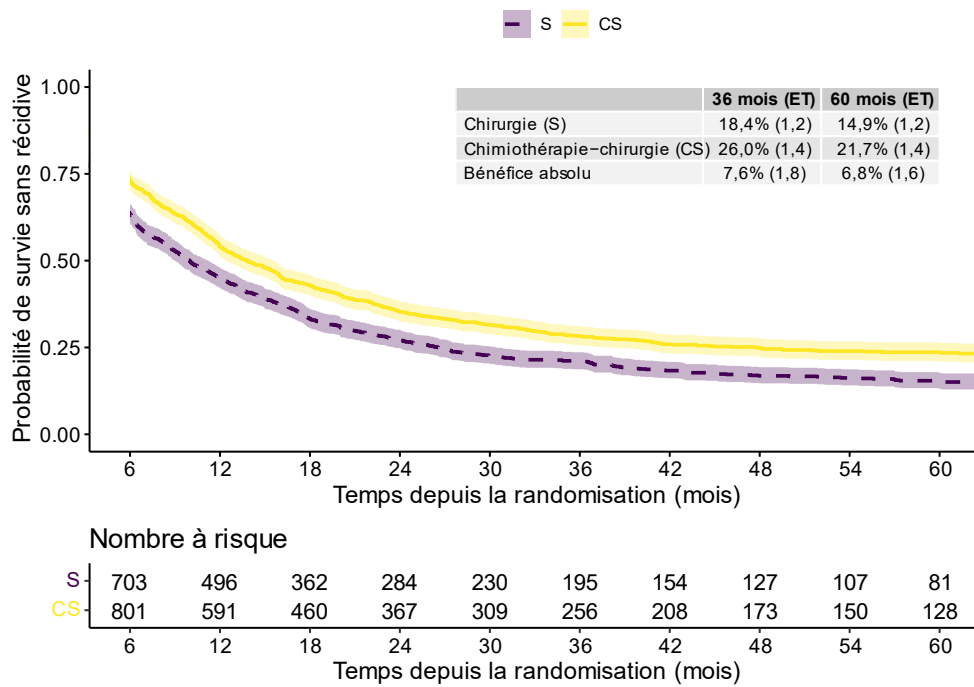


Figure 20 : Courbe de survie sans récurrence S contre CS

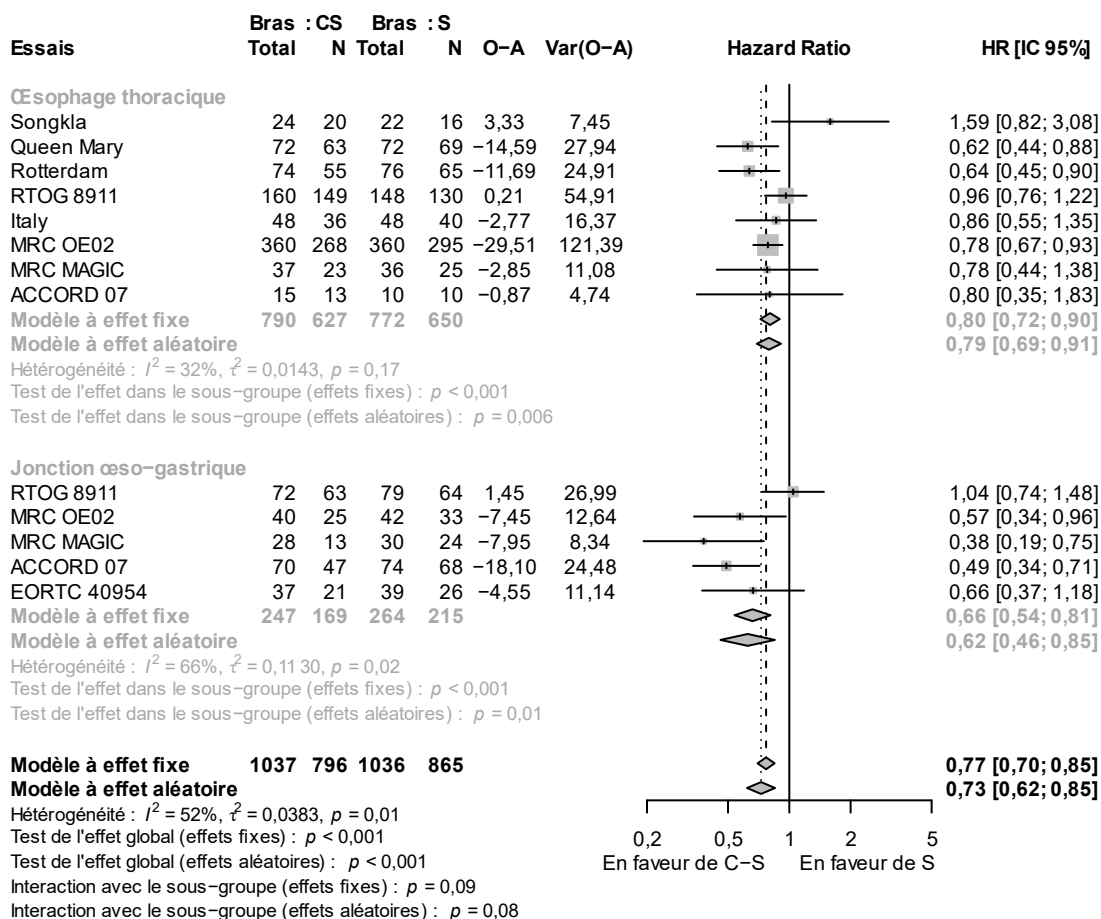
Un « landmark » de 6 mois a été appliqué

S : Chirurgie, CS : Chimiothérapie-Chirurgie, ET : écart-type

L'analyse selon la localisation tumorale suggérait une différence d'effet ( $p_{\text{interaction}} = 0,08$ ). L'ET était :  $HR_{\text{aléatoire}} = 0,79[0,69-0,91]$  pour les tumeurs situées dans l'œsophage thoracique et  $HR_{\text{aléatoire}} = 0,62[0,46-0,85]$  pour celles situées à la jonction œsogastrique. Cette différence relative correspondait à une dRMST à 5 ans de +3,8 [IC95% +0,7 ; +6,9] mois pour les lésions localisées dans l'ET contre +9,8 [IC95% +4,0 ; +15,5] mois pour celles à la JOG.

L'analyse selon les sous-groupes histologiques ne suggérait pas de différence d'effet ( $p_{\text{interaction}} = 0,31$ ). L'ET était de  $HR_{\text{aléatoire}} = 0,83[0,68-1,00]$  pour les carcinomes épidermoïdes et  $HR_{\text{aléatoire}} = 0,69[0,57-0,83]$  pour les adénocarcinomes.

Les résultats ne suggéraient aucune interaction avec l'âge en catégorielle ( $p = 0,10$ ) ou continue ( $p = 0,96$ ) ou le sexe ( $p = 0,54$ ).



**Figure 21** : Forest plot de l'effet sur la survie sans récurrence de CS contre S selon la localisation tumorale

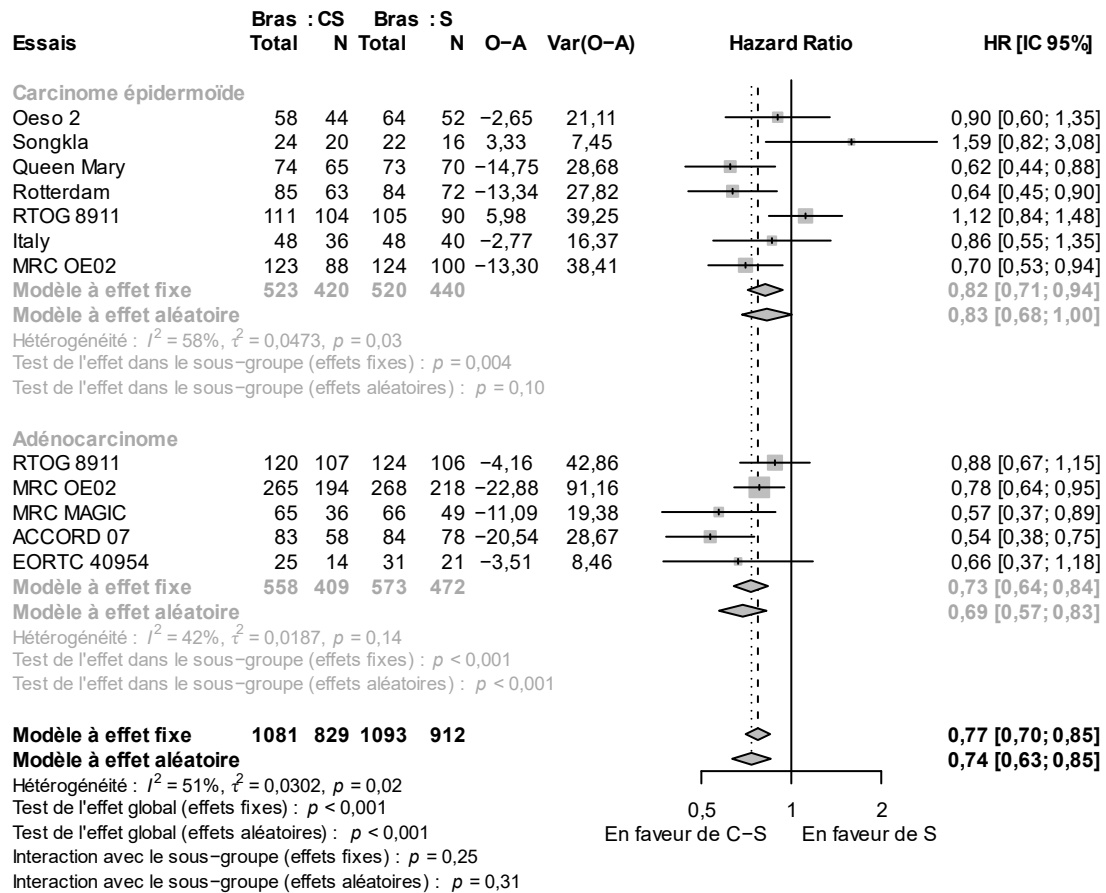


Figure 22 : Forest plot de l'effet sur la survie sans récurrence de CS contre S selon le type histologique

### Risques compétitifs / modes de récurrence

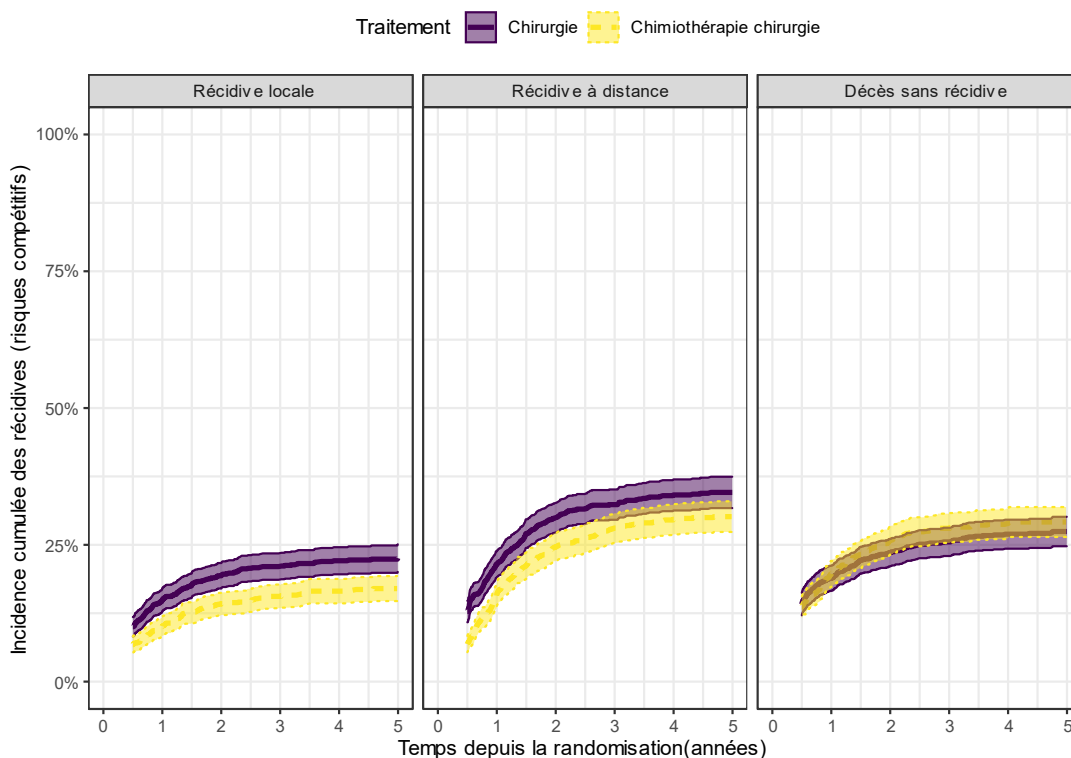
Les risques compétitifs étaient analysables dans 10 essais (2 225 patients) (Tableau annexe 2, page 159). Le mode de récurrence le plus fréquent était la récurrence à distance (ou locale et à distance synchrone) survenant chez 31,4% des patients. Dix-neuf pourcents des patients récidivaient localement.

Les récurrences locales étaient moins fréquentes dans le groupe CS avec un sHR = 0,73[0,60-0,88],  $p = 0,001$ . Cet effet protecteur ne variait pas avec la localisation tumorale ( $p = 0,87$ ) ou le type histologique ( $p = 0,97$ ).

Les récurrences à distance étaient également moins fréquentes dans le groupe CS avec un sHR = 0,82[0,71-0,95],  $p = 0,01$ . De même, cet effet protecteur ne variait pas avec la localisation tumorale ( $p = 0,55$ ) ou le type histologique ( $p = 0,22$ ).

Les décès sans récurrence n'étaient pas influencés par le traitement sHR = 1,12[0,96-1,29],  $p = 0,17$  dans la population globale. Une interaction était mise en évidence avec la localisation anatomique ( $p_{interaction} = 0,009$ ). La CS faisait varier le risque de décéder sans avoir récidivé de sHR = 1,23[1,03-1,46] pour les tumeurs situées dans l'ET contre HR = 0,68[0,46-1,02] pour celles situées à la JOG. Il existait également une interaction avec le sous type histologique ( $p_{interaction} = 0,006$ ) avec un effet délétère de la chimiothérapie pour les carcinomes épidermoïdes HR = 1,38[1,12-1,70] contre HR = 0,90[0,73-1,12] pour les adénocarcinomes.

Les courbes d'incidences cumulées sont présentées dans la **Figure 23**.



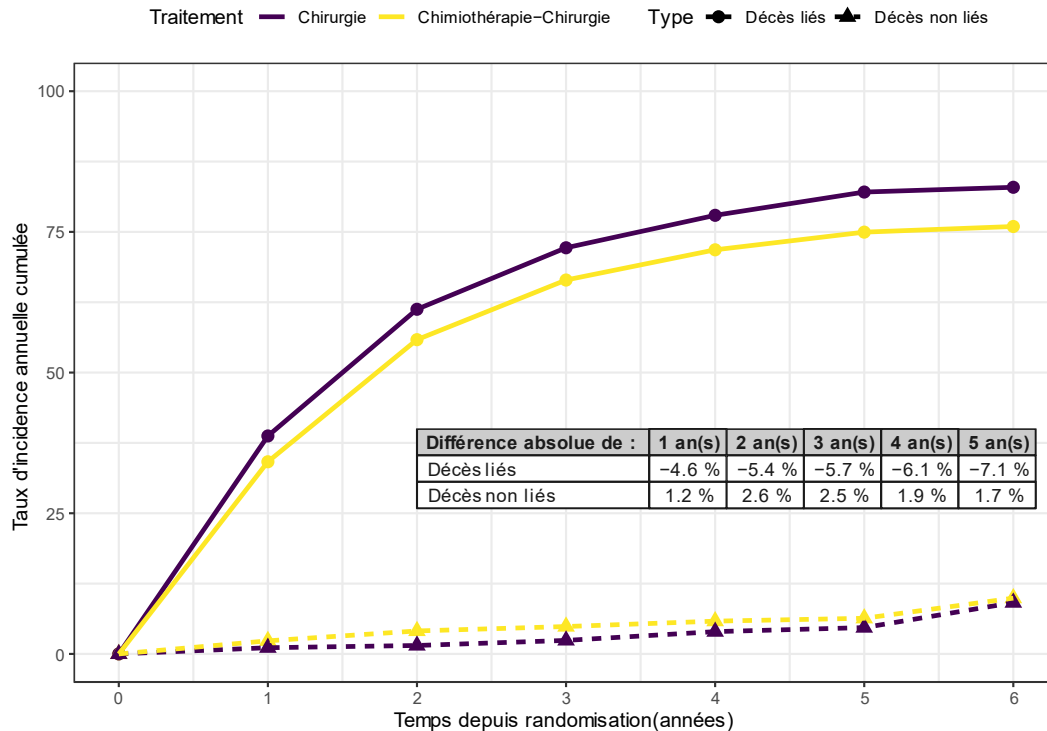
**Figure 23** : Incidences cumulées des trois risques compétitifs dans la comparaison chirurgie contre chimiothérapie chirurgie

La zone en couleur correspond à l'intervalle de confiance à 95%

### *Mortalité liée et non liée au cancer*

Les causes spécifiques de décès étaient analysables dans 10 essais (2 225 patients).

Au cours du suivi, 1 690 (76%) décès liés au cancer et 92 (4,1%) décès non liés au cancer sont survenus. L'analyse des causes spécifiques de décès montrait que la CS conférait un risque de décès lié au cancer plus faible  $HR = 0,76[0,65-0,90]$ ,  $p = 0,004$ . Il n'existait pas de différence dans la mortalité non liée au cancer  $HR = 1,45[0,76-2,76]$ ,  $p = 0,37$ . Les différences absolues de taux de mortalité à 5 ans étaient de  $-7,1 (\pm 1,8) \%$  pour les décès liés au cancer et de  $+1,7 (\pm 1,7) \%$  pour les décès non liés (**Figure 24**).



**Figure 24 :** Incidence annuelle cumulée des décès causes spécifiques dans la comparaison chirurgie contre chimiothérapie chirurgie

### Qualité de la résection

L'information sur la qualité de la chirurgie était disponible dans 13 essais (2 349 patients). (Tableau annexe 4, page 161) Le taux moyen de résection R0 était de 67,0% avec des taux variant dans les études de 27% à 87%. La chimiothérapie n'était pas associée à un meilleur taux de résection R0  $RR_{aléatoire} = 1,06[0,96-1,16]$  ;  $p = 0,26$ . L'hétérogénéité était importante ( $I^2 = 55\%$ ,  $\tau^2 = 0.0125$ ,  $p = 0.009$ ). Ce RR ne variait ni avec le type histologique ( $p = 0,61$ ), ni avec la localisation tumorale ( $p = 0,89$ ).

### Morbidité et mortalité postopératoires

La mortalité postopératoire était disponible dans 9 études (1 955 patients). (Tableau annexe 5, page 162) Le taux moyen de mortalité était de 7,0% avec des taux par essai allant de 3,1% à 13%. Aucune différence n'était observée entre les groupes S et CS :  $RR_{fixe} = 1,12[0,81-1,55]$ ,  $p = 0,50$ .

La morbidité postopératoire était disponible dans 9 études (1 900 patients). (Tableau annexe 6, page 163) Le taux moyen de morbidité grave était de 39,2% avec des taux par études allant de 4,3% à 50%. Une telle variation indique probablement un recueil très hétérogène de ce paramètre notamment dans les études réalisées avant sa standardisation. Aucune différence n'était observée entre les groupes  $RR_{fixe} = 1,05[0,90-1,22]$  ;  $p = 0,57$ .

### Toxicités

Des informations sur la toxicité étaient rarement disponibles. Quand elles l'étaient, les toxicités étaient rarement gradées. La toxicité la plus souvent rapportée était celle sur la numération sanguine (leucocytes et plaquettes), la plus rare, la toxicité cutanée. Les toxicités les plus fréquemment rencontrées étaient celles sur les neutrophiles (23,1%), les

lymphocytes (13,7%) et la mucite (13,7%). Toutes les autres toxicités survenaient chez moins de 10% des patients.

**Tableau 19 :** Toxicités observées dans les essais comparant chimiothérapie-chirurgie à chirurgie

Toxicités	Nombre d'essais (patients)*	N avec toxicité/N total dans le bras chimiothérapie	Taux
Neutrophiles	5 (909)	104/451	23,1%
Leucocytes	7 (1107)	73/533	13,7%
Mucite	6 (976)	65/476	13,7%
Nausée	6 (985)	47/487	9,7%
Tube digestif haut	2 (142)	4/72	5,6%
Plaquettes	7 (1107)	28/533	5,3%
Anémie	6 (938)	22/444	5,0%
Œsophage	3 (609)	11/305	3,6%
Diarrhée	6 (985)	17/486	3,5%
Cardiaque	5 (854)	10/426	2,3%
Créatinine	4 (387)	4/193	2,1%
Pulmonaire	4 (340)	2/159	1,3%
Cutanée	5 (518)	1/255	0,4%
Syndrome main-pied	2 (122)	0/60	0,0%

\* Nombre d'essais ayant rapporté des informations sur cette toxicités et comportant moins de 20% de données manquantes

## Analyse en sous-groupes et de sensibilité

Les analyses de sensibilité prévues par le protocole montraient des résultats identiques à l'analyse principale. Notamment l'exclusion de l'essai Songkla que l'on pouvait considérer comme « aberrant » ne permettait ni de corriger l'hétérogénéité dans l'analyse de la SG ( $I^2 = 40\%$ ,  $\tau^2 = 0,019$ ,  $p = 0,07$ ), ni celle de la SSR ( $I^2 = 54\%$ ,  $\tau^2 = 0,0317$ ,  $p = 0,03$ ). Les résultats n'étaient pas non plus modifiés par l'exclusion des deux essais ayant inclus moins de 60 patients (Songkla et MD Anderson) :  $HR_{aléatoire} = 0,81[0,70-0,93]$  pour la SG et  $HR_{aléatoire} = 0,72[0,62-0,82]$  pour la SSR.

Le protocole prévoyait également une analyse selon le type de chimiothérapie prévu dans l'essai : 5-FU/Cisplatine contre les autres protocoles. Cette analyse ne montrait pas de différence d'effet sur la survie globale ( $p_{interaction} = 0,57$ ) ou la survie sans récurrence ( $p_{interaction} = 0,45$ ).

Enfin, une analyse de sensibilité non prévue par le protocole a été réalisée. Cette analyse vérifiait que les interactions observées persistaient si on restreignait l'analyse aux essais ayant inclus des patients dans tous les sous-groupes. Quatre essais (RTOG8911, MRC

OE02, MRC MAGIC et ACCORD 07) avaient inclus à la fois des tumeurs localisées dans l'ET et la JOG. L'interaction entre l'effet traitement et la localisation tumorale sur la survie globale était comparable à celle de l'analyse principale ( $p_{\text{interaction}} = 0,05$ ), ainsi que celle sur la SSR ( $p_{\text{interaction}} = 0,05$ ). De même, deux essais avaient inclus à la fois des CE et des ADK (RTOG8911 et MRC OE02). Les interactions étaient également comparables à celles de l'analyse principale :  $p_{\text{interaction}} = 0,28$  pour la SG et  $p_{\text{interaction}} = 0,62$  pour la SSR.



## MA par paire 2 : chirurgie seule contre chimioradiothérapie suivie de chirurgie

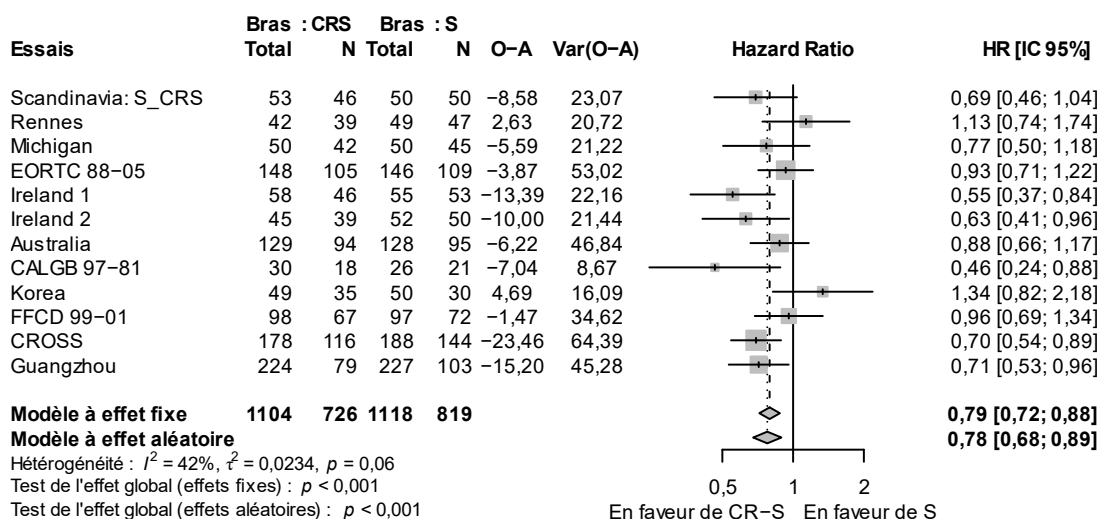
### Effet sur la survie globale

#### Effet en population générale

Les informations de SG étaient disponibles pour 2 222 patients. Au cours du suivi, 1 545 évènements de SG ont été observés. La radio chimiothérapie préopératoire était associée à une meilleure SG  $HR_{aléatoire} = 0,78 [0,68; 0,89]$ . Il existait de l'hétérogénéité ( $I^2 = 42\%$  ;  $\tau^2 = 0,023$  ;  $p = 0,06$ ) (**Figure 25**).

Ce bénéfice relatif se traduisait par un bénéfice absolu en SG de +8,3% à 3 ans passant de 38,0% à 46,3% et de +8,5% à 5 ans passant de 28,4% à 37% (**Figure 26**). La dRMST à 5 ans était de 3,9 mois ( $\pm 1,4$ ) passant de 27,0 à 31,4 mois.

Le suivi médian était de 6,4 [6,1-6,9] ans.



**Figure 25** : Forest plot de l'effet de la CRS vs S sur la survie globale dans la population générale

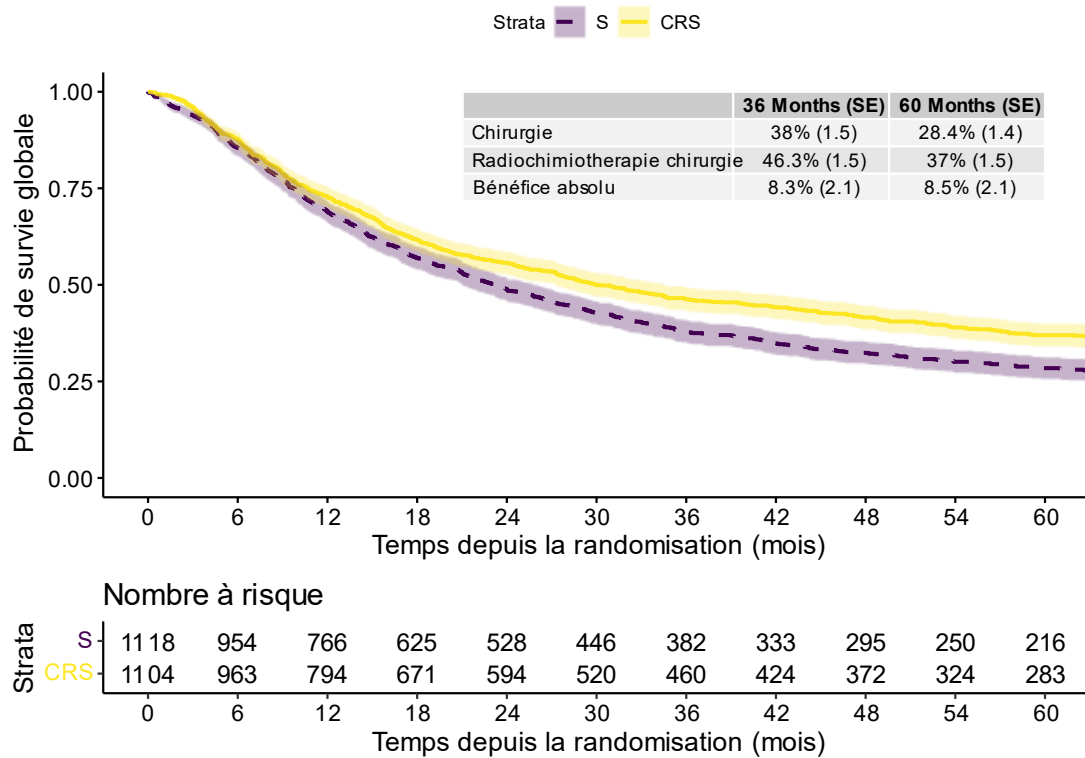


Figure 26 : Courbes de survie globale dans la comparaison CRS contre S

### Effet en fonction des sous-groupes de participants

L'analyse selon la localisation tumorale suggérait une modification de l'ET ( $p_{interaction} = 0,04$ ). L'effet de la CRS était de  $HR_{aléatoire} = 0,79[0,69-0,91]$  pour les tumeurs situées dans l'ET et de  $HR_{aléatoire} = 0,51[0,33-0,79]$  pour celles situées à la JOG (Figure 27). Il n'y avait toutefois que 108 patients avec une tumeur de la JOG.

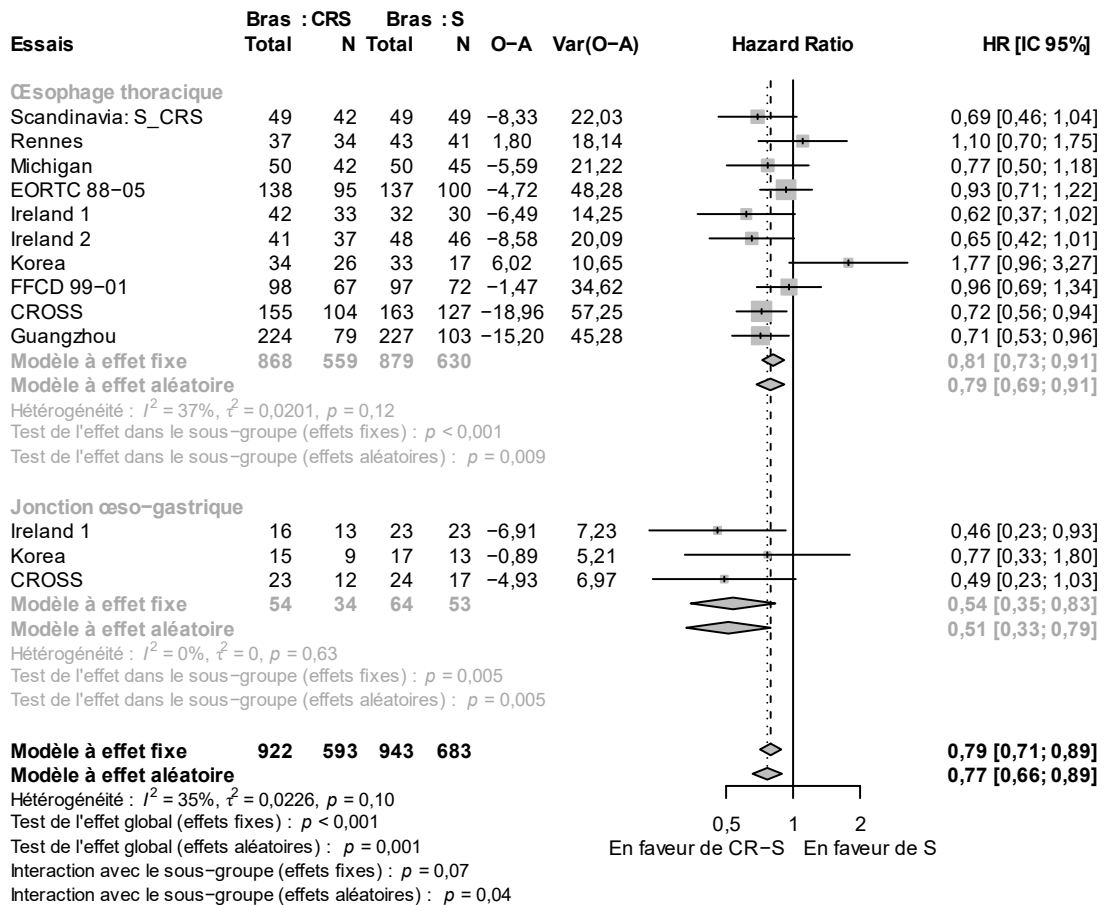
On n'observait pas non plus d'interaction avec le type histologique ( $p_{interaction} = 0,78$ ). L'ET était :  $HR_{aléatoire} = 0,79[0,67-0,94]$  pour les CE et de  $HR_{aléatoire} = 0,74[0,61-0,90]$  pour les ADK (Figure 28).

L'analyse selon le sexe suggérait également une variation de l'ET ( $p_{interaction} = 0,003$ ) avec un effet de la CRS plus marquée chez les femmes  $HR_{aléatoire} = 0,57[0,43-0,74]$  que chez les hommes  $HR_{aléatoire} = 0,86[0,73-1,01]$  (Figure annexe 1, page 164). Ceci se traduisait par une dRMST à 5 ans de 9,93 [5,95 ; 13,92] mois chez les femmes contre 2,27[-0,85 ; + 5,38] mois chez les hommes.

Quand l'âge était considéré comme une variable qualitative à quatre classes, une interaction significative était observée ( $p_{interaction} = 0,01$ ). L'effet était maximal dans la classe d'âge 65-69 ans ( $HR_{aléatoire} = 0,52[0,40-0,69]$ ) et minimal dans la classe 70 et plus  $HR_{aléatoire} = 1,00 [0,74-1,34]$ . L'effet était de  $HR_{aléatoire} = 0,85[0,73-0,99]$  chez les moins de 60 ans et de  $HR_{aléatoire} = 0,77[0,56-1,06]$  chez les 60-64 ans (Figure annexe 2, page 165). Ce résultat tranchait avec celui obtenu quand l'âge était considéré comme une variable

continue (modèle IPD en une étape), où l'on n'observait pas d'interaction ( $p_{\text{interaction}} = 0,61$ ) (Figure annexe 3, page 167).

Le T et le N du TNM préopératoire était manquant dans près de 50% des essais rendant difficile son analyse.



*Figure 27 : Forest plot de l'effet de CRS contre S sur la survie globale selon la localisation tumorale*

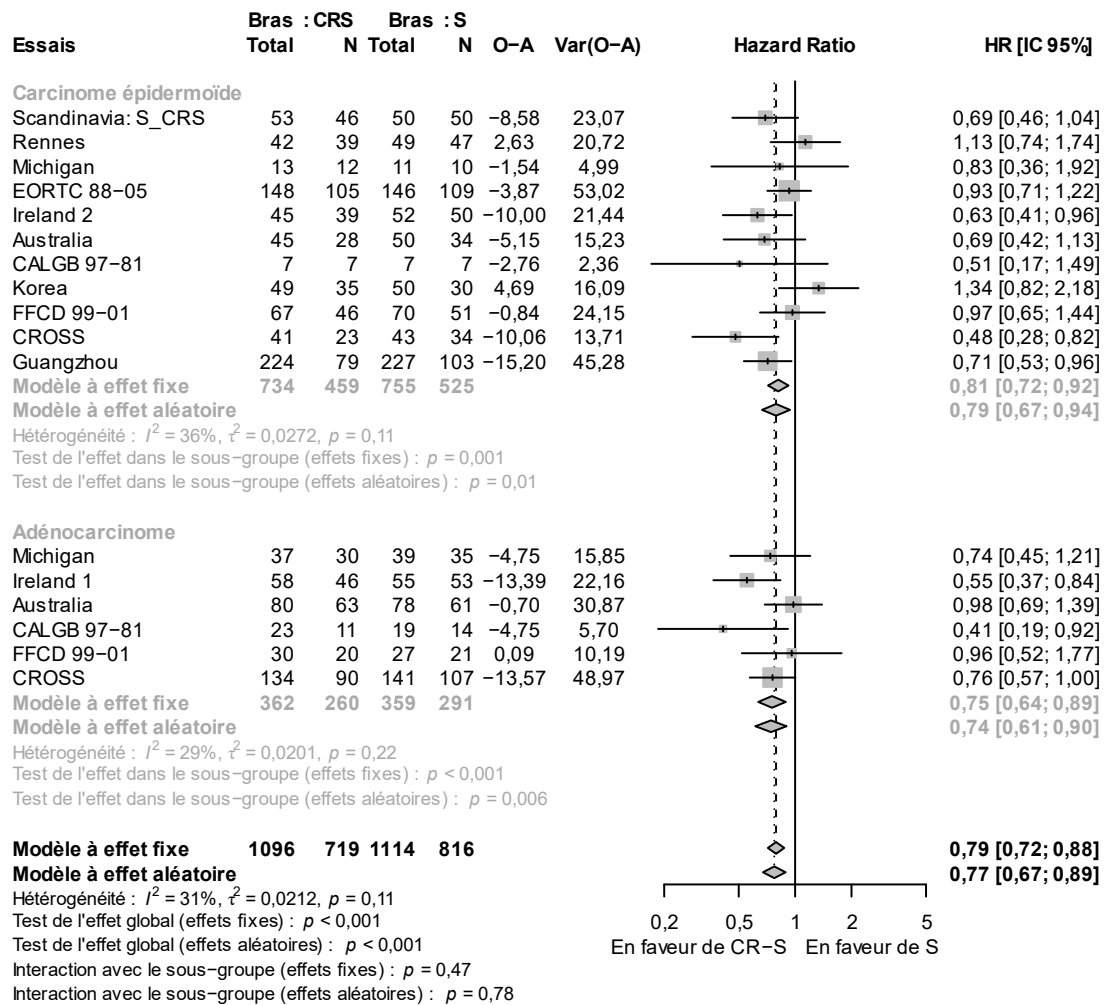


Figure 28: Forest plot de l'effet de CRS contre S sur la survie globale selon le type histologique

## Effet sur les autres critères

### Survie sans récurrence

La SSR était disponible dans neuf essais (1 909 patients). Au cours du suivi, 1 340 évènements de SSR ont été observés. La CRS était associée à une meilleure SSR  $HR_{aléatoire} = 0,81 [0,71-0,92]$ . Il existait de l'hétérogénéité ( $I^2 = 48\%$  ;  $\tau^2 = 0,03$  ;  $p = 0,05$ ) (Figure 29).

Le bénéfice absolu de SSR à 3 ans était de +8,5% passant de 34,7% à 43,1% et de +8,5% à 5 ans passant de 27,5% à 36%. La dRMST était de +4,2( $\pm 1,8$ ) mois (Figure 30).

L'analyse selon la localisation tumorale ne suggérait pas d'interaction ( $p_{interaction} = 0,42$ ). L'effet était :  $HR_{aléatoire} = 0,81 [0,70-0,93]$  pour les lésions situées dans l'ET et de  $HR_{aléatoire} = 0,62 [0,36-1,07]$  pour celles situées sur la JOG (Figure 31). On n'observait pas non plus d'arguments pour une variation de l'ET avec le type histologique ( $p_{interaction} = 0,51$ ). L'effet de la CRS était :  $HR_{aléatoire} = 0,85 [0,71-1,01]$  pour les CE et de  $HR_{aléatoire} = 0,77 [0,62-0,94]$  pour les ADK (Figure 32).

Il n'existait pas d'interaction avec l'âge considéré en continu ( $p_{interaction} = 0,32$ ) ou en catégoriel ( $p = 0,11$ ). Comme pour l'OS, l'effet de la CRS semblait varier en fonction du

sexe ( $p_{interaction} = 0,03$ ). L'ET était plus marqué chez les femmes  $HR_{aléatoire} = 0,57[0,42-0,79]$  que chez les hommes  $HR_{aléatoire} = 0,85 [0,73-0,98]$ . Cette différence correspondait à une dRMST à 5 ans de 15,75 [9,08 ; 22,42] mois chez les femmes contre 2,26 [-2,5 ; 7,03] mois chez les hommes.

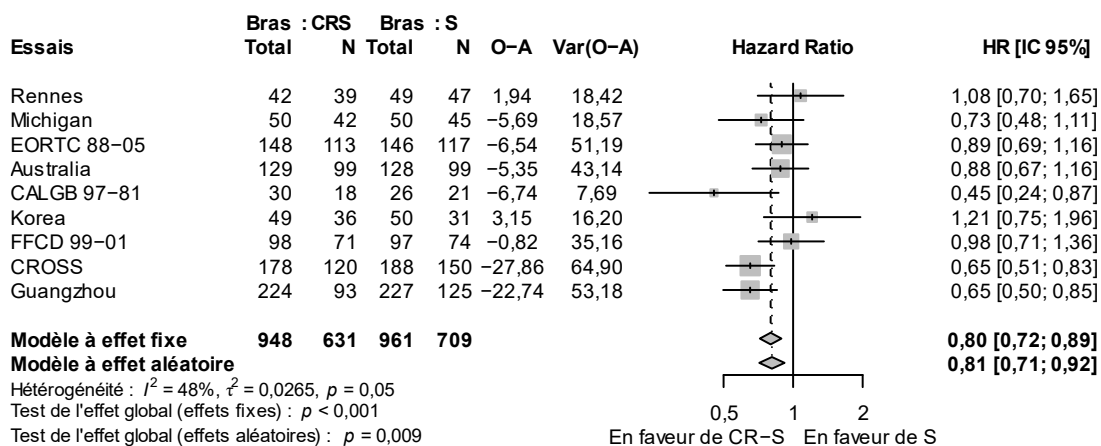


Figure 29 : Forest plot de l'effet de CRS contre S sur la survie sans récurrence dans la population globale

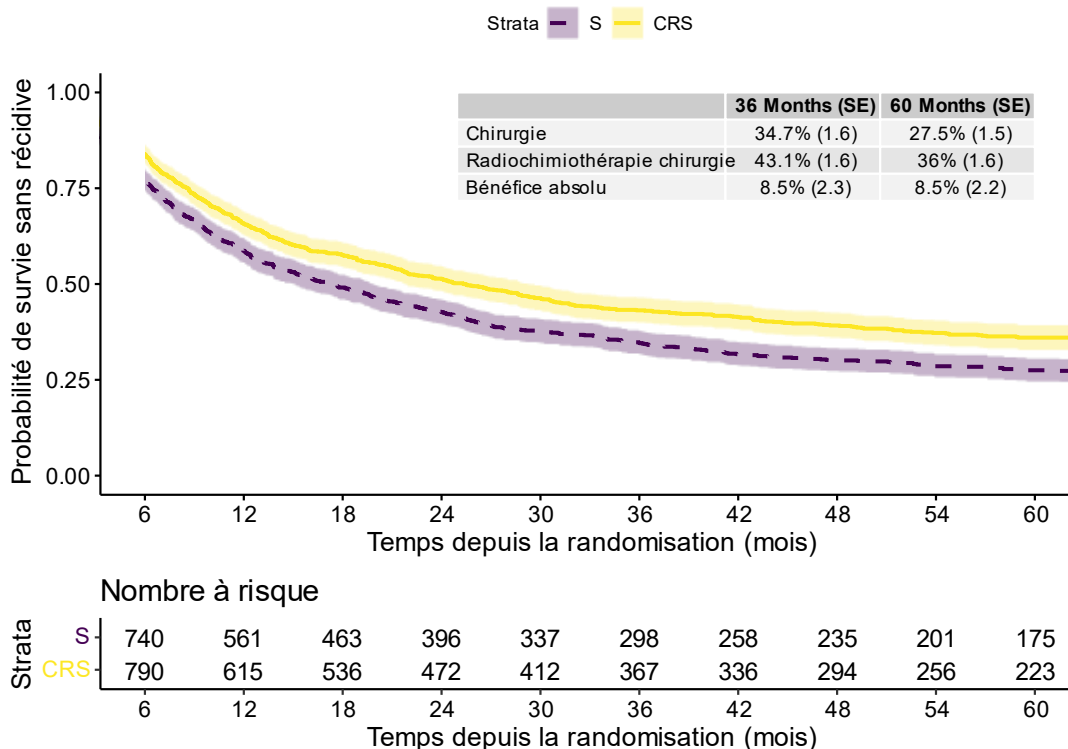


Figure 30 : Courbes de survies sans récurrence dans la comparaison CRS contre S

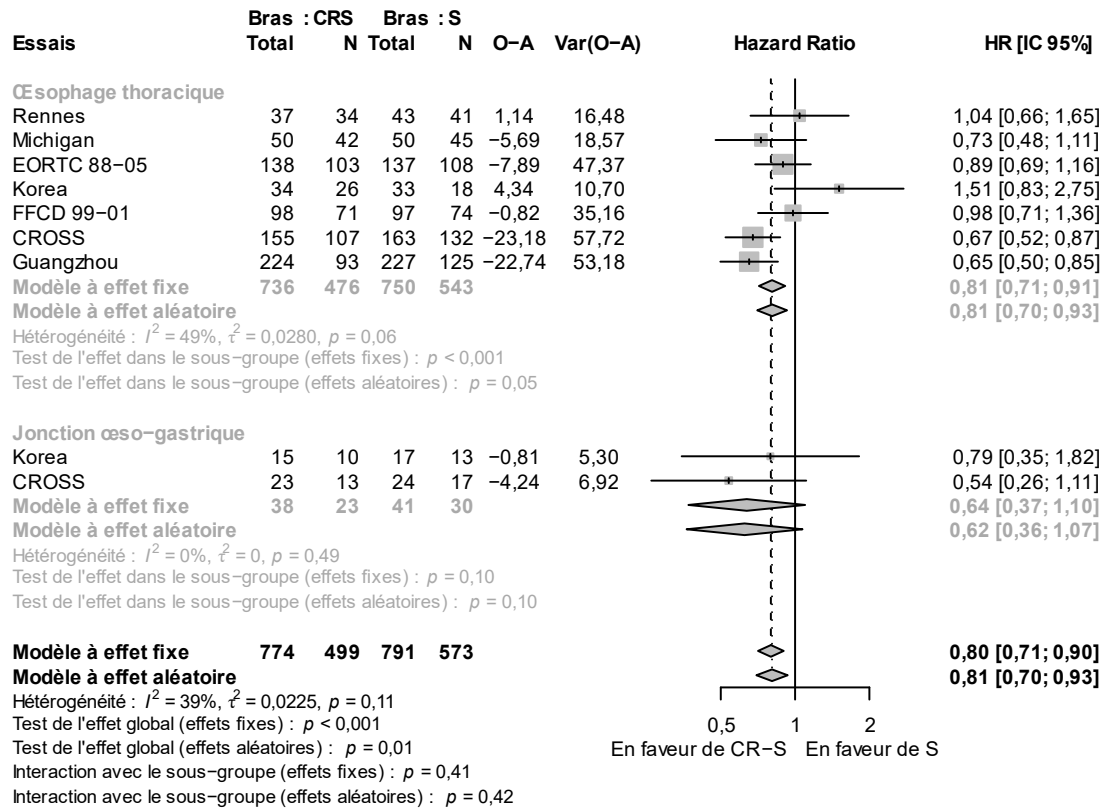


Figure 31 : Forest plot de l'effet de CRS contre S sur la survie sans récurrence selon la localisation tumorale

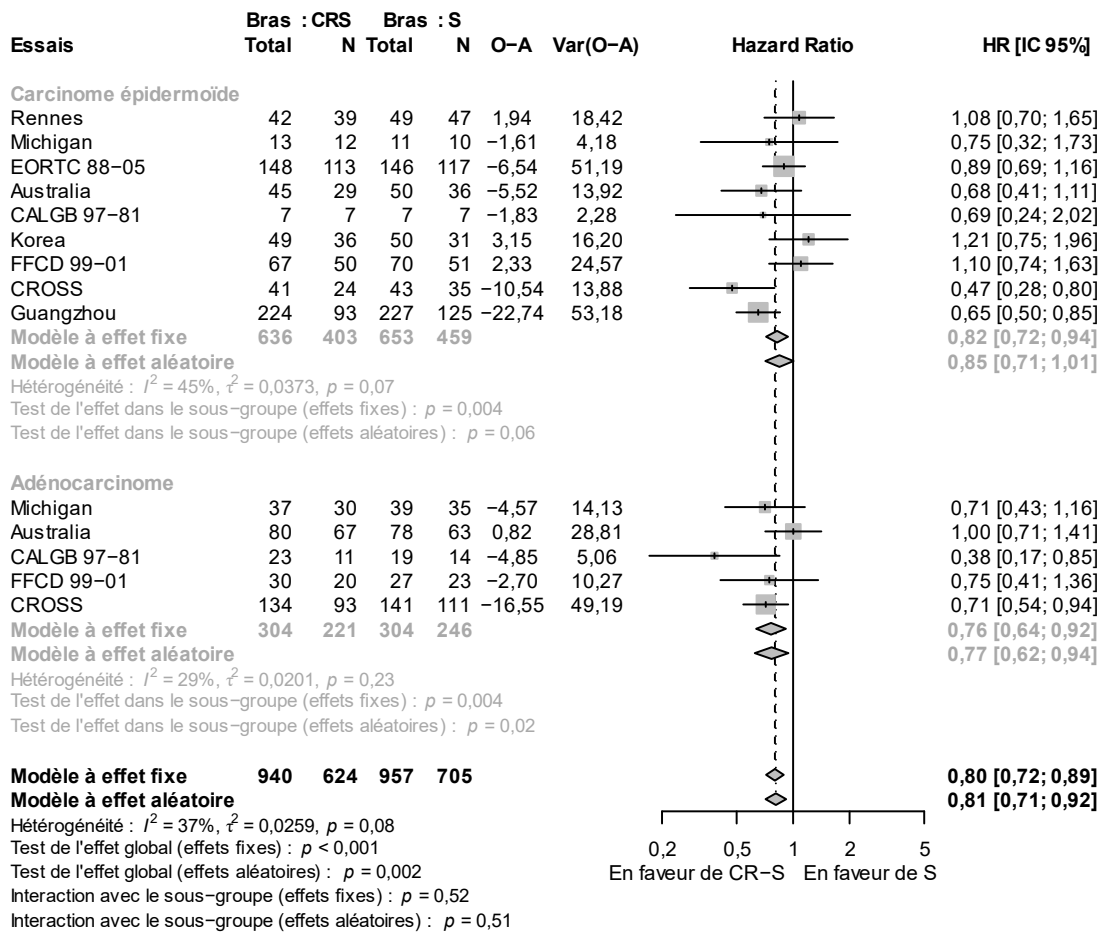


Figure 32 : Forest plot de l'effet de CRS contre S sur la survie sans récurrence selon le type histologique

### Risques compétitifs / modes de récurrence

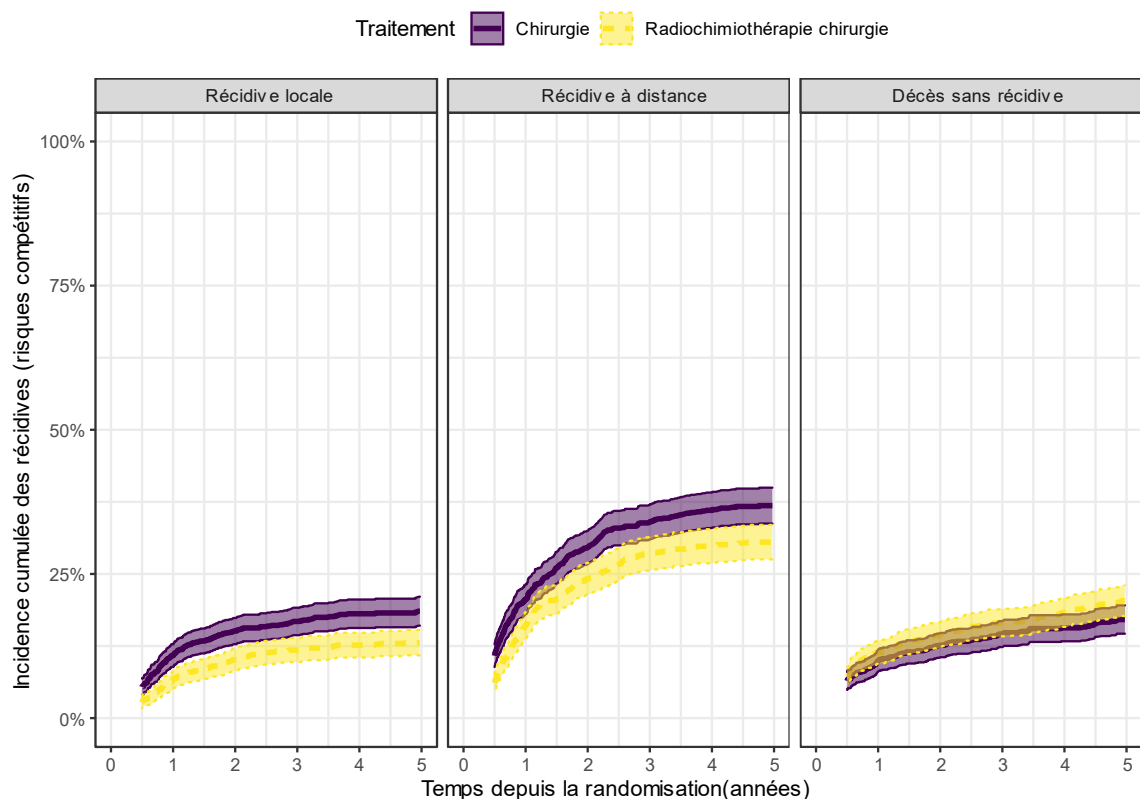
Le site de récurrence était disponible dans neuf essais (1 909 patients). (Tableau annexe 2, page 159)

Le mode de récurrence le plus fréquent était la récurrence à distance (ou locale et à distance synchrone) survenant chez 34% des patients. Seize pourcents des patients avaient une récurrence locale. La **Figure 33** montre l'incidence cumulée des trois risques compétitifs au cours du temps selon le traitement.

La CRS diminuait le risque de récurrence locale sHR = 0,67[0,54-0,84] ( $p = 0,001$ ) avec un effet équivalent dans les deux localisations anatomiques ( $p_{\text{interaction}} = 0,45$ ) et les deux histologies ( $p_{\text{interaction}} = 0,82$ ).

La CRS diminuait également le risque de récurrence à distance sHR = 0,80[0,68-0,93] ( $p = 0,004$ ) avec un effet équivalent dans les deux localisations anatomiques ( $p = 0,70_{\text{interaction}}$ ) et les deux histologies ( $p = 0,85_{\text{interaction}}$ ).

Au cours du suivi, 397 décès sans récurrence ont été observés, dont 215 dans le groupe CRS et 182 dans le groupe S. Un effet délétère de la CRS sur le décès sans récurrence était observé sHR = 1,24[1,02-1,51] ( $p = 0,03$ ). Cet effet ne variait pas avec la localisation anatomique ( $p_{\text{interaction}} = 0,85$ ) ou le type histologique ( $p_{\text{interaction}} = 0,61$ ).



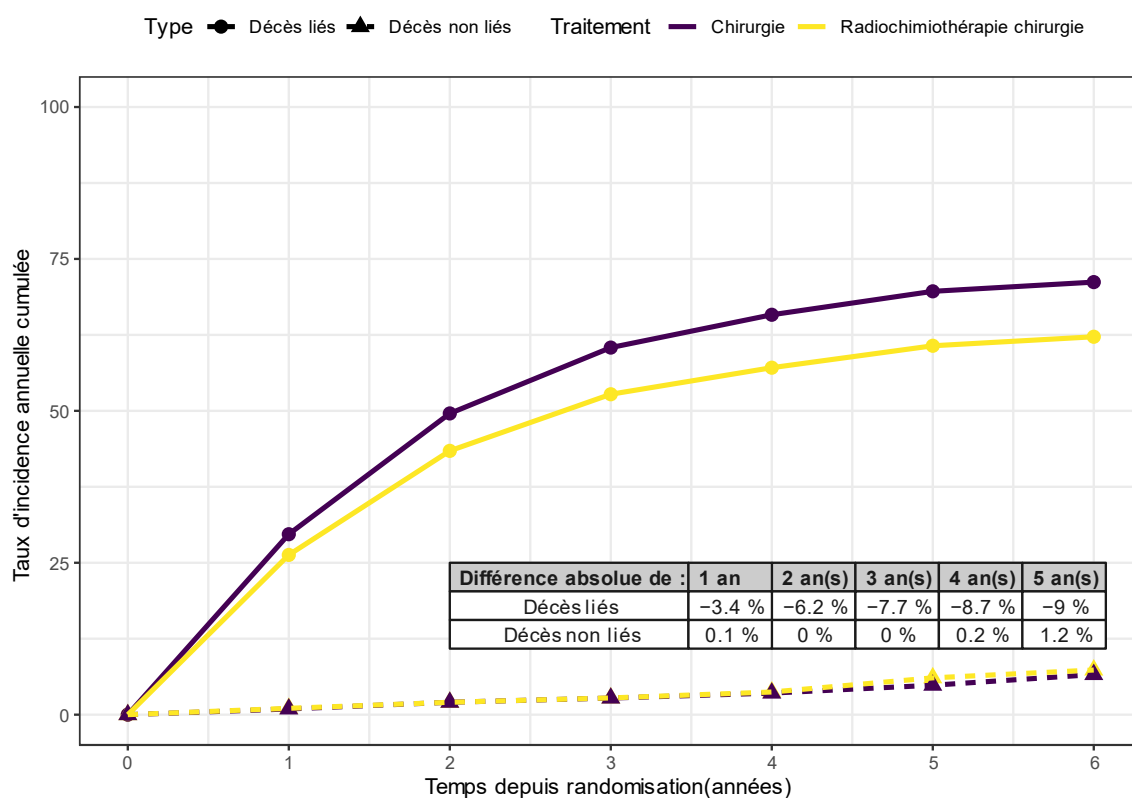
**Figure 33** : Incidence cumulée des trois risques compétitifs dans la comparaison CRS contre S



### Mortalité liée et non liée au cancer

Les causes de décès étaient disponibles dans neuf essais (1 909 patients). La plupart des patients étaient décédés d'une cause considérée comme liée au cancer (62%)

La **Figure 34** montre l'incidence cumulée des décès liés et non liés au cancer selon le traitement reçu. La CRS était responsable d'une diminution des décès liés au cancer  $HR_{\text{fixe}} = 0,83[0,74-0,94]$ . Cette différence relative correspondait à un bénéfice absolu de -9% à 5 ans. Cet effet ne variait pas avec la localisation ( $p = 0,14$ ) ou le type histologique ( $p = 0,64$ ). Il n'y avait pas de différence de décès non liés au cancer  $HR_{\text{fixe}} = 1,28[0,66-2,47]$ .



**Figure 34** : Incidence cumulée des décès spécifiques dans la comparaison CRS contre S

### Qualité de la résection

La qualité de la résection chirurgicale était disponible dans 10 essais (1975 patients) si l'on excluait les deux essais Irlandais qui avaient 94% de données manquantes. (Tableau annexe 4, page 161). Le taux moyen global de résections à visée curatrice était de 76% et s'échelonnait de 43% à 86% selon les essais. La chimioradiothérapie n'améliorait pas ce taux  $RR_{\text{aléatoire}} = 1,04[0,92-1,18]$ . L'analyse en sous-groupe n'apportait pas d'argument pour une variation de l'effet avec la localisation tumorale ( $p_{\text{interaction}} = 0,80$ ) ou le type histologique ( $p_{\text{interaction}} = 0,86$ ).

### *Morbidité et mortalité postopératoire*

La mortalité postopératoire était disponible dans neuf essais (1 724 patients). (Tableau annexe 5, page 162) Le taux global de mortalité postopératoire était de 4,0% et variait de 0,7% à 7,7% selon l'essai. La CRS n'augmentait pas le risque de mortalité postopératoire  $RR_{\text{fixe}} = 1,34[0,80-2,23]$ .

La morbidité postopératoire était décrite dans six essais (1410 patients). (Tableau annexe 6, page 163) Le taux global de morbidité était de 50,1%, variant de 22% à 73% selon l'essai. La CRS n'augmentait pas le risque de morbidité postopératoire  $RR_{\text{fixe}} = 1,06[0,97-1,16]$ .

### *Toxicités*

Des informations sur les toxicités étaient rapportées de façon variable selon les essais. Certaines toxicités étaient rapportées par sept essais et d'autres par un seul. La toxicité la plus fréquemment observée était celle sur les neutrophiles (26,2%), suivie de celle sur les leucocytes (22,4%). La troisième toxicité la plus fréquente était la toxicité œsophagienne 8,6%. Toutes les autres toxicités étaient observées chez moins de 5% des patients (**Tableau 20**).

*Tableau 20 : Toxicités dans la comparaison CRS contre S*

Type de toxicité	Nombre d'essai*	Patients inclus dans ces essais*	N avec toxicités	N dans le bras RCS	Pourcentage
Neutrophiles	5	1203	150	573	26,2
Leucocytes	7	1596	171	762	22,4
Œsophage	3	556	22	257	8,6
Plaquettes	7	1596	31	761	4,1
Anémie	4	837	16	396	4,0
Nausées	5	1397	23	664	3,5
Mucite	3	737	4	348	1,1
Diarrhée	5	1397	3	664	0,5
Cardiaque	4	752	1	338	0,3
Créatinine	3	836	1	406	0,2
Syndrome Main-Pieds	3	557	0	266	0,0
Autres	2	286	0	130	0,0
Pulmonaire	2	457	0	209	0,0
Cutanée	2	457	0	209	0,0
Tube digestif supérieur	1	91	0	42	0,0

\* Les essais avec moins de 20% de données non-manquantes étaient exclus

## Analyse en sous-groupes et de sensibilité

Les analyses de sensibilité ne montraient pas de résultats en discordance avec ceux de l'analyse principale.

Seul l'essai CALGB 97-81 pourrait être considéré comme présentant un résultat aberrant. Après exclusion de cet essai, le résultat observé dans l'analyse principale n'était pas modifié aussi bien pour la SG ( $HR_{\text{aléatoire}} = 0,79[0,69-0,91]$ ) que pour la SSR ( $HR_{\text{aléatoire}} = 0,81[0,71-0,94]$ ). L'hétérogénéité baissait légèrement en SG ( $I^2 = 38\%$ ,  $\tau^2 = 0,02$ ,  $p = 0,1$ ) et en SSR ( $I^2 = 38\%$ ,  $\tau^2 = 0,02$ ,  $p = 0,14$ ). Cet essai était par ailleurs le seul essai à avoir inclus moins de 60 patients.

La plupart des essais avaient utilisé une chimiothérapie à base de 5-FU Cisplatine. Seuls quatre (Scandinavia, EORTC 88-05, CROSS et Guangzhou) avaient utilisé un autre protocole. Le type de chimiothérapie ne semblait pas faire varier l'ET sur la SG ( $p_{\text{interaction}} = 0,47$ ). Les résultats étaient moins clairs sur la SSR ( $p_{\text{interaction}} = 0,06$ ). Les chimiothérapies non basées sur le 5FU Cisplatine semblaient plus efficaces  $HR_{\text{aléatoire}} = 0,81[0,71-0,92]$  que celles utilisant ce protocole  $HR_{\text{aléatoire}} = 0,89[0,76-1,05]$ . Ces résultats doivent être interprétés avec prudence car le groupe d'essais ayant utilisé un autre protocole est principalement constitué des deux essais les plus récents. Des modifications d'autres paramètres (comme les techniques d'irradiation) pourraient également expliquer cette différence.

L'interactions suggérée avec le sexe se produisait pour une variable dont toutes les modalités étaient présentes dans tous les essais.

## MA par paire 3 : chimiothérapie suivie de chirurgie contre chimioradiothérapie suivie de chirurgie

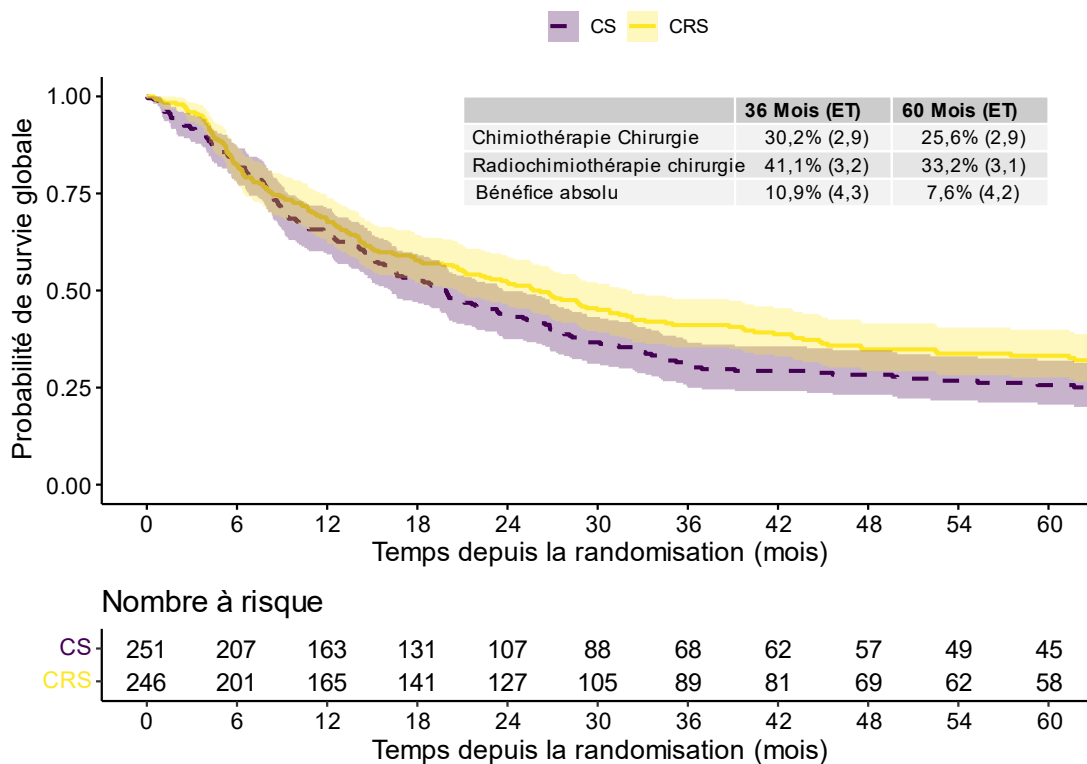
### Effet sur la survie globale

#### Effet en population générale

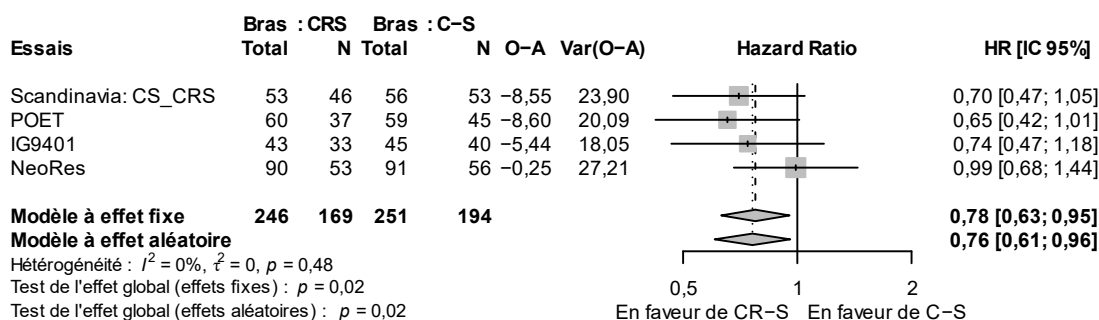
Les données de SG étaient disponibles pour 497 patients. Au cours du suivi, 363 décès sont survenus. La chimioradiothérapie était associée à une meilleure survie globale que la chimiothérapie  $HR_{\text{fixe}} = 0,78[0,63-0,95]$ ,  $p = 0,02$  (**Figure 36**). L'hétérogénéité était faible ( $I^2 = 0\%$ ,  $\tau^2 = 0$ ,  $p = 0,48$ ).

Le bénéfice absolu de SG à 3 ans était de +10,9% passant de 30,2% à 41,1%, et de +7,6% ( $\pm 4,2$ ) à 5 ans passant de 25,6% à 33,2%. La d RMST à 5 ans était de 4,2( $\pm 1,9$ ) mois passant de 23,4 à 30,5 mois (**Figure 35**).

Le suivi médian était de 8,0 [6,2-9,6] ans.



**Figure 35** : Courbes de survie globale de CS contre CRS



**Figure 36** : Forest plot de l'effet de CS contre CRS sur la survie globale dans la population globale

### Effet en fonction des sous-groupes de participants

Dans cette troisième MA, la puissance était limitée. Les analyses en sous-groupes doivent être interprétées avec prudence.

L'analyse de l'ET selon la localisation ne montrait pas d'argument pour une interaction significative ( $p_{\text{interaction}} = 0,47$ ). L'ET était de  $HR_{\text{fixe}} = 0,75[0,60-0,93]$  pour les tumeurs situées dans l'ET et de  $HR_{\text{fixe}} = 1,00[0,52-1,94]$  pour celles situées à la JOG (**Figure 37**).

L'effet ne semblait pas varier significativement non plus avec le type histologique ( $p_{\text{interaction}} = 0,35$ ). Il était de  $HR_{\text{fixe}} = 0,69[0,49-0,97]$  pour les CE et de  $HR_{\text{fixe}} = 0,83[0,64-1,08]$  pour les ADK (**Figure 38**).

L'ET ne semblait pas varier avec l'âge, aussi bien en continu ( $p_{\text{interaction}} = 0,27$ ) qu'en qualitatif ( $p_{\text{interaction}} = 0,12$ ). L'ET variait avec le sexe ( $p_{\text{interaction}} = 0,005$ ). L'effet était plus marqué chez les femmes  $HR_{\text{fixe}} = 0,45[0,25-0,79]$  que chez les hommes  $HR_{\text{fixe}} = 0,89[0,71-1,11]$  (Figure annexe 5, page 168). Cette différence correspondait à une dRMST de 13,75 [4,85 ; 22,65] mois chez les femmes contre 1,31 [-2,42 ; 5,05] mois chez les hommes.

## Effet sur les autres critères

### Survie sans récurrence

Des données de SSR étaient disponibles dans trois essais (388 patients).

Au cours du suivi 272 événements de SSP ont été observés.

La CRS était associée à une meilleure SSP que la CS  $HR_{\text{fixe}} = 0,77[0,60-0,98]$ . Il n'y avait pas d'hétérogénéité (**Figure 39**).

Le bénéfice en SSR à 3 ans était de +9,6% passant de 33,7% à 43,7% et de +9,2 ( $\pm 4,9$ ) % à 5 ans, passant de 27,8% à 37,0%. La dRMST à 5 ans était de +4,0 ( $\pm 2,4$ ) mois passant de 27,7 à 31,9 mois (**Figure 40**).

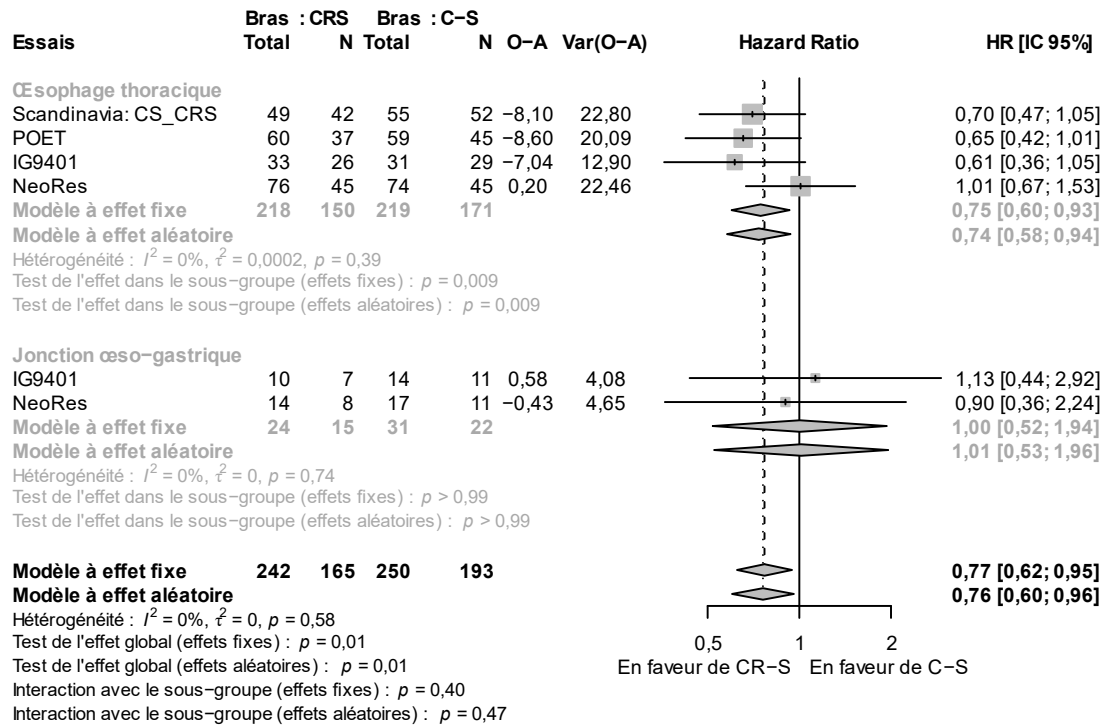


Figure 37 : Forest plot de l'effet de CS contre CRS sur la survie globale en fonction de la localisation tumorale

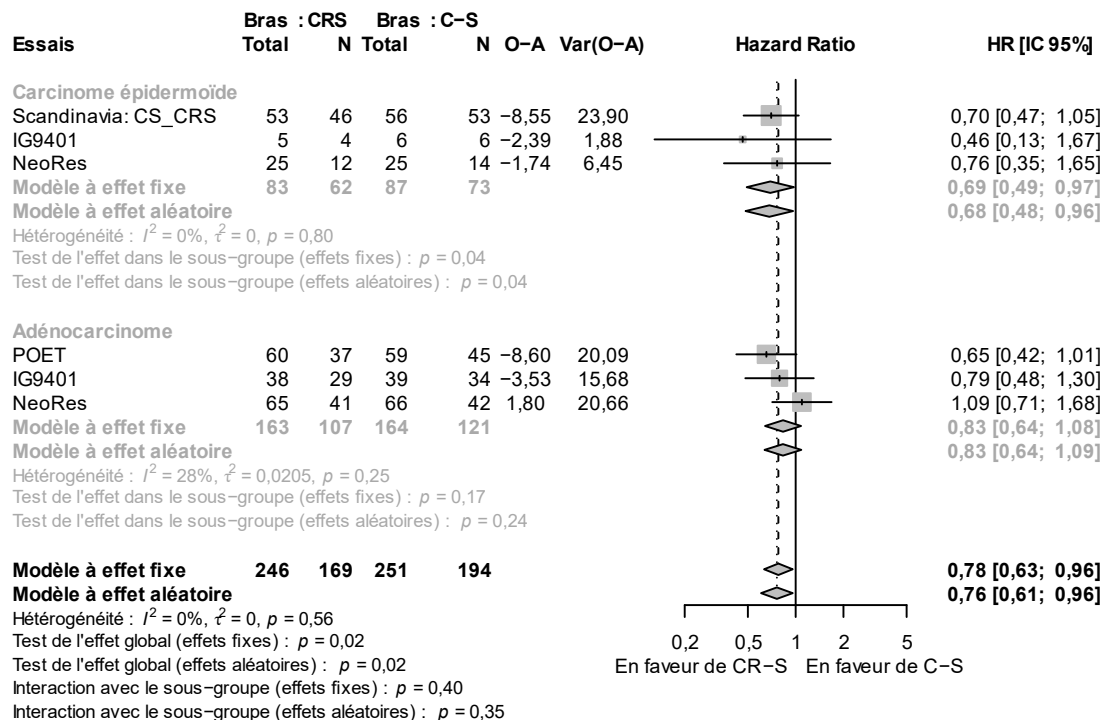


Figure 38 : Forest plot de l'effet de CS contre CRS sur la survie globale selon le type histologique

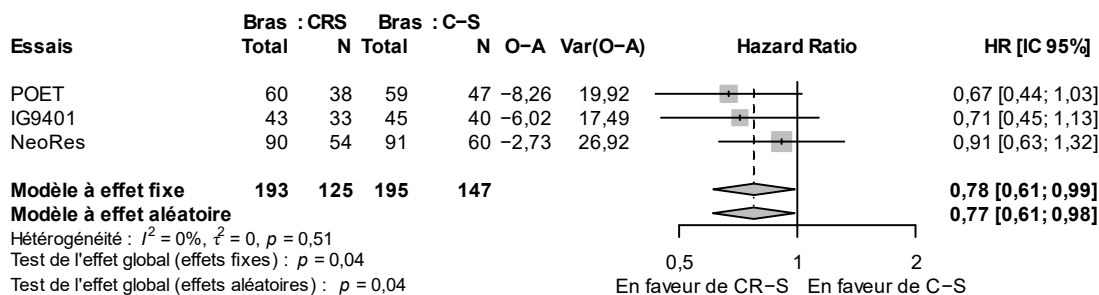


Figure 39 : Forest plot de l'effet de CS contre CRS sur la survie sans récurrence dans la population globale

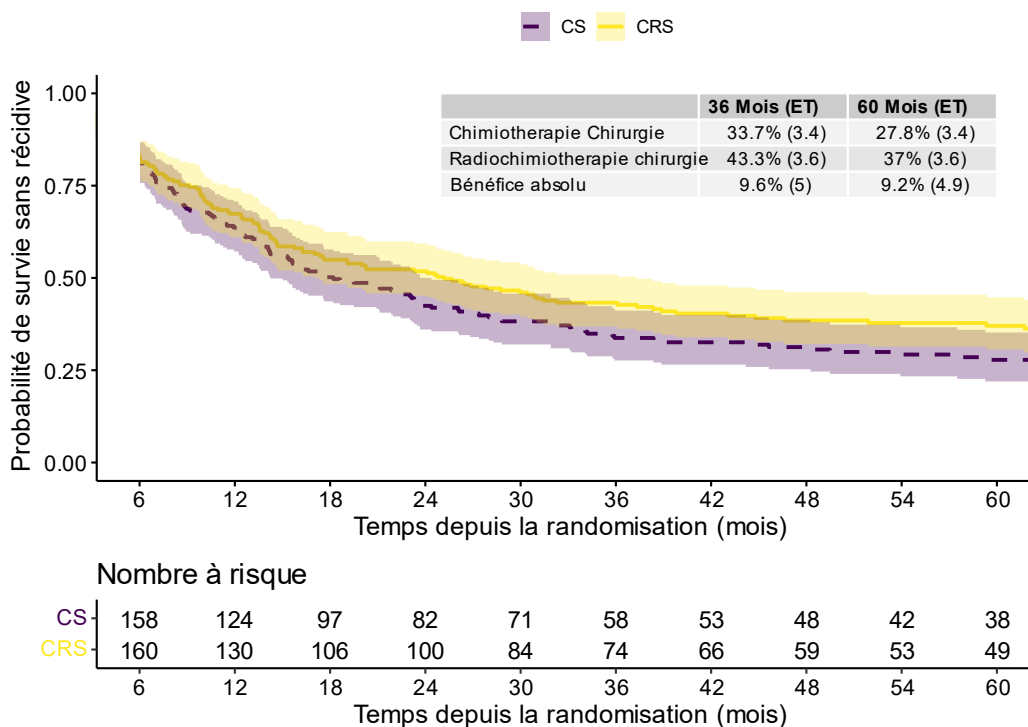


Figure 40 : Courbes de survie sans récurrence dans la comparaison CS contre CRS

L'analyse en sous-groupe ne suggérait pas de variation de l'ET avec la localisation tumorale ( $p_{interaction} = 0,63$ ) (Figure 41) ou le type histologique ( $p_{interaction} = 0,50$ ) (Figure 42).

L'effet ne semblait pas varier avec l'âge en catégorie ( $p_{interaction}=0,76$ ) ou l'âge en continu ( $p_{interaction}=0,66$ ). L'analyse de l'ET selon le sexe montrait une différence moins évidente que pour la SG ( $p_{interaction}=0,13$ ). L'ET traitement semblait légèrement plus grand pour les femmes  $HR_{fixe}=0,45[0,21-0,97]$  que pour les hommes  $HR_{fixe}=0,84[0,65-1,09]$ . Toutefois cette comparaison reposait sur un nombre faible d'évènements particulièrement chez les femmes (33 évènements).

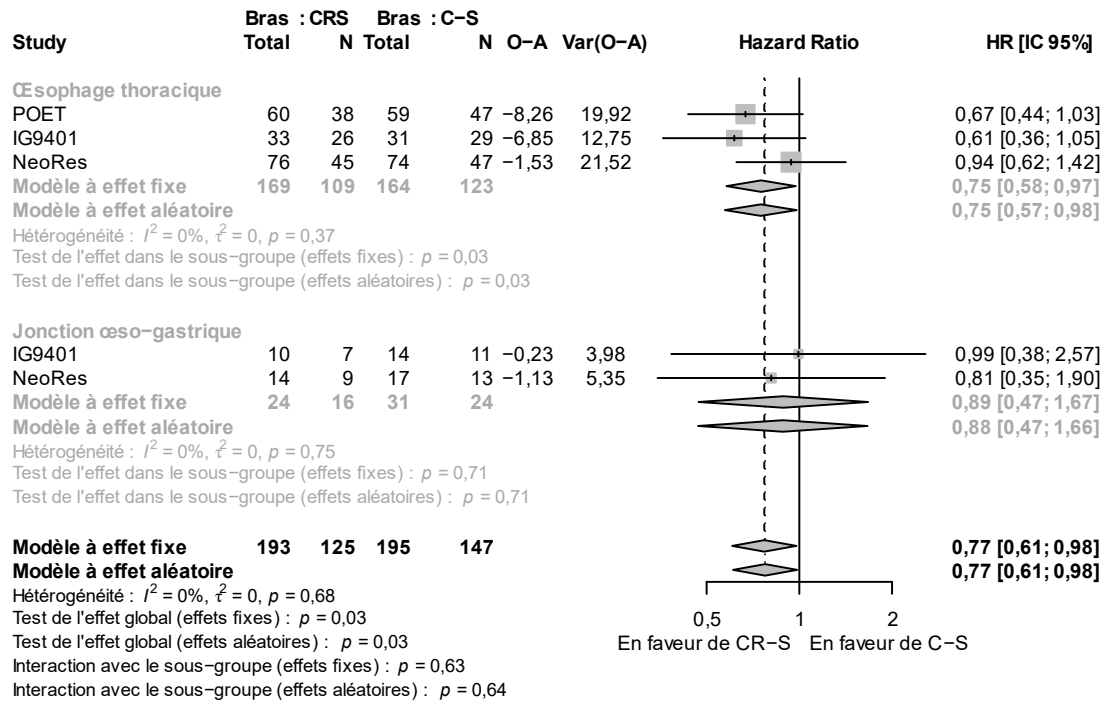


Figure 41 : Forest plot de l'effet de CS contre CRS sur la survie sans récurrence selon la localisation tumorale

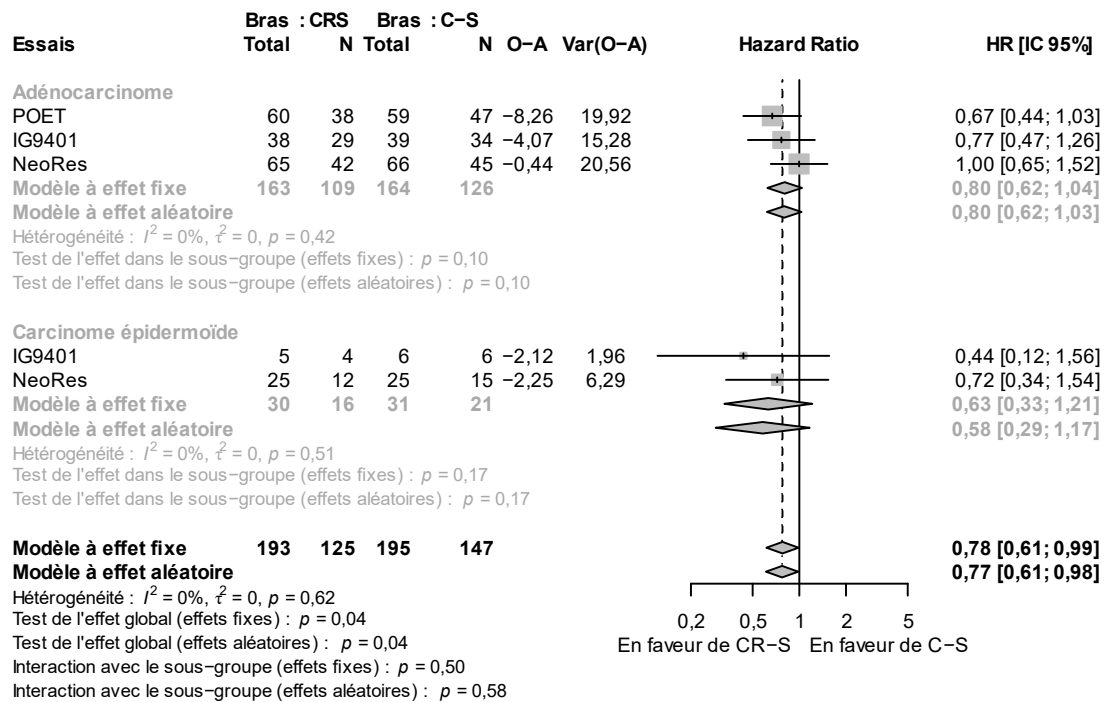


Figure 42 : Forest plot de l'effet de CS contre CRS sur la survie sans récurrence selon le type histologique

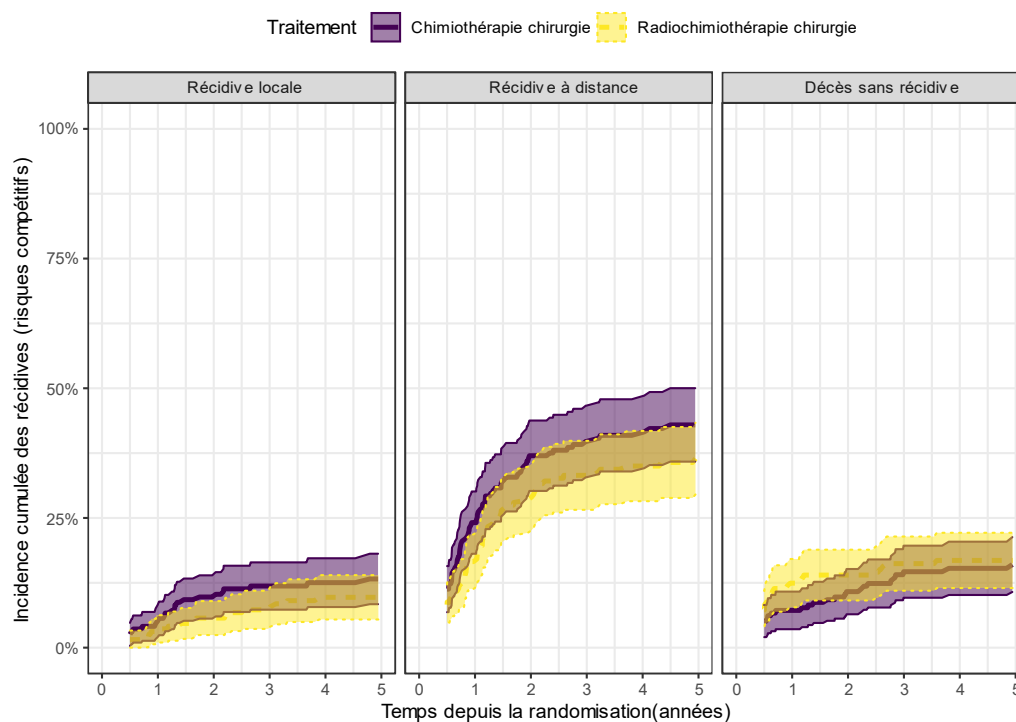


### *Risques compétitifs/ modes de récurrence*

Des informations détaillées sur les récurrences étaient disponibles dans 3 essais (388 patients). (Tableau annexe 2, page 159)

Le principal site de récurrence était à distance ou locale et à distance synchrone (40%). Onze pourcents des patients avaient une récurrence locale. La CRS n'était pas significativement associée à moins de récurrence locale (sHR = 0,74[0,41-1,35],  $p = 0,34$ ), récurrence à distance (sHR = 0,79[0,58-1,08],  $p = 0,14$ ) ou décès sans récurrence (sHR = 0,94[0,61-1,47],  $p = 0,82$ ) que la CS. Il n'y avait d'hétérogénéité dans aucune des trois comparaisons ( $p=0,48$ ,  $p = 0,86$  et  $p = 0,61$ ). La **Figure 43** montre l'incidence cumulée de ces trois événements au cours du temps selon le traitement reçu.

La localisation tumorale n'était pas en interaction avec le traitement pour le risque de récurrence locale ( $p_{interaction} = 0,65$ ), le risque de récurrence à distance ( $p_{interaction} = 0,68$ ) ou le décès sans récurrence ( $p_{interaction} = 0,68$ ). Le type histologique ne l'était pas non plus avec un de ces trois événements ( $p_{interaction} = 0,93$ ,  $p_{interaction} = 0,98$  et  $p_{interaction} = 0,56$ ).



**Figure 43** : Incidences cumulées des trois risques compétitifs dans la comparaison CS contre CRS

### Mortalité liée et non liée au cancer

Des informations détaillées sur les causes de décès étaient disponibles dans 3 essais (388 patients).

La plupart des patients décédaient de cause liée au cancer (69%). La CRS semblait être associée à moins de décès liés  $HR_{\text{fixe}} = 0,78[0,61-1,01]$  à la limite de la significativité  $p = 0,06$ . Ce bénéfice ne variait pas avec la localisation tumorale ( $p_{\text{interaction}} = 0,14$ ) ou le type histologique ( $p_{\text{interaction}} = 0,31$ ). La CRS n'était pas associée à un risque différent de décès non liés  $HR_{\text{fixe}} = 0,89[0,21-3,81]$ ,  $p = 0,82$  mais cette analyse ne reposait que sur 25 évènements.

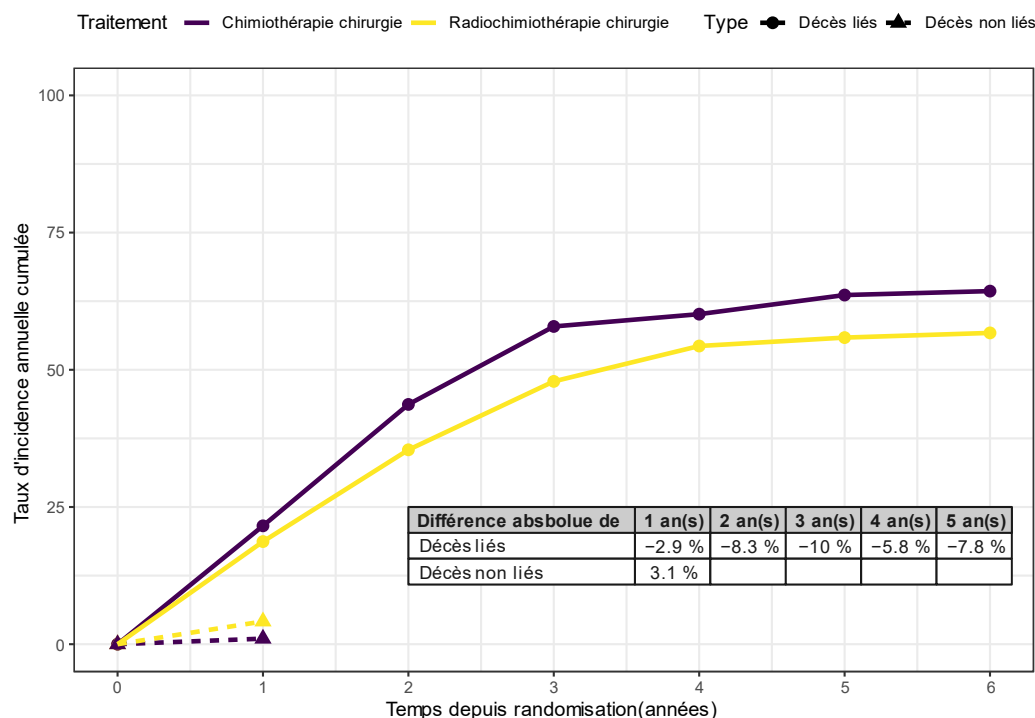


Figure 44 : Incidence des causes spécifiques de décès dans la comparaison CS contre CRS

### Qualité de la résection

Des informations sur la qualité de la résection étaient disponibles dans tous les quatre essais (493 patients). (Tableau annexe 4, page 161). Le taux de résection curative global était de 73,4% et variait de 47% à 86% selon l'essai. L'administration de CRS n'améliorait pas la qualité de la résection chirurgicale par rapport à CS :  $RR_{\text{fixe}} = 1,04[0,95-1,13]$ ,  $p = 0,45$ .

Il n'existait pas d'argument pour une variation de cet ET avec la localisation tumorale ( $p_{\text{interaction}} = 0,64$ ).

L'ET semblait potentiellement varier avec le type histologique ( $p_{\text{interaction}} = 0,06$ ). L'ET était plus marqué dans les CE ( $RR = 1,19[1,00-1,42]$ ) que dans les ADK ( $RR = 0,98[0,88-1,09]$ ).

### *Morbidité et mortalité postopératoires*

Des taux de mortalités postopératoires étaient disponibles dans deux essais (237 patients, 8 décès). (Tableau annexe 5, page 162) Le taux global était de 3,4% et variait de 0% à 4,4%. L'administration de CRS n'augmentait pas, par rapport à CS, le risque de mortalité postopératoire de façon significative  $RR_{\text{fixe}} = 2,62[0,61-11,22]$ ,  $p = 0,20$ .

Des informations sur la morbidité postopératoire n'étaient disponibles que dans l'essai IG9401 (pour 81 patients). (Tableau annexe 6, page 163) Dans cet unique essai, la morbidité ne variait pas selon le type de traitement ( $RR = 0,99[0,67-1,46]$ ).

### *Toxicités*

L'essai IG9401 est le seul à avoir fourni des données de toxicité (88 patients). La toxicité la plus fréquente était celle sur l'œsophage 5,7%, suivie des nausées (4,5%), de la toxicité cardiaque (4,5%) et celle sur la fonction rénale (3,4%). Toutes les autres toxicités observées (Neutrophiles, pulmonaires) touchaient moins de 3% des patients. Aucune de ces toxicités ne semblait varier avec le traitement administré de façon significative, mais cette analyse repose sur un nombre très faible de patients.

### *Analyses en sous-groupes et de sensibilités*

Aucun essai ne semblait être clairement un outlier. L'essai NéoRes présentait un résultat légèrement différent de celui des autres, mais avec un large recouvrement des intervalles de confiance et l'absence d'hétérogénéité évidente. De même, aucun des essais n'avait inclus moins de 60 patients.

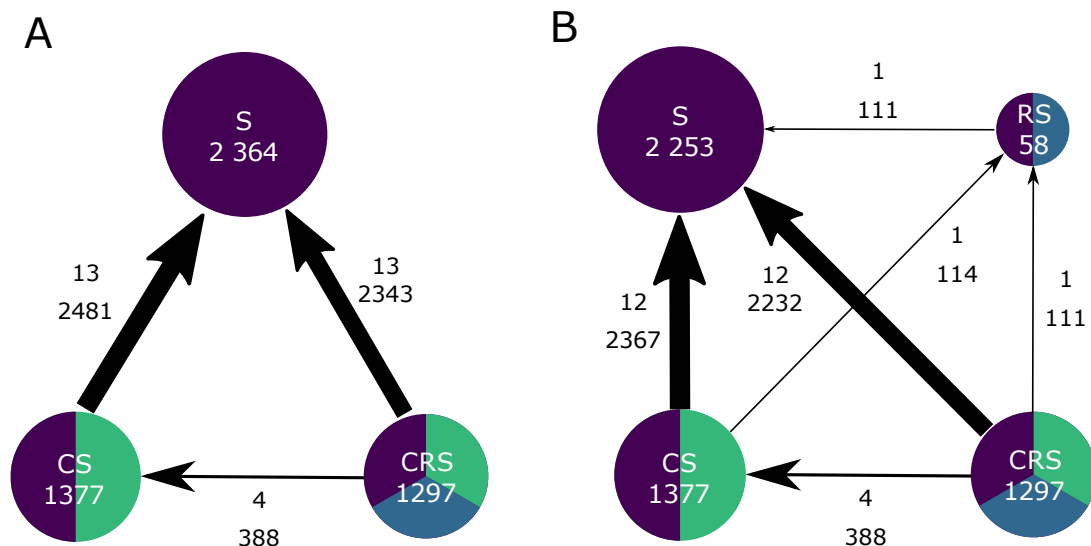
Le protocole prévoyait l'analyse selon le type de chimiothérapie 5-FU-Cisplatine contre les autres. Seul l'essai Scandinavia avait utilisé une chimiothérapie à base de Cisplatine-Bléomycine au lieu du 5-FU cisplatine. Il n'y avait pas d'interaction entre le type de chimiothérapie et la SG ( $p_{\text{interaction}} = 0,31$ ). L'essai Scandinavia n'ayant pas fourni de données de DFS, il n'était pas possible de tester l'interaction pour ce critère. Le protocole prévoyait également une analyse selon que la chimiothérapie utilisée dans le groupe CS était la même ou non que dans le groupe CRS. Tous les essais ayant utilisé la même chimiothérapie, cette analyse n'était pas nécessaire.

## Méta-Analyse en réseau

### Caractéristiques du réseau et des participants

#### *Géométrie du réseau*

Le réseau avait une géométrie déséquilibrée avec la plupart de l'information contenue dans les comparaisons S contre CS et S contre CRS et peu d'information dans la comparaison CS contre CRS qui est la comparaison cliniquement la plus pertinente.



**Figure 45 :** Réseau de MANATEC-02 : (A) comme un réseau standard, (B) comme un réseau additif

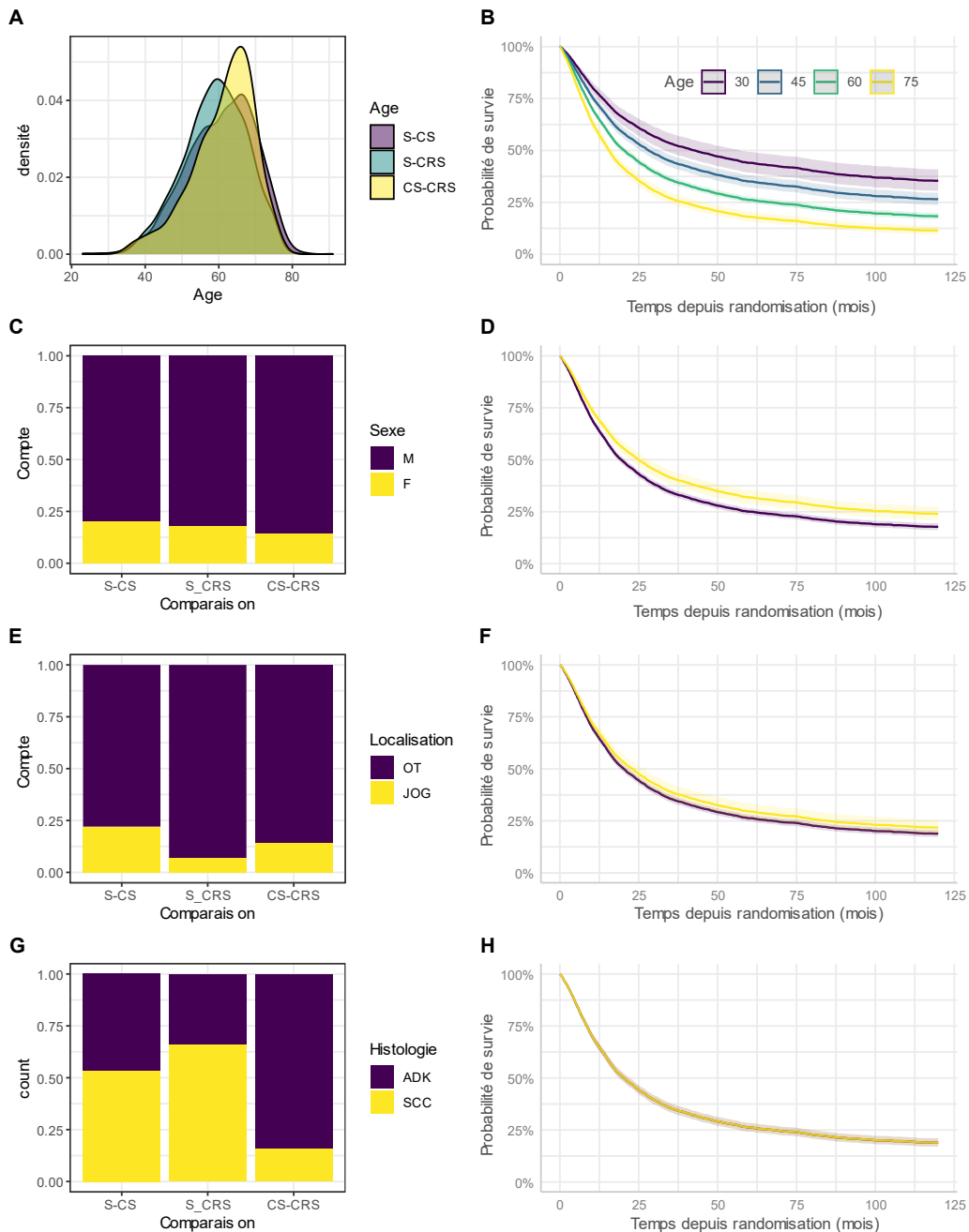
Chaque cercle représente un traitement dont la surface est proportionnelle au nombre de patients l'ayant reçu. Ce nombre est noté dans le cercle. Chaque flèche représente une comparaison. Les nombres notés au-dessus de la flèche sont : le nombre d'essais et le nombre de patients pour cette comparaison. L'épaisseur de la flèche est proportionnelle au nombre de patients. Le sens de la flèche indique le sens de la comparaison :  $X \rightarrow Y$  signifie X contre Y. Dans la configuration B les patients traités par RS proviennent du quatrième bras de l'essai Scandinavia.

La seconde particularité de ce réseau est qu'il pouvait être vu comme un modèle additif avec trois composantes S, C et R. Tous les protocoles comprenaient S, soit seul, soit associé. Ce traitement était donc considéré comme la référence dans les modèles. Dans les essais ayant utilisé la CRS on observait que la partie chimiothérapie était très proche, quand elle n'était pas strictement équivalente entre les patients ayant reçu CS et CRS. En revanche, le faible nombre de patients ayant eu RS rendait difficile l'estimation de l'interaction C\*R et donc la recherche d'une synergie ou d'un antagonisme entre les C et R.

#### *Hypothèse de transitivité*

L'analyse des critères d'inclusion des patients dans les différents essais montrait que ceux-ci sont très proches. Tous les essais utilisaient également la chirurgie dans tous les bras. Les malades étaient donc tout opérables et similaires sur ce point. Tous les essais comportaient également au moins un bras recevant de la chimiothérapie, les contre-indications à ce traitement ne peuvent donc pas biaiser les résultats. Les essais de la comparaison S contre CS n'utilisaient pas de radiothérapie. Un malade ayant une contre-

indication à ce traitement n'aurait donc pas pu être inclus dans ces essais, mais aurait pu l'être dans les autres. Toutefois les contre-indications à ce traitement sont très rares, principalement dominées par un antécédent de radiothérapie du même espace anatomique. Cette dernière est de plus le plus souvent utilisée pour traiter un autre cancer, ce qui constitue souvent un critère d'exclusion pour un essai thérapeutique. L'hypothèse de transitivité est donc tout à fait plausible dans notre réseau.



**Figure 46:** Distribution entre les comparaisons et effet sur la survie globale de l'âge (A et B), le sexe (C et D) la localisation tumorale (E et F) et le type histologique (G et H)

S : chirurgie, CS : chimiothérapie chirurgie, CRS : chimioradiothérapie chirurgie, M : Homme, F : femme, OT : œsophage thoracique, JOG : jonction oesogastrique, ADK : adénocarcinome, SCC : Carcinome épidermoïde

### *Distribution des modificateurs dans le réseau*

Les variables potentiellement liées aux critères de jugement n'étaient pas distribuées de façon identique dans toutes les comparaisons. L'âge moyen dans les comparaisons S-CS, S-CRS et CS-CRS était de 62(54, 68), 60 (54, 65) et 63 (56, 67) ans respectivement ( $p < 0,001$ ). Les taux d'hommes étaient de 80%, 82% et 84% respectivement ( $p = 0,093$ ). Les taux de tumeur localisées dans l'œsophage thoracique étaient de 78%, 93% et 89%,  $p < 0,001$ . Le taux de CE était de 53%, 67% et 34%, ( $p < 0,001$ ). Certaines de ces variables étaient également liées à la survie (âge :  $p < 0,0001$  ; histologie :  $p = 0,98$  ; localisation :  $p = 0,06$  ; sexe :  $p < 0,0001$ ) (**Figure 46**). De tels résultats plaident pour l'ajustement des modèles sur ces variables.

## Effet des traitements sur la survie globale

### *Effet en population générale*

Les résultats étaient ajustés sur l'âge, le sexe, la localisation tumorale et le type histologique et combinaient les 20 jeux de données imputées.

Dans le modèle standard, la CS et la CRS amélioraient la SG par rapport à S seul :  $HR_{aléatoire} = 0,86[0,75 - 0,99]$ ,  $p = 0,03$  et  $HR_{aléatoire} = 0,77[0,68 - 0,87]$ ,  $p < 0,001$ , respectivement. La comparaison CRS contre CS avait un  $HR_{aléatoire} = 0,90[0,74 - 1,09]$ ,  $p = 0,27$ . Il existait de l'hétérogénéité dans le réseau  $Q = 50,23$ ,  $p = 0,03$ ,  $I^2 = 48,2\%$  [19,0% ; 66,9%].

Dans le modèle additif, adjoindre une chimiothérapie préopératoire à la chirurgie, ajusté sur la radiothérapie, améliorait, à la limite de la significativité, la survie globale  $HR_{aléatoire} = 0,89[0,78-1,00]$ ,  $p = 0,06$ . Adjoindre de la radiothérapie, ajusté sur la chimiothérapie, améliorait la survie globale  $HR_{aléatoire} = 0,83[0,73-0,96]$ ,  $p = 0,03$ . La combinaison chimioradiothérapie améliorait donc également la survie globale avec  $HR_{aléatoire} = 0,74[0,66-0,84]$ ,  $p < 0,001$ .

L'effet de l'ajout de la chimiothérapie était plus variable que celui de l'ajout de la radiothérapie (écart-type de l'effet aléatoire 0,16 contre 0,05).

Le suivi médian était de 6,2 [6,0-6,4] ans.

### *Effets en fonction des sous-groupes de participants*

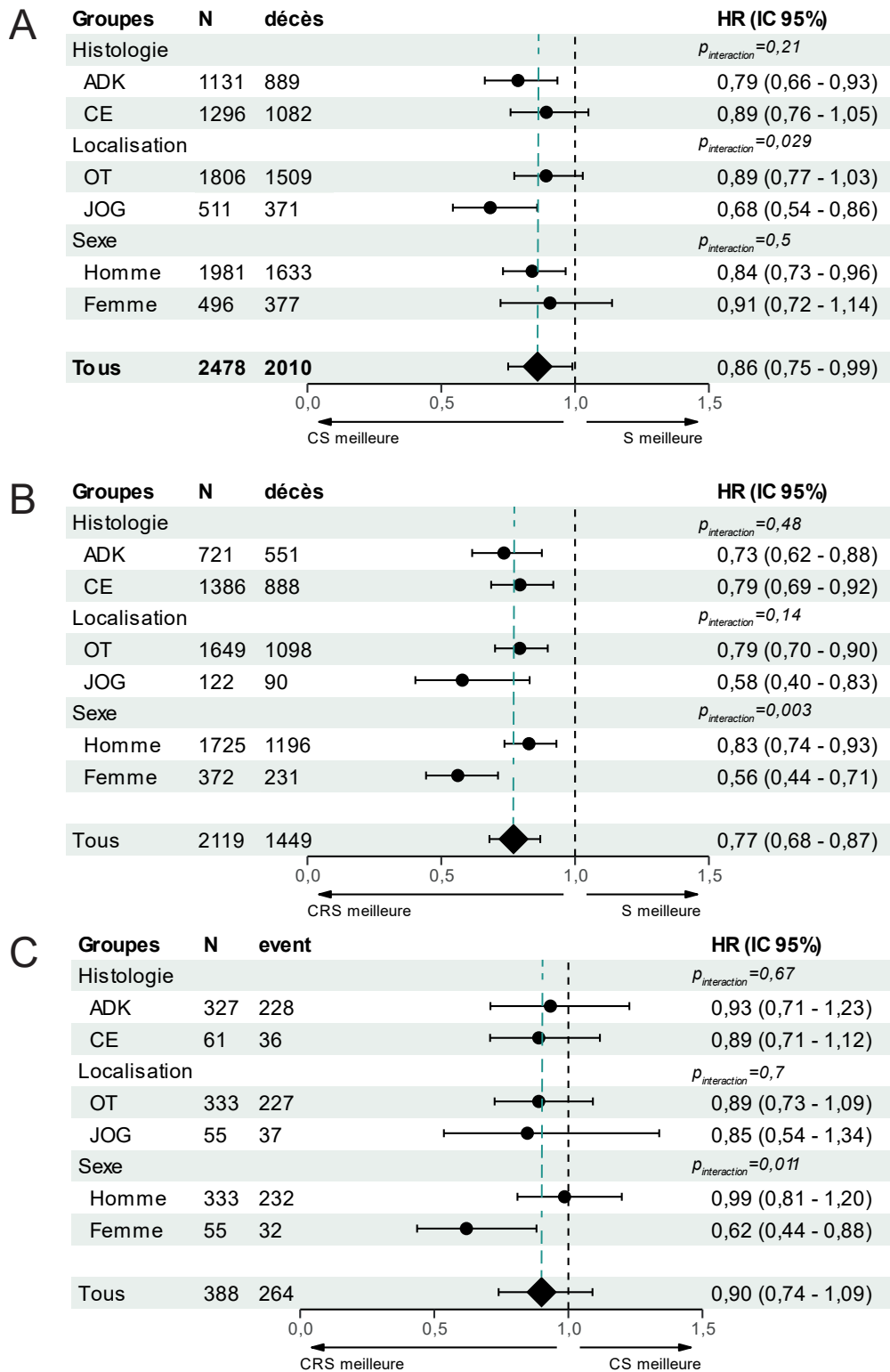
Dans le modèle standard, il existait une interaction significative avec la localisation tumorale dans la comparaison CS contre S ( $p_{interaction} = 0,029$ ), mais pas dans CRS contre S ( $p_{interaction} = 0,14$ ) ou dans CRS contre CS ( $p_{interaction} = 0,7$ ). Dans le modèle additif, l'analyse de l'effet en fonction de la localisation tumorale suggérait une interaction avec l'effet de la chimiothérapie ( $p_{interaction} = 0,005$ ), mais pas avec celui de la radiothérapie ( $p_{interaction} = 0,24$ ) ou de la combinaison ( $p_{interaction} = 0,38$ ). L'effet de la chimiothérapie était plus prononcé pour les tumeurs situées à la JOG  $HR_{aléatoire} = 0,68[0,55-0,84]$  que celles situées dans l'ET  $HR_{aléatoire} = 0,93[0,82-1,06]$ .

Dans le modèle standard on n'observait d'interaction significative dans aucune des comparaisons ( $p_{interaction} = 0,21$  ;  $p_{interaction} = 0,48$  ;  $p_{interaction} = 0,67$  ; pour CS-S, CRS-S et CRS-CS respectivement). Dans le modèle additif, il n'existait pas non plus d'interaction

significative entre le type histologique et l'effet de la chimiothérapie ( $p_{\text{interaction}} = 0,089$ ) ou l'effet de la radiothérapie ( $p_{\text{interaction}} = 0,12$ ) ou de la combinaison ( $p_{\text{interaction}} = 0,82$ ).

Dans le modèle standard on n'observait pas d'argument pour une interaction entre l'âge (codé en continu) et les comparaisons CS-S, CRS-S et CRS-CS ( $p_{\text{interaction}} = 0,78$ ,  $p_{\text{interaction}} = 0,71$  et  $p_{\text{interaction}} = 0,77$  respectivement). Dans le modèle additif, il n'y avait pas non plus d'argument pour une interaction avec l'effet de la chimiothérapie ( $p_{\text{interaction}} = 0,34$ ), l'effet de radiothérapie ( $p_{\text{interaction}} = 0,61$ ) ou de leur combinaison ( $p_{\text{interaction}} = 0,8$ ).

Dans le modèle standard on observait une interaction significative avec le sexe dans la comparaison CS contre S ( $p_{\text{interaction}} = 0,003$ ) et CRS contre CS ( $p_{\text{interaction}} = 0,011$ ) mais pas dans CS contre S ( $p_{\text{interaction}} = 0,21$ ). Dans le modèle additif, le sexe semblait faire varier l'effet de la radiothérapie ( $p_{\text{interaction}} = 0,015$ ), mais pas celui de la chimiothérapie ( $p_{\text{interaction}} = 0,62$ ). En combinant les deux traitements, l'interaction restait significative ( $p_{\text{interaction}} = 0,004$ ). L'effet de la radiothérapie était plus marqué chez les femmes  $HR_{\text{aléatoire}} = 0,66[0,53-0,84]$  que chez les hommes  $HR_{\text{aléatoire}} = 0,88[0,77-1,02]$ . (**Figure 47**).



**Figure 47 :** Effets des modificateurs de l'effet sur la survie globale pour : (A) la comparaison CS-S, (B) la comparaison CRS-S (C) la comparaison CRS-CS dans le modèle standard

Les résultats sont ceux de l'analyse poolée des 20 jeux de données imputées. Les modèles sont tous ajustés sur la chimiothérapie, la radiothérapie, l'âge, le sexe, la localisation tumorale et le type histologique. HR : Hazard Ratio, IC : intervalle de confiance, ADK : adénocarcinome, CE : carcinome épidermoïde, ET : œsophage thoracique, JOG : jonction œsogastrique. Les effectifs (N/décès) sont ceux des données non manquantes.



## Effet des traitements sur les autres critères

### *Survie sans récurrence*

Le « landmark » de 6 mois était appliqué. Les modèles étaient ajustés sur l'âge, le sexe, la localisation tumorale et le type histologique et combinait les 20 jeux de données imputées.

Dans le modèle standard, la CS et la CRS amélioraient la SSR par rapport à S seule :  $HR_{aléatoire} = 0,79[0,69-0,91]$ ,  $p = 0,001$  et  $HR_{aléatoire} = 0,78[0,68-0,88]$ ,  $p < 0,001$ , respectivement. La CRS contre CS avait un  $HR_{aléatoire} = 0,98[0,80-1,21]$ ,  $p = 0,85$ . Il existait de l'hétérogénéité avec  $Q = 41,46$ ,  $p = 0,032$  et  $I^2 = 51,8\%$  [20,5% ; 70,7%].

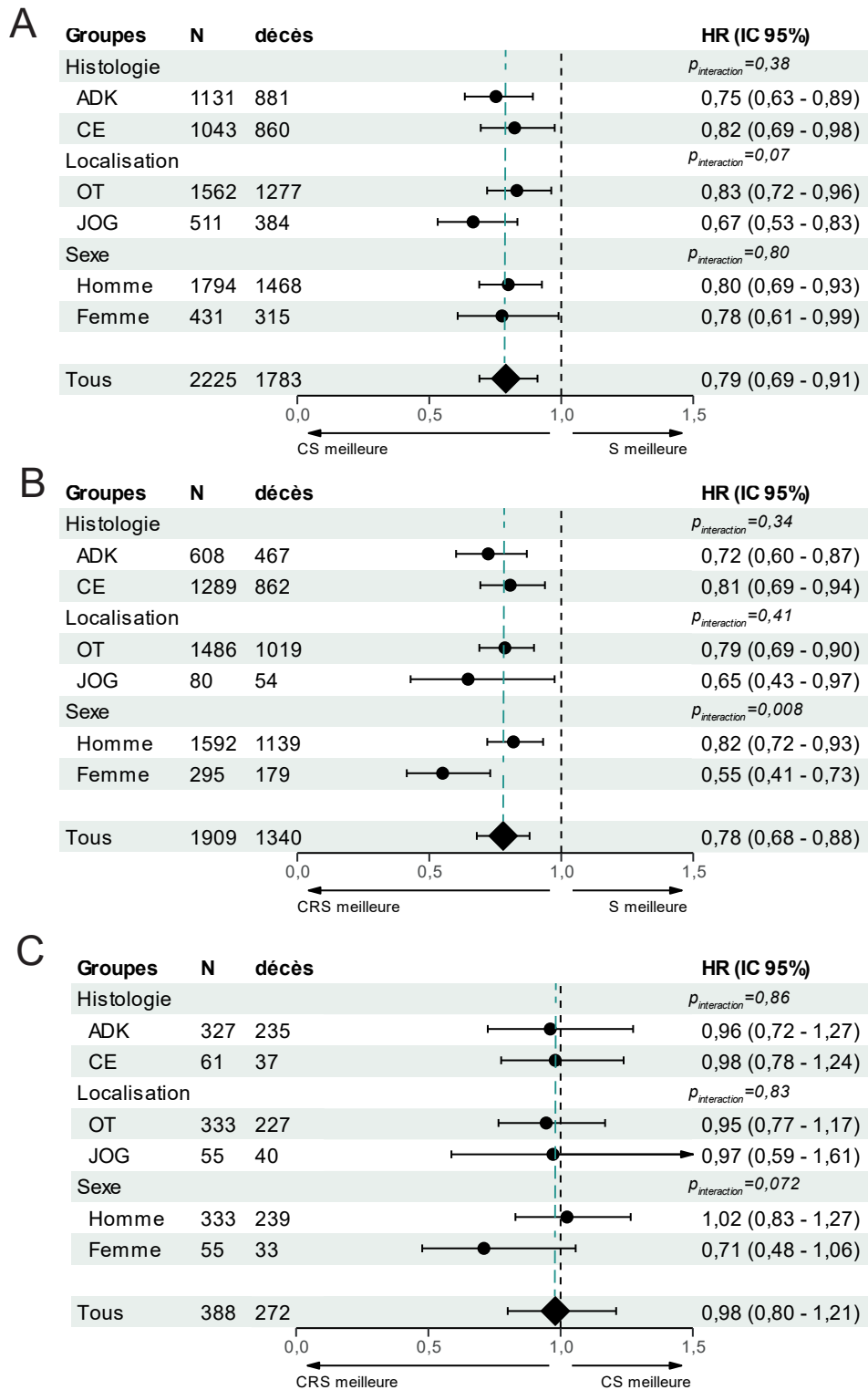
Dans le modèle additif, après ajustement sur la présence d'une radiothérapie, adjoindre une chimiothérapie préopératoire à la chirurgie améliorait la SSR  $HR_{aléatoire} = 0,80[0,71-0,91]$ ,  $p < 0,001$ . Ajusté sur la présence d'une chimiothérapie, adjoindre une radiothérapie n'améliorait pas de façon significative la SSR  $HR_{aléatoire} = 0,92[0,78-1,10]$ ,  $p = 0,28$ . La combinaison chimioradiothérapie préopératoire améliorait la SSR  $HR_{aléatoire} = 0,74[0,65-0,84]$ ,  $p < 0,001$ .

L'analyse selon la localisation tumorale montrait une variation de l'effet dans la comparaison CS-S chimiothérapie ( $p_{interaction} = 0,07$ ) mais pas dans la comparaison CRS-S ( $p_{interaction} = 0,41$ ) ou CRS-CS ( $p_{interaction} = 0,83$ ). L'effet de la CS était plus marqué pour les tumeurs localisées à la JOG  $HR_{aléatoire} = 0,67[0,53-0,83]$  que celles localisées dans l'ET  $HR_{aléatoire} = 0,79[0,69-0,90]$ . (**Figure 48**)

Il n'y avait pas d'arguments pour une variation des effets avec le type histologique dans les trois comparaisons ( $p_{interaction} = 0,38$ ,  $p_{interaction} = 0,34$ ,  $p_{interaction} = 0,86$  respectivement). (**Figure 48**)

Il n'y avait pas non plus d'argument suggérant une variation de l'effet du traitement dans selon l'âge ( $p_{interaction} = 0,61$ ,  $p_{interaction} = 0,60$ ,  $p_{interaction} = 0,32$  respectivement).

Il ne semblait pas avoir d'interaction entre le sexe et l'effet traitement dans la comparaison CS-S ( $p_{interaction} = 0,80$ ). En revanche, l'effet du traitement semblait influencé par le sexe dans les comparaisons CRS-S ( $p_{interaction} = 0,008$ ), ou CRS-CS ( $p_{interaction} = 0,072$ ). Dans les deux cas l'effet était plus marqué chez les femmes que chez les hommes. (**Figure 48**)



**Figure 48 :** Effets des modificateurs de l'effet sur la survie sans récurrence pour : (A) la comparaison CS-S, (B) la comparaison CRS-S (C) la comparaison CRS-CS dans le modèle standard

Les résultats sont ceux de l'analyse poolée des 20 jeux de données imputées. Les essais sans information de DFS étaient exclus. Les modèles sont tous ajustés sur la chimiothérapie, la radiothérapie, l'âge, le sexe, la localisation tumorale et le type histologique. HR : Hazard Ratio, IC : intervalle de confiance, ADK : adénocarcinome, CE : carcinome épidermoïde, OT : œsophage thoracique, JOG : jonction œsogastrique. Les effectifs (N/décès) sont ceux des données non manquantes.

### *Risques compétitifs / modes de récurrence*

Les résultats étaient obtenus par combinaison des 20 jeux de données imputées. Les modèles étaient ajustés sur l'âge, le sexe, le type histologique et la localisation tumorale. Après exclusion des essais n'ayant fourni aucune donnée de ce type, l'analyse portait sur 22 essais et 4 522 patients. Le « landmark » de 6 mois était appliqué.

Au cours du suivi, 773 récurrences locales ont été observées (Tableau annexe 2, page 159). Dans le modèle standard, une diminution du risque de récurrence locale était observée dans les comparaisons CS contre S et CRS contre S : sHR = 0,74 [0,62 – 0,88] < 0,001 et 0,65 [0,53 – 0,81], p < 0,001 respectivement mais pas dans CRS contre CS sHR = 0,88 [0,66 – 1,18], p = 0,41. Dans le modèle additif, ajusté sur la présence d'une radiothérapie, ajouter une chimiothérapie permettait de diminuer le risque de récurrence local sHR = 0,74 [0,62-0,88], p=0,001. Ajusté sur la présence d'une chimiothérapie, ajouter de la radiothérapie ne semblait pas modifier le risque de récurrence locale sHR = 0,88 [0,68-1,15], p = 0,36. La combinaison de ces deux traitements permettait de diminuer le risque sHR = 0,65 [0,53-0,81], p<0,001.

Au cours du suivi, 1 502 récurrences à distance ou locales et à distance synchrones ont été observées (Tableau annexe 2, page 159). Dans le modèle standard, une diminution du risque de récurrence à distance était observée dans les comparaisons CS contre S et CRS contre S : sHR = 0,86 [0,75 – 0,98] < 0,021 et 0,77 [0,67 – 0,89], p < 0,001 respectivement mais pas dans CRS contre CS sHR = 0,90 [0,73 – 1,11], p = 0,33. Dans le modèle additif, ajusté sur la présence d'une radiothérapie, ajouter une chimiothérapie permettait de diminuer le risque de sHR = 0,86 [0,75-0,98], p = 0,021. Ajusté sur la présence d'une chimiothérapie, ajouter de la radiothérapie ne semblait pas modifier le risque de récurrence à distance sHR = 0,90 [0,76-1,07], p = 0,25. La combinaison de ces deux traitements permettait de diminuer le risque sHR = 0,77 [0,67-0,89], p<0,001.

Au cours du suivi, 1 119 décès sans récurrence ont été observés (Tableau annexe 2, page 159). Dans le modèle standard, seule la comparaison CRS contre S montrait une augmentation significative du risque de décès sans récurrence sHR = 1,21 [1,01-1,45], p=0,034. Une augmentation non significative était observée dans les comparaisons CS contre S et CRS contre CS : 1,12 [0,97-1,29], p = 0,12 et 1,08 [0,86-1,38], p = 0,51, respectivement. Dans le modèle additif, ajouter une chimiothérapie ou une radiothérapie, ne semblait pas modifier le risque de décès sans récurrence sHR = 1,12 [0,97-1,29], p = 0,12 et 1,08 [0,88-1,34], p = 0,46 respectivement. Toutefois, la combinaison de ces deux traitements augmentait ce risque sHR = 1,21 [1,01-1,45], p = 0,034.

### *Mortalité liée et non liée au cancer*

Les résultats étaient obtenus par combinaison des 20 jeux de données imputées. Les modèles étaient ajustés sur l'âge, le sexe, le type histologique et la localisation tumorale. Après exclusion des essais n'ayant fourni aucune donnée de ce type, l'analyse portait sur 22 essais et 4 522 patients.

Au cours du suivi, 3 119 décès liés au cancer ont été observés. Dans le modèle standard, comparées à S, CS et CRS diminuaient le risque de décès liés au cancer : RR<sub>aléatoire</sub> = 0,83 [0,72 – 0,95], p = 0,007 et RR<sub>aléatoire</sub> = 0,76 [0,68 – 0,86], p < 0,001. La comparaisons CRS contre CS avait un RR<sub>aléatoire</sub> = 0,92 [0,76 – 1,12], p = 0,42. Dans le modèle additif, ajusté

sur la présence d'une radiothérapie, ajouter une chimiothérapie permettait de diminuer le risque de décès lié au cancer  $HR_{aléatoire} = 0,87[0,77-0,99]$ ,  $p = 0,033$ . Ajusté sur la présence d'une chimiothérapie, ajouter une radiothérapie diminuait également le risque de décès lié au cancer  $HR_{aléatoire} = 0,83[0,72-0,96]$ ,  $p = 0,011$ . La combinaison des deux avait un  $HR_{aléatoire} = 0,72[0,63-0,82]$ ,  $p < 0,001$ .

Au cours du suivi, 188 décès non liés au cancer ont été observés. Dans le modèle standard, une augmentation significative du risque de décès non lié au cancer n'était observée dans aucune des comparaisons ( $p = 0,57$  ;  $p = 0,88$  et  $p = 0,72$  ; pour CS contre S, CRS contre S et CRS contre CS respectivement). Dans le modèle additif, on n'observait pas non plus d'argument pour une variation du risque de décès non liés au cancer avec l'ajout d'une chimiothérapie  $HR_{aléatoire} = 1,13[0,69-1,86]$ ,  $p = 0,63$ , ou l'ajout d'une radiothérapie  $HR_{aléatoire} = 0,94[0,52-1,70]$ ,  $p = 0,84$ . La combinaison de ces deux traitements ne semblait pas non plus modifier ce risque  $HR_{aléatoire} = 1,06[0,94-1,20]$ ,  $p = 0,34$ .

### *Qualité de la résection*

Les résultats étaient donnés sur la combinaison des données imputées, un  $RR > 1$  indiquant une plus grande probabilité d'avoir une résection R0 et donc favorable au traitement. Dans le modèle standard aucune des comparaisons ne montrait d'effet sur la qualité de la résection ( $p = 0,82$  ;  $p = 0,3$  ;  $p = 0,46$  ; pour CS contre S, CRS contre S et CRS contre CS respectivement). Dans le modèle additif l'ajout d'une chimiothérapie n'augmentait pas la probabilité d'obtenir une résection R0  $RR_{aléatoire} = 1,05 [0,73-1,53]$ ,  $p = 0,79$ . De même, l'ajout d'une radiothérapie ne modifiait pas cette probabilité  $RR_{aléatoire} = 1,18 [0,85-1,66]$ ,  $p = 0,32$ . La combinaison des deux traitements n'atteignait pas non plus la significativité statistique  $RR_{aléatoire} = 1,25 [0,87-1,79]$ ,  $p = 0,23$ .

### *Morbidité et mortalité postopératoire*

Dans ces analyses un  $RR > 1$  indique une plus grande probabilité d'observer une morbidité ou une mortalité (défavorable au traitement).

Dans le modèle standard, l'analyse du risque de morbidité postopératoire ne montrait pas d'arguments pour une influence des traitements dans les trois comparaisons ( $RR_{aléatoire} = 1,17 [0,90-1,52]$ ,  $p = 0,23$  ;  $RR_{aléatoire} = 1,09 [0,82-1,44]$ ,  $p = 0,55$  ;  $RR_{aléatoire} = 0,93 [0,66-1,31]$ ,  $p = 0,16$  pour CS contre S, CRS contre S et CRS contre CS respectivement). Dans le modèle additif, ce risque ne variait pas non plus avec l'ajout d'une chimiothérapie ( $RR_{aléatoire} = 1,16[0,91-1,47]$ ,  $p = 0,22$ ), ou d'une radiothérapie ( $RR_{aléatoire} = 0,97[0,68-1,38]$ ,  $p = 0,86$ ). La combinaison des deux traitements avait un  $RR_{aléatoire} = 1,13[0,89-1,43]$ ,  $p = 0,33$ .

Dans le modèle standard, une augmentation du risque de mortalité postopératoire proche de la significativité n'était observée que dans la comparaison CRS contre S :  $RR_{aléatoire} = 1,64 [0,95 - 2,83]$ ,  $p = 0,076$ . Dans les comparaisons CS contre S et CRS contre CS on n'observait pas d'arguments francs pour une augmentation de ce risque ( $RR_{aléatoire} = 1,22 [0,63-2,35]$ ,  $p = 0,56$  et  $RR_{aléatoire} = 1,35 [0,54-3,35]$ ,  $p = 0,14$  respectivement). Dans le modèle additif, on ne mettait pas en évidence d'argument pour une variation du risque de mortalité postopératoire avec l'ajout d'une chimiothérapie ( $RR_{aléatoire} = 1,30[0,70-2,44]$ ,  $p = 0,41$ ), ou l'ajout d'une radiothérapie ( $RR_{aléatoire} = 1,09[0,47-2,53]$ ,  $p = 0,83$ ). La combinaison des deux traitements montrait  $RR_{aléatoire} = 1,43[0,79-2,56]$ ,  $p = 0,23$ .

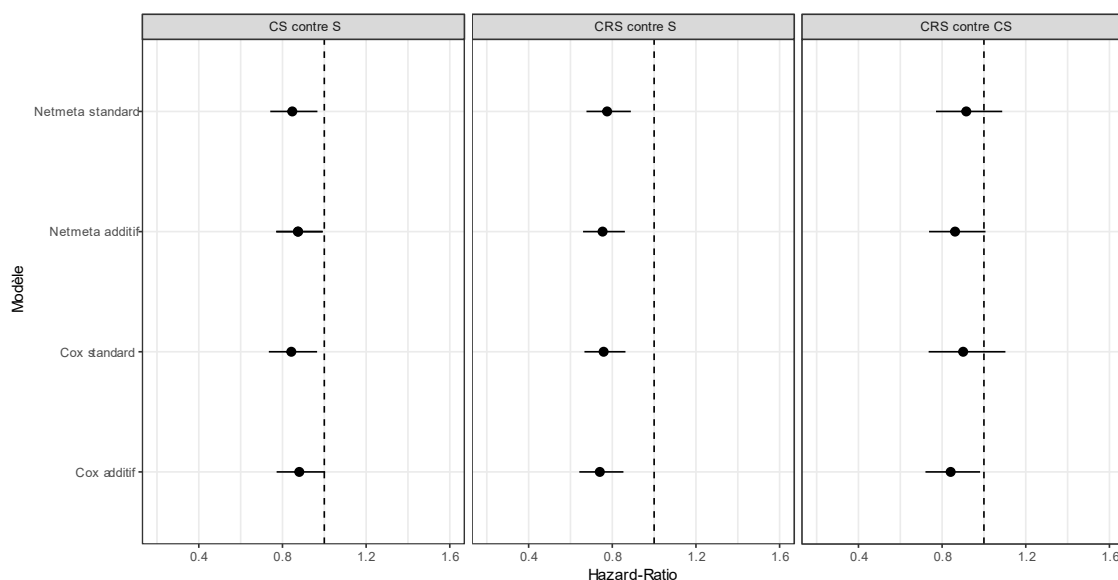
## Analyses de sensibilité

### *Analyses de sensibilité sur la modélisation du réseau*

Deux analyses de sensibilité portant sur la modélisation du réseau ont été réalisées : une première sur la façon de calculer les paramètres, la seconde sur une formulation alternative du réseau de traitement. Ces analyses étaient réalisés sur la survie globale.

Pour la première analyse de sensibilité nous avons utilisé la méthode basée sur les contrastes et la théorie des graphes. Cette méthode fréquentiste a été démontrée comme équivalente aux modèles Bayésien. Le package R `netmeta` s'est récemment enrichi de la fonction `netcomb` permettant de réaliser également des réseaux additifs en deux étapes.(28) Les données non ajustées et non imputées étaient utilisées. Les résultats étaient très proches de ceux du modèle de Cox. Les contrastes CS-S CRS-S et CRS-CS dans la modélisation standard étaient respectivement 0,85 [0,74-0,97], 0,78[0,68-0,89] et 0,92[0,77-1.09]. Dans la modélisation additive, l'ajout d'une chimiothérapie, l'ajout d'une radiothérapie et de leur combinaison étaient respectivement de :  $HR_{aléatoire} = 0,87[0,77-0,99]$ ,  $HR_{aléatoire} = 0,86[0,74-1,01]$  et  $HR_{aléatoire} = 0,75[0,66-0,86]$ .

Pour la seconde analyse de sensibilité, le réseau était analysé comme un réseau standard sans ajustement et sans utilisation des données imputées. Les résultats étaient très proches de ceux de l'analyse principale. Le modèle de Cox standard estimait le contraste S contre CS à  $HR_{aléatoire} = 0,84 [0,73-0,97]$ . Le contraste CRS contre S était estimé à  $HR_{aléatoire} = 0,76 [0,67-0,86]$ . Enfin le contraste CRS contre CS était estimé à  $HR_{aléatoire} = 0,90 [0,74-1,10]$ . La **Figure 49** résume le résultat des quatre approches dans l'évaluation des contrastes.



**Figure 49** : Résultats des analyses de sensibilité sur les quatre façons d'aborder le réseau pour la survie globale

### *Exploration de l'hypothèse de cohérence*

Seul le modèle standard reposait sur l'hypothèse de cohérence. Étant donné sa formulation, le modèle additif ne repose pas sur l'hypothèse de cohérence comme dans une MAR classique.

La méthode du « node-splitting » consiste à comparer les estimateurs basés uniquement sur les données directes à ceux basés uniquement sur les données indirectes. La fonction `netmeta` permet d'obtenir directement ces estimateurs. Pour le contraste CRS contre CS on obtenait un  $HR_{direct}=0,80[0,58-1,08]$  et un  $HR_{indirect}=0,98[0,79-1,20]$ , ce qui correspond à un effet traitement estimé légèrement plus fort sur les données directes. Il est également possible de tester l'incohérence en testant l'hétérogénéité entre les « design » (design by treatment interaction model). On trouvait une valeur de Q à 1,14 ( $p=0,29$ ).

Une seconde approche, basée sur les données individuelles, consiste à introduire dans le modèle un paramètre testant l'incohérence. Si le réseau est cohérent, ce paramètre devrait être estimé comme nul, ou du moins avec un intervalle de confiance comprenant la valeur 0. Un grand intérêt de cette approche est de pouvoir ajuster les résultats sur des covariables et/ou les interactions, afin de voir si tout ou partie de l'incohérence provient de leur effet. Sur les données non ajustées, on observait une incohérence estimée à  $-0,23[-0,57 ; 0,11]$  sans que celle-ci atteigne la significativité statistique ( $p=0,19$ ). Après ajustement sur les covariables, cette incohérence diminuait à  $-0,1[-0,5 ; 0,31]$ ,  $p = 0,60$ . Ce résultat suggère que l'incohérence observée sur les données non ajustées est au moins en partie due à une différence de distribution des modificateurs ou aux interactions.

## Chapitre 4. Discussion et perspectives

### Résultats principaux

#### Survie globale en population générale

Pour la survie globale, les méta-analyses deux à deux et l'analyse en réseau ont confirmé les résultats observés dans la plupart des essais et les AD-MA de ces essais. Prises séparément, la chimiothérapie et la chimioradiothérapie sont toutes les deux supérieures à la chirurgie seule sur la survie globale. Comme le préconisent déjà les recommandations Européennes et Américaines, la chirurgie d'emblée ne devrait donc plus être proposée à un patient présentant un cancer de l'œsophage localement avancé. La magnitude de ce bénéfice peut en revanche paraître décevante. La chimiothérapie améliorait la SG avec un  $HR_{aléatoire} = 0,83[0,72-0,96]$  dans la comparaison deux à deux et  $HR_{aléatoire} = 0,86[0,75-0,99]$  dans l'analyse en réseau. Une telle différence relative correspondait à un dRMST à 5 ans de 3,3 mois seulement. De même, la CRS améliorait la SG avec un  $HR_{aléatoire} = 0,77[0,68-0,87]$  dans l'analyse deux à deux et  $HR_{aléatoire} = 0,74[0,65-0,85]$  dans le réseau. Cette différence correspondait à une dRMST à 5 ans de 3,9 mois seulement.

En ce qui concerne la comparaison CS contre CRS, les conclusions sont plus délicates. L'analyse deux à deux montrait une amélioration de la SG avec un  $HR_{fixe} = 0,78[0,63-0,95]$ . Dans le réseau standard on observait un effet moins prononcé :  $HR_{aléatoire} = 0,90[0,74 - 1,09]$ . Enfin dans le modèle additif, ajouter R à un protocole comprenant déjà CS améliorait la SG avec un  $HR_{aléatoire} = 0,83[0,73-0,96]$ . L'analyse de l'incohérence du réseau, montrait une incohérence présente mais limitée, n'atteignant pas le seuil de significativité dans les deux méthodes utilisées. Bien que la fixation d'un seuil arbitraire à 5% puisse faire penser que ces résultats sont discordants, on observe que les estimations du bénéfice sont en fait très proches. Si le bénéfice de la CRS par rapport à la CS existe, celui-ci est probablement assez limité.

Nos résultats sont proches des résultats préliminaires de l'essai en cours NéoAegis rapportés sous forme de présentation au congrès de l'ASCO 2021. (91) Cet essai, non inclus dans MANATEC-02 car en cours de recrutement, a randomisé 362 patients pour recevoir, soit un protocole CS basé sur du FLOT, soit un protocole de CRS de type CROSS. Initialement construit pour montrer une supériorité de la CRS sur la CS de 10% en termes de SG, l'objectif a été changé pour de la non-infériorité avec une marge de 5% à la première analyse intermédiaire. Les résultats préliminaires ne montrent pas de supériorité de la CRS sur S à 3 ans :  $HR 1,02 (IC95\% 0,74-1,42)$ . Toutefois ceux-ci ne sont basés que sur une partie de la population incluse, les inclusions étant arrêtées depuis décembre 2020. Les résultats finaux seront disponibles en 2025. De tels résultats semblent compatible avec un bénéfice faible nécessitant un très grand échantillon pour être démontré.

#### Survie globale dans les sous-groupes

Un des grands intérêts de la IPD-MA était de pouvoir rechercher une interaction entre l'effet des traitements et les sous-groupes. Le modificateur le plus probable était le type histologique. Cette supposition reposait sur l'analogie avec les tumeurs histologiquement proches (cancers ORL pour les CE, cancers gastriques pour les ADK) et les observations de certains essais.

Dans les méta-analyses deux à deux, nous n'avons pas observé d'interaction avec le type histologique, ni dans la comparaison S-CS ( $p_{\text{interaction}}=0,26$ ), ni dans S-CRS ( $p_{\text{interaction}}=0,78$ ), ni dans CS-CRS ( $p_{\text{interaction}}=0,35$ ). Ces résultats étaient confirmés dans l'analyse en réseau pour l'ajout d'une chimiothérapie ( $p_{\text{interaction}}=0,089$ ) ou l'ajout d'une radiothérapie ( $p_{\text{interaction}}=0,12$ ). Nos résultats ne plaident donc pas en faveur de la recommandation d'un protocole différent en fonction du type histologique. La différence d'effet traitement observée dans certains essais, notamment CROSS,(80) n'est donc pas confirmée sur nos données. Parmi les cinq essais en cours, aucun n'a choisi d'inclure les deux types histologiques. Récemment, Kamarajah (92) a rapporté les résultats d'une étude rétrospective portant sur 11 167 ADK et 2 367 CE de la *National Cancer Database* américaine avec des résultats discordants avec les nôtres. Dans cette étude, la CRS n'apportait pas de bénéfice de survie par rapport à la CS pour les ADK = 1,04[0,95-1,14] mais un bénéfice pour les CE HR = 0,78[0,62-0,97]. Les biais de confusion étaient contrôlés par un score de propension et une analyse multivariable avec les limites habituelles des études de registre. Cette étude de forte puissance permettait d'avoir une estimation de l'effet du traitement hors essai thérapeutique, mais ces résultats doivent être interprétés avec précaution, le niveau de preuve étant inférieur à celui d'un essai randomisé.

Le second modificateur supposé était la localisation de la tumeur. En effet, la JOG est une zone charnière entre deux territoires. A quelques centimètres près, une même tumeur peut donc être considérée et traitée comme un cancer de l'œsophage ou comme un cancer de l'estomac. La classification de Siewert est actuellement la plus utilisée pour guider ce choix.(2) De façon simplifiée, les tumeurs Siewert 1 sont actuellement traitées comme des cancers de l'œsophage, les Siewert 3 comme des cancers de l'estomac et la question n'est pas tranchée pour les Siewert 2. Dans les méta-analyses CS contre S et CRS contre S, des arguments pour un effet différentiel existaient ( $p_{\text{interaction}}=0,07$  et  $p_{\text{interaction}}=0,04$ ). L'effet semblait plus important pour les tumeurs situées à la JOG que dans l'ET avec une dRMST à 5 ans de sept mois contre deux. La méta-analyse CRS contre CS ne montrait pas d'argument dans ce sens ( $p_{\text{interaction}}=0,47$ ). L'analyse en réseau confirmait ceci avec une interaction dans la comparaison CS-S ( $p_{\text{interaction}}=0,029$ ), mais pas dans CRS-S ( $p_{\text{interaction}}=0,14$ ) ou dans CRS-CS ( $p_{\text{interaction}}=0,7$ ). Outre la randomisation qui avait équilibré les sous-types histologiques, les résultats obtenus dans le réseau sont également ajustés sur le type histologique, rendant peu probable un biais de confusion lié à ce facteur. Nos résultats suggèrent donc que les tumeurs de la JOG sont plus chimio-sensibles que celles de l'œsophage. De telles constatations sont en accord avec les données récentes concernant le cancer de l'estomac. Dans cette pathologie, le traitement de référence est actuellement la chimiothérapie péri-opératoire. Celle-ci semble supérieure à la chirurgie seule dès les stades précoces de la maladie (T1b).

L'âge ne semblait pas modifier l'effet du traitement. Si l'amplitude de l'âge des patients inclus (entre 23 et 91 ans) suggérait un large éventail, en pratique 50% d'entre eux avaient en réalité entre 54 et 67 ans (IQR). Aucune des trois méta-analyses deux à deux ne montrait d'interaction avec l'âge considéré comme une variable continue ( $p_{\text{interaction}}=0,98$ ,  $p_{\text{interaction}}=0,61$  et  $p_{\text{interaction}}=0,27$  pour S-CS, S-CRS et CS-CRS respectivement). Seule la méta-analyse CRS contre S montrait une interaction avec l'âge comme une variable qualitative à 4 classes, mais cette constatation semblait être davantage liée aux seuils choisis qu'à un véritable effet. L'analyse en réseau confirmait ceci avec l'absence d'interaction avec l'ajout d'une chimiothérapie ( $p_{\text{interaction}}=0,34$ ) ou celui d'une radiothérapie ( $p_{\text{interaction}}=0,61$ ).



Dans la MA S contre CS, le sexe ne semblait pas influencer l'effet traitement ( $p=0,43$ ). En revanche, dans les MA S-CRS et CS-CRS, on observait une interaction significative ( $p_{\text{interaction}} = 0,003$  et  $p_{\text{interaction}} = 0,005$  respectivement). Dans ces deux MA, le bénéfice était supérieur chez les femmes. Dans la comparaison CRS contre S, cette différence se traduisait par un bénéfice en dRMST à 5 ans de 9,9 mois chez les femmes contre 2,3 chez les hommes. Dans la comparaison CRS contre CS, les bénéfices étaient de 13,7 mois contre 1,31 mois respectivement. L'analyse en réseau confirmait cette interaction absente pour l'ajout d'une chimiothérapie ( $p_{\text{interaction}} = 0,62$ ) mais présente pour l'ajout d'une radiothérapie ( $p_{\text{interaction}} = 0,015$ ). A notre connaissance, une interaction avec le sexe n'avait jamais été observée dans les néoplasies, soit anatomiquement (ORL)(34,93), soit fonctionnellement soit histologiquement (rectum(94), anus(95)) avoisinantes. La randomisation ayant probablement équilibré les autres facteurs de confusion potentiels tel qu'une éventuelle infection HPV, aucune hypothèse physiopathologique n'était privilégiée. En plus de son effet prédictif, le sexe avait aussi un effet pronostique. La survie à 5 ans était de 33% chez les femmes contre 25% chez les hommes.

Enfin, les derniers modificateurs que nous aurions souhaité étudier étaient les T et N du TNM. L'absence de données est surprenante compte tenu de l'importance de ces paramètres dans le choix du traitement.

## Survie sans récurrence, sites de récurrence et décès liés au cancer

Le bénéfice observé en SG se retrouvait également en survie sans récurrence. Celui-ci ne variait pas selon le type histologique. Il faut toutefois noter que l'utilisation d'un landmark, rendu nécessaire par les séquences de traitements, limitaient l'utilisation de l'information disponible dans les six premiers mois. Dans l'analyse en risque compétitif, cet effet se retrouvait aussi bien en termes de récurrence local qu'en termes de récurrence à distance. De façon logique, compte tenu des modes d'action, la chimiothérapie permettait de réduire à la fois les récurrences locales et à distance. De façon plus surprenante, la radiothérapie, aussi bien dans le modèle standard que le modèle additif, ne semblait pas diminuer le risque de récurrence local loco-régional. Mais s'il existe un doute sur la qualité du recueil de ce paramètre et si les récurrences locales et à distance synchrones ont été classées comme récurrence à distance, ces résultats sont décevants. A l'instar de ce que l'on observe dans les cancers du rectum, on aurait pu penser que même en l'absence d'effet sur la SG un effet sur le risque de récurrence local aurait été présent.

Dans toutes les méta-analyses par paires et dans le réseau, le type de traitement ne semblait pas influencer la qualité de la résection chirurgicale. Ces résultats doivent être interprétés avec précaution en raison d'un risque d'hétérogénéité entre les essais pour ce critère n'ayant fait que récemment l'objet d'effort de standardisation. Si cette information se confirmait, ceci irait à l'encontre des conceptions actuelles(96), dans lesquelles la chimioradiothérapie est préférentiellement utilisée en cas de doute sur la qualité de la résection. On pourrait toutefois arguer que de tels malades n'ont peut-être pas été inclus dans les essais présents puisque présentant une contre-indication relative à la chirurgie d'emblée.

Concernant les complications péri-opératoire induites, ces deux traitements néoadjuvants n'étaient pas responsables d'une hausse de la mortalité ou de la morbidité postopératoire. Si l'analyse de la mortalité postopératoire est probablement assez robuste (critère facile à recueillir), celle de la morbidité doit être prise avec plus de précautions. En effet, la forte

variation d'une étude à l'autre fait porter le doute quant à l'exhaustivité et l'homogénéité du recueil dans tous les essais.

L'analyse des risques compétitifs dans le réseau montrait un risque de décès sans récurrence supérieur chez les patients traités par CRS que ceux traités par S. Dans les méta-analyses par paire, on retrouvait cet effet dans la comparaison CRS contre S. Cet effet « délétère » doit être mis en perspective avec deux observations. Premièrement, cette surmortalité était, au global, très largement compensée par la diminution des décès après récurrence. Deuxièmement, une telle observation est classique en risque compétitif puisqu'éliminer des causes de mortalité par cancer « permet de mourir d'autre chose ». Toutefois, elle pourrait aussi être le témoin d'une toxicité tardive des traitements. Nous n'avons pas été en mesure de récupérer plus d'information sur ces toxicités tardives. L'analyse des causes spécifique de décès ne montrait pas d'argument dans ce sens, car les décès non liés n'étaient pas augmentés. Toutefois les décès non liés au cancer étaient très rares, car la méthode de Peto est très restrictive, rendant faible la puissance pour analyser ce critère.

### Estimation par un modèle de Cox IPD en une étape

Dans cette étude, nous avons utilisé le modèle de Cox mixte pour estimer les paramètres de la NMA basée sur les IPD. Nous n'avons pas rencontré de problème de convergence. Dans l'analyse de sensibilité les résultats étaient très proches de l'approche IPD en deux étapes basée sur les contrastes proposée par Rücker (27) confirmant que notre méthodologie était robuste. L'avantage de notre modèle était que celui-ci permettait plus facilement d'introduire des interactions et de l'ajustement sur d'autres facteurs. La combinaison des jeux de données imputées s'effectuait naturellement selon les règles proposées par Rubin.(97) L'autre approche concurrente était l'approche Bayésienne.(19) Celle-ci a les mêmes avantages que notre modèle et peut faire appel à différents modèles. Un travail ultérieur pourrait être de confirmer que les résultats de cette approche sont équivalents au notre et de permettre une comparaison des temps de calcul.

La plupart des hypothèses statistiques sous-jacentes à nos méthodes ont été vérifiées. Toutefois, vérifier l'hypothèse de proportionnalité des risques dans un modèle de Cox mixte est plus difficile à tester si le traitement est également le siège d'un effet aléatoire. En cas de suspicion de violation manifeste de cette hypothèse, des extensions du modèle de Cox ou du modèle de Poisson permettraient d'obtenir des estimations plus fiables.

Dans cette étude nous avons également exploré l'utilisation d'un modèle additif à la place du modèle standard. Cette modélisation était particulièrement adaptée à notre réseau où les traitements se combinent plus qu'ils ne sont des alternatives les uns aux autres. Les résultats des deux approches étaient très proches. Lors de la rédaction du protocole, ces approches étaient encore au début de leur développement. Aussi, nous avons choisi d'exclure de la NMA les essais ayant comparés la radiothérapie suivie de chirurgie à la chirurgie seule, car la radiothérapie n'est plus délivrée sans chimiothérapie depuis de nombreuses années. Ces essais nous auraient finalement été utiles puisqu'ils auraient permis d'estimer de façon plus optimale le modèle additif et de rechercher une synergie (ou un antagonisme) entre l'effet de la chimiothérapie et celui de la radiothérapie.

## Forces de l'étude

Notre étude présente plusieurs forces qui doivent être mises en valeur. Il s'agit, à l'heure actuelle, de la seule IPD-NMA sur le sujet. Nous avons été en mesure de récupérer les IPD de la plupart des essais. La disponibilité des IPD nous a permis d'analyser l'interaction avec les sous-groupes, ce qui constitue une des questions les plus importantes à l'heure actuelle. L'échantillon de grande taille permettait d'avoir la puissance pour identifier des interactions cliniquement pertinentes. De même, l'utilisation des méthodes de NMA nous a permis de mieux analyser la comparaison cliniquement la plus pertinente (CS-CRS) en combinant informations directes et indirectes. La méthodologie statistique que nous avons employée est rigoureuse, et fait appel aux derniers développements disponibles (imputations multiples, modèle IPD en une étape, réseaux additifs etc...). Plusieurs analyses de sensibilité ont été réalisées. Leurs résultats concordant avec l'analyse principale confirment que nos résultats ne sont pas dépendants du mode de représentation du réseau ou engendrés par des essais atypiques.

## Limites de l'étude

Plusieurs limites doivent être reconnues. Premièrement la période d'inclusion est longue, s'étendant sur plus de 20 ans. Au cours de cette période la prise en charge générale des patients s'est modifiée, ce qui pourrait être responsable de biais. Toutefois l'utilisation d'un modèle avec un risque de base « aléatoire » permet de limiter ce phénomène en permettant à chaque essai d'avoir son propre risque de base. De même au cours de cette période, les traitements se sont potentiellement modifiés. Même si toutes les chimiothérapies comprenaient un sel de platine, la seconde molécule, le nombre de cycles ou la dose ont pu varier au cours du temps. De même les doses, fractionnement et étalement de radiothérapie n'étaient pas constant. L'utilisation d'un modèle avec un effet traitement aléatoire a permis d'en limiter les conséquences en laissant varier l'effet traitement d'un essai à l'autre autour d'une valeur qui peut être considérée comme l'effet moyen du traitement. La totalité des essais n'a pas pu être récupérée, certains auteurs ne nous ayant jamais répondu. Toutefois, en raison de la taille de ces essais largement inférieure à celles des essais inclus, il est peu probable que l'inclusion de ceux-ci aurait pu être à l'origine d'une modification substantielle des résultats. De plus, une étude récente de notre équipe a montré que la qualité des essais des équipes « injoignable » était souvent moins bonne que celle des autres.(98)

Bien que notre étude présente globalement une puissance satisfaisante, celle-ci reste limitée pour la principale comparaison d'intérêt : CS-CRS. Pour cette comparaison, seulement 497 patients (10.0%) avaient été inclus dans la comparaison directe contre 4 488 (90%) dans les comparaisons indirectes qui apportent moins d'informations.

La distribution des modificateurs n'était pas équilibrée. En effet, le cancer de l'œsophage touche principalement les hommes. En pratique, seulement 20% des patients inclus étaient des femmes. De même, ces essais ont principalement inclus des cancers localisés dans l'œsophage thoracique et seulement 22% des tumeurs étaient localisées au niveau de la JOG. Les interactions observées avec le sexe et la localisation doivent donc être interprétées à la lumière de cette limite.

La qualité du recueil des critères de jugement peut également être mise en question. Si notre critère principal de jugement (la SG) est très robuste et probablement très bien renseigné, il n'en est pas de même pour tous les autres critères. Concernant la SSR et les localisations des récurrences, plusieurs essais n'ont pas rapporté de données sur ce critère. Si l'utilisation d'un landmark était nécessaire compte tenu des séquences de traitement, celui est potentiellement responsable d'une perte partielle d'information en compressant temporellement tous les événements précoces. Dans notre étude, 34% des événements de SSR sont survenus avant ce landmark. De même, diagnostiquer une récurrence nécessite un protocole de surveillance régulier des patients et des examens complémentaires performants. Il n'est pas certain que ce fut le cas de façon identique dans tous les essais.

Les autres critères doivent être interprétés avec encore plus de prudence. Notamment, la morbidité postopératoire n'a fait l'objet d'un véritable effort de standardisation qu'assez récemment(46) et n'est passé dans la pratique courante que plus tardivement. Dans cette étude, la variation du taux de morbidité postopératoire de moins de 5% à plus de 50% selon les essais, suggère que le recueil de cette donnée et/ou la classification n'a pas été identique. Il en va de même pour la qualité de la résection chirurgicale. Des travaux récents ont bien démontré l'intérêt d'une standardisation de l'examen anatomopathologique afin de le rendre reproductible. Ceci est d'autant plus vrai après traitement néoadjuvant et pourrait expliquer en partie les résultats décevants sur ce critère.

Enfin, les informations de toxicité ne sont probablement pas exploitables au même niveau que les classifications utilisées de nos jours.(47) Beaucoup d'essais ne nous ont pas fourni de données de toxicité. Quand c'était le cas, le grade de la toxicité était souvent manquant. Ces données doivent donc être considérées comme descriptives.

## Perspectives

Plusieurs essais thérapeutiques sont en cours (**Tableau 10**, page 67). Une fois les résultats de ceux-ci disponibles, ils pourraient être intégrés progressivement au réseau selon le principe de la « living network meta-analysis ».(99–101) Dans cette méthode, une veille régulière est assurée pour mettre à jour les résultats. Dans notre modèle fréquentiste, une difficulté pourrait être de contrôler le risque de première espèce.

D'un point de vue clinique, on pourrait considérer le résultat de ces traitements néoadjuvants comme décevant. En effet, même si la significativité statistique est très fortement atteinte, il n'en va pas de même pour la significativité clinique. Dans les MA deux à deux, où les méthodes de RMST sont plus facilement applicables, nous avons observé, à l'horizon cinq ans, qu'adjoindre à la chirurgie une chimiothérapie prolongeait la survie de 3,3 mois et une chimioradiothérapie de 3,9 mois en moyenne. Cet effet semble insuffisant, surtout quand il est mis en balance avec la durée d'administration de ces deux traitements qui est du même ordre de grandeur. De plus, le taux de survivants à 5 ans reste faible, de l'ordre de 30% selon les groupes, ce qui place ce cancer comme une des néoplasies les plus léthales pour une forme non métastatique. Les données préliminaires présentées à l'ASCO en 2021 de l'essai NeoAegis(91) qui utilise des protocoles de traitement que l'on considère comme le plus performant à l'heure actuelle (FLOT/CROSS) montre un taux de survie à 3 ans de l'ordre de 55%, supérieur à celui que nous avons observé, laissant entrevoir un peu d'espoir. De même, l'étude en population de Kamarajah(92) montrait un taux de survie à 5 ans de 36%.

Une autre piste d'amélioration pourrait être l'avènement de nouvelles molécules de chimiothérapie. Les études que nous avons incluses utilisaient toutes des protocoles de chimiothérapies conventionnelles. Le développement à partir de la période 2005-2010 des thérapies ciblées, puis à partir de la période 2015-2020 de l'immunothérapie pourrait apporter de nouvelles pistes capables d'améliorer ces résultats à long terme. Aucune n'est pourtant actuellement validée, même en phase métastatique. Le cetuximab a été testé dans une étude randomisée de phase 2/3 et semblait délétère.(102) Les premiers résultats avec les anti-angiogéniques sont dans l'ensemble décevants.(103) Deux protocoles d'immunothérapie (nivolumab et pembrolizumab) ont été testés en phase avancée avec des gains sur la survie globale.(104,105) Toutefois, le protocole de traitement par pembrolizumab nécessitait une expression de PD1 de 10 ou plus qui n'était présente que chez un patient sur deux. Le trastuzumab qui est indiqué pour les cancers de l'estomac surexprimant Her2 en phase avancée n'a pas encore été validé dans des stratégies péri-opératoire, mais un essai est en cours.(106)

D'un point de vue statistique, les données de la méta-analyse ont été utilisées pour évaluer la survie sans récurrence comme critère de substitution de la survie globale et feront l'objet d'une communication séparée.

## Chapitre 5. Conclusions

Dans toutes les analyses, la chirurgie d'emblée présentait une survie inférieure aux protocoles multimodaux. Elle ne devrait donc plus être proposée.

La chimiothérapie et la chimioradiothérapie préopératoires apportaient un bénéfice par rapport à la chirurgie d'emblée. Selon la méthode d'analyse, adjoindre de la radiothérapie (modèle additif) ou faire une séquence CRS plutôt que CS (modèle standard) semblait apporter un bénéfice limité de survie globale et ne diminuait pas le risque de récurrence locale.

Le type histologique ne faisait varier l'efficacité du traitement dans aucune comparaison et aucun modèle. Nos données ne sont pas en faveur d'une personnalisation du traitement en fonction de l'histologie. La localisation tumorale semblait faire varier l'efficacité avec un effet plus marqué pour les tumeurs localisées à la JOG que dans l'OT mais dans le même sens.

Le bénéfice était observé sur les différents paramètres oncologiques de survie (SG, SSR, survie spécifique, récurrence locale et à distance) mais n'améliorait pas le taux de R0. Bien que les données doivent être interprétées avec prudence, le gain de survie ne se faisait pas au prix d'une augmentation de la morbi-mortalité postopératoire.

D'un point de vue statistique, notre modèle basé sur les IPD en une étape permettait d'estimer les deux modèles avec des estimations proches de celles de l'approche en deux étapes basée sur les contrastes. Ce modèle permettait en outre d'estimer des interactions d'ajuster les estimations sur des facteurs de confusion et de profiter de l'imputation multiple.

# Références

1. Rouvière Henri. Anatomie humaine descriptive et topographique . Tome II . Tronc / H. Rouvière,... 9e éd. révisée. Paris: Masson et Cie; 1962. 1163 p.
2. Siewert JR, Hölscher AH, Becker K, Gössner W. [Cardia cancer: attempt at a therapeutically relevant classification]. *Chir Z Alle Geb Oper Medizen*. janv 1987;58(1):25-32.
3. Cancers [Internet]. [cité 7 août 2021]. Disponible sur: <https://www.santepubliquefrance.fr/maladies-et-traumatismes/cancers>
4. Globocan 2012 - Home [Internet]. [cité 29 mars 2016]. Disponible sur: <http://globocan.iarc.fr/Default.aspx>
5. Leong T, Smithers BM, Haustermans K, Michael M, Gebski V, Miller D, et al. TOPGEAR: A Randomized, Phase III Trial of Perioperative ECF Chemotherapy with or Without Preoperative Chemoradiation for Resectable Gastric Cancer: Interim Results from an International, Intergroup Trial of the AGITG, TROG, EORTC and CCTG. *Ann Surg Oncol*. 2017;24(8):2252-8.
6. Tang H, Tan L, Shen Y, Wang H, Lin M, Feng M, et al. CMISG1701: A multicenter prospective randomized phase III clinical trial comparing neoadjuvant chemoradiotherapy to neoadjuvant chemotherapy followed by minimally invasive esophagectomy in patients with locally advanced resectable esophageal squamous cell. *BMC Cancer* [Internet]. 2017;17(1). Disponible sur: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85021368329&doi=10.1186%2Fs12885-017-3446-7&partnerID=40&md5=4abd2399dc6fefe9849641017efc5a70>
7. Hoeppe J, Lordick F, Brunner T, Glatz T, Bronsert P, Rothling N, et al. ESOPEC: prospective randomized controlled multicenter phase III trial comparing perioperative chemotherapy (FLOT protocol) to neoadjuvant chemoradiation (CROSS protocol) in patients with adenocarcinoma of the esophagus (NCT02509286). *BMC Cancer*. juill 2016;16:503-503.
8. Reynolds JV, Preston SR, O'Neill B, Baeksgaard L, Griffin SM, Mariette C, et al. ICORG 10-14: NEOadjuvant trial in Adenocarcinoma of the oEsophagus and oesophagoGastric junction International Study (Neo-AEGIS). *BMC CANCER*. 2017;17.

9. Huang Y, Wang H, Luo G, Zhang Y, Wang L, Li K. A systematic review and network meta-analysis of neoadjuvant therapy combined with surgery for patients with resectable esophageal squamous cell carcinoma. *Int J Surg Lond Engl*. févr 2017;38:41-7.
10. Ronellenfitch U, Schwarzbach M, Hofheinz R, Kienle P, Kieser M, Slinger TE, et al. Perioperative chemo(radio)therapy versus primary surgery for resectable adenocarcinoma of the stomach, gastroesophageal junction, and lower esophagus. *Cochrane Database Syst Rev*. 31 mai 2013;(5):CD008107.
11. Kidane B, Coughlin S, Vogt K, Malthaner R. Preoperative chemotherapy for resectable thoracic esophageal cancer. *Cochrane Database Syst Rev*. 19 mai 2015;(5):CD001556.
12. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med*. 30 août 2002;21(16):2313-24.
13. Simmonds M, Stewart G, Stewart L. A decade of individual participant data meta-analyses: A review of current practice. *Contemp Clin Trials*. nov 2015;45(Pt A):76-83.
14. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br J Cancer*. janv 1977;35(1):1-39.
15. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis*. avr 1985;27(5):335-71.
16. Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, Stewart LA. Statistical analysis of individual participant data meta-analyses: a comparison of methods and recommendations for practice. *PloS One*. 2012;7(10):e46042.
17. Bowden J, Tierney JF, Simmonds M, Copas AJ, Higgins JP. 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. *Res Synth Methods*. sept 2011;2(3):150-62.
18. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med*. 20 déc 2010;29(29):3046-67.
19. Freeman SC, Carpenter JR. Bayesian one-step IPD network meta-analysis of time-to-event data using Royston-Parmar models. *Res Synth Methods*. déc 2017;8(4):451-64.
20. McGilchrist CA. REML estimation for survival models with frailty. *Biometrics*. mars 1993;49(1):221-5.
21. Sargent DJ. A general framework for random effects survival analysis in the Cox proportional hazards setting. *Biometrics*. déc 1998;54(4):1486-97.
22. Crowther MJ, Riley RD, Staessen JA, Wang J, Gueyffier F, Lambert PC. Individual patient data meta-analysis of survival data using Poisson regression models. *BMC Med Res Methodol*. 23 mars 2012;12:34.



## Références

23. Tu Y-K. Use of generalized linear mixed models for network meta-analysis. *Med Decis Mak Int J Soc Med Decis Mak.* oct 2014;34(7):911-8.
24. Fisher DJ, Copas AJ, Tierney JF, Parmar MKB. A critical review of methods for the assessment of patient-level interactions in individual participant data meta-analysis of randomized trials, and guidance for practitioners. *J Clin Epidemiol.* sept 2011;64(9):949-67.
25. Hua H, Burke DL, Crowther MJ, Ensor J, Tudur Smith C, Riley RD. One-stage individual participant data meta-analysis models: estimation of treatment-covariate interactions must avoid ecological bias by separating out within-trial and across-trial information. *Stat Med.* 28 2017;36(5):772-89.
26. Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res.* juin 2008;17(3):279-301.
27. Rücker G. Network meta-analysis, electrical networks and graph theory. *Res Synth Methods.* déc 2012;3(4):312-24.
28. Rücker G, Petropoulou M, Schwarzer G. Network meta-analysis of multicomponent interventions. *Biom J Biom Z.* mai 2020;62(3):808-21.
29. Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. *Value Health J Int Soc Pharmacoeconomics Outcomes Res.* oct 2008;11(5):956-64.
30. Pignon JP, Hill C. Meta-analyses of randomised clinical trials in oncology. *Lancet Oncol.* août 2001;2(8):475-82.
31. Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing the consistency assumption by exploring treatment by covariate interactions in mixed treatment comparison meta-analysis: individual patient-level covariates versus aggregate trial-level covariates. *Stat Med.* 20 déc 2012;31(29):3840-57.
32. Pignon J-P, le Maître A, Maillard E, Bourhis J, MACH-NC Collaborative Group. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* juill 2009;92(1):4-14.
33. Bourhis J, Overgaard J, Audry H, Ang KK, Saunders M, Bernier J, et al. Hyperfractionated or accelerated radiotherapy in head and neck cancer: a meta-analysis. *Lancet Lond Engl.* 2 sept 2006;368(9538):843-54.
34. Lacas B, Bourhis J, Overgaard J, Zhang Q, Grégoire V, Nankivell M, et al. Role of radiotherapy fractionation in head and neck cancers (MARCH): an updated meta-analysis. *Lancet Oncol.* 2017;18(9):1221-37.
35. Ollier E, Blanchard P, Teuff GL, Michiels S. Penalized Poisson model for network meta-analysis of individual patient time-to-event data. *ArXiv210300069 Stat [Internet].* 26 févr 2021 [cité 3 mai 2021]; Disponible sur: <http://arxiv.org/abs/2103.00069>

36. Jackson D, Law M, Stijnen T, Viechtbauer W, White IR. A comparison of seven random-effects models for meta-analyses that estimate the summary odds ratio. *Stat Med.* 30 2018;37(7):1059-85.
37. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med.* 30 mars 2010;29(7-8):932-44.
38. Günhan BK, Friede T, Held L. A design-by-treatment interaction model for network meta-analysis and meta-regression with integrated nested Laplace approximations. *Res Synth Methods.* juin 2018;9(2):179-94.
39. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Stat Med.* 30 juin 2009;28(14):1861-81.
40. Salanti G, Marinho V, Higgins JPT. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *J Clin Epidemiol.* août 2009;62(8):857-64.
41. Donegan S, Welton NJ, Tudur Smith C, D'Alessandro U, Dias S. Network meta-analysis including treatment by covariate interactions: Consistency can vary across covariate values. *Res Synth Methods.* déc 2017;8(4):485-95.
42. Tonin FS, Borba HH, Mendes AM, Wiens A, Fernandez-Llimos F, Pontarolo R. Description of network meta-analysis geometry: A metrics design study. *PloS One.* 2019;14(2):e0212650.
43. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. *JAMA.* 28 avr 2015;313(16):1657.
44. Sposto R, Stablein D, Carter-Campbell S. A partially grouped logrank test. *Stat Med.* 30 mars 1997;16(6):695-704.
45. Early Breast Cancer Trialists' Collaborative Group. Treatment of early breast cancer / Early Breast Cancer Trialists' Collaborative Group (EBCTCG). New York: Oxford University Press; 1990. (Oxford medical publications).
46. Dindo D, Demartines N, Clavien P-A. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg.* août 2004;240(2):205-13.
47. Common Terminology Criteria for Adverse Events (CTCAE) | Protocol Development | CTEP [Internet]. [cité 26 oct 2020]. Disponible sur: [https://ctep.cancer.gov/protocolDevelopment/electronic\\_applications/ctc.htm](https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm)
48. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 15 juin 2002;21(11):1539-58.
49. Lueza B, Rotolo F, Bonastre J, Pignon J-P, Michiels S. Bias and precision of methods for estimating the difference in restricted mean survival time from an individual patient

## Références

- data meta-analysis. *BMC Med Res Methodol* [Internet]. 29 mars 2016 [cité 22 déc 2017];16. Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4812643/>
50. Rotolo F, Michiels S. Testing the treatment effect on competing causes of death in oncology clinical trials. *BMC Med Res Methodol*. 29 mai 2014;14:72.
  51. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol*. 31 juill 2015;15:58.
  52. Roth JA, Pass HI, Flanagan MM, Graeber GM, Rosenberg JC, Steinberg S. Randomized clinical trial of preoperative and postoperative adjuvant chemotherapy with cisplatin, vindesine, and bleomycin for carcinoma of the esophagus. *J Thorac Cardiovasc Surg*. août 1988;96(2):242-8.
  53. Nygaard K, Hagen S, Hansen HS, Hatlevoll R, Hultborn R, Jakobsen A, et al. Pre-operative radiotherapy prolongs survival in operable esophageal carcinoma: a randomized, multicenter study of pre-operative radiotherapy and chemotherapy. The second Scandinavian trial in esophageal cancer. *World J Surg*. déc 1992;16(6):1104-9; discussion 1110.
  54. Maipang T, Vasinanukorn P, Petpichetchian C, Chamroonkul S, Geater A, Chansawwaang S, et al. Induction chemotherapy in the treatment of patients with carcinoma of the esophagus. *J Surg Oncol*. juill 1994;56(3):191-7.
  55. Law S, Fok M, Chow S, Chu KM, Wong J. Preoperative chemotherapy versus surgical therapy alone for squamous cell carcinoma of the esophagus: a prospective randomized trial. *J Thorac Cardiovasc Surg*. août 1997;114(2):210-7.
  56. Boonstra JJ, Kok TC, Wijnhoven BP, van Heijl M, van Berge Henegouwen MI, Ten Kate FJ, et al. Chemotherapy followed by surgery versus surgery alone in patients with resectable oesophageal squamous cell carcinoma: long-term results of a randomized controlled trial. *BMC Cancer*. 19 mai 2011;11:181.
  57. Kelsen DP, Winter KA, Gunderson LL, Mortimer J, Estes NC, Haller DG, et al. Long-term results of RTOG trial 8911 (USA Intergroup 113): a random assignment trial comparison of chemotherapy followed by surgery compared with surgery alone for esophageal cancer. *J Clin Oncol Off J Am Soc Clin Oncol*. août 2007;25(24):3719-25.
  58. Kelsen DP, Ginsberg R, Pajak TF, Sheahan DG, Gunderson L, Mortimer J, et al. Chemotherapy followed by surgery compared with surgery alone for localized esophageal cancer. *N Engl J Med*. 31 déc 1998;339(27):1979-84.
  59. Ancona E, Ruol A, Santi S, Merigliano S, Sileni VC, Koussis H, et al. Only pathologic complete response to neoadjuvant chemotherapy improves significantly the long term survival of patients with resectable esophageal squamous cell carcinoma: final report of a randomized, controlled trial of preoperative chemotherapy versus surgery alone. *Cancer*. 1 juin 2001;91(11):2165-74.
  60. Allum WH, Stenning SP, Bancewicz J, Clark PI, Langley RE. Long-term results of a randomized trial of surgery with or without preoperative chemotherapy in esophageal cancer. *J Clin Oncol Off J Am Soc Clin Oncol*. 20 oct 2009;27(30):5062-7.

61. Medical Research Council Oesophageal Cancer Working Group. Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. *Lancet Lond Engl*. 18 mai 2002;359(9319):1727-33.
62. Cunningham D, Allum WH, Stenning SP, Thompson JN, Van de Velde CJH, Nicolson M, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med*. 6 juill 2006;355(1):11-20.
63. Ychou M, Boige V, Pignon J-P, Conroy T, Bouché O, Lebreton G, et al. Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: an FNCLCC and FFCO multicenter phase III trial. *J Clin Oncol Off J Am Soc Clin Oncol*. 1 mai 2011;29(13):1715-21.
64. Schuhmacher C, Gretschel S, Lordick F, Reichardt P, Hohenberger W, Eisenberger CF, et al. Neoadjuvant chemotherapy compared with surgery alone for locally advanced cancer of the stomach and cardia: European Organisation for Research and Treatment of Cancer randomized trial 40954. *J Clin Oncol Off J Am Soc Clin Oncol*. déc 2010;28(35):5210-8.
65. Schlag PM. Randomized trial of preoperative chemotherapy for squamous cell cancer of the esophagus. The Chirurgische Arbeitsgemeinschaft Fuer Onkologie der Deutschen Gesellschaft Fuer Chirurgie Study Group. *Arch Surg Chic Ill* 1960. déc 1992;127(12):1446-50.
66. Wang XL, Wu GX, Zhang MD, Guo M, Zhang H, Sun XF. A favorable impact of preoperative FPLC chemotherapy on patients with gastric cardia cancer. *Oncol Rep*. 2000;7(2):241-4.
67. Wang C, Ding T, Chang L. [A randomized clinical study of preoperative chemotherapy for esophageal carcinoma]. *Zhonghua Zhong Liu Za Zhi*. mai 2001;23(3):254-5.
68. Baba M, Natsugoe S, Shimada M, Nakano S, Kusano C, Fukumoto T, et al. Prospective evaluation of preoperative chemotherapy in resectable squamous cell carcinoma of the thoracic esophagus. *Dis Esophagus Off J Int Soc Dis Esophagus*. 2000;13(2):136-41.
69. Le Prise E, Etienne PL, Meunier B, Maddern G, Ben Hassel M, Gedouin D, et al. A randomized study of chemotherapy, radiation therapy, and surgery versus surgery for localized squamous cell carcinoma of the esophagus. *Cancer*. 1 avr 1994;73(7):1779-84.
70. Urba SG, Orringer MB, Turrisi A, Iannettoni M, Forastiere A, Strawderman M. Randomized trial of preoperative chemoradiation versus surgery alone in patients with locoregional esophageal carcinoma. *J Clin Oncol Off J Am Soc Clin Oncol*. 15 janv 2001;19(2):305-13.
71. Bosset JF, Gignoux M, Triboulet JP, Tiret E, Manton G, Elias D, et al. Chemoradiotherapy followed by surgery compared with surgery alone in squamous-cell cancer of the esophagus. *N Engl J Med*. 17 juill 1997;337(3):161-7.
72. Walsh TN, Noonan N, Hollywood D, Kelly A, Keeling N, Hennessy TP. A comparison of multimodal therapy and surgery for esophageal adenocarcinoma. *N Engl J Med*. 15 août 1996;335(7):462-7.

## Références

73. Walsh TN, Grennell M, Mansoor S, Kelly A. Neoadjuvant treatment of advanced stage esophageal adenocarcinoma increases survival. *Dis Esophagus Off J Int Soc Dis Esophagus*. 2002;15(2):121-4.
74. Bass GA, Furlong H, O'Sullivan KE, Hennessy TPJ, Walsh TN. Chemoradiotherapy, with adjuvant surgery for local control, confers a durable survival advantage in adenocarcinoma and squamous cell carcinoma of the oesophagus. *Eur J Cancer Oxf Engl* 1990. avr 2014;50(6):1065-75.
75. Burmeister BH, Smithers BM, GebSKI V, Fitzgerald L, Simes RJ, Devitt P, et al. Surgery alone versus chemoradiotherapy followed by surgery for resectable cancer of the oesophagus: a randomised controlled phase III trial. *Lancet Oncol*. sept 2005;6(9):659-68.
76. Tepper J, Krasna MJ, Niedzwiecki D, Hollis D, Reed CE, Goldberg R, et al. Phase III trial of trimodality therapy with cisplatin, fluorouracil, radiotherapy, and surgery compared with surgery alone for esophageal cancer: CALGB 9781. *J Clin Oncol Off J Am Soc Clin Oncol*. 1 mars 2008;26(7):1086-92.
77. Lee J-L, Park SI, Kim S-B, Jung H-Y, Lee GH, Kim J-H, et al. A single institutional phase III trial of preoperative chemotherapy with hyperfractionation radiotherapy plus surgery versus surgery alone for resectable esophageal squamous cell carcinoma. *Ann Oncol Off J Eur Soc Med Oncol*. juin 2004;15(6):947-54.
78. Mariette C, Dahan L, Mornex F, Maillard E, Thomas P-A, Meunier B, et al. Surgery alone versus chemoradiotherapy followed by surgery for stage I and II esophageal cancer: final analysis of randomized controlled phase III trial FFCO 9901. *J Clin Oncol Off J Am Soc Clin Oncol*. 10 août 2014;32(23):2416-22.
79. van Hagen P, Hulshof MCCM, van Lanschot JJB, Steyerberg EW, van Berge Henegouwen MI, Wijnhoven BPL, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med*. mai 2012;366(22):2074-84.
80. Shapiro J, van Lanschot JJB, Hulshof MCCM, van Hagen P, van Berge Henegouwen MI, Wijnhoven BPL, et al. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. *Lancet Oncol*. sept 2015;16(9):1090-8.
81. YANG H, FU J, LIU M, FANG W, WANG J, CHEN Y, et al. [A multi-centered randomized controlled study of neo-adjuvant chemoradiotherapy followed by surgery versus surgery alone for locally advanced squamous cell carcinoma of esophagus: an interim analysis]. *Zhonghua Yi Xue Za Zhi*. avr 2012;92(15):1028-32.
82. Apinop C, Puttisak P, Preecha N. A prospective study of combined therapy in esophageal cancer. *Hepatogastroenterology*. août 1994;41(4):391-3.
83. An F, Huang J, Xie Y, Chen S, Rong T. [A prospective study of combined chemoradiotherapy followed by surgery in the treatment of esophageal carcinoma]. *Zhonghua Zhong Liu Za Zhi*. juill 2003;25(4):376-9.

84. Lv J, Cao XF, Zhu B, Ji L, Tao L, Wang DD. Long-term efficacy of perioperative chemoradiotherapy on esophageal squamous cell carcinoma. *World J Gastroenterol.* 7 avr 2010;16(13):1649-54.
85. Peng L, Xie T-P, Han Y-T, Lang J-Y, Li T, Fu B-Y, et al. Randomized controlled study on preoperative concurrent chemoradiotherapy versus surgery alone for esophageal squamous cell carcinoma. *Tumor.* 2008;28(7):620-2.
86. Zhao Q, Li Y, Wang J, Zhang J, Qiao X, Tan B, et al. Concurrent Neoadjuvant Chemoradiotherapy for Siewert II and III Adenocarcinoma at Gastroesophageal Junction. *Am J Med Sci.* juin 2015;349(6):472-6.
87. Stahl M, Walz MK, Stuschke M, Lehmann N, Meyer H-J, Riera-Knorrenschild J, et al. Phase III comparison of preoperative chemotherapy compared with chemoradiotherapy in patients with locally advanced adenocarcinoma of the esophagogastric junction. *J Clin Oncol Off J Am Soc Clin Oncol.* 20 févr 2009;27(6):851-6.
88. Stahl M, Walz MK, Riera-Knorrenschild J, Stuschke M, Sandermann A, Bitzer M, et al. Preoperative chemotherapy versus chemoradiotherapy in locally advanced adenocarcinomas of the oesophagogastric junction (POET): Long-term results of a controlled randomised trial. *Eur J Cancer Oxf Engl* 1990. août 2017;81:183-90.
89. Burmeister BH, Thomas JM, Burmeister EA, Walpole ET, Harvey JA, Thomson DB, et al. Is concurrent radiation therapy required in patients receiving preoperative chemotherapy for adenocarcinoma of the oesophagus? A randomised phase II trial. *Eur J Cancer Oxf Engl* 1990. févr 2011;47(3):354-60.
90. Klevebro F, Alexandersson von Döbeln G, Wang N, Johnsen G, Jacobsen A-B, Friesland S, et al. A randomized clinical trial of neoadjuvant chemotherapy versus neoadjuvant chemoradiotherapy for cancer of the oesophagus or gastro-oesophageal junction. *Ann Oncol Off J Eur Soc Med Oncol.* avr 2016;27(4):660-7.
91. Reynolds JV, Preston SR, O'Neill B, Lowery MA, Baeksgaard L, Crosby T, et al. Neo-AEGIS (Neoadjuvant trial in Adenocarcinoma of the Esophagus and Esophago-Gastric Junction International Study): Preliminary results of phase III RCT of CROSS versus perioperative chemotherapy (Modified MAGIC or FLOT protocol). (NCT01726452). *J Clin Oncol.* 20 mai 2021;39(15\_suppl):4004-4004.
92. Kamarajah SK, Phillips AW, Ferri L, Hofstetter WL, Markar SR. Neoadjuvant chemoradiotherapy or chemotherapy alone for oesophageal cancer: population-based cohort study. *Br J Surg.* 30 avr 2021;108(4):403-11.
93. Blanchard P, Lee A, Marguet S, Leclercq J, Ng WT, Ma J, et al. Chemotherapy and radiotherapy in nasopharyngeal carcinoma: an update of the MAC-NPC meta-analysis. *Lancet Oncol.* juin 2015;16(6):645-55.
94. Abraha I, Aristei C, Palumbo I, Lupattelli M, Trastulli S, Cirocchi R, et al. Preoperative radiotherapy and curative surgery for the management of localised rectal carcinoma. *Cochrane Database Syst Rev.* 3 oct 2018;2018(10):CD002102.

## Références

95. Werner RN, Gaskins M, Avila Valle G, Budach V, Koswig S, Mosthaf FA, et al. State of the art treatment for stage I to III anal squamous cell carcinoma: A systematic review and meta-analysis. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* avr 2021;157:188-96.
96. Mariette C. What is the optimal neoadjuvant treatment for locally advanced oesophageal adenocarcinoma? *Ann Oncol Off J Eur Soc Med Oncol.* 1 mars 2017;28(3):447-50.
97. Rubin DB. *Multiple Imputation for Nonresponse in Surveys.* Hoboken (New Jersey): John Wiley & Sons; 2004. 320 p. (Wiley Classics Library).
98. Fayard F, Petit C, Lacas B, Pignon JP. Impact of missing individual patient data on 18 meta-analyses of randomised trials in oncology: Gustave Roussy experience. *BMJ Open.* 13 août 2018;8(8):e020499.
99. Créquit P, Trinquart L, Yavchitz A, Ravaud P. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: the example of lung cancer. *BMC Med.* 20 janv 2016;14:8.
100. Vandvik PO, Brignardello-Petersen R, Guyatt GH. Living cumulative network meta-analysis to reduce waste in research: A paradigmatic shift for systematic reviews? *BMC Med.* 29 mars 2016;14:59.
101. Créquit P, Martin-Montoya T, Attiche N, Trinquart L, Vivot A, Ravaud P. Living network meta-analysis was feasible when considering the pace of evidence generation. *J Clin Epidemiol.* avr 2019;108:10-6.
102. Crosby T, Hurt CN, Falk S, Gollins S, Mukherjee S, Staffurth J, et al. Chemoradiotherapy with or without cetuximab in patients with oesophageal cancer (SCOPE1): a multicentre, phase 2/3 randomised trial. *Lancet Oncol.* juin 2013;14(7):627-37.
103. Ku GY, Bains MS, Park DJ, Janjigian YY, Rusch VW, Rizk NP, et al. Phase II study of bevacizumab and preoperative chemoradiation for esophageal adenocarcinoma. *J Gastrointest Oncol.* déc 2016;7(6):828-37.
104. Kojima T, Shah MA, Muro K, Francois E, Adenis A, Hsu C-H, et al. Randomized Phase III KEYNOTE-181 Study of Pembrolizumab Versus Chemotherapy in Advanced Esophageal Cancer. *J Clin Oncol Off J Am Soc Clin Oncol.* 10 déc 2020;38(35):4138-48.
105. Kato K, Cho BC, Takahashi M, Okada M, Lin C-Y, Chin K, et al. Nivolumab versus chemotherapy in patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy (ATTRACTION-3): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol.* nov 2019;20(11):1506-17.
106. Wagner AD, Grabsch HI, Mauer M, Marreaud S, Caballero C, Thuss-Patience P, et al. EORTC-1203-GIICG - the « INNOVATION »-trial: Effect of chemotherapy alone versus chemotherapy plus trastuzumab, versus chemotherapy plus trastuzumab plus pertuzumab, in the perioperative treatment of HER2 positive, gastric and gastroesophageal junction adenocarcinoma on pathologic response rate: a randomized phase II-intergroup trial of the EORTC-Gastrointestinal Tract Cancer Group, Korean Cancer Study Group and Dutch Upper GI-Cancer group. *BMC Cancer.* 24 mai 2019;19(1):494.

# Annexes

## Partie simulation



Tableau annexe 1 : Résultats des simulations pour les deux effets traitements estimés directement

Configuration 1																									Configuration 2									
					A-C						B-C																							
					IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta							
coef	$\sigma$	$\tau$	s	param	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES								
-0,5	0,01	0,01	1	<55	-0,003	0,103	NA	NA	0,001	0,053	-0,007	0,104	NA	NA	0	0,052	-0,004	0,095	NA	NA	-0,002	0,047	0,001	0,095	NA	NA	0,001	0,046						
-0,5	0,01	0,01	1	55-60	-0,001	0,108	NA	NA	0,001	0,053	0,006	0,106	NA	NA	0	0,052	-0,006	0,098	NA	NA	-0,002	0,047	-0,003	0,1	NA	NA	0,001	0,046						
-0,5	0,01	0,01	1	60-65	-0,004	0,107	NA	NA	0,001	0,053	-0,004	0,11	NA	NA	0	0,052	0,001	0,094	NA	NA	-0,002	0,047	0,003	0,096	NA	NA	0,001	0,046						
-0,5	0,01	0,01	1	>65	0,001	0,103	NA	NA	0,001	0,053	0	0,098	NA	NA	0	0,052	0	0,09	NA	NA	-0,002	0,047	0,002	0,087	NA	NA	0,001	0,046						
-0,5	0,01	0,01	1	TE <sub>age=60</sub>	-0,001	0,053	0,001	0,055	0,001	0,053	-0,001	0,052	0	0,057	0	0,052	-0,002	0,047	-0,001	0,05	-0,002	0,047	0,001	0,046	0,001	0,049	0,001	0,046						
-0,5	0,01	0,01	1	vTE <sub>age+1</sub>	0	0,007	0,004	0,109	0	0	0	0,006	-0,006	0,115	0	0	0	0,006	0,002	0,097	0	0	0	0,006	0,003	0,096	0	0						
-0,5	0,01	0,01	2	<55	0,004	0,122	NA	NA	0,161	0,062	-0,003	0,107	NA	NA	0,001	0,055	0,002	0,105	NA	NA	0,16	0,053	0	0,095	NA	NA	0,003	0,048						
-0,5	0,01	0,01	2	55-60	-0,003	0,125	NA	NA	0,044	0,062	-0,003	0,106	NA	NA	0,001	0,055	0,001	0,115	NA	NA	0,043	0,053	-0,004	0,1	NA	NA	0,003	0,048						
-0,5	0,01	0,01	2	60-65	0,006	0,125	NA	NA	-0,031	0,062	0	0,109	NA	NA	0,001	0,055	0	0,111	NA	NA	-0,032	0,053	0,002	0,099	NA	NA	0,003	0,048						
-0,5	0,01	0,01	2	>65	0,002	0,115	NA	NA	-0,148	0,062	0	0,104	NA	NA	0,001	0,055	0,006	0,097	NA	NA	-0,149	0,053	0,003	0,09	NA	NA	0,003	0,048						
-0,5	0,01	0,01	2	TE <sub>age=60</sub>	0,002	0,061	0,006	0,065	0,007	0,062	-0,001	0,054	0,001	0,058	0,001	0,055	0,002	0,053	0,005	0,056	0,006	0,053	0,001	0,047	0,003	0,05	0,003	0,048						
-0,5	0,01	0,01	2	vTE <sub>age+1</sub>	0	0,008	-0,002	0,126	-0,016	0	0	0,007	-0,005	0,114	0	0	0	0,007	-0,003	0,107	-0,016	0	0	0,006	0	0,102	0	0						
-0,5	0,01	0,01	3	<55	-0,035	0,081	NA	NA	0,037	0,063	-0,004	0,242	NA	NA	0,003	0,052	-0,037	0,077	NA	NA	0,062	0,054	-0,028	0,157	NA	NA	-0,017	0,048						
-0,5	0,01	0,01	3	55-60	-0,004	0,135	NA	NA	-0,08	0,063	0,002	0,169	NA	NA	0,003	0,052	-0,002	0,12	NA	NA	-0,054	0,054	0,009	0,137	NA	NA	-0,017	0,048						
-0,5	0,01	0,01	3	60-65	-0,011	0,175	NA	NA	-0,155	0,063	-0,002	0,117	NA	NA	0,003	0,052	-0,005	0,146	NA	NA	-0,13	0,054	0	0,109	NA	NA	-0,017	0,048						
-0,5	0,01	0,01	3	>65	-0,032	0,254	NA	NA	-0,272	0,063	0,002	0,065	NA	NA	0,003	0,052	-0,01	0,162	NA	NA	-0,246	0,054	0	0,063	NA	NA	-0,017	0,048						
-0,5	0,01	0,01	3	TE <sub>age=60</sub>	-0,002	0,083	-0,026	1,108	-0,117	0,063	0,001	0,074	0,016	0,871	0,003	0,052	0,001	0,066	0,041	0,543	-0,092	0,054	0,002	0,062	0,04	0,546	-0,017	0,048						
-0,5	0,01	0,01	3	vTE <sub>age+1</sub>	0	0,008	-0,004	0,138	-0,016	0	0	0,007	-0,002	0,109	0	0	0	0,006	0,004	0,068	-0,016	0	0	0,006	-0,005	0,068	0	0						
-0,5	0,01	0,1	1	<55	-0,001	0,108	NA	NA	0,003	0,062	0,001	0,11	NA	NA	0,005	0,065	0,008	0,1	NA	NA	0,003	0,056	0,007	0,097	NA	NA	0,004	0,055						
-0,5	0,01	0,1	1	55-60	0,001	0,109	NA	NA	0,003	0,062	0	0,111	NA	NA	0,005	0,065	0,001	0,102	NA	NA	0,003	0,056	0,003	0,104	NA	NA	0,004	0,055						
-0,5	0,01	0,1	1	60-65	0,001	0,112	NA	NA	0,003	0,062	0,008	0,117	NA	NA	0,005	0,065	-0,002	0,101	NA	NA	0,003	0,056	0,005	0,1	NA	NA	0,004	0,055						
-0,5	0,01	0,1	1	>65	0,001	0,109	NA	NA	0,003	0,062	0,002	0,108	NA	NA	0,005	0,065	-0,001	0,091	NA	NA	0,003	0,056	-0,003	0,094	NA	NA	0,004	0,055						
-0,5	0,01	0,1	1	TE <sub>age=60</sub>	0,001	0,061	0,004	0,066	0,003	0,062	0,003	0,064	0,003	0,069	0,005	0,065	0,002	0,056	0,004	0,057	0,003	0,056	0,003	0,055	0,003	0,057	0,004	0,055						
-0,5	0,01	0,1	1	vTE <sub>age+1</sub>	0	0,007	-0,005	0,129	0	0	0	0,007	0,003	0,131	0	0	0	0,006	0,002	0,111	0	0	0	0,006	-0,003	0,112	0	0						
-0,5	0,01	0,1	2	<55	0,012	0,126	NA	NA	0,164	0,069	0	0,114	NA	NA	0,003	0,062	0,002	0,108	NA	NA	0,16	0,06	-0,001	0,102	NA	NA	0	0,055						
-0,5	0,01	0,1	2	55-60	0,001	0,13	NA	NA	0,048	0,069	0	0,112	NA	NA	0,003	0,062	0,004	0,118	NA	NA	0,043	0,06	-0,005	0,105	NA	NA	0	0,055						
-0,5	0,01	0,1	2	60-65	0	0,127	NA	NA	-0,028	0,069	0,006	0,113	NA	NA	0,003	0,062	0,001	0,112	NA	NA	-0,032	0,06	-0,002	0,101	NA	NA	0	0,055						
-0,5	0,01	0,1	2	>65	0,007	0,115	NA	NA	-0,145	0,069	-0,006	0,107	NA	NA	0,003	0,062	0	0,103	NA	NA	-0,149	0,06	-0,004	0,093	NA	NA	0	0,055						
-0,5	0,01	0,1	2	TE <sub>age=60</sub>	0,005	0,07	0,009	0,074	0,01	0,069	0	0,061	0,002	0,066	0,003	0,062	0,002	0,061	0,005	0,065	0,005	0,06	-0,003	0,056	0	0,059	0	0,055						
-0,5	0,01	0,1	2	vTE <sub>age+1</sub>	0	0,007	-0,002	0,147	-0,016	0	0	0,007	0	0,13	0	0	0	0,007	0,003	0,122	-0,016	0	0	0,006	0,002	0,119	0	0						
-0,5	0,01	0,1	3	<55	-0,038	0,085	NA	NA	0,036	0,068	-0,007	0,252	NA	NA	0,003	0,062	-0,03	0,085	NA	NA	0,065	0,063	-0,019	0,163	NA	NA	-0,018	0,056						
-0,5	0,01	0,1	3	55-60	-0,008	0,139	NA	NA	-0,081	0,068	-0,001	0,168	NA	NA	0,003	0,062	-0,009	0,128	NA	NA	-0,052	0,063	-0,002	0,137	NA	NA	-0,018	0,056						
-0,5	0,01	0,1	3	60-65	-0,005	0,178	NA	NA	-0,157	0,068	-0,004	0,123	NA	NA	0,003	0,062	-0,008	0,146	NA	NA	-0,128	0,063	0,003	0,114	NA	NA	-0,018	0,056						
-0,5	0,01	0,1	3	>65	-0,029	0,256	NA	NA	-0,273	0,068	0,002	0,072	NA	NA	0,003	0,062	-0,001	0,155	NA	NA	-0,244	0,063	-0,001	0,07	NA	NA	-0,018	0,056						
-0,5	0,01	0,1	3	TE <sub>age=60</sub>	-0,001	0,087	-0,004	1,21	-0,119	0,068	-0,003	0,082	-0,033	1,044	0,003	0,062	0,003	0,071	-0,017	0,607	-0,09	0,063	0,002	0,067	-0,016	0,612	-0,018	0,056						
-0,5	0,01	0,1	3	vTE <sub>age+1</sub>	0	0,008	-0,001	0,151	-0,016	0	0	0,007	0,004	0,13	0	0	0	0,006	-0,003	0,076	-0,016	0	0	0,005	0,002	0,077	0	0						

Configuration 1																	Configuration 2														
					A-C						B-C											A-C					B-C				
coef	$\sigma$	$\tau$	s	param	IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta				
					Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	
-0,5	0,1	0,01	1	<55	0,001	0,104	NA	NA	0,007	0,053	0,004	0,109	NA	NA	0,003	0,055	0,001	0,093	NA	NA	0,002	0,049	-0,002	0,088	NA	NA	0,001	0,047			
-0,5	0,1	0,01	1	55-60	0,005	0,114	NA	NA	0,007	0,053	0,002	0,107	NA	NA	0,003	0,055	0,005	0,094	NA	NA	0,002	0,049	0,001	0,097	NA	NA	0,001	0,047			
-0,5	0,1	0,01	1	60-65	0,007	0,112	NA	NA	0,007	0,053	-0,007	0,113	NA	NA	0,003	0,055	0,002	0,099	NA	NA	0,002	0,049	0,003	0,094	NA	NA	0,001	0,047			
-0,5	0,1	0,01	1	>65	0,008	0,098	NA	NA	0,007	0,053	0,006	0,098	NA	NA	0,003	0,055	0	0,091	NA	NA	0,002	0,049	0,001	0,09	NA	NA	0,001	0,047			
-0,5	0,1	0,01	1	TE <sub>age=60</sub>	0,006	0,053	0,007	0,057	0,007	0,053	0,002	0,055	0,002	0,058	0,003	0,055	0,002	0,049	0,002	0,051	0,002	0,049	0,001	0,047	0,002	0,05	0,001	0,047			
-0,5	0,1	0,01	1	vTE <sub>age+1</sub>	0	0,007	-0,004	0,112	0	0	0	0,006	0,003	0,107	0	0	0	0,006	0	0,101	0	0	0	0,006	-0,004	0,101	0	0			
-0,5	0,1	0,01	2	<55	0	0,121	NA	NA	0,159	0,06	0,002	0,107	NA	NA	0,002	0,051	0,007	0,105	NA	NA	0,158	0,054	0,007	0,095	NA	NA	0,003	0,049			
-0,5	0,1	0,01	2	55-60	0,001	0,119	NA	NA	0,042	0,06	-0,003	0,113	NA	NA	0,002	0,051	-0,01	0,111	NA	NA	0,041	0,054	-0,003	0,102	NA	NA	0,003	0,049			
-0,5	0,1	0,01	2	60-65	-0,002	0,125	NA	NA	-0,034	0,06	0,001	0,109	NA	NA	0,002	0,051	0,002	0,104	NA	NA	-0,034	0,054	0	0,1	NA	NA	0,003	0,049			
-0,5	0,1	0,01	2	>65	0,001	0,114	NA	NA	-0,15	0,06	0,001	0,099	NA	NA	0,002	0,051	0,002	0,102	NA	NA	-0,151	0,054	-0,004	0,089	NA	NA	0,003	0,049			
-0,5	0,1	0,01	2	TE <sub>age=60</sub>	0	0,06	0,002	0,064	0,004	0,06	0,001	0,052	0,003	0,056	0,002	0,051	0,001	0,054	0,002	0,059	0,004	0,054	0	0,05	0,002	0,052	0,003	0,049			
-0,5	0,1	0,01	2	vTE <sub>age+1</sub>	0	0,008	-0,003	0,132	-0,016	0	0	0,007	0,003	0,117	0	0	0	0,007	-0,002	0,109	-0,016	0	0	0,006	0,002	0,098	0	0			
-0,5	0,1	0,01	3	<55	-0,04	0,077	NA	NA	0,031	0,063	-0,015	0,244	NA	NA	0	0,054	-0,036	0,076	NA	NA	0,064	0,055	-0,018	0,163	NA	NA	-0,017	0,048			
-0,5	0,1	0,01	3	55-60	-0,009	0,139	NA	NA	-0,085	0,063	0,001	0,156	NA	NA	0	0,054	-0,001	0,122	NA	NA	-0,052	0,055	0,002	0,131	NA	NA	-0,017	0,048			
-0,5	0,1	0,01	3	60-65	-0,019	0,191	NA	NA	-0,161	0,063	0	0,123	NA	NA	0	0,054	-0,001	0,143	NA	NA	-0,128	0,055	0,002	0,106	NA	NA	-0,017	0,048			
-0,5	0,1	0,01	3	>65	-0,034	0,262	NA	NA	-0,278	0,063	-0,001	0,067	NA	NA	0	0,054	-0,006	0,153	NA	NA	-0,244	0,055	0	0,065	NA	NA	-0,017	0,048			
-0,5	0,1	0,01	3	TE <sub>age=60</sub>	-0,005	0,086	0	1,057	-0,123	0,063	-0,002	0,074	-0,005	0,9	0	0,054	0,003	0,066	0,048	0,521	-0,09	0,055	0,002	0,063	0,048	0,522	-0,017	0,048			
-0,5	0,1	0,01	3	vTE <sub>age+1</sub>	0	0,008	0	0,132	-0,016	0	0	0,006	0,001	0,112	0	0	0	0,006	0,005	0,065	-0,016	0	0	0,006	-0,006	0,065	0	0			
-0,5	0,1	0,1	1	<55	0,009	0,104	NA	NA	0,002	0,06	-0,005	0,108	NA	NA	0	0,061	0,005	0,096	NA	NA	0,004	0,057	0,002	0,099	NA	NA	0,004	0,056			
-0,5	0,1	0,1	1	55-60	-0,003	0,116	NA	NA	0,002	0,06	-0,003	0,116	NA	NA	0	0,061	0,005	0,107	NA	NA	0,004	0,057	0,006	0,101	NA	NA	0,004	0,056			
-0,5	0,1	0,1	1	60-65	-0,001	0,111	NA	NA	0,002	0,06	0,002	0,115	NA	NA	0	0,061	0,001	0,104	NA	NA	0,004	0,057	0,001	0,101	NA	NA	0,004	0,056			
-0,5	0,1	0,1	1	>65	-0,002	0,102	NA	NA	0,002	0,06	-0,001	0,106	NA	NA	0	0,061	0,003	0,095	NA	NA	0,004	0,057	0,008	0,09	NA	NA	0,004	0,056			
-0,5	0,1	0,1	1	TE <sub>age=60</sub>	0,001	0,06	0,002	0,066	0,002	0,06	-0,002	0,061	-0,003	0,066	0	0,061	0,004	0,057	0,004	0,06	0,004	0,057	0,004	0,056	0,004	0,059	0,004	0,056			
-0,5	0,1	0,1	1	vTE <sub>age+1</sub>	0	0,006	0,005	0,127	0	0	0	0,007	0,002	0,129	0	0	0	0,006	-0,004	0,114	0	0	0	0,006	-0,003	0,109	0	0			
-0,5	0,1	0,1	2	<55	0,004	0,128	NA	NA	0,159	0,069	0	0,109	NA	NA	0,004	0,061	0,003	0,112	NA	NA	0,163	0,062	0,002	0,095	NA	NA	0,004	0,055			
-0,5	0,1	0,1	2	55-60	-0,003	0,123	NA	NA	0,042	0,069	-0,001	0,113	NA	NA	0,004	0,061	0,005	0,116	NA	NA	0,046	0,062	0,004	0,103	NA	NA	0,004	0,055			
-0,5	0,1	0,1	2	60-65	0	0,122	NA	NA	-0,033	0,069	0,004	0,116	NA	NA	0,004	0,061	0,011	0,118	NA	NA	-0,03	0,062	0	0,107	NA	NA	0,004	0,055			
-0,5	0,1	0,1	2	>65	0,002	0,118	NA	NA	-0,15	0,069	0,002	0,109	NA	NA	0,004	0,061	0,002	0,105	NA	NA	-0,146	0,062	0,001	0,094	NA	NA	0,004	0,055			
-0,5	0,1	0,1	2	TE <sub>age=60</sub>	0,001	0,069	0,003	0,074	0,004	0,069	0,001	0,061	0,004	0,066	0,004	0,061	0,005	0,062	0,008	0,065	0,008	0,062	0,002	0,055	0,004	0,058	0,004	0,055			
-0,5	0,1	0,1	2	vTE <sub>age+1</sub>	0	0,007	0,004	0,142	-0,016	0	0	0,007	-0,001	0,136	0	0	0	0,007	0	0,125	-0,016	0	0	0,006	0,003	0,114	0	0			
-0,5	0,1	0,1	3	<55	-0,036	0,087	NA	NA	0,039	0,072	0,007	0,237	NA	NA	0,004	0,057	-0,035	0,082	NA	NA	0,066	0,062	-0,021	0,165	NA	NA	-0,015	0,056			
-0,5	0,1	0,1	3	55-60	0,001	0,14	NA	NA	-0,077	0,072	0,003	0,164	NA	NA	0,004	0,057	0,009	0,122	NA	NA	-0,051	0,062	0,005	0,138	NA	NA	-0,015	0,056			
-0,5	0,1	0,1	3	60-65	-0,003	0,19	NA	NA	-0,153	0,072	0,004	0,124	NA	NA	0,004	0,057	0,001	0,142	NA	NA	-0,126	0,062	0,002	0,114	NA	NA	-0,015	0,056			
-0,5	0,1	0,1	3	>65	-0,004	0,251	NA	NA	-0,269	0,072	0,001	0,069	NA	NA	0,004	0,057	-0,002	0,161	NA	NA	-0,243	0,062	0,005	0,069	NA	NA	-0,015	0,056			
-0,5	0,1	0,1	3	TE <sub>age=60</sub>	0,009	0,092	0,022	1,16	-0,115	0,072	0,004	0,078	-0,038	1,066	0,004	0,057	0,007	0,072	-0,016	0,63	-0,088	0,062	0,002	0,069	-0,021	0,632	-0,015	0,056			
-0,5	0,1	0,1	3	vTE <sub>age+1</sub>	0	0,008	0,002	0,144	-0,016	0	0	0,006	0,005	0,133	0	0	0	0,006	-0,003	0,079	-0,016	0	0	0,006	0,003	0,079	0	0			
-0,2	0,01	0,01	1	<55	0,002	0,099	NA	NA	-0,001	0,051	-0,004	0,1	NA	NA	0,001	0,051	0	0,092	NA	NA	0	0,046	-0,003	0,096	NA	NA	0,001	0,049			
-0,2	0,01	0,01	1	55-60	-0,007	0,102	NA	NA	-0,001	0,051	-0,002	0,102	NA	NA	0,001	0,051	-0,004	0,091	NA	NA	0	0,046	0,003	0,097	NA	NA	0,001	0,049			

Annexes

Configuration 1																	Configuration 2														
					A-C						B-C											A-C					B-C				
coef	$\sigma$	$\tau$	s	param	IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1	EES	Biais	EES	Biais	EES	IPD Poisson1	EES	Biais	EES	Biais	EES			
					Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES			
-0,2	0,01	0,01	1	60-65	-0,002	0,104	NA	NA	-0,001	0,051	0,003	0,111	NA	NA	0,001	0,051	0,003	0,094	NA	NA	0	0,046	0,005	0,097	NA	NA	0,001	0,049			
-0,2	0,01	0,01	1	>65	-0,001	0,098	NA	NA	-0,001	0,051	0,005	0,096	NA	NA	0,001	0,051	0,001	0,087	NA	NA	0	0,046	0,002	0,088	NA	NA	0,001	0,049			
-0,2	0,01	0,01	1	TE <sub>age=60</sub>	-0,002	0,051	-0,002	0,055	-0,001	0,051	0,001	0,05	0,001	0,055	0,001	0,051	0	0,046	0	0,049	0	0,046	0,001	0,049	0,002	0,051	0,001	0,049			
-0,2	0,01	0,01	1	vTE <sub>age+1</sub>	0	0,006	0,001	0,108	0	0	0	0,006	0	0,106	0	0	0	0,006	0,002	0,101	0	0	0	0,006	0,006	0,092	0	0			
-0,2	0,01	0,01	2	<55	0,002	0,1	NA	NA	0,066	0,052	0	0,102	NA	NA	0	0,05	0,002	0,094	NA	NA	0,063	0,048	-0,001	0,09	NA	NA	0	0,047			
-0,2	0,01	0,01	2	55-60	0,007	0,109	NA	NA	0,019	0,052	-0,001	0,101	NA	NA	0	0,05	0,002	0,1	NA	NA	0,017	0,048	0	0,099	NA	NA	0	0,047			
-0,2	0,01	0,01	2	60-65	0,001	0,111	NA	NA	-0,011	0,052	-0,003	0,101	NA	NA	0	0,05	0,001	0,099	NA	NA	-0,014	0,048	0,001	0,099	NA	NA	0	0,047			
-0,2	0,01	0,01	2	>65	0,005	0,102	NA	NA	-0,058	0,052	-0,002	0,097	NA	NA	0	0,05	0,002	0,088	NA	NA	-0,06	0,048	-0,005	0,087	NA	NA	0	0,047			
-0,2	0,01	0,01	2	TE <sub>age=60</sub>	0,004	0,051	0,004	0,056	0,004	0,052	-0,001	0,05	-0,001	0,054	0	0,05	0,002	0,048	0,002	0,051	0,002	0,048	-0,001	0,047	0	0,049	0	0,047			
-0,2	0,01	0,01	2	vTE <sub>age+1</sub>	0	0,007	-0,002	0,113	-0,006	0	0	0,006	0,004	0,107	0	0	0	0,006	-0,002	0,1	-0,006	0	0	0,006	0,002	0,092	0	0			
-0,2	0,01	0,01	3	<55	-0,015	0,066	NA	NA	0,014	0,053	-0,004	0,227	NA	NA	0,001	0,05	-0,017	0,066	NA	NA	0,02	0,05	-0,017	0,155	NA	NA	-0,011	0,047			
-0,2	0,01	0,01	3	55-60	0,003	0,121	NA	NA	-0,033	0,053	0,001	0,157	NA	NA	0,001	0,05	0	0,108	NA	NA	-0,027	0,05	-0,003	0,126	NA	NA	-0,011	0,047			
-0,2	0,01	0,01	3	60-65	0,001	0,158	NA	NA	-0,063	0,053	-0,003	0,118	NA	NA	0,001	0,05	-0,009	0,129	NA	NA	-0,057	0,05	-0,003	0,102	NA	NA	-0,011	0,047			
-0,2	0,01	0,01	3	>65	-0,009	0,214	NA	NA	-0,11	0,053	0,001	0,062	NA	NA	0,001	0,05	-0,01	0,143	NA	NA	-0,103	0,05	-0,004	0,062	NA	NA	-0,011	0,047			
-0,2	0,01	0,01	3	TE <sub>age=60</sub>	0,002	0,073	0,006	0,981	-0,048	0,053	-0,001	0,072	0,055	0,887	0,001	0,05	-0,003	0,059	-0,024	0,506	-0,042	0,05	-0,004	0,059	-0,02	0,504	-0,011	0,047			
-0,2	0,01	0,01	3	vTE <sub>age+1</sub>	0	0,007	0	0,122	-0,006	0	0	0,006	-0,007	0,11	0	0	0	0,005	-0,003	0,063	-0,006	0	0	0,005	0,002	0,063	0	0			
-0,2	0,01	0,1	1	<55	0,004	0,105	NA	NA	0,002	0,06	0,004	0,104	NA	NA	0,004	0,06	0,002	0,094	NA	NA	0,002	0,054	0,004	0,089	NA	NA	0,003	0,051			
-0,2	0,01	0,1	1	55-60	0	0,109	NA	NA	0,002	0,06	0,001	0,109	NA	NA	0,004	0,06	0,003	0,101	NA	NA	0,002	0,054	0,005	0,101	NA	NA	0,003	0,051			
-0,2	0,01	0,1	1	60-65	-0,007	0,109	NA	NA	0,002	0,06	0,002	0,111	NA	NA	0,004	0,06	-0,001	0,097	NA	NA	0,002	0,054	-0,001	0,097	NA	NA	0,003	0,051			
-0,2	0,01	0,1	1	>65	0,002	0,102	NA	NA	0,002	0,06	0,001	0,1	NA	NA	0,004	0,06	0,001	0,087	NA	NA	0,002	0,054	-0,001	0,088	NA	NA	0,003	0,051			
-0,2	0,01	0,1	1	TE <sub>age=60</sub>	0	0,059	0,003	0,064	0,002	0,06	0,002	0,06	0,003	0,065	0,004	0,06	0,001	0,054	0,002	0,056	0,002	0,054	0,002	0,051	0,002	0,054	0,003	0,051			
-0,2	0,01	0,1	1	vTE <sub>age+1</sub>	0	0,006	0	0,127	0	0	0	0,006	-0,002	0,129	0	0	0	0,006	-0,003	0,113	0	0	0	0,005	-0,002	0,116	0	0			
-0,2	0,01	0,1	2	<55	0,009	0,108	NA	NA	0,069	0,061	0,004	0,109	NA	NA	0,002	0,06	-0,003	0,098	NA	NA	0,063	0,056	-0,002	0,092	NA	NA	0,001	0,055			
-0,2	0,01	0,1	2	55-60	0,008	0,117	NA	NA	0,022	0,061	-0,006	0,113	NA	NA	0,002	0,06	0,004	0,104	NA	NA	0,017	0,056	0	0,099	NA	NA	0,001	0,055			
-0,2	0,01	0,1	2	60-65	0,005	0,112	NA	NA	-0,008	0,061	0,003	0,114	NA	NA	0,002	0,06	0	0,098	NA	NA	-0,013	0,056	-0,002	0,097	NA	NA	0,001	0,055			
-0,2	0,01	0,1	2	>65	0,004	0,105	NA	NA	-0,055	0,061	0	0,099	NA	NA	0,002	0,06	0,004	0,094	NA	NA	-0,06	0,056	0,001	0,097	NA	NA	0,001	0,055			
-0,2	0,01	0,1	2	TE <sub>age=60</sub>	0,006	0,061	0,006	0,065	0,007	0,061	0	0,06	0,002	0,063	0,002	0,06	0,001	0,056	0,001	0,058	0,002	0,056	-0,001	0,055	0,002	0,057	0,001	0,055			
-0,2	0,01	0,1	2	vTE <sub>age+1</sub>	0	0,006	-0,001	0,13	-0,006	0	0	0,007	0	0,127	0	0	0	0,006	0,002	0,116	-0,006	0	0	0,006	0,001	0,117	0	0			
-0,2	0,01	0,1	3	<55	-0,016	0,073	NA	NA	0,014	0,06	-0,001	0,225	NA	NA	0,003	0,059	-0,01	0,07	NA	NA	0,028	0,054	-0,008	0,154	NA	NA	-0,006	0,054			
-0,2	0,01	0,1	3	55-60	-0,005	0,122	NA	NA	-0,033	0,06	0,002	0,166	NA	NA	0,003	0,059	0,003	0,111	NA	NA	-0,019	0,054	0,001	0,135	NA	NA	-0,006	0,054			
-0,2	0,01	0,1	3	60-65	0,002	0,161	NA	NA	-0,063	0,06	0,003	0,12	NA	NA	0,003	0,059	0,005	0,134	NA	NA	-0,049	0,054	0,001	0,115	NA	NA	-0,006	0,054			
-0,2	0,01	0,1	3	>65	0,011	0,227	NA	NA	-0,11	0,06	0	0,069	NA	NA	0,003	0,059	-0,002	0,146	NA	NA	-0,096	0,054	0	0,067	NA	NA	-0,006	0,054			
-0,2	0,01	0,1	3	TE <sub>age=60</sub>	0,005	0,078	0,032	1,091	-0,048	0,06	0,003	0,079	-0,029	1,002	0,003	0,059	0,005	0,063	0,005	0,586	-0,034	0,054	0	0,067	0,005	0,584	-0,006	0,054			
-0,2	0,01	0,1	3	vTE <sub>age+1</sub>	0	0,007	0,004	0,136	-0,006	0	0	0,006	0,004	0,125	0	0	0	0,005	0	0,074	-0,006	0	0	0,006	0	0,073	0	0			
-0,2	0,1	0,01	1	<55	-0,001	0,1	NA	NA	-0,001	0,052	-0,004	0,105	NA	NA	0	0,052	-0,005	0,088	NA	NA	-0,001	0,044	-0,008	0,094	NA	NA	0	0,045			
-0,2	0,1	0,01	1	55-60	-0,004	0,106	NA	NA	-0,001	0,052	-0,003	0,107	NA	NA	0	0,052	0,002	0,093	NA	NA	-0,001	0,044	0,004	0,094	NA	NA	0	0,045			
-0,2	0,1	0,01	1	60-65	-0,001	0,105	NA	NA	-0,001	0,052	-0,003	0,111	NA	NA	0	0,052	-0,003	0,094	NA	NA	-0,001	0,044	0,001	0,092	NA	NA	0	0,045			
-0,2	0,1	0,01	1	>65	-0,002	0,098	NA	NA	-0,001	0,052	0,005	0,094	NA	NA	0	0,052	0,004	0,088	NA	NA	-0,001	0,044	0	0,087	NA	NA	0	0,045			

Configuration 1																	Configuration 2											
					A-C						B-C						A-C					B-C						
coef	$\sigma$	$\tau$	s	param	IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta		IPD Poisson1		AD metareg		AD-netmeta	
					Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES	Biais	EES
-0,2	0,1	0,01	1	TE <sub>age=60</sub>	-0,002	0,052	-0,002	0,057	-0,001	0,052	-0,001	0,052	0	0,054	0	0,052	-0,001	0,044	-0,001	0,047	-0,001	0,044	-0,001	0,045	0	0,047	0	0,045
-0,2	0,1	0,01	1	vTE <sub>age+1</sub>	0	0,006	0	0,111	0	0	0	0,006	0,005	0,104	0	0	0	0,006	-0,003	0,094	0	0	0	0,006	0,003	0,09	0	0
-0,2	0,1	0,01	2	<55	0,007	0,103	NA	NA	0,066	0,054	0	0,103	NA	NA	0	0,052	0	0,093	NA	NA	0,061	0,049	0,003	0,091	NA	NA	0,003	0,047
-0,2	0,1	0,01	2	55-60	-0,001	0,111	NA	NA	0,02	0,054	-0,005	0,104	NA	NA	0	0,052	-0,005	0,1	NA	NA	0,014	0,049	0,004	0,097	NA	NA	0,003	0,047
-0,2	0,1	0,01	2	60-65	0,002	0,11	NA	NA	-0,01	0,054	0,001	0,108	NA	NA	0	0,052	-0,001	0,095	NA	NA	-0,016	0,049	0,001	0,094	NA	NA	0,003	0,047
-0,2	0,1	0,01	2	>65	0,011	0,1	NA	NA	-0,057	0,054	-0,001	0,097	NA	NA	0	0,052	0,003	0,089	NA	NA	-0,063	0,049	0,002	0,089	NA	NA	0,003	0,047
-0,2	0,1	0,01	2	TE <sub>age=60</sub>	0,005	0,053	0,005	0,058	0,005	0,054	-0,001	0,052	0,001	0,056	0	0,052	-0,001	0,049	-0,002	0,052	-0,001	0,049	0,002	0,047	0,003	0,05	0,003	0,047
-0,2	0,1	0,01	2	vTE <sub>age+1</sub>	0	0,007	0,006	0,109	-0,006	0	0	0,006	-0,002	0,108	0	0	0	0,006	0	0,098	-0,006	0	0	0,006	0,002	0,096	0	0
-0,2	0,1	0,01	3	<55	-0,017	0,068	NA	NA	0,011	0,052	0,001	0,229	NA	NA	-0,002	0,05	-0,016	0,065	NA	NA	0,022	0,048	-0,013	0,152	NA	NA	-0,008	0,045
-0,2	0,1	0,01	3	55-60	-0,002	0,114	NA	NA	-0,036	0,052	-0,004	0,158	NA	NA	-0,002	0,05	-0,001	0,109	NA	NA	-0,024	0,048	0,002	0,13	NA	NA	-0,008	0,045
-0,2	0,1	0,01	3	60-65	-0,005	0,163	NA	NA	-0,066	0,052	-0,005	0,113	NA	NA	-0,002	0,05	-0,001	0,131	NA	NA	-0,055	0,048	0	0,105	NA	NA	-0,008	0,045
-0,2	0,1	0,01	3	>65	-0,007	0,219	NA	NA	-0,113	0,052	-0,003	0,062	NA	NA	-0,002	0,05	-0,005	0,148	NA	NA	-0,101	0,048	0	0,06	NA	NA	-0,008	0,045
-0,2	0,1	0,01	3	TE <sub>age=60</sub>	0,001	0,071	-0,007	0,916	-0,051	0,052	-0,003	0,072	0,01	0,866	-0,002	0,05	-0,001	0,059	-0,027	0,486	-0,04	0,048	0,001	0,057	-0,022	0,489	-0,008	0,045
-0,2	0,1	0,01	3	vTE <sub>age+1</sub>	0	0,007	-0,001	0,114	-0,006	0	0	0,006	-0,001	0,108	0	0	0	0,006	-0,003	0,061	-0,006	0	0	0,005	0,003	0,061	0	0
-0,2	0,1	0,1	1	<55	0,007	0,108	NA	NA	0,005	0,061	-0,001	0,104	NA	NA	0,002	0,059	0,002	0,095	NA	NA	0,003	0,055	0,002	0,093	NA	NA	0,002	0,054
-0,2	0,1	0,1	1	55-60	0,008	0,112	NA	NA	0,005	0,061	0,008	0,105	NA	NA	0,002	0,059	-0,002	0,102	NA	NA	0,003	0,055	0,002	0,099	NA	NA	0,002	0,054
-0,2	0,1	0,1	1	60-65	0,002	0,105	NA	NA	0,005	0,061	-0,002	0,107	NA	NA	0,002	0,059	0,008	0,101	NA	NA	0,003	0,055	0,003	0,102	NA	NA	0,002	0,054
-0,2	0,1	0,1	1	>65	0,002	0,106	NA	NA	0,005	0,061	0,001	0,104	NA	NA	0,002	0,059	0	0,093	NA	NA	0,003	0,055	0,001	0,09	NA	NA	0,002	0,054
-0,2	0,1	0,1	1	TE <sub>age=60</sub>	0,005	0,061	0,005	0,064	0,005	0,061	0,001	0,059	0,001	0,065	0,002	0,059	0,002	0,055	0,002	0,058	0,003	0,055	0,002	0,054	0,002	0,055	0,002	0,054
-0,2	0,1	0,1	1	vTE <sub>age+1</sub>	0	0,006	0,002	0,115	0	0	0	0,006	0,005	0,13	0	0	0	0,006	-0,003	0,114	0	0	0	0,006	0	0,109	0	0
-0,2	0,1	0,1	2	<55	0,003	0,118	NA	NA	0,064	0,062	-0,003	0,106	NA	NA	0	0,059	0,005	0,098	NA	NA	0,064	0,056	-0,002	0,097	NA	NA	0,001	0,055
-0,2	0,1	0,1	2	55-60	-0,001	0,116	NA	NA	0,018	0,062	0,001	0,108	NA	NA	0	0,059	0,001	0,102	NA	NA	0,018	0,056	0,001	0,101	NA	NA	0,001	0,055
-0,2	0,1	0,1	2	60-65	0,008	0,111	NA	NA	-0,013	0,062	-0,002	0,11	NA	NA	0	0,059	0,002	0,103	NA	NA	-0,013	0,056	0,001	0,1	NA	NA	0,001	0,055
-0,2	0,1	0,1	2	>65	0	0,107	NA	NA	-0,059	0,062	-0,001	0,106	NA	NA	0	0,059	0,003	0,094	NA	NA	-0,059	0,056	0,002	0,094	NA	NA	0,001	0,055
-0,2	0,1	0,1	2	TE <sub>age=60</sub>	0,003	0,062	0,002	0,066	0,002	0,062	-0,001	0,059	0	0,064	0	0,059	0,003	0,057	0,002	0,06	0,002	0,056	0	0,055	0,002	0,059	0,001	0,055
-0,2	0,1	0,1	2	vTE <sub>age+1</sub>	0	0,007	0,006	0,128	-0,006	0	0	0,007	-0,001	0,135	0	0	0	0,006	-0,002	0,116	-0,006	0	0	0,006	0	0,12	0	0
-0,2	0,1	0,1	3	<55	-0,016	0,073	NA	NA	0,013	0,06	-0,004	0,238	NA	NA	0,002	0,06	-0,014	0,073	NA	NA	0,025	0,055	-0,012	0,16	NA	NA	-0,007	0,053
-0,2	0,1	0,1	3	55-60	0,003	0,126	NA	NA	-0,034	0,06	-0,002	0,161	NA	NA	0,002	0,06	0,002	0,11	NA	NA	-0,022	0,055	0,001	0,132	NA	NA	-0,007	0,053
-0,2	0,1	0,1	3	55-60	-0,002	0,16	NA	NA	-0,064	0,06	0,002	0,117	NA	NA	0,002	0,06	0,003	0,132	NA	NA	-0,052	0,055	0,002	0,11	NA	NA	-0,007	0,053
-0,2	0,1	0,1	3	>65	-0,01	0,229	NA	NA	-0,11	0,06	0,001	0,072	NA	NA	0,002	0,06	0,002	0,151	NA	NA	-0,099	0,055	0,001	0,067	NA	NA	-0,007	0,053
-0,2	0,1	0,1	3	TE <sub>age=60</sub>	0,002	0,079	-0,004	1,051	-0,049	0,06	-0,001	0,079	-0,009	0,991	0,002	0,06	0,003	0,064	0,003	0,58	-0,037	0,055	0	0,064	-0,002	0,575	-0,007	0,053
-0,2	0,1	0,1	3	vTE <sub>age+1</sub>	0	0,007	-0,001	0,131	-0,006	0	0	0,006	0,001	0,123	0	0	0	0,006	0	0,073	-0,006	0	0	0,005	0,001	0,072	0	0

Ttt : coefficient de l'effet traitement,  $\sigma$  : variabilité inter-essai du risque de base,  $\tau$  : variabilité inter-essai de l'effet traitement, s : scénario, param : paramètre estimé, EES : erreur empirique standard, TE<sub>age=60</sub> : effet traitement à l'âge moyen de 60 ans, vTE<sub>age+1</sub> : variation de l'effet traitement pour chaque année d'âge supplémentaire. NA : non applicable

## Partie méta-analyse

### Advisory board

**Pierre Thirion**

Department of Radiation Oncology, St. Luke's Hospital,  
Dublin, Ireland.

Courriel : Pierre@thirion.ie

**Jayne Tierney**

MRC Clinical Trials Unit at UCL  
London, UK

E-mail: jayne.tierney@ucl.ac.uk

**Val Gebski**

National Health and Medical Research Council Clinical Trials Centre,  
University of Sydney,  
Sydney, NSW, Australia

Courriel : val@ctc.usyd.edu.au

**Bryan H Burmeister**

Division of Cancer Services,  
Princess Alexandra Hospital, University of Queensland,  
Brisbane, Australia.

Courriel : bryan.burmeister@health.qld.gov.au

**Xavier Paoletti**

Department of biostatistics and epidemiology  
Gustave Roussy Cancer Campus  
Villejuif, France

Courriel : Xavier.paoletti@gustaveroussy.fr

**Johanna van Sandick**

Department of Surgical Oncology  
The Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital,  
Amsterdam, The Netherlands

E-mail: j.v.sandick@nki.nl

**Jianhua Fu**

Department of Thoracic Surgery  
Sun Yat-Sen University Cancer Center,  
Guangzhou, China

E-mail: fujh@sysucc.org.cn

## Equations de recherche

### Pubmed

(((((("Esophagogastric Junction"[MeSH] OR "Cardia"[MeSH] OR oesophago-gastric junction OR gastroesophageal junction) AND (neoplasms OR cancer OR tumor OR adenocarcinoma OR carcinoma OR epidermoid OR squamous cell)) OR ("Esophageal Neoplasms"[MeSH] AND "Stomach Neoplasms"[MeSH]) OR ("Esophageal Neoplasms"[MeSH] OR (esophag\* AND (neoplasms OR cancer OR tumor OR adenocarcinoma OR carcinoma OR squamous cell)))) AND (((chemotherapy OR drug therapy OR cisplatinum OR carboplatin OR oxaliplatin OR Bléomycine OR mytomycin c OR methotrexate OR 5-Fluorouracile OR hydroxyurea OR vindesine OR vinblastine OR vinorelbine OR vincristine OR taxane\* OR paclitaxel OR docetaxel OR gemcitabine OR vepesid OR VP-16 OR VP16 OR irinotecan OR FOLFOX OR T-FOX) OR (radio-chemotherapy OR chemoradiotherapy OR chemo-radiotherapy OR chemoradiotherapy OR radiotherapy)) AND (neoadjuvant OR neo-adjuvant OR induction OR preoperative OR perioperative))) AND (esophag\*[Title] OR oesophag\*[Title] OR junction[Title] OR cardia[Title]) AND (((randomized controlled trial[Publication Type] OR clinical trial, phase iii[Publication Type] OR clinical trial, phase iv[Publication Type] OR clinicaltrials.gov[Secondary Source ID] OR isrcn[Secondary Source ID]) OR randomized controlled trials as topic[MeSH Terms]) AND ((random OR randomised OR randomized) AND (trial\* OR study OR studies)) OR (rct OR rcts)) AND ("1990"[Date - Publication] : "2017"[Date - Publication]))

### Web of science

#6 #5 AND #4 AND #3 AND #2 AND #1  
*DocType=All document types; Language=All languages;*

#5 (TS=(random\*))  
*DocType=All document types; Language=All languages;*

(TS=(neoadjuvant OR (neo adjuvant) OR preoperative OR pre operative OR perioperative  
 #4OR (peri operative) OR induction))  
*DocType=All document types; Language=All languages;*

(TS=(chemotherapy OR (drug therapy) OR chemoradiation OR chemoradiotherapy OR  
 #3chemoradiotherapy OR radio-chemotherapy OR pharmacotherapy))  
*DocType=All document types; Language=All languages;*

(TS=(cancer\* OR carcinoma\* OR adenocarcinoma\* OR malignan\* OR tumor\* OR  
 #2tumour\* OR neoplasm OR (squamous)))  
*DocType=All document types; Language=All languages;*

(TS= (oesophag\* OR esophagi\* OR cardia OR (gastro-oesophag\* junction) OR (gastro-  
 #1esophag\* junction) OR (gastrooesophag\* junction) OR (gastroesophag\* junction) OR  
 (Esophagogastric Junction) OR (oesophago-gastric junction)))  
*DocType=All document types; Language=All languages;*

### Scopus

(TITLE-ABS-KEY ( oesophag\* OR esophag\* OR cardia OR "gastro-oesophag\* junction" OR "gastro-esophag\*  
 junction" OR "gastrooesophag\* junction" OR "gastroesophag\* junction" OR "Esophagogastric Junction" OR  
 "oesophago-gastric junction" ) AND TITLE-ABS-KEY ( cancer\* OR carcinoma\* OR adenocarcinoma\* OR  
 malignan\* OR tumor\* OR tumour\* OR neoplasm\* ) AND TITLE-ABS-KEY ( squamous OR epidermoid OR  
 "undifferentiated carcinoma" OR adenocarcinoma\* OR carcinoma\* ) AND TITLE-ABS-KEY ( chemotherapy  
 OR chemoradiation OR "drug therapy" OR chemoradiotherapy OR chemoradiotherapy OR radio-chemotherapy  
 OR pharmacotherapy ) AND TITLE-ABS-KEY ( neoadjuvant OR "neo adjuvant" OR preoperative OR "pre  
 operative" OR perioperative OR "peri operative" OR induction ) AND TITLE-ABS-KEY ( random OR

## Annexes

randomise OR randomize OR randomised OR randomized OR rct OR rcts ) AND TITLE-ABS-KEY ( trial OR trials OR study OR studies ) ) AND PUBYEAR > 1989

## Repartitions des critères secondaires dans les essais

Tableau annexe 2 : Répartition des modes de récurrences (risques compétitifs) dans les essais

	Vivant sans récurrence	Récurrence locale	Récurrence à distance	Récurrence locale et à distance	Décès sans récurrence	Total
<b>Chirurgie contre Chimiothérapie-Chirurgie</b>						
Oeso 2	26 (21%)	12 (9,8%)	8 (6,6%)	9 (7,4%)	67 (55%)	122 (100%)
Songkla	10 (22%)	5 (11%)	6 (13%)	2 (4,3%)	23 (50%)	46 (100%)
Queen Mary	12 (8,2%)	35 (24%)	38 (26%)	16 (11%)	46 (31%)	147 (100%)
Rotterdam	34 (20%)	48 (28%)	37 (22%)	20 (12%)	30 (18%)	169 (100%)
RTOG 8911	53 (11%)	136 (29%)	162 (35%)	42 (9,0%)	74 (16%)	467 (100%)
Italy	20 (21%)	30 (31%)	15 (16%)	8 (8,3%)	23 (24%)	96 (100%)
MRC OE02	181 (23%)	127 (16%)	151 (19%)	53 (6,6%)	290 (36%)	802 (100%)
MRC MAGIC	46 (35%)	13 (9,9%)	14 (11%)	2 (1,5%)	56 (43%)	131 (100%)
ACCORD 07	32 (19%)	20 (12%)	70 (41%)	20 (12%)	27 (16%)	169 (100%)
EORTC 40954	29 (38%)	3 (3,9%)	17 (22%)	16 (21%)	11 (14%)	76 (100%)
<b>Sous total</b>	<b>443 (20%)</b>	<b>429 (19%)</b>	<b>518 (23%)</b>	<b>188 (8,4%)</b>	<b>647 (29%)</b>	<b>2,225 (100%)</b>
<b>Chirurgie contre Chimioradiothérapie-Chirurgie</b>						
Rennes	5 (5,5%)	31 (34%)	11 (12%)	5 (5,5%)	39 (43%)	91 (100%)
Michigan	13 (13%)	20 (20%)	39 (39%)	15 (15%)	13 (13%)	100 (100%)
EORTC 88-05	64 (22%)	97 (33%)	51 (17%)	21 (7,1%)	61 (21%)	294 (100%)
Australia	59 (23%)	40 (16%)	93 (36%)	44 (17%)	21 (8,2%)	257 (100%)
CALGB 97-81	17 (30%)	5 (8,9%)	15 (27%)	2 (3,6%)	17 (30%)	56 (100%)
Korea	32 (32%)	13 (13%)	18 (18%)	7 (7,1%)	29 (29%)	99 (100%)
FFCD 99-01	50 (26%)	21 (11%)	30 (15%)	22 (11%)	72 (37%)	195 (100%)
CROSS	96 (26%)	26 (7,1%)	89 (24%)	59 (16%)	96 (26%)	366 (100%)
Guangzhou	233 (52%)	47 (10%)	94 (21%)	28 (6,2%)	49 (11%)	451 (100%)
<b>Sous total</b>	<b>569 (30%)</b>	<b>300 (16%)</b>	<b>440 (23%)</b>	<b>203 (11%)</b>	<b>397 (21%)</b>	<b>1 909 (100%)</b>
<b>Chimiothérapie-Chirurgie contre Chimioradiothérapie-Chirurgie</b>						
POET	34 (29%)	17 (14%)	21 (18%)	18 (15%)	29 (24%)	119 (100%)
IG9401	15 (17%)	13 (15%)	40 (45%)	10 (11%)	10 (11%)	88 (100%)
NeoRes	67 (37%)	14 (7,7%)	35 (19%)	29 (16%)	36 (20%)	181 (100%)
<b>Sous total</b>	<b>116 (30%)</b>	<b>44 (11%)</b>	<b>96 (25%)</b>	<b>57 (15%)</b>	<b>75 (19%)</b>	<b>388 (100%)</b>
<b>Total général</b>	<b>1 128 (25%)</b>	<b>773 (17%)</b>	<b>1 054 (23%)</b>	<b>448 (9,9%)</b>	<b>1 119 (25%)</b>	<b>4 522 (100%)</b>



Tableau annexe 3: Cause spécifique de décès dans les essais

	Vivant	Décès lié au cancer	Décès non lié au cancer	Total
<b>Chirurgie contre Chimiothérapie-Chirurgie</b>				
<b>Oeso 2</b>	26 (21%)	91 (75%)	5 (4,1%)	122 (100%)
<b>Songkla</b>	10 (22%)	36 (78%)	0 (0%)	46 (100%)
<b>Queen Mary</b>	16 (11%)	105 (71%)	26 (18%)	147 (100%)
<b>Rotterdam</b>	36 (21%)	122 (72%)	11 (6,5%)	169 (100%)
<b>RTOG 8911</b>	66 (14%)	373 (80%)	28 (6,0%)	467 (100%)
<b>Italy</b>	24 (25%)	62 (65%)	10 (10%)	96 (100%)
<b>MRC OE02</b>	147 (18%)	654 (82%)	1 (0,1%)	802 (100%)
<b>MRC MAGIC</b>	47 (36%)	84 (64%)	0 (0%)	131 (100%)
<b>ACCORD 07</b>	37 (22%)	124 (73%)	8 (4,7%)	169 (100%)
<b>EORTC 40954</b>	34 (45%)	39 (51%)	3 (3,9%)	76 (100%)
<b>Sous-total</b>	<b>443 (20%)</b>	<b>1,690 (76%)</b>	<b>92 (4,1%)</b>	<b>2,225 (100%)</b>
<b>Chirurgie contre Chimioradiothérapie-Chirurgie</b>				
<b>Rennes</b>	5 (5,5%)	78 (86%)	8 (8,8%)	91 (100%)
<b>Michigan</b>	13 (13%)	83 (83%)	4 (4,0%)	100 (100%)
<b>EORTC 88-05</b>	80 (27%)	195 (66%)	19 (6,5%)	294 (100%)
<b>Australia</b>	68 (26%)	179 (70%)	10 (3,9%)	257 (100%)
<b>CALGB 97-81</b>	17 (30%)	39 (70%)	0 (0%)	56 (100%)
<b>Korea</b>	34 (34%)	62 (63%)	3 (3,0%)	99 (100%)
<b>FFCD 99-01</b>	56 (29%)	130 (67%)	9 (4,6%)	195 (100%)
<b>CROSS</b>	106 (29%)	260 (71%)	0 (0%)	366 (100%)
<b>Guangzhou</b>	269 (60%)	164 (36%)	18 (4,0%)	451 (100%)
<b>Sous-total</b>	<b>648 (34%)</b>	<b>1 190 (62%)</b>	<b>71 (3,7%)</b>	<b>1 909 (100%)</b>
<b>Chimiothérapie-Chirurgie contre Chimioradiothérapie-Chirurgie</b>				
<b>POET</b>	37 (31%)	78 (66%)	4 (3,4%)	119 (100%)
<b>IG9401</b>	15 (17%)	63 (72%)	10 (11%)	88 (100%)
<b>NeoRes</b>	72 (40%)	98 (54%)	11 (6,1%)	181 (100%)
<b>Sous-total</b>	<b>124 (32%)</b>	<b>239 (62%)</b>	<b>25 (6,4%)</b>	<b>388 (100%)</b>
<b>Total général</b>	<b>1 215 (27%)</b>	<b>3 119 (69%)</b>	<b>188 (4,2%)</b>	<b>4 522 (100%)</b>

Tableau annexe 4:Qualité de la résection chirurgicale dans les essais

	Résection non R0	Résection R0	Inconnue	Total
<b>Chirurgie contre Chimiothérapie-Chirurgie</b>				
MD Anderson	7 (19%)	29 (81%)	0 (0%)	36 (100%)
Scandinavia : S_CS	66 (62%)	39 (37%)	1 (0,9%)	106 (100%)
Scandinavia : RS_CRS	56 (50%)	48 (43%)	7 (6,3%)	111 (100%)
Oeso 2	47 (39%)	75 (61%)	0 (0%)	122 (100%)
Songkla	11 (24%)	35 (76%)	0 (0%)	46 (100%)
Queen Mary	60 (41%)	87 (59%)	0 (0%)	147 (100%)
Rotterdam	47 (28%)	96 (57%)	26 (15%)	169 (100%)
RTOG 8911	184 (39%)	283 (61%)	0 (0%)	467 (100%)
Italy	24 (25%)	72 (75%)	0 (0%)	96 (100%)
MRC OE02	132 (16%)	581 (72%)	89 (11%)	802 (100%)
MRC MAGIC	93 (71%)	35 (27%)	3 (2,3%)	131 (100%)
ACCORD 07	40 (24%)	128 (76%)	1 (0,6%)	169 (100%)
EORTC 40954	8 (11%)	66 (87%)	2 (2,6%)	76 (100%)
<b>Total</b>	<b>775 (31%)</b>	<b>1,574 (64%)</b>	<b>129 (5,2%)</b>	<b>2,478 (100%)</b>
<b>Chirurgie contre Chimioradiothérapie-Chirurgie</b>				
Scandinavia : S_CRS	58 (56%)	44 (43%)	1 (1,0%)	103 (100%)
Rennes	34 (37%)	47 (52%)	10 (11%)	91 (100%)
Michigan	12 (12%)	86 (86%)	2 (2,0%)	100 (100%)
EORTC 88-05	76 (26%)	206 (70%)	12 (4,1%)	294 (100%)
Ireland 1	6 (5,3%)	0 (0%)	107 (95%)	113 (100%)
Ireland 2	6 (6,2%)	0 (0%)	91 (94%)	97 (100%)
Australia	50 (19%)	206 (80%)	1 (0,4%)	257 (100%)
CALGB 97-81	10 (18%)	42 (75%)	4 (7,1%)	56 (100%)
Korea	26 (26%)	73 (74%)	0 (0%)	99 (100%)
FFCD 99-01	30 (15%)	158 (81%)	7 (3,6%)	195 (100%)
CROSS	106 (29%)	260 (71%)	0 (0%)	366 (100%)
Guangzhou	62 (14%)	389 (86%)	0 (0%)	451 (100%)
<b>Total</b>	<b>476 (21%)</b>	<b>1,511 (68%)</b>	<b>235 (11%)</b>	<b>2,222 (100%)</b>
<b>Chimiothérapie-Chirurgie contre Chimioradiothérapie-Chirurgie</b>				
Scandinavia : CS_CRS	56 (51%)	51 (47%)	2 (1,8%)	109 (100%)
POET	35 (29%)	83 (70%)	1 (0,8%)	119 (100%)
IG9401	15 (17%)	72 (82%)	1 (1,1%)	88 (100%)
NeoRes	25 (14%)	156 (86%)	0 (0%)	181 (100%)
<b>Total</b>	<b>131 (26%)</b>	<b>362 (73%)</b>	<b>4 (0,8%)</b>	<b>497 (100%)</b>
<b>Total général</b>	<b>1 382 (27%)</b>	<b>3 447 (66%)</b>	<b>368 (7,1%)</b>	<b>5 197 (100%)</b>

Tableau annexe 5 : Mortalité postopératoire dans les essais

	Pas de décès	Décès	Inconnu	Total
<b>Chirurgie contre Chimiothérapie-Chirurgie</b>				
MD Anderson	0 (0%)	0 (0%)	36 (100%)	36 (100%)
Scandinavia : S_CS	0 (0%)	0 (0%)	106 (100%)	106 (100%)
Scandinavia : RS_CRIS	0 (0%)	0 (0%)	111 (100%)	111 (100%)
Oeso 2	87 (71%)	16 (13%)	19 (16%)	122 (100%)
Songkla	43 (93%)	3 (6,5%)	0 (0%)	46 (100%)
Queen Mary	141 (96%)	6 (4,1%)	0 (0%)	147 (100%)
Rotterdam	161 (95%)	7 (4,1%)	1 (0,6%)	169 (100%)
RTOG 8911	404 (87%)	13 (2,8%)	50 (11%)	467 (100%)
Italy	93 (97%)	3 (3,1%)	0 (0%)	96 (100%)
MRC OE02	672 (84%)	76 (9,5%)	54 (6,7%)	802 (100%)
MRC MAGIC	0 (0%)	0 (0%)	131 (100%)	131 (100%)
ACCORD 07	154 (91%)	9 (5,3%)	6 (3,6%)	169 (100%)
EORTC 40954	64 (84%)	3 (3,9%)	9 (12%)	76 (100%)
<b>Sous-total</b>	<b>1 819 (73%)</b>	<b>136 (5,5%)</b>	<b>523 (21%)</b>	<b>2 478 (100%)</b>
<b>Chirurgie contre Chimioradiothérapie-Chirurgie</b>				
Scandinavia : S_CRIS	0 (0%)	0 (0%)	103 (100%)	103 (100%)
Rennes	69 (76%)	7 (7,7%)	15 (16%)	91 (100%)
Michigan	96 (96%)	4 (4,0%)	0 (0%)	100 (100%)
EORTC 88-05	260 (88%)	22 (7,5%)	12 (4,1%)	294 (100%)
Ireland 1	102 (90%)	5 (4,4%)	6 (5,3%)	113 (100%)
Ireland 2	81 (84%)	10 (10%)	6 (6,2%)	97 (100%)
Australia	0 (0%)	0 (0%)	257 (100%)	257 (100%)
CALGB 97-81	55 (98%)	1 (1,8%)	0 (0%)	56 (100%)
Korea	0 (0%)	1 (1,0%)	98 (99%)	99 (100%)
FFCD 99-01	188 (96%)	7 (3,6%)	0 (0%)	195 (100%)
CROSS	357 (98%)	9 (2,5%)	0 (0%)	366 (100%)
Guangzhou	448 (99%)	3 (0,7%)	0 (0%)	451 (100%)
<b>Sous-total</b>	<b>1 656 (75%)</b>	<b>69 (3,1%)</b>	<b>497 (22%)</b>	<b>2 222 (100%)</b>
<b>Chimiothérapie-Chirurgie contre Chimioradiothérapie-Chirurgie</b>				
Scandinavia : CS_CRIS	0 (0%)	0 (0%)	109 (100%)	109 (100%)
POET	0 (0%)	0 (0%)	119 (100%)	119 (100%)
IG9401	81 (92%)	0 (0%)	7 (8,0%)	88 (100%)
NeoRes	148 (82%)	8 (4,4%)	25 (14%)	181 (100%)
<b>Sous-total</b>	<b>229 (46%)</b>	<b>8 (1,6%)</b>	<b>260 (52%)</b>	<b>497 (100%)</b>
<b>Total</b>	<b>3 704 (71%)</b>	<b>213 (4,1%)</b>	<b>1 280 (25%)</b>	<b>5 197 (100%)</b>

Tableau annexe 6 : Morbidité grave postopératoire dans les essais

	Pas de morbidité	Morbidité	Inconnue	Total
<b>Chirurgie contre Chimiothérapie-Chirurgie</b>				
MD Anderson	0 (0%)	0 (0%)	36 (100%)	36 (100%)
Scandinavia : S_CS	0 (0%)	0 (0%)	106 (100%)	106 (100%)
Scandinavia : RS_CRS	0 (0%)	0 (0%)	111 (100%)	111 (100%)
Oeso 2	42 (34%)	61 (50%)	19 (16%)	122 (100%)
Songkla	44 (96%)	2 (4,3%)	0 (0%)	46 (100%)
Queen Mary	62 (42%)	78 (53%)	7 (4,8%)	147 (100%)
Rotterdam	0 (0%)	0 (0%)	169 (100%)	169 (100%)
RTOG 8911	272 (58%)	145 (31%)	50 (11%)	467 (100%)
Italy	65 (68%)	31 (32%)	0 (0%)	96 (100%)
MRC OE02	439 (55%)	307 (38%)	56 (7,0%)	802 (100%)
MRC MAGIC	55 (42%)	61 (47%)	15 (11%)	131 (100%)
ACCORD 07	121 (72%)	42 (25%)	6 (3,6%)	169 (100%)
EORTC 40954	56 (74%)	17 (22%)	3 (3,9%)	76 (100%)
<b>Total</b>	<b>1 156 (47%)</b>	<b>744 (30%)</b>	<b>578 (23%)</b>	<b>2 478 (100%)</b>
<b>Chirurgie contre Chimioradiothérapie-Chirurgie</b>				
Scandinavia : S_CRS	0 (0%)	0 (0%)	103 (100%)	103 (100%)
Rennes	45 (49%)	30 (33%)	16 (18%)	91 (100%)
Michigan	78 (78%)	22 (22%)	0 (0%)	100 (100%)
EORTC 88-05	149 (51%)	133 (45%)	12 (4,1%)	294 (100%)
Ireland 1	0 (0%)	0 (0%)	113 (100%)	113 (100%)
Ireland 2	0 (0%)	0 (0%)	97 (100%)	97 (100%)
Australia	0 (0%)	0 (0%)	257 (100%)	257 (100%)
CALGB 97-81	0 (0%)	0 (0%)	56 (100%)	56 (100%)
Korea	0 (0%)	0 (0%)	99 (100%)	99 (100%)
FFCD 99-01	83 (43%)	92 (47%)	20 (10%)	195 (100%)
CROSS	97 (27%)	269 (73%)	0 (0%)	366 (100%)
Guangzhou	241 (53%)	171 (38%)	39 (8,6%)	451 (100%)
<b>Total</b>	<b>693 (31%)</b>	<b>717 (32%)</b>	<b>812 (37%)</b>	<b>2 222 (100%)</b>
<b>Chimiothérapie-Chirurgie contre Chimioradiothérapie-Chirurgie</b>				
Scandinavia : CS_CRS	0 (0%)	0 (0%)	109 (100%)	109 (100%)
POET	0 (0%)	0 (0%)	119 (100%)	119 (100%)
IG9401	36 (41%)	45 (51%)	7 (8,0%)	88 (100%)
NeoRes	0 (0%)	0 (0%)	181 (100%)	181 (100%)
<b>Total</b>	<b>36 (7,2%)</b>	<b>45 (9,1%)</b>	<b>416 (84%)</b>	<b>497 (100%)</b>
<b>Total</b>	<b>1 885 (36%)</b>	<b>1 506 (29%)</b>	<b>1 806 (35%)</b>	<b>5 197 (100%)</b>

## Figures supplémentaires des méta-analyses deux à deux

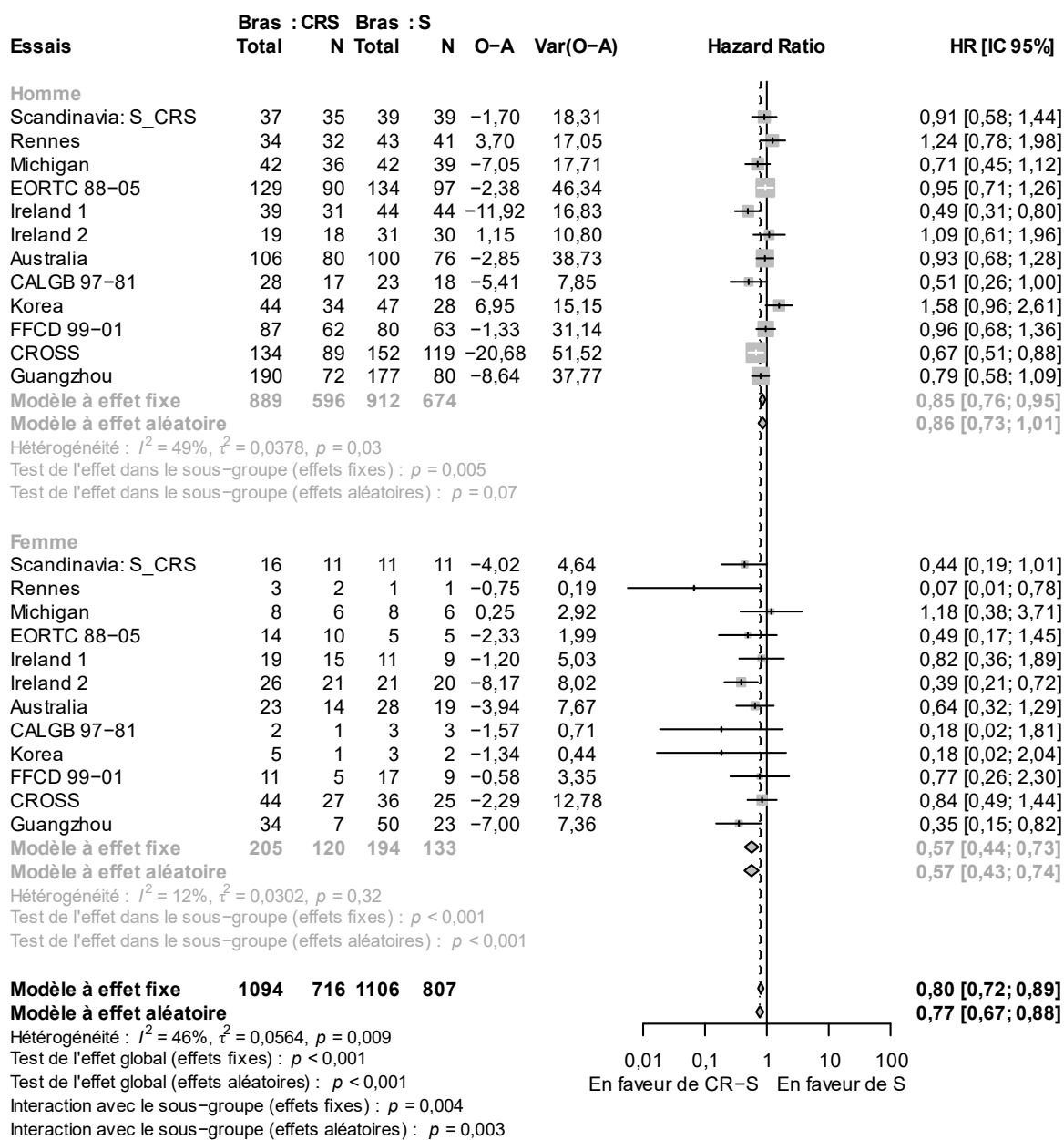


Figure annexe 1 : Forest plot de l'effet de CRS contre S sur la survie globale selon le sexe

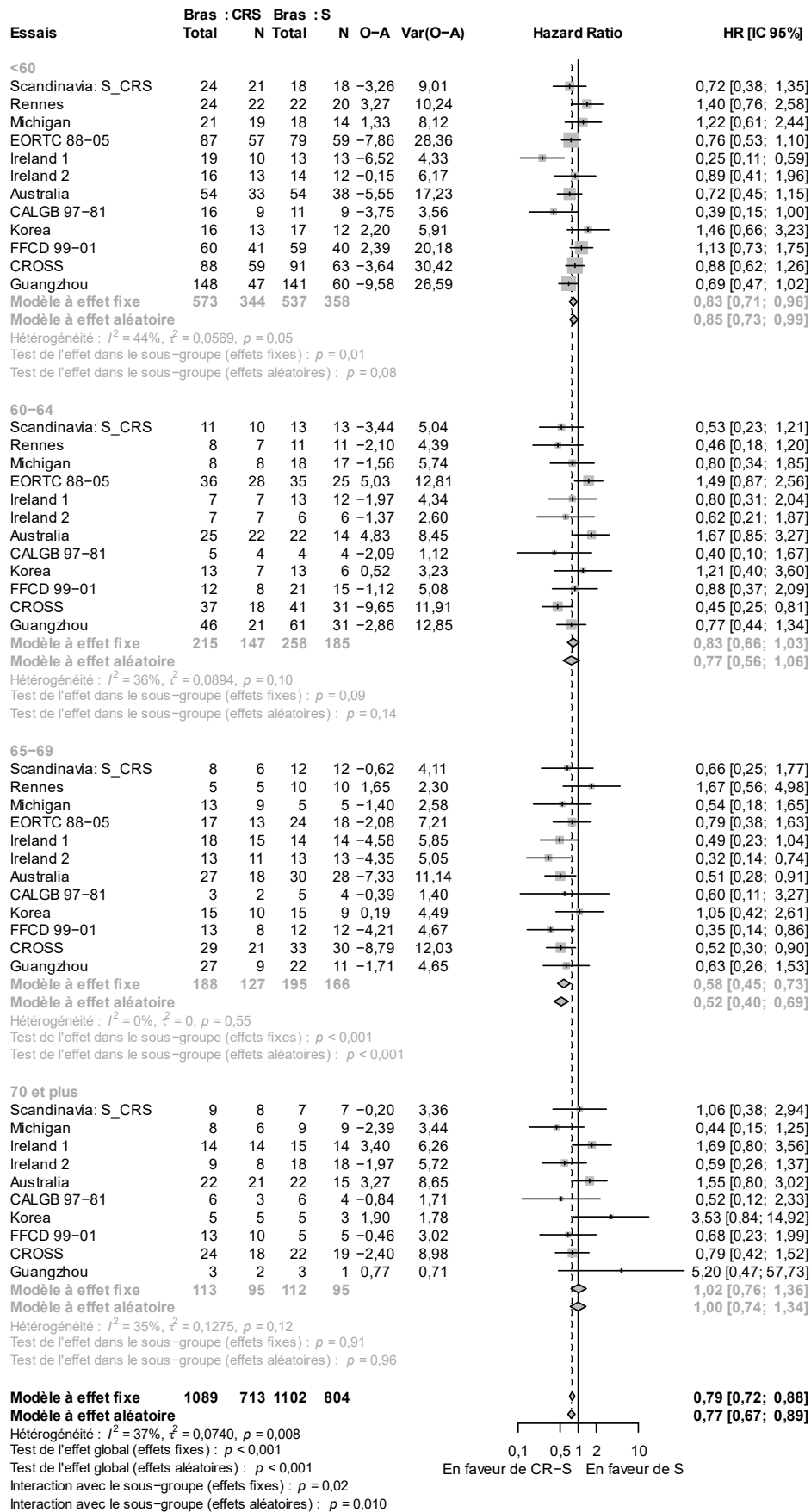


Figure annexe 2 : Forest plot de l'effet de la CRS contre S sur la survie globale selon l'âge

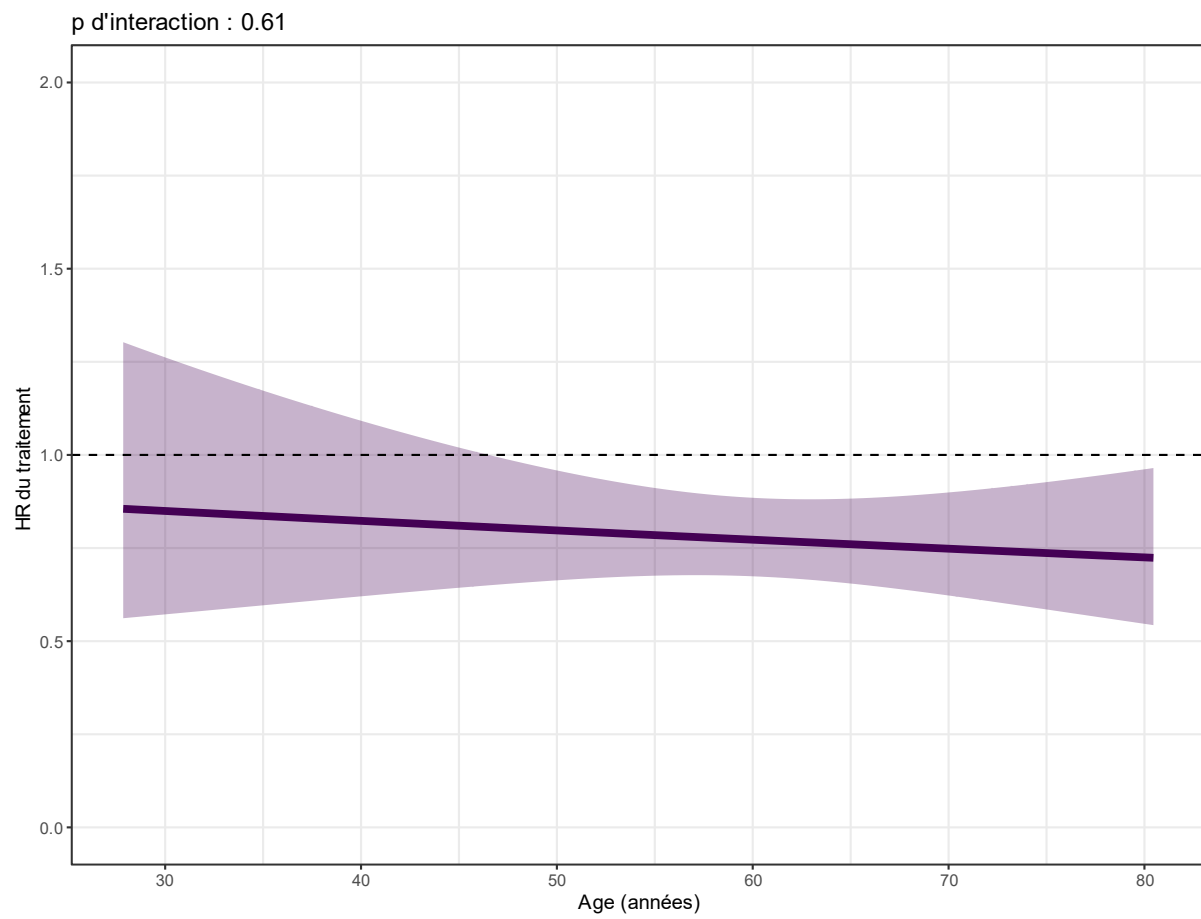


Figure annexe 3 : Interaction entre l'effet de la CRS contre S et l'âge (modèle IPD en une étape)

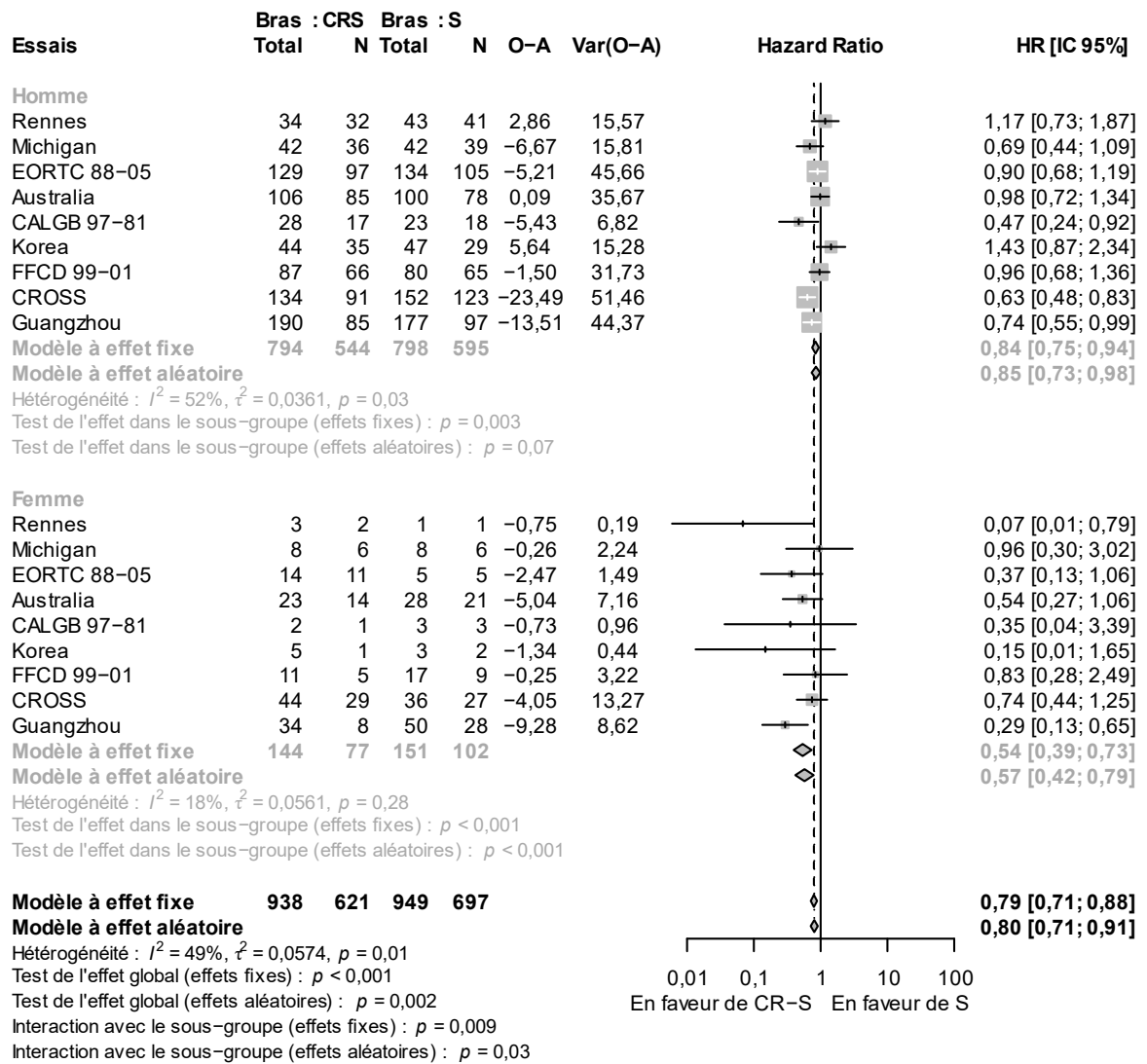


Figure annexe 4 : Forest plot de l'effet de CRS contre S sur la SSR selon le sexe



Annexes

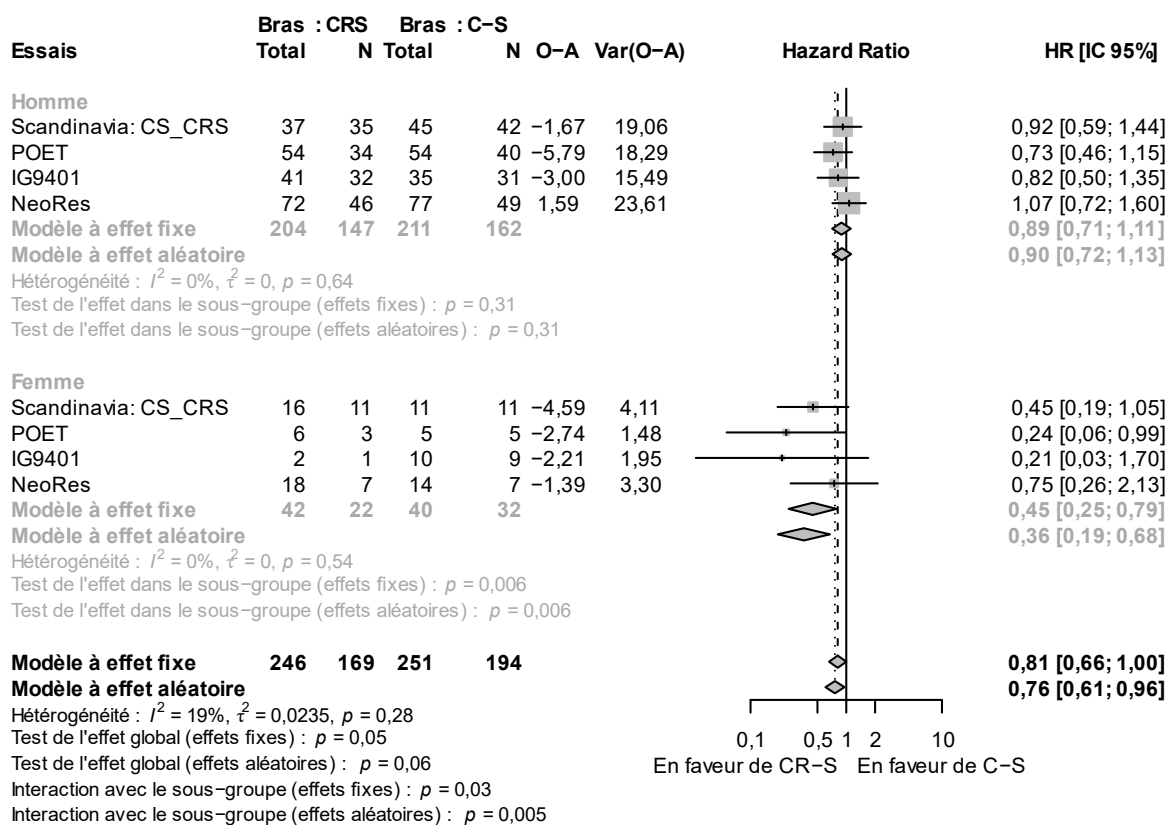


Figure annexe 5 : Forest plot de l'effet de la CRS contre CS sur la survie globale en fonction du sexe



**Titre :** *Meta analyse des traitements néoadjuvants des cancers de l'œsophage et de la jonction oeso-gastrique*

**Mots clés :** cancer oesophage, jonction oeso-gastrique, méta-analyse, neoadjuvant, interaction statistique

**Résumé :** Les cancers de l'œsophage restent grevés d'un pronostic à long terme sombre. Des stratégies multimodales associant différemment chirurgie (S), chimiothérapie (C) et radiothérapie (R) ont été développées afin d'améliorer ce pronostic. Ce cancer est également hétérogène en raison de la présence de deux histologies différentes (adénocarcinomes et carcinomes épidermoïdes) et deux localisations anatomiques (œsophage thoracique et jonction œsogastrique). La problématique était donc l'étude de l'interaction entre l'effet traitement et des modificateurs dans une situation clinique comportant plus de deux traitements. La méta-analyse en réseau sur données individuelles d'essais randomisés était donc la méthodologie la plus appropriée.

La première partie de la thèse a permis l'étude des performances d'un modèle permettant d'évaluer une méta-analyse en réseau sur données individuelles en une étape à critère de jugement de type survie sous paradigme fréquentiste.

L'accent était porté sur une situation où de l'interaction avec un modificateur était présente.

La seconde partie de la thèse portait sur la réalisation d'une analyse sur données individuelles. Trois méta-analyses par paire comparant les trois stratégies actuelles S contre CS, S contre CRS et CS contre CRS ont été réalisées. Cette analyse a été complétée par une méta-analyse en réseau utilisant toute l'information disponible. Le critère de jugement principal était la survie globale. Les critères secondaires étaient la survie sans récurrence, le mode de récurrence (sous forme de risques compétitifs) les causes de décès, la qualité de la résection et la morbi-mortalité postopératoire. Pour chaque analyse une recherche d'interaction entre l'effet traitement et le type histologique ou la localisation tumorale était recherchée.

**Title:** *Meta-analysis of neoadjuvant treatment of esophageal and gastroesophageal junction carcinoma*

**Keywords:** esophagus cancer, gastro-esophageal junction, meta-analysis, neoadjuvant, statistical interaction

**Abstract:** Esophagus cancer still carries a dismal prognosis. Multimodal treatment strategies usually combine surgery (S), chemotherapy (C) and radiotherapy (R) to obtain better chances of survival. Esophagus cancer is also a heterogeneous entity with two histological subtypes (adenocarcinoma and squamous cell carcinoma) and two anatomical locations (thoracic esophagus and gastroesophageal junction). The challenge was the evaluation of the interaction between the treatment effect and modifiers in a clinical situation where more than two treatments are available. Individual participant data network meta-analysis was therefore the optimal methodology.

The first part of the thesis focused on studying the performance of a model for one step individual participant data network meta-analysis for a time to event outcome under the frequentist paradigm.

Special emphasis was given to the situation in which an interaction between the treatment effect and a modifier is present.

The second part of the thesis dealt with the analysis of the individual participant data of available studies. Three pairwise meta-analyses comparing the three most frequent therapeutic sequences were conducted : S vs. CS, S vs. CRS and CS vs. CRS. These analyses were completed by a network meta-analysis combining all the available information in a one-step design. Primary outcome was overall survival. Secondary outcomes were disease-free survival, type of recurrence (local or distant or death without recurrence with competing risks), cause of death, quality of the surgical resection and postoperative morbidity and mortality. For each analysis an interaction between the treatment effects and the histological subtypes or the anatomical location were sought.