



**HAL**  
open science

# Study of the occurrence of multiple iatrogenic events in long-term survivors of childhood cancer

Thibaud Charrier

► **To cite this version:**

Thibaud Charrier. Study of the occurrence of multiple iatrogenic events in long-term survivors of childhood cancer. Cancer. Université Paris-Saclay, 2024. English. NNT: 2024UPASR030. tel-04938759

**HAL Id: tel-04938759**

**<https://theses.hal.science/tel-04938759v1>**

Submitted on 10 Feb 2025

**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.

# Study of the occurrence of multiple iatrogenic events in long-term survivors of childhood cancers

*Etude de la survenue d'évènements iatrogènes  
multiples des traitements des cancers pédiatriques*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 570, Santé Publique (EDSP)

Spécialité de doctorat: Biostatistiques et data sciences

Graduate School : Santé publique, Référent : Faculté de médecine

Thèse préparée dans les unités de recherche CESP (Université Paris-Saclay, UVSQ, Gustave Roussy, Inserm) et U900 (Institut Curie, Inserm), sous la direction du Dr **Rodrigue Allodji** et la co-direction du Pr **Aurélien Latouche**

Thèse soutenue à Paris Saclay, le 16 décembre 2024, par

**Thibaud CHARRIER**

## Composition du jury

### Membres du jury avec voix délibérative

<b>Pierre BLANCHARD</b> PU-PH, Université Paris Saclay	Président
<b>Gérard MICHEL</b> PU-PH, Aix Marseille Université	Rapporteur & Examineur
<b>Etienne DANTAN</b> Docteur (HDR), Nantes Université	Rapporteur & Examineur
<b>Mounia HOCINE</b> Docteur (HDR), Conservatoire national des arts et métiers	Examinatrice
<b>Karen LEFFONDRE</b> PU, ISPED - Université de Bordeaux	Examinatrice

**Titre:** Etude de la survenue d'évènements iatrogènes multiples des traitements des cancers pédiatriques

**Mots clés:** Anciens patients de cancers pédiatriques; évènements iatrogènes multiples; seconds cancers; maladies cardiaques; années de vie perdues; variable dépendant du temps

**Résumé:** Le taux de survie à 5 ans des cancers de l'enfant atteint aujourd'hui 80 %. Les survivants à 5 ans de cancer pédiatrique (CCS) ont cependant un risque accru d'occurrence de nombreux évènements. Beaucoup d'études ont été réalisées pour comprendre les liens entre ces évènements iatrogènes et les traitements pédiatriques, et un taux élevé d'évènements iatrogènes multiples a été observé. Cependant, ce taux d'évènements iatrogènes multiples est méconnu. Ce projet de thèse avait pour objectif d'étudier les causes et conséquences de ce taux élevé.

La recherche présentée dans cette thèse est basée sur les données de la French Childhood Cancer Survivors Study, une cohorte rétrospective à suivi prospectif de 7670 CCS diagnostiqués avant l'âge de 21 ans entre 1945 et 2000 en France. La cohorte contient des données détaillées sur les traitements des cancers pédiatriques, et un important travail de validation des évènements iatrogènes a été réalisé. Nous avons utilisé la méthode de landmark combiné avec des pseudo-observations, afin de pouvoir étudier les effets des seconds cancers sur l'incidence cumulée d'évènement cardiaque et les années de vie perdues des patients.

Nous avons constaté une augmentation de risque instantané (csHR: 2.1, 95% CI: 1,5-2,9) et de l'incidence cumulée d'évènement cardiaque (CD) (+3,8%, 95% CI: 0,5-7,1) après un second cancer. Nous avons aussi constaté un impact de la radiothérapie, des seconds cancers, et des CD sur l'espérance de vie des patients, avec respectivement 6, 10,5 et 7,7 années de vie perdues à 16 ans. Nous avons trouvé un effet nul de l'interaction entre second cancer et CD, mais notre analyse de simulation a montré que ce résultat est biaisé par la corrélation entre seconds cancers et CD.

En conclusion, dans cette thèse nous avons démontré que le risque accru de multi-morbidité chez les CCS est partiellement attribuable à l'occurrence d'un premier évènements iatrogène. Nous avons aussi montré un fort impact des seconds cancers et évènements cardiaques sur l'espérance de vie des patients. Ces résultats nous amènent à recommander une forte surveillance des patients développant l'une de ces maladies, et de poursuivre la recherche sur la multi-morbidité chez les CCS qui apparaît complexe et forte de conséquences pour les patients.

**Title:** Study of the occurrence of multiple iatrogenic events in long-term survivors of childhood cancers

**Keywords:** Survivors of childhood cancers; Multiple iatrogenic events; Subsequent tumours; Cardiac Disease; life years lost; time dependent covariate

**Abstract:** 5-years survival of childhood cancer exceeds 80% today. Nonetheless, 5-year childhood cancer survivors (CCS) are at increased risk of health-related late effects. Many studies have been conducted to investigate the link between those late effects and childhood cancer treatments. Furthermore, an elevated number of patients experiencing multiple iatrogenic event was observed. However, little is known about the causes of this multi-morbidity. This thesis aimed to study the causes et consequences of this elevated multi-morbidity.

Research presented in this thesis is based on the French Childhood Cancer Survivors Study, a retrospective cohort with prospective follow-up, following 7,670 CCS diagnosed in France before 21 years old between 1945 and 2000. This cohort contains detailed data regarding childhood cancer treatments, and the iatrogenic event observed were validated by trained professionals. We used the landmark method combined with pseudo-observations to study the consequences of subsequent malignant neoplasm on the cumulative incidence of cardiac disease, and the life years lost by CCS.

We found a subsequent malignant neoplasm (SMN) to cause a two-fold increase in the cause-specific hazard of cardiac disease (CD) and a 3.8% increase of its cumulative incidence. We found the use of radiotherapy to treat childhood cancer, the occurrence of SMN, and of CD to impact the life expectancy of CCS, with respectively 6, 10.5, and 7.7 life years lost at 16 years old. We found a nul effect of the interaction of SMN and CD on the life expectancy, although our simulation study suggested this was biased by the correlation between SMN and CD.

In conclusion, we demonstrated that the increased risk of multi-morbidity among CCS is partially attributable to the occurrence of a first iatrogenic event. We also showed that subsequent malignant neoplasm and cardiac disease have a strong impact on the life expectancy of CCS. Therefore, we recommend to keep following in details the CCS developing either disease, and to pursue further research on the multi-morbidity among CCS which appears to be complex and consequential for patients.

# Résumé Substantiel

**Titre :** Etude de la survenue d'évènements iatrogènes multiples des traitements des cancers pédiatriques

**Mots clés :** Anciens patients de cancers pédiatriques ; évènements iatrogènes multiples ; seconds cancers ; maladies cardiaques ; années de vie perdues ; variable dépendant du temps

## Contexte

La survie des enfants atteints de cancer s'est nettement améliorée ces dernières décennies, grâce aux progrès thérapeutiques. Le taux de survie à 5 ans dépasse actuellement 80% dans la plupart des pays à revenu élevé (Steliarova-Foucher et al. 2004 ; Botta et al. 2022). On estime aujourd'hui qu'en Europe entre 300 000 et 500 000 personnes sont des anciens patients de cancer pédiatrique, et cette population continue de croître (Whelan and Alva 2018).

Cependant, cette survie accrue chez ces patients a mis en évidence de nombreuses séquelles des traitements reçus dans l'enfance. Ces séquelles affectent tous les organes et systèmes corporels, à des degrés variables (Robison and Hudson 2014). Parmi ces séquelles se trouvent notamment les seconds cancers, les maladies cardiaques, les diabètes, et les maladies rénales chroniques (Oeffinger et al. 2006 ; Haddy et al. 2016 ; Reulen et al. 2010).

De nombreux travaux ont étudiés les liens entre les traitements de cancer pédiatriques et ces effets secondaires. Ainsi, la radiothérapie et la chimiothérapie sont reconnus comme responsables de la principale augmentation du risque de certains effets secondaires dans cette population (Hudson et al. 2021). Des effets de niveau de doses ont été trouvés, ainsi que des différences selon les organes exposés à la radiothérapie. Des différences ont aussi été trouvées en fonction de l'âge au diagnostic et du sexe. Aujourd'hui, des travaux sont menés pour explorer les effets des faibles doses de radiothérapie et des chimiothérapie cytotoxiques (Chounta, Allodji, et al. 2023). Les facteurs génétiques sont aussi étudiés, afin de raffiner les profils de risques des patients



et permettre un suivi à long terme personnalisé (Ducos et al. 2023 ; Aba et al. 2023).

Des travaux ont aussi mis en évidence un risque de multi-morbidité accrue dans cette population (Oeffinger et al. 2006 ; Moskowitz et al. 2019). Cette multi-morbidité peut s'expliquer par des facteurs de risques communs entre les effets secondaires, la radiothérapie et la chimiothérapie augmentant le risque de nombreux évènements. Aucun travail n'a cherché à quantifier les causes de multi-morbidité, notamment pour savoir si l'occurrence d'un effet secondaire augmenterait le risque d'occurrence d'un autre. Cette multi-morbidité a des conséquences sur la vie des patients, et certaines études ont aussi suggéré un effet de celle-ci sur la mortalité des patients (Moskowitz et al. 2019).

L'objectif principal de cette thèse est d'étudier la survenue d'évènements iatrogènes multiples des traitements des cancers pédiatriques.

### **La cohorte et les données**

Pour atteindre nos objectifs, nous avons utilisé les données de la cohorte French Childhood Cancer Survivors Study (FCCSS). La FCCSS est une cohorte rétrospective à suivi prospectif qui comprend 7670 anciens patients de cancer pédiatrique diagnostiqués pour un cancer solide ou un lymphome avant l'âge de 21 ans entre 1946 et 2000 dans cinq centres de lutte contre le cancer en France et ayant survécu au moins 5 ans après leur diagnostic. Le suivi prospectif de cette cohorte a été réalisé à l'aide de plusieurs sources, notamment les dossiers médicaux, des questionnaires auto-déclarés, les consultations de suivi à long terme, et les données du Système National des Données de Santé. Ce suivi a permis d'identifier et valider les effets secondaires dans cette cohorte. Les données disponibles sur les traitements de cancer pédiatriques incluent les doses de radiothérapie pour chacun des organes du corps, les doses cumulées de chimiothérapie par drogue en  $\text{mg}/\text{m}^2$ , le sexe, l'âge au diagnostic de cancer pédiatrique, et le type de cancer pédiatrique selon la classification internationale pour les cancers pédiatriques - 3e édition.

## Quantifier les conséquences d'effets secondaires

Une part importante du travail mené a consisté à rechercher des méthodes appropriées pour répondre aux objectifs posés. La difficulté réside notamment dans l'analyse de données de survie, qui sont incomplètement observées à cause de mécanismes connus comme la censure à droite et la troncation à gauche. Les données observées en analyse de survie sont les temps d'évènement  $T$ . La censure à droite arrive lorsque le temps  $T$  n'est pas observé, mais qu'il est observé que l'évènement n'a pas eu lieu avant le temps de censure  $C$ . La troncation à gauche arrive lorsqu'un patient n'est pas observé avant un temps  $T_L$ , et qu'ainsi le temps  $T$  est nécessairement supérieur à ce  $T_L$ . Deux quantités sont souvent utilisés pour étudier ce type de données, à savoir le risque instantané d'évènement, et l'incidence cumulée qui est la probabilité d'avoir l'évènement avant un temps  $t$ . Le risque instantané est la quantité la plus utilisée pour les modèles de régression pour la simplicité d'utilisation des modèles, mais les interprétations sur l'incidence cumulée sont préférées. Dans le cas simple d'un unique évènement et de variable à temps fixe (ie. ne changeant pas au cours du temps), les modèles sur le risque instantané donnent des informations sur l'effet des variables sur l'incidence cumulée.

La seconde difficulté réside dans la nature de nos variables d'intérêts. Nos variables d'intérêts sont les effets secondaires observés, tel que les seconds cancers. La valeur de ces variables changeant au cours du temps, celles-ci sont des variables dépendantes du temps. De plus, celles-ci appartiennent à la sous catégorie de variables internes dépendantes du temps, car internes au patient (Kalbfleisch and Prentice 2002). Malheureusement, les modèles de régression sur le risque instantané ne peuvent pas être utilisés pour informer de l'effet d'une variable interne dépendant du temps sur l'incidence cumulée (Austin, Latouche, and Fine 2020). Nous avons donc recherché une méthode adaptée à une interprétation de l'effet d'une variable interne dépendant du temps sur l'incidence cumulée, cette quantité étant une probabilité et donc de plus grand intérêt que le risque instantané.

La solution utilisée au cours de la thèse combine la méthode de landmark et les pseudo-

observations. La méthode de landmark consiste en choisir un temps  $l$  et analyser les données conditionnellement à la survie au temps  $l$ , en remplaçant les variables dépendants du temps  $X(t)$  par les variables à temps fixe ayant pour valeur  $X(l)$  (Anderson, Cain, and Gelber 1983; Cortese and Andersen 2010). Les pseudo-observations sont une méthode utilisée en analyse de survie pour remplacer le temps d'évènement qui est incomplètement observé, par la contribution du patient à la l'estimation de la quantité d'intérêt avec un estimateur jackknife (Per Kragh Andersen, Klein, and Rosthøj 2003; Graw, Gerds, and Schumacher 2009). Ces pseudo-observations sont ensuite utilisées comme variable dépendante dans un modèle statistique classique, et permettent de réaliser des régressions sur des quantités autrement difficiles à estimer en survie. Notamment, nous avons utilisé les pseudo-observations pour estimer des effets sur l'incidence cumulée d'évènement – qui est la probabilité d'avoir l'évènement avant un temps  $t$ , et les années de vie perdues.

### **Les principales contributions**

#### **L'effet d'un second cancer sur le risque d'évènement cardiaque**

La première contribution a été d'étudier le risque d'évènement cardiaque après un second cancer. L'objectif premier était de mettre en évidence qu'un effet secondaire pouvait augmenter le risque d'occurrence d'un autre effet secondaire et ainsi que la multi-morbidité dans notre population dépasse le simple cas d'augmentation simultanée par les traitements de cancer pédiatriques. Le deuxième objectif était de quantifier l'éventuelle augmentation de risque d'évènement cardiaque, afin d'améliorer les recommandations de suivi à long terme des patients.

Utilisant la cohorte FCCSS et la méthode de landmark, nous avons quantifié l'effet d'un second cancer sur la probabilité d'avoir un évènement cardiaque avant un temps  $t$  en prenant en compte le décès comme risque compétitif. Nous avons pu ajusté nos modèles de régression sur les traitements de cancer pédiatriques, afin de prendre en compte les facteurs de risque communs entre seconds cancers et évènements cardiaques. Les données de doses de radiothérapie

disponibles dans la FCCSS sont rarement disponibles en pratique clinique, nous avons donc aussi estimé l'effet d'un second cancer en ajustant uniquement pour les informations habituellement disponibles en pratique clinique – à savoir l'utilisation ou non de radiothérapie et/ou chimiothérapie. Nous avons estimé qu'un second cancer augmentait la probabilité d'évènement cardiaque de 3,8% (95% CI : 0,5% - 7,1%) au maximum, après ajustement complet sur les doses de radiothérapie et chimiothérapie.

Nous avons aussi cherché à identifier un effet des traitements des seconds cancers. Nous n'avons pas accès à ces données de traitement, nous avons donc décidé d'utiliser les données de type de second cancer, qui sont fortement associés aux informations de traitement. Nous n'avons pas trouvé de différences concluantes sur la probabilité d'évènement cardiaque, probablement dû au faible nombre de seconds cancers observées par sous-groupe de type créés selon la cardio-toxicité des traitements probablement reçus. Nous avons toutefois observé que les points d'estimation suggéraient un effet plus important des cancers du sein (radiothérapie au coeur), sarcomes osseux, et des cancers des tissus mous (anthracyclines) comparés aux autres types de cancers. Finalement, nous avons aussi regardé un effet du temps depuis le second cancer sur la probabilité d'évènement cardiaque. L'idée sous-jacente est que le risque de décès est plus élevé dans les premières années après un évènement cardiaque, et que ce changement dans le risque de décès pourrait influencer l'effet d'un second cancer sur la probabilité d'évènement cardiaque. Nous n'avons pas conclu à une différence d'effet au cours du temps, de nouveau du au faible nombre d'évènements observés par catégorie.

Nous avons aussi réalisé ces analyses pour quantifier l'effet d'un second cancer sur le risque instantané d'évènement cardiaque à l'aide d'un modèle de Cox. Après ajustement pour les doses de radiothérapie et chimiothérapie, nous avons pu trouver que l'occurrence d'un second cancer multiplie par 2,1 (95% CI : 1,5 - 2,9) le risque instantané d'évènement cardiaque, et par 5,6 (95% CI : 4,7 - 6,6) celui de décès. Nous avons de nouveau trouvé des différences par type de cancer, avec les sarcomes osseux et cancers des tissus mous multipliant par 2,3 (95% CI :

1,2 - 4,4) le risque instantané d'évènement cardiaque, les cancers du sein par 1,7 (95% CI : 1,0 - 3,0), et les autres cancers par 1,4 (95% CI : 1,0 - 1,9). De même, nous avons trouvé des différences selon le temps depuis le second cancer. Dans les 5 premières années après le second cancer, le risque instantané d'évènement cardiaque est multiplié par 2,5 (95% CI : 1,5 - 4,1) et celui de décès par 10,4 (95% CI : 8,8 - 12,5), alors que plus de 5 ans après le second cancer ces risques sont multipliés respectivement par 1,9 (95% CI : 1,3 - 2,8) et 2,9 (95% CI : 2,3 - 3,6). Ainsi, nous avons montré qu'un second cancer augmente le risque instantané et la probabilité d'occurrence d'évènement cardiaque, prouvant ainsi que la multi-morbidité chez les anciens patients de cancer pédiatrique ne se limite pas à des facteurs de risque partagés. Nous avons aussi obtenu des résultats interprétables dans un contexte de santé publique et clinique.

### **Les années de vie perdues**

La deuxième contribution a été d'étudier les années de vie perdues par les traitements de cancer pédiatriques, les effets secondaires, et la co-occurrence d'effets secondaires. Cette contribution avait deux objectifs, le premier de quantifier la mortalité chez les anciens patients de cancer pédiatrique sur une métrique communicable dans un contexte de santé publique, et la deuxième d'étudier l'effet propre de la multi-morbidité sur la mortalité.

Nous avons utilisé la FCCSS et la méthode de landmark pour étudier les années de vie perdues jusqu'à 60 ans, ainsi que les années de vie perdues dans les dix années après le temps de landmark. Ainsi, nous avons pu fournir des résultats interprétables et communicables tout en étudiant les changements des facteurs de risque au cours de la vie des patients.

Nous avons ainsi montré que la radiothérapie est le traitement pédiatrique qui contribue le plus aux années de vie perdues avec 6,0 années de vie perdues (95% CI : 4,7 - 7,3) au temps de landmark 16 ans. Nous avons aussi montré que les seconds cancers et évènements cardiaques contribuaient considérablement aux années de vie perdues, avec 10,5 (95% CI : 5,4 - 15,7) années de vie perdues par seconds cancers et 7,7 (95% CI : 1,1 - 14,3) par évènement cardiaque

au temps de landmark 16 ans. Au temps de landmark 32 ans, nous avons estimé respectivement 4,3 (95% CI : 1,8 - 6,7) et 6,4 (95% CI : 2,5 - 10,3) années de vie perdues.

Nous avons trouvé une augmentation des années de vie perdues par radiothérapie au cours de la vie des patients. Nous n'avons pas observé de changement au cours de la vie pour les années de vie perdues par second cancer, évènements cardiaques, et âge au diagnostic de cancer pédiatrique.

Nous n'avons pas trouvé d'effets propre de la multi-morbidité sur les années de vie perdues des anciens patients de cancer pédiatrique. Nous soupçonnons notre faible prise en compte de la réduction du risque de mortalité par second cancer au cours du temps depuis le second cancer d'avoir biaisé notre estimation des années de vie par multi-morbidité vers un effet protecteur de celle-ci. Ce soupçon a été l'élément motivateur de l'analyse de simulation constituant la troisième contribution de cette thèse.

### **Simulation de la sensibilité du landmark aux coefficients dépendant du temps**

La troisième contribution a été une analyse de simulation étudiant la sensibilité de la méthode de landmark aux coefficients dépendant du temps de variable dépendant du temps, motivée par les résultats trouvées sur les années de vie perdues.

Soit  $X_1(t) = \mathbb{1}(\text{second cancer observé avant } t)$ ,  $X_2(t) = \mathbb{1}(\text{évènement cardiaque observé avant } t)$  des variables binaires dépendant du temps et changeant une seule fois. Si l'effet de  $X_1(t)$  sur le décès décroît avec le temps depuis son temps de pas, alors l'estimateur  $\widehat{\beta}_1^{LM}$  acquis par landmark est affecté par la différence entre le temps de landmark  $l$ , et le temps de pas de  $X_1(t)$ . Lorsque nous avons estimé l'effet de  $X_1(t)$  et  $X_2(t)$ , nous nous sommes aussi posé la question d'un effet de  $X_3(t) = \mathbb{1}(X_1(t) \& X_2(t))$ . Toutefois, les deux groupes de patients tel que  $X_1(l) = 1, X_2(l) = 0$  et  $X_1(l) = 1, X_2(l) = 1$  ont une distribution de temps de pas de  $X_1(l)$  différentes, avec la première en moyenne plus proche de  $l$  que la seconde. En conséquence, l'effet de  $X_1(t)$  est plus faible dans le groupe  $X_1(l) = 1, X_2(l) = 1$ , et l'effet estimé de  $X_3(t)$  est biaisé

par cette différence. Nous avons étudié l'impact de cet effet.

A l'écriture de ce manuscrit, les résultats préliminaires suggèrent que cet effet est faible dans notre population.

### **Discussion générale**

Les analyses sur la cohorte FCCSS ont permis de mettre en évidence un fort effet des seconds cancer sur les maladies cardiaques. Nous avons aussi mis en évidence que les effets secondaires des traitements de cancer pédiatriques avaient un impact marqué sur l'espérance de vie des patients. Finalement, nous avons réalisé une analyse de simulation complète pour étudier la stabilité des résultats obtenus par analyse de landmark dans notre configuration.

Ces travaux suggère la besoin de mener une étude approfondie de la multi-morbidité chez les anciens patients de cancer pédiatrique. Des questions restent ouvertes, tel que la contribution des traitements de second cancer à l'augmentation d'évènements cardiaque observés. Les patterns de multi-morbidités chez les patients les plus âgés nécessiteront aussi d'être étudié lorsque les cohortes auront un suivi suffisamment long, afin de comparer à la multi-morbidité en population gériatrique générale.

Les résultats fournis par cette thèse contribueront ainsi à améliorer les recommandations internationales pour la surveillance des survivants de cancer pédiatrique, notamment pour le risque d'évènement cardiaque.

## Acknowledgements

Tout d'abord, je tiens à remercier mes directeurs de thèse, Docteur Rodrigue Allodji et Professeur Aurélien Latouche pour leur confiance, attention et soutien pendant ces trois années de thèse.

Je tiens à remercier chacun des membres du jury pour avoir accepté de juger ce travail. Je remercie le Docteur Etienne Dantan et le Professeur Gérard Michel qui m'ont fait l'honneur d'être rapporteurs de ma thèse. Au Professeur Pierre Blanchard, au Professeure Karen Leffondré et au Docteur Mounia Hocine, je vous remercie d'avoir accepté d'examiner ce travail.

Je souhaite ensuite remercier l'ensemble de l'équipe Epidémiologie des Radiations, et l'équipe Méthodes statistiques pour la médecine personnalisée pour leur accueil et soutien. Merci à Boris et Brice pour leur conseils scientifiques. Merci à Giao, Vincent S, et Amel Belhout pour leur aide avec les données de la FCCSS. Merci à Françoise pour sa grande disponibilité au sein de l'équipe et son aide avec les tâches administratives.

Un grand merci à tous les (ex/)doctorants que j'ai côtoyés pendant ces trois ans: May, Daniel, Stefania, Naïla, Claire, Vincent Z, Rivalin, Beatriz, et Antoine pour votre écoute et soutien moral sans faille. Merci à Paul, Maxence, Juliette, Marc, Valentine, et Mary pour les rires et nombreux jeux pendant cette thèse.

A mes parents, un grand merci pour votre soutien et encouragement pendant ces trois ans de thèse ! Bien sûr à Adrien, Annabelle, Anne-Marie et Jacqueline.

Et finalement, merci à Camille pour avoir rendu cette thèse aussi simple à vivre.





## Communications and supervisions

### Published Articles:

Thibaud Charrier, Nadia Haddy, Boris Schwartz, Neige Journy, Brice Fresneau, Charlotte Demoor-Goldschmidt, Ibrahima Diallo, Aurore Surun, Isabelle Aerts, François Doz, Vincent Souchard, Giao Vu-Bezin, Anne Laprie, Sarah Lemler, Véronique Letort, Carole Rubino, Stéphanie Chounta, Florent de Vathaire, Aurélien Latouche, and Rodrigue S. Allodji. Increased Cardiac Risk After a Second Malignant Neoplasm Among Childhood Cancer Survivors: A FCCSS Study. *J Am Coll Cardiol CardioOnc.* 2023 Dec,5 (6) 792–803. 10.1016/j.jacc.2023.07.008

Thibaud Charrier, Nadia Haddy, Brice Fresneau, Boris Schwartz, Neige Journy, Charlotte Demoor-Goldschmidt, Ibrahima Diallo, Isabelle Aerts, François Doz, Vincent Souchard, Giao Vu-Bezin, Anne Laprie, Sarah Lemler, Véronique Letort, Carole Rubino, Kaniav Kamary, Naïla Myriam Aba, Claire Ducos, Médéa Locquet, Florent de Vathaire, Rodrigue S. Allodji, Aurélien Latouche. Life Years Lost by Childhood Cancer Treatment and Health Related Late Effects Among Childhood Cancer Survivors. *Cancer Epidemiology.* 2024 Dec,102692. 10.1016/j.canep.2024.102692

### Articles Under Preparation:

Thibaud Charrier, Aurélien Latouche, Rodrigue S. Allodji. Landmark Estimation of Life Years Lost with Time-Dependent Covariates, Time-Varying Effect, and their Interaction. *Journal undetermined.* Under writing.

### Oral Presentation in Conferences

Journées de Biostatistique du G.D.R. “Statistique et Santé”, la Société française de biométrie et du groupe “Biopharmacie et Santé” de la Société française de statistiques, Rennes, France, 17th - 18th November 2022. Increased Cardiac Risk After a Second Malignant Neoplasm Among Childhood Cancer Survivors: A FCCSS Study

**Posters in Conferences:**

Radiation Research Society 2022 Annual Meeting, Waikoloa Village, Hawaii, USA, 16th - 18th October 2022. Increased Cardiac Risk After a Second Malignant Neoplasm Among Childhood Cancer Survivors: A FCCSS Study.

International Society of Radiation Epidemiology and Dosimetry 1st meeting , Sitges, Spain, 16th - 18th May 2023. Increased Cardiac Risk After a Second Malignant Neoplasm Among Childhood Cancer Survivors: A FCCSS Study.

International Society for Clinical Biostatistics Annual Conference, Milan, Italy, 27th - 31st August 2023. Increased Cardiac Risk After a Second Malignant Neoplasm Among Childhood Cancer Survivors: A FCCSS Study.

**Oral Presentations in Seminars:**

L'Assemblée Générale et la Journée Scientifique Annuelle du CESP, Faculté de médecine, Kremlin-Bicêtre, France, 13th November 2023. Increased Cardiac Risk After a Second Malignant Neoplasm Among Childhood Cancer Survivors: A FCCSS Study.

Séminaire d'unité U900 Curie

**Research Supervision:**

Mohamed AIT-ALI: Master 2 internship (March - August 2024), Sorbonne University. Exploration des méthodes de classes latentes pour les parcours de soins des survivants de cancer pédiatrique.

**Other publications in which the doctoral student has contributed:**

Stefania Chounta, Rodrigue Allodji, Maria Vakalopoulou, Mahmoud Bentriou, Duyen Thi Do, Florent De Vathaire, Ibrahima Diallo, Brice Fresneau, Thibaud Charrier, Vincent Zossou, Stergios Christodoulidis, Sarah Lemler, Veronique Letort Le Chevalier, Dosiomics-Based Prediction of Radiation-Induced Valvulopathy after Childhood Cancer. *Cancers (Basel)* 2023 Jun 8;15(12):3107 10.3390/cancers15123107

# Contents

<b>Résumé Substantiel</b>	<b>III</b>
<b>Acknowledgements</b>	<b>XI</b>
<b>Communications and supervisions</b>	<b>XIII</b>
<b>Contents</b>	<b>XV</b>
<b>List of Tables</b>	<b>XXIII</b>
<b>List of Figures</b>	<b>XXV</b>
<b>Abbreviations</b>	<b>XXXI</b>
<b>1 Introduction to Childhood Cancer Survivors</b>	<b>1</b>
1.1 Childhood Cancer Overview . . . . .	1
1.1.1 Childhood Cancer Incidence . . . . .	1
1.2 Childhood Cancer Mortality . . . . .	2
1.3 Childhood Cancer Treatment . . . . .	3
1.4 Childhood Cancer Late Effects . . . . .	4
1.5 Multi-Morbidity after Childhood Cancer . . . . .	5
1.5.1 Definitions . . . . .	6
1.5.2 Multiple Events in Childhood Cancer Survivors . . . . .	6
1.5.3 Plausible Pathways . . . . .	7
1.5.4 Death, a Competing Event to Multi-Morbidity . . . . .	8
1.6 Thesis Objectives . . . . .	9

<b>2</b>	<b>Thesis Data: The French Childhood Cancer Survivors Study</b>	<b>11</b>
2.1	History of the French Childhood Cancer Survivors Study . . . . .	11
2.2	Previous use . . . . .	11
2.3	Legal Authorization . . . . .	12
2.4	Childhood Cancer Treatment . . . . .	12
2.5	Iatrogenics Events Identification and Validation . . . . .	13
2.6	Vital Status and Iatrogenics Events . . . . .	13
2.7	Patients' characteristics . . . . .	14
2.7.1	Overall Description . . . . .	17
2.7.2	Changes Across Years of Childhood Cancer Diagnosis . . . . .	17
2.8	Morbidities and Mortality . . . . .	18
2.9	Multi-Morbidity . . . . .	20
2.9.1	Multiple Events . . . . .	20
2.9.2	Hazard Functions . . . . .	21
<b>3</b>	<b>Survival analysis and extensions to account for multiple type of events and time-dependent exposure</b>	<b>23</b>
3.1	Event History Data . . . . .	23
3.2	Survival Analysis . . . . .	25
3.2.1	Non Parametric Estimators for Survival Analysis . . . . .	28
3.2.2	Life Years Lost by a Categorical Covariate . . . . .	29
3.2.3	Regression Model for the Hazard . . . . .	30
3.2.4	Pseudo Observations for Survival Analysis . . . . .	31
3.3	Multi-State Models . . . . .	33
3.3.1	Non Parametric Estimators for Multi-States . . . . .	36
3.3.2	Proportional Cause-Specific Hazard Models . . . . .	37

3.3.3	Pseudo Observations for Multi-State . . . . .	39
3.3.4	Life Years Lost by Cause and Life Years Lost by Covariate . . . . .	40
3.4	Time-Dependent Covariates . . . . .	41
3.4.1	Proportional Hazard Models . . . . .	41
3.4.1.1	Time-Dependent Covariates without Competing Risk . . . . .	42
3.4.1.2	Time-Dependent Covariates with Competing Risk . . . . .	43
3.4.2	Landmark . . . . .	44
3.4.2.1	The Concept of Landmark . . . . .	44
3.4.2.2	Landmark for Time-Dependent Covariates with Competing Risks . . . . .	45
3.4.2.3	Landmark to Study Time-Varying Coefficients . . . . .	46
3.4.2.4	Dynamic Pseudo Observations . . . . .	47
3.4.2.4.1	Pseudo-Observations with a Single Landmark Time Point . . . . .	47
3.4.2.4.2	Pseudo-Observations with Landmark Super Models . . . . .	48
3.4.2.4.3	Pseudo Observations with Left Truncation . . . . .	49
3.4.3	Modelling a Transient State with a Time-Dependent Covariate . . . . .	50
<b>4</b>	<b>Cardiac Disease Risk after a Second Malignant Neoplasm</b>	<b>53</b>
4.1	Context . . . . .	53
4.2	Methods . . . . .	54
4.2.1	Data . . . . .	54
4.2.2	Metrics of Interest . . . . .	55
4.2.2.1	Cumulative Incidence . . . . .	55
4.2.2.2	Cause Specific Hazard . . . . .	56
4.2.3	Childhood Cancer Information . . . . .	56
4.2.4	SMN Type . . . . .	58

4.2.5	Time Dependent Effect of SMN . . . . .	58
4.3	Results . . . . .	59
4.3.1	Population at Landmark Time . . . . .	59
4.3.2	Cumulative Incidence . . . . .	64
4.3.2.1	Effect of any SMN occurrence . . . . .	64
4.3.2.2	Effect by Type of SMN . . . . .	66
4.3.2.3	Effect of SMN within 5 Years and Later . . . . .	71
4.3.3	Cox Model . . . . .	71
4.3.3.1	Effect of any SMN occurrence . . . . .	71
4.3.3.2	Effect by Type of SMN . . . . .	77
4.3.3.3	Effect of SMN within 5 Years and Later . . . . .	77
4.3.4	Effects of Childhood Cancer Treatments . . . . .	80
4.4	Discussion . . . . .	80
4.4.1	Increased Risk of Cardiac Disease After a Second Malignant Neoplasm . . . . .	80
4.4.2	The confounding of childhood cancer treatments . . . . .	83
4.4.3	Changes of SMN effect on CD risk over time. . . . .	84
4.4.4	The effect of cardio-toxic treatments of SMN. . . . .	84
4.4.5	Limitations . . . . .	85
4.4.6	Perspectives . . . . .	86
4.5	Summary . . . . .	86
<b>5</b>	<b>Life Years Lost by Iatrogenic Events</b>	<b>87</b>
5.1	Context . . . . .	87
5.2	Methods . . . . .	88
5.2.1	Data . . . . .	88
5.2.2	Life Years Lost . . . . .	89

5.2.2.1	Covariates . . . . .	89
5.2.2.2	Life Years Lost with Left Truncation and Right Censoring . . . . .	90
5.2.2.3	Pseudo-Observations to Fit the Coefficients . . . . .	91
5.2.3	Life Years Lost by Cardiac Death . . . . .	91
5.3	Results . . . . .	92
5.3.1	Life Years Lost up to 60 Years Old . . . . .	92
5.3.1.1	Life Years Lost by Childhood Cancer Treatment . . . . .	92
5.3.1.2	Life Years Lost by Iatrogenic Events . . . . .	93
5.3.1.3	Life Years Lost by Multi-Morbidity . . . . .	95
5.3.2	Life Years Lost up to 10 Years after Landmark Time . . . . .	96
5.3.2.1	Life Years Lost by Childhood Cancer Treatment . . . . .	96
5.3.2.2	Life Years Lost by Iatrogenic Events . . . . .	97
5.3.2.3	Life Years Lost by Multi-Morbidity . . . . .	97
5.3.3	Life Years By Cardiac Death . . . . .	98
5.4	Discussion . . . . .	100
5.4.1	Important Contributors to Life Years Lost . . . . .	100
5.4.2	The Advantage of the Life Years Lost Quantity . . . . .	102
5.4.3	Cause Specific Life Years Lost . . . . .	102
5.4.4	Limitations . . . . .	103
5.4.5	Perspectives . . . . .	104
5.5	Summary . . . . .	104

<b>6</b>	<b>Landmark Estimation of Life Years Lost with Time-Dependent Covariates, Time-Varying Effect, and their Interaction</b>	<b>105</b>
6.1	Introduction . . . . .	105
6.2	Simulation Scenarios . . . . .	108



6.2.1	Simulating SMN and CD . . . . .	108
6.2.2	Simulating Death . . . . .	109
6.2.3	Summary of simulation procedure . . . . .	110
6.2.4	Estimated Models . . . . .	111
6.3	Simulation Results . . . . .	111
6.3.1	Estimates of Life Years Lost by SMN . . . . .	111
6.3.2	Estimates of Life Years Lost by Multi-morbidity . . . . .	114
6.4	Expected coefficients . . . . .	117
6.4.1	Landmark time = 0 . . . . .	117
6.4.2	Landmark time > 0 . . . . .	117
6.5	Discussion . . . . .	119
6.5.1	Preliminary Results . . . . .	119
6.5.2	Design of the Simulations . . . . .	119
6.5.3	Work in Progress . . . . .	120
6.6	Summary . . . . .	120
<b>7</b>	<b>General Discussion and Conclusion</b>	<b>121</b>
7.1	Main Findings . . . . .	121
7.2	Strengths and Limitations . . . . .	123
7.3	Clinical Implications . . . . .	124
7.4	Future Research Perspectives . . . . .	125
7.5	Conclusions . . . . .	125
	<b>References</b>	<b>127</b>

<b>Appendices</b>	<b>143</b>
<b>A Anthracyclines Doses Equivalent</b>	<b>143</b>
<b>B Details of FCCSS data</b>	<b>145</b>
B.1 Radiotherapy by Childhood Cancer Type . . . . .	145
B.2 Changes in Patients Characteristics by Year of Childhood Cancer Diagnosis . .	147
<b>C Published Works</b>	<b>149</b>



## List of Tables

2.1	Description of FCCSS patients . . . . .	14
4.1	Covariates combinations used to study SMN effect on CD. . . . .	57
4.2	Description of the population at landmark times on the patient's age scale. . .	60
4.3	Description of the population at landmark times on the time since diagnosis scale. . . . .	62
5.1	Covariates combinations used to study life years lost among childhood cancer survivors . . . . .	90
A.1	Ponderations used to compute cumulative dose of anthracyclines. Doses are in $\text{mg}/\text{m}^2$ . . . . .	143
B.1	Percent of FCCSS patient treated with radiotherapy by type of childhood cancer.	145
B.2	Percent of FCCSS patients treated with high doses of radiotherapy at the brain and heart, by type of childhood cancer. . . . .	145



## List of Figures

1.1	Worldwide proportion of cancer type by age group, 2001–10. (Steliarova-Foucher et al., 2017) . . . . .	2
1.2	Overall survival, by year of cancer diagnosis, among cancer patients who were diagnosed before the age of 20 years. Data from the Surveillance, Epidemiology, and End Results program of the US National Cancer Institute. (Robison et al., 2014) . . . . .	3
1.3	Range of health-related and quality-of-life outcomes among long-term survivors of childhood and adolescent cancers. (Robison et al., 2014) . . . . .	5
2.1	Survival functions of the most common iatrogenic events in the FCCSS. . . . .	14
2.2	Cumulative Incidence of iatrogenic events, stratified by the use of radiotherapy and/or chemotherapy during childhood cancer treatment. Death is included as a competing event. . . . .	19
2.3	Cumulative incidence of death, stratified by the use of radiotherapy and/or chemotherapy during childhood cancer treatment. . . . .	20
2.4	Hazards Functions estimated using kernel . . . . .	22
3.1	Representation of observed data in survival analysis, with $T$ the time of occurrence of event, and $C$ right censoring. . . . .	24
3.2	Representation of interval censored data, with $T_l, T_r$ the time at which patients are observed, $T$ the event and $C$ the right censoring. (Sabathé, 2015) . . . . .	25
3.3	Two states survival model. . . . .	26
3.4	RMST up to $\tau = 10$ . . . . .	27

3.5	Life Years Lost up to 10 years by the binary covariate X is the red area. . . . .	29
3.6	Examples of competing risk models . . . . .	33
3.7	Illustrative example of the Illness-Death model. . . . .	34
3.8	Motivating example for the Phd of a multi-state setting with two absorbing states and one transient state. . . . .	34
4.1	Effect of a SMN on the cumulative incidence of CD. . . . .	64
4.2	Additive effect of each type of SMN on the cumulative incidence of Cardiac Disease. All landmark times are years since childhood cancer diagnosis. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT ( $\text{mg}/\text{m}^2$ ). . . . .	67
4.3	Additive effect of each type of SMN on the cumulative incidence of Cardiac Disease. All landmark times are patients' age in years. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT ( $\text{mg}/\text{m}^2$ ). . . . .	68
4.4	Additive effect of each type of SMN on the cumulative incidence of death. All landmark times are years since childhood cancer diagnosis. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT ( $\text{mg}/\text{m}^2$ ). . . . .	69
4.5	Additive effect of each type of SMN on the cumulative incidence of death. All landmark times are patients' age in years. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT ( $\text{mg}/\text{m}^2$ ). . . . .	70
4.6	Additive effect of Second Malignant Neoplasm based on time since its diagnosis on the cumulative incidence of Cardiac Disease. All landmark times are years since childhood cancer diagnosis. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT ( $\text{mg}/\text{m}^2$ ). . . . .	72

4.7 Additive effect of Second Malignant Neoplasm based on time since its diagnosis on the cumulative incidence of Cardiac Disease. All landmark times are patients' age in years. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>). . . . . 73

4.8 Additive effect of Second Malignant Neoplasm based on time since its diagnosis on the cumulative incidence of death. All landmark times are years since childhood cancer diagnosis. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>). . . . . 74

4.9 Additive effect of Second Malignant Neoplasm based on time since its diagnosis on the cumulative incidence of death. All landmark times are patients' age in years. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>). . . . . 75

4.10 Multiplicative effect of SMN on the cause specific hazard of CD. The error bar is the 95% confidence interval. . . . . 76

4.11 Multiplicative effect of SMN on the cause specific hazard of death. The error bar is the 95% confidence interval. . . . . 76

4.12 Multiplicative effect of each type of Second Malignant Neoplasm on the cause specific hazard of Cardiac Disease. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>). . . . . 78

4.13 Multiplicative effect of each type of Second Malignant Neoplasm on the cause specific hazard of death. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>). . . . . 79

4.14 Multiplicative effect of Second Malignant Neoplasm based on time since its diagnosis on the cause specific hazard of Cardiac Disease. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>). . . . . 81



4.15	Multiplicative effect of Second Malignant Neoplasm based on time since its diagnosis on the cause specific hazard of death. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m <sup>2</sup> ). . . . .	82
5.1	Life Years Lost by all cause death up to 60 years old, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, and age at childhood cancer diagnosis. . . . .	93
5.2	Life Years Lost by all cause death up to 60 years old, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, age at childhood cancer diagnosis, and the studied iatrogenic event. . . . .	94
5.3	All cause death Life Years Lost up to 60 years old by two-ways interaction of subsequent malignant neoplasm and cardiac disease, conditional on survival at each landmark time. Black dotted lines represent the 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, age at childhood cancer diagnosis, subsequent malignant neoplasm, and cardiac disease. . . . .	95
5.4	Life Years Lost within 10 years of landmark time, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, and age at childhood cancer diagnosis. . . . .	96

5.5 Life Years Lost within 10 years of landmark time, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, age at childhood cancer diagnosis, and the studied iatrogenic event. . . . . 98

5.6 All cause death Life Years Lost within 10 years of landmark time old by two-ways interaction of subsequent malignant neoplasm and cardiac disease. Black dotted lines represent the 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, age at childhood cancer diagnosis, subsequent malignant neoplasm, and cardiac disease. . . . . 99

5.7 Life Years Lost by cardiac death up to 60 years old, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, and age at childhood cancer diagnosis. . . . . 100

5.8 Life Years Lost by all cause death up to 60 years old, conditional on survival at each landmark time. Error bars represent 95% confidence interval. Red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, Sex, and age at diagnosis. . . . . 101

6.1 Multi-State model showing the transitions to death after multi-morbidity in orange, and the transitions to death after a single morbidity in blue. . . . . 106

6.2 Density of SMN times, conditional on survival up to 40 years old and diagnosis of SMN before 40 years old. Lines are stratified based on diagnosis of a CD before 40 years old. . . . . 107

6.3 Life Years Lost by SMN, when the true effect of SMN on the survival function is constant. Each column corresponds to a different configuration of covariates adjusted on (either only SMN, both SMN and CD, or SMN and CD and their interaction). Each row corresponds to a combination of the true effect of CD on the survival function (either null or constant) and the landmark time at which estimation is done (20, 30, 40). . . . . 112

6.4 Life Years Lost by SMN, when the true effect of SMN on the survival function is decreasing. Each column corresponds to a different configuration of covariates adjusted on (either only SMN, both SMN and CD, or SMN and CD and their interaction). Each row corresponds to a combination of the true effect of CD on the survival function (either null or constant) and the landmark time at which estimation is done (20, 30, 40). . . . . 113

6.5 Life Years Lost by interaction of SMN and CD, when adjusting for SMN and CD, when the true effect of this interaction on survival is null. Columns correspond to the true effect of CD on survival. Rows correspond to the landmark time at which estimation is done (20, 30, 40, 50). . . . . 115

6.6 Life Years Lost by interaction of SMN and CD, when adjusting for SMN and CD, when the true effect of this interaction on survival is constant. Columns correspond to the true effect of CD on survival. Rows correspond to the landmark time at which estimation is done (20, 30, 40, 50). . . . . 116

B.1 Age at Childhood Cancer Diagnosis among FCCSS patients by Year of Childhood Cancer Diagnosis . . . . . 147

B.2 Use of Radiotherapy and Chemotherapy among FCCSS patients by Year of Childhood Cancer Diagnosis . . . . . 148

## Abbreviations

CCS: Childhood Cancer Survivors

CD: Cardiac Disease

CI: Confidence Interval

CRF: Chronic Renal Failure

csHR: Cause-Specific Hazard Ratio

CT: Chemotherapy

FCCSS: French Childhood Cancer Survivors Study

LYL: Life Years Lost

PcsH: Proportional Cause-Specific Hazard

PH: Proportional Hazard

RT: Radiotherapy

SMN: Second Malignant Neoplasm



# 1 - Introduction to Childhood Cancer Survivors

In this chapter, the epidemiology of childhood cancer is introduced, along with details of the multi-morbidity among childhood cancer survivors, and the research objectives of this thesis.

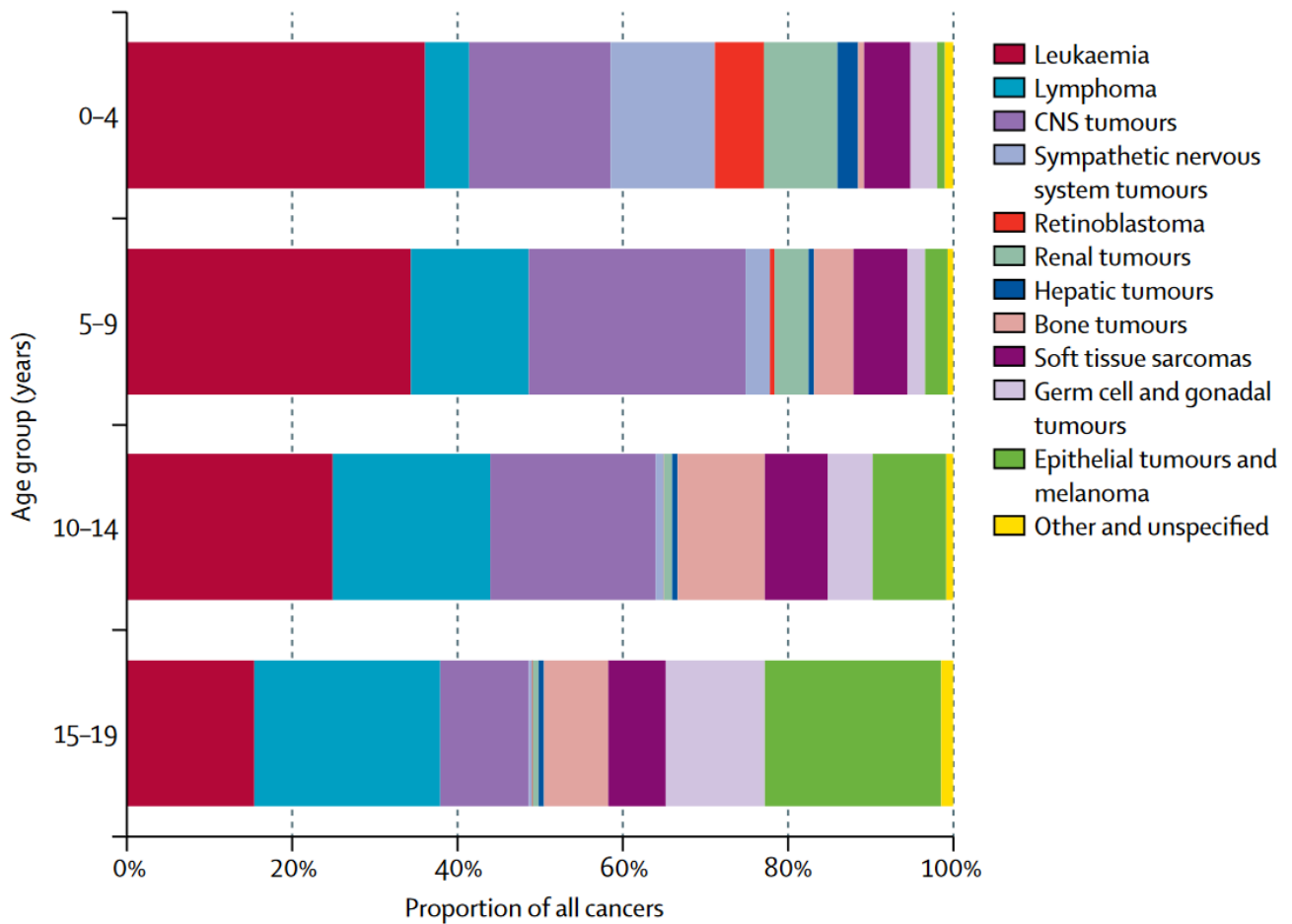
## 1.1 . Childhood Cancer Overview

### 1.1.1 . Childhood Cancer Incidence

Childhood Cancer affected over 1800 children every year in France in 2020. This corresponds to an incidence of 161 per million-years. In total, 0.23% of all children will experience a cancer during childhood in France (“RNCE, Registre National Des Cancers de l’enfant” 2024). Data from the SEER program showed a worldwide age standardized incidence of 178.0 per million children per year, corresponding to 360,114 children diagnosed with cancer worldwide in 2015 (Johnston et al. 2021). Disparities exist as access to health care vary across countries and regions, and not all countries have childhood cancer registries. All of this makes cancer the leading non-accidental cause of death among children. There exists a sex difference, with boys developing slightly more cancer than girls to a ratio of 1.1, although this ratio varies for each cancer type and by region (“RNCE, Registre National Des Cancers de l’enfant” 2024).

Childhood cancers are rare cancers that differ from adult cancers by type and location. The 3rd edition of International Classification of Childhood Cancer (ICCC-3) is the current classification of childhood cancers and differs from adult cancer classification. ICCC-3 divides childhood cancer in 12 major groups, further divided in 47 subgroups and mostly classify using cancer morphology, topology, and development (Steliarova-Foucher et al. 2005). Incidence of each tumor type varies strongly as patients age, with leukemia being 36.1% of all cancer diagnosis among 0-4 years old, but only 15.4% among 15-19 years old. Similarly, lymphomas are 5.3% of

cancer diagnosis among 0-4 years old, and 22.5% among 15-19 years old. Figure 1.1 shows the worldwide proportion of all cancer by age group.

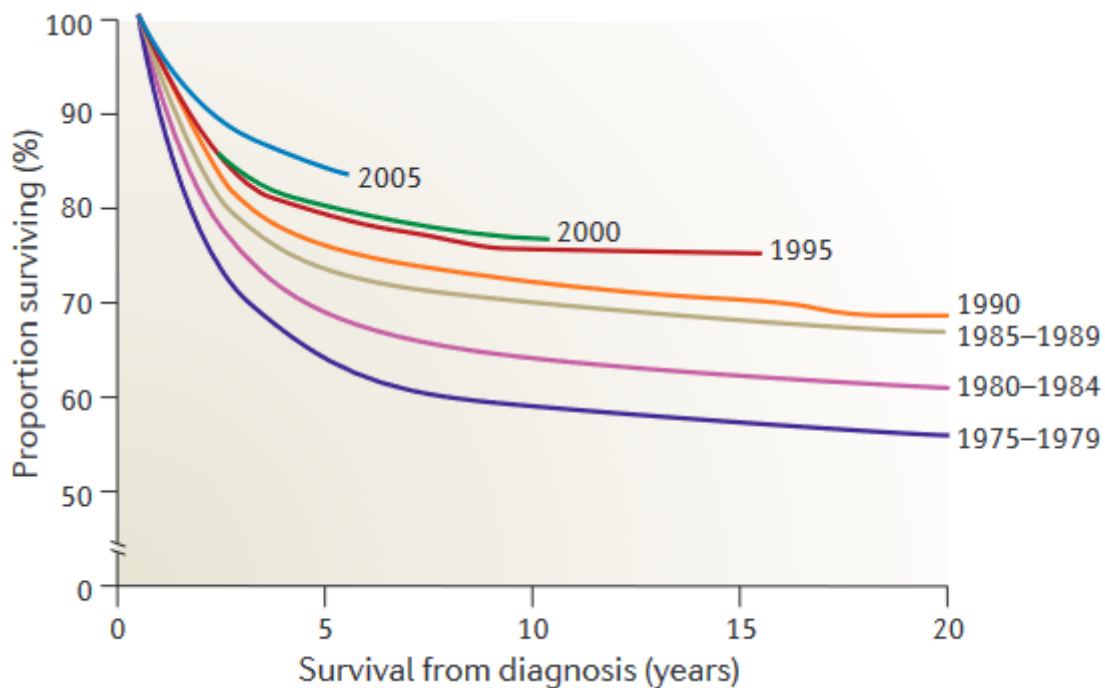


**Figure 1.1:** Worldwide proportion of cancer type by age group, 2001–10. (Steliarova-Foucher et al., 2017)

## 1.2 . Childhood Cancer Mortality

Due to major clinical progress, the European global 5-year survival rate of childhood cancer went up from 44% (95% CI: 44-45) in the 1970s to 64% (95% CI: 64-65) in the 1980s, 74% (95% CI: 73-74) in the 1990s and 81% (95% CI: 81-82) in 2010-2014 (Steliarova-Foucher et al. 2004; Botta et al. 2022), with most 5-year childhood cancer survivors becoming long term survivors (Figure 1.2). The long term survival rate increased in pair with the 5-year survival

rate. Important disparities remain based on sex, cancer type, subtype, histology, location, age at diagnosis, world location, and socio-economic status (Ssenyonga et al. 2022). Indeed, Central Nervous System Cancers has one of the worst 5-year survival rate in Europe, with 58.2% (95% CI: 52.9-63.0) in 2005-2007, compared to Hodgkin’s lymphoma which has one of the highest 5-year survival rate in Europe, of 95.7% (95% CI: 89.5-98.1) in 2005-2007 (Gatta et al. 2014).



**Figure 1.2:** Overall survival, by year of cancer diagnosis, among cancer patients who were diagnosed before the age of 20 years. Data from the Surveillance, Epidemiology, and End Results program of the US National Cancer Institute. (Robison et al., 2014)

### 1.3 . Childhood Cancer Treatment

Treatments of childhood cancer has evolved a lot since the 1940s. In the 40s and 50s, radiotherapy was widely used. Chemotherapy started being heavily used in the 60s. Radiotherapy and chemotherapy became increasingly used together in the 80s. Since the 2000s, immunotherapy is being increasingly used to treat childhood cancers. Doses of each treatments have been re-



duced over time, as better equipment was made available (for radiotherapy), knowledge about sequels of high doses increased, and trials investigating minimum dose required to treat were done (Gibson et al. 2018).

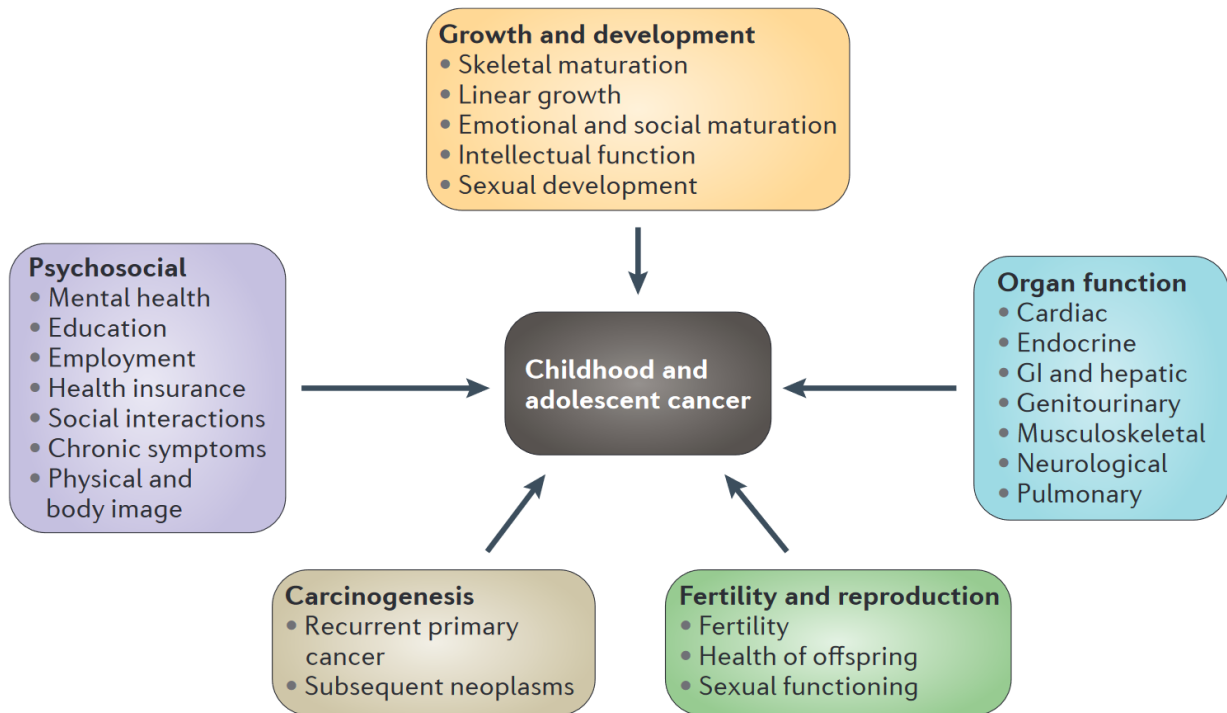
Those treatments and their intensities are not only strongly associated with eras of treatments, but also with type of cancers (Hudson et al. 2021). For instance, radiotherapy is commonly used to treat CNS tumors, often with relatively high doses (Appendix B).

#### **1.4 . Childhood Cancer Late Effects**

As a consequence of increased childhood cancer survival, the childhood cancer survivors population is growing and aging. This shone a light on long term consequences of childhood cancer treatment, with a very wide scope of treatment related late effects which can affect almost any organ or body system to varying degrees (Figure 1.3). Those treatment related late effects are also called iatrogenic events. Subsequent Malignant Neoplasms, Cardiac Diseases, fertility issues, mental health, social inclusion, and body growth have all been found to be affected by childhood cancer treatments. The probability of chronic health condition of any grade from diagnosis to 25 years after was found to be 66.8%, and 73.4% 30 years after diagnosis (Oeffinger et al. 2006). This probability is the cumulative incidence, which we properly define later in Section 3.3. The cumulative incidence for chronic health condition of grade 3 or higher was found to be 33.1% 25 years after diagnosis, and 42.4% 30 years after diagnosis (Oeffinger et al. 2006). Childhood cancer iatrogenic events are known to affect life expectancy of patients, with survivors developing haematological conditions before 32 years losing 19.93 (95% CI: 15.33 - 27.34) life of years (Chang et al. 2022).

In the USA and Europe many cohorts were built to study those iatrogenic effects, with a shared purpose of understanding them and guiding the long started and still undergoing harm reduction process of childhood cancer treatments. Results from those cohorts have improved

the understanding of each treatment consequences, notably regarding radiotherapy doses and drugs of chemotherapy. Those results are also used to update guidelines of care for childhood cancer survivors (Michel et al. 2019; Hudson et al. 2021).



**Figure 1.3:** Range of health-related and quality-of-life outcomes among long-term survivors of childhood and adolescent cancers. (Robison et al., 2014)

## 1.5 . Multi-Morbidity after Childhood Cancer

In this subsection we will discuss multi-morbidity, what it is, how it occurs in childhood cancer survivors and the FCCSS, and plausible explanations for it. A description of multi-morbidity in the cohort used is in Section 2.9

### **1.5.1 . Definitions**

Multi-morbidity is a topic of great importance for patients' health and quality of life. It is often studied in the geriatric population, psychiatric patients, and in Health Related Quality of Life studies after cancer (Yu et al. 2024; Davies et al. 2022). In those contexts, what counts as a morbidity covers a lot of different events types and severity.

In this manuscript, we will adopt a stricter definition of what constitutes a morbidity by restricting it to severe iatrogenic events. We do this to restrict our study to the consequences of what differentiates childhood cancer survivors from the general population, and to only include clinically validated events as lower severity events are more prone to a detection bias.

To summarize, multi-morbidity here designs the occurrences of two or more severe iatrogenic events for a patient. Investigations done in this thesis are restricted to Second Malignant Neoplasms (SMN), Cardiac Diseases (CD), Diabetes, and Chronic Renal Failures (CRF).

### **1.5.2 . Multiple Events in Childhood Cancer Survivors**

Childhood cancer survivors are at a higher risk of experiencing multiple chronic health conditions than their siblings. Adjusting for sex, age, socio-economic factors, and race Oeffinger et al. (2006) showed a staggering difference, with 37.6% of childhood cancer survivors experiencing 2 chronic health conditions compared with only 13.1% of the controls doing so. Similarly, they found that a difference in the risk of experiencing three or more chronic health conditions, as 23.8% of childhood cancer survivors do compared to 5.4% of the controls. Those results show that childhood cancer survivors are frailer than their siblings. It falls short of providing an explanation for it, because the authors didn't explore whether those results are solely driven by increased risk of each chronic health condition by childhood cancer treatment. Others found that by age 50 childhood cancer survivors experienced an average of 4.7 (95% CI: 4.6-4.9) chronic health conditions graded as severe or higher, and 17.1 (95% CI: 16.2-18.0) chronic health conditions of any grade (respectively 2.3 (95% CI: 1.9-2.7) and 9.2 (95% CI:

7.9-10.6) for community controls) (Bhakta et al. 2017).

The Mean Cumulative Count was previously used to show that childhood cancer survivors treated with radiotherapy have a higher average number of subsequent neoplasm than childhood cancer survivors treated without radiotherapy (Dong et al. 2015). Those results mostly show that childhood cancer survivors are at high risk of multiple subsequent neoplasm, which haven't been studied.

The number of hospitalizations was also found to be higher among childhood cancer survivors compared to general population controls, and the use of radiotherapy was found to be associated with a higher number of hospitalizations (Bejarano-Quisoboni et al. 2022). Those results can't be linked with certainty to multi-morbidity, because patients can be hospitalized numerous times for the same chronic health condition.

### **1.5.3 . Plausible Pathways**

Here, we discuss mechanisms that could explain how multi-morbidity works among CCS.

First and foremost, it is of note that many iatrogenic events have shared risk factors. For instance, radiotherapy at the heart greatly increase the risk of both breast cancer and cardiac disease (Chounta, Lemler, et al. 2023; Jaworski et al. 2013). This is an example of childhood cancer specific shared risk factor. Others shared risk factors unrelated to childhood cancer exist. For instance, smoking and air pollution are associated with both lung cancer and cardiac diseases (Groot and Munden 2012; Ambrose and Barua 2004). We will refer to such as generic shared risk factors.

Secondly, patients suffering a first morbidity are on average having a higher risk of morbidity than those morbidity-free, because of unobserved risk factors. In this thesis, this includes environmental risk factors, genetics risk factors, and patient-specific frailty. We will refer to such as unobserved shared risk factors.

Lastly, comes the late effect of the morbidity, which can be further separated in two. Firstly,

the frailty induced by the morbidity itself. For instance, surviving a heart failure often comes with a massive drawback regarding patient's overall health, weakening their resistance to other morbidities (Wohlfahrt et al. 2023). We will refer to such as morbidity induced risk factors.

Secondly, the treatment induced frailty. The same way childhood cancer treatments have iatrogenic effects, the treatment of, say, subsequent neoplasms have their own iatrogenic effects which further increase the risk of patients (Boudoulas et al. 2022) . We will refer to such as morbidity's treatments risk factors.

We deem all those mechanisms plausible. It is of note that some can be distinguished with appropriate data (e.g. morbidity risk factors and morbidity's treatments risk factors) but others can't, such as morbidity induced risk factors and patient-specific frailty without making strong assumption on the form of both of those (Heckman 2001; Feller 1991; Heckman and Borjas 1980; Kessing, Olsen, and Andersen 1999; Pénichoux 2012)

We presented a simple version of each mechanisms, assuming each exist independently. Not only is it possible that those mechanisms co-exist, but they might also interact and feed into each other.

Another important aspect is the shift in consequences of morbidity as patients age. Indeed, younger patients tend to recover better and suffer less physical consequences than older patients in the short term (Peters, Mendoza Schulz, and Reuss-Borst 2016). This can be a mediation factor lowering the consequences of morbidity among CCS, due to their early onset – which is still bad overall.

#### **1.5.4 . Death, a Competing Event to Multi-Morbidity**

We want to highlight that all previously discussed mechanisms of multi-morbidity also apply to risk of death.

For instance, Moskowitz et al. (2019) investigated increased risk of death by breast cancer among CCS compared to de novo breast cancer patients. They did not find a marked difference

of breast cancer mortality between CCS and controls with de novo breast cancer (csHR: 1.3, 95% CI: 0.9-2.0). However, they found that CCS are at higher risk of death by any cause, with a csHR of 2.4 (95% CI: 1.7-3.2) after adjusting for breast cancer treatment. This increase is further marked when looking at death of non-breast cancer health-related causes, as the csHR is then 5.5 (95% CI: 3.4-9.0). This highlights a mortality problem among CCS, and the needs to look at the treatments received during childhood to assess the mortality risk after a iatrogenic event – as the mortality after morbidity is higher than among non CCS.

Accounting for competing risks in survival analysis is necessary to transpose results on the cumulative incidence scale to real life population (T. Therneau, Crowson, and Atkinson 2024). Indeed, ignoring those competing risks inflates the censored population and wrongfully increases the estimated cumulative incidence.

This is why CCS are at both increased risk of further morbidity after a first morbidity, by the mechanisms explained in Section 1.5.3, and at lowered risk of further morbidity after a first morbidity as their mortality risk increases. Guidelines for analyzing and reporting competing risks recommend to report results on competing risks, even those not of interest in the study, to provide a complete view of mechanisms at play in the presence of competing risks (Aurelien Latouche et al. 2013).

## **1.6 . Thesis Objectives**

Now that the context is set, we will discuss the objectives and works of the thesis. The global goal was to explain the sequence of iatrogenic events among childhood cancer survivors along with their mechanisms. In order to do so, we used the cohort FCCSS (see Section 2), and:

1. Studied the impact of a Second Malignant Neoplasm diagnosis on the risk of cardiac disease among childhood cancer survivors, while adjusting for childhood cancer treatments (Section 4).

2. Studied the impact of Second Malignant Neoplasm, cardiac disease, chronic renal failure, diabetes, and co-occurrence of those iatrogenic events on the life expectancy of childhood cancer survivors across patients' life (Section 5).

We also performed a simulation study to:

3. Investigate the estimation bias of a multi-morbidity effect via the landmark method, when death by each morbidity decreases with time since their diagnosis (Section 6).

In total, this thesis contains 7 chapters. Section 1 introduces the topic of childhood cancer survivors and the goal of the thesis. Section 2 describe the data source used in the thesis, namely the cohort FCCSS. Section 3 provides a general discussion on the topic of time-dependent covariates in survival analysis, and a detailed discussion on the landmark method used across this thesis. Section 4 describes a study we did on the FCCSS investigating the impact of Second Malignant Neoplasm on the cumulative incidence and instantaneous risk of cardiac disease. Section 5 describes a study we did on the FCCSS, investigation the Life Years Lost by Second Malignant Neoplasm, Cardiac Disease, Chronical Renal Failure, and diabetes among childhood cancer survivors. Section 6 describes a simulation study we did investigating the validity of the landmark method to study time-dependent covariates with time-varying coefficients on life years lost. Section 7 discusses the main results and takeaways of the thesis, along with a generic discussion about the strengths and limitations of this work. It concludes this manuscript with recommendations for future works and a general conclusion.

## **2 - Thesis Data: The French Childhood Cancer Survivors Study**

In this chapter the data used across the manuscript is introduced and the characteristics of cohort's patients are detailed.

### **2.1 . History of the French Childhood Cancer Survivors Study**

At first, the French Childhood Cancer Survivors Study (FCCSS) was the french part of the retrospective cohort EURO2K which included patients treated from 1945 to 1986 in five french centers: Institut Gustave Roussy, Institut Curie, Institut Jean Godinot, Centre Claudius Régaud, and Centre Antoine Lacassagne. It was later extended to include patients diagnosed from 1986 to 2000 in two french centers: Institut Gustave Roussy and Institut Curie. Nowadays, the FCCSS is a retrospective cohort with a prospective follow-up including 7,670 5-year childhood cancer survivors treated for a solid cancer or lymphoma before 2000 in one the five centers mentioned previously. Of those 7,670, one was later determined to have been diagnosed with a leukemia and was therefore removed from the analysis conducted in this thesis, leading to a cohort of 7,669. The median follow-up is of 30 years [min: 5, max: 74]

### **2.2 . Previous use**

This cohort has previously been used for three main things. Firstly, to study the long term effects of childhood cancer treatments. This includes looking at the consequences of radiotherapy and chemotherapy use for each iatrogenic events, as well as investigating the doses of radiotherapy or cytotoxic chemotherapy inducing harm on the patients (Haddy et al. 2016; Chounta, Lemler, et al. 2023; Chounta, Allodji, et al. 2023; Allodji et al. 2015, 2019). Secondly, to study



differences with the general population on a wide range of criteria, such as cost of healthcare and hospitalizations (Schwartz et al. 2024; Agnès Dumas, De Vathaire, and Vassal 2016; A. Dumas et al. 2015; Bejarano-Quisoboni et al. 2022, 2023). Thirdly to investigate the genetics and genomics characteristics influencing radiotherapy and chemotherapy consequences in order to provide better personalized health care (Aba et al. 2023; Ducos et al. 2023).

### **2.3 . Legal Authorization**

The FCCSS received approval from a National Committee on Ethics and the French National Agency Regulating Data Protection (agreement nos. 902287 and 12038829). Written informed consent was obtained from all patients, parents, or guardians in accordance with national research ethics requirements.

### **2.4 . Childhood Cancer Treatment**

Chemotherapy, surgery, and radiotherapy information for childhood cancer treatment were abstracted from medical records. Chemotherapy exposure was defined as receipt (or not) of any chemotherapy agent, anthracyclines, alkylating , or platinum agents. For each chemotherapy class, cumulative doses were computed in  $\text{mg}/\text{m}^2$  by summing chemotherapy doses across cycles. Anthracycline cumulative dose was computed using doxorubicin equivalents in order to account for the disparity in cardio-toxicity by  $\text{mg}/\text{m}^2$  of each drug (Feijen et al. 2019) (see Appendix A for details of the computation).

For external beam radiotherapy radiation, dose distributions to the heart were retrospectively reconstructed on patient specific voxel phantoms, considering individual patient treatment information. This information included treatment machine, type of radiation, beam energy, irradiation technique, field size and shape, gantry and collimator angles, use of accessories, target volume location, and total delivered dose. This retrospective reconstruction was necessary,

because computed tomography scans were not used on many patients. More details on the methodology and dosimetry software package was published previously (Veres et al. 2014). The use of voxel phantoms and careful delineation of organs allowed the computation of mean radiation dose at the heart, brain, thyroid, lungs, and other organs.

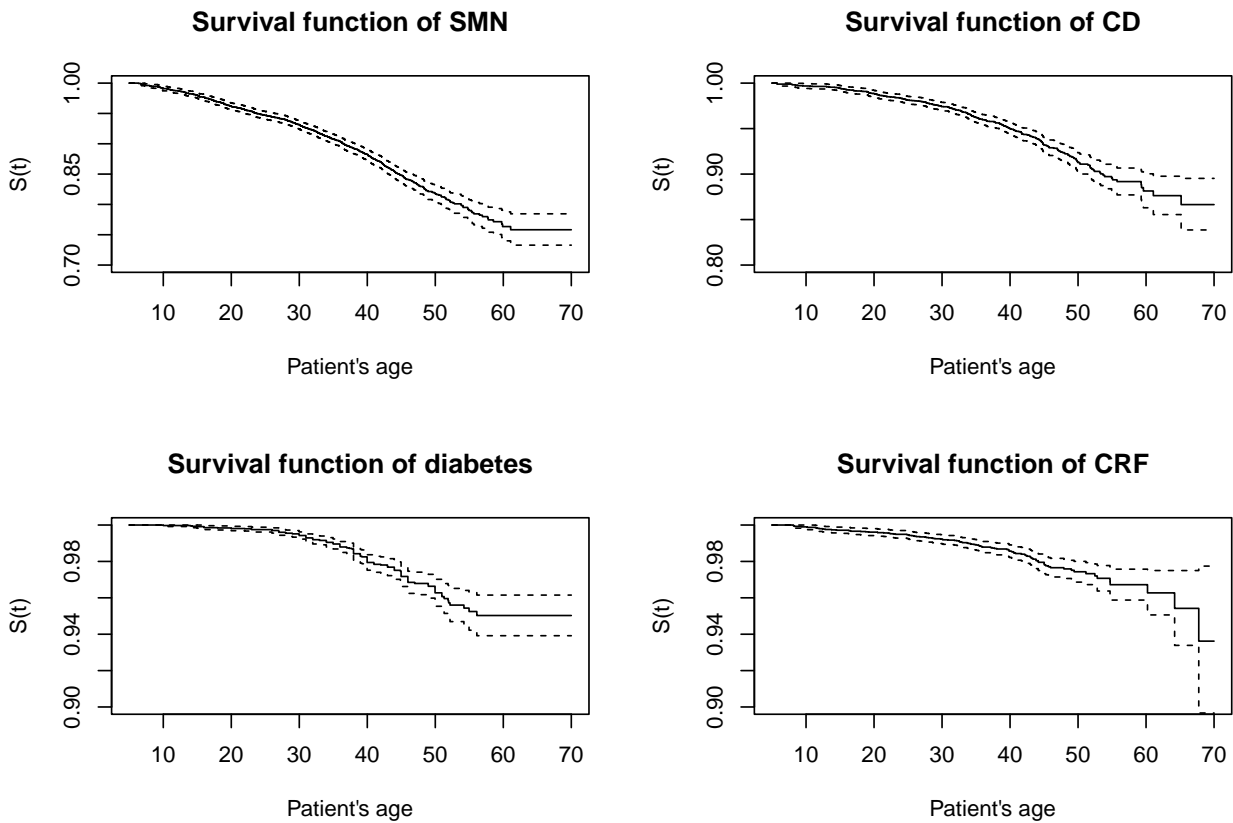
## **2.5 . Iatrogenics Events Identification and Validation**

Clinical and epidemiological follow-up, including self-administered questionnaires and cohort linkage with the French Hospital Database and Health Insurance Information System, was performed to identify the occurrence of iatrogenic effects. For patients treated at Gustave Roussy, clinic long-term follow-up was also performed. SMNs and non-cancerous diseases were identified through these different sources and subsequently validated by a trained and experienced clinical research associate. Validation was based on medical, pathology, or radiological reports obtained from the treating centers or from referring doctors, regardless of the data source used for first identification. Late events were graded according to the Common Terminology Criteria for Adverse Events (version 4.03). In this thesis, late events of interest were Second Malignant Neoplasms, Cardiac Diseases, diabetes, chronic renal failures, and death. Unless specified otherwise, future reference to cardiac disease events refer to severe cardiac disease events (grade  $\geq 3$ ), based on the consideration that non-severe cardiac disease events may be self-reported and could cause reporting bias in the data.

## **2.6 . Vital Status and Iatrogenics Events**

Vital status of all patients and causes of death for deceased patients were obtained from CépiDC. Causes of death were available for 451 out of the 1496 deceased patients. Using the grade of validated cardiac diseases, we were able to identify 51 deaths by cardiac disease. The four iatrogenic events used during my thesis are Second Malignant Neoplasm, Cardiac Diseases,

Diabetes, and Chronic Renal Failures. FCCSS patients experienced 828, 379, 139, 112 of each (10.8%, 4.9%, 1.8%, 1.5%). Survival functions for each iatrogenic events (Figure 2.1) show a high marginal risk of SMN and CD, with events starting before adulthood for SMN, and earlier than observed in the general population for CD.



**Figure 2.1:** Survival functions of the most common iatrogenic events in the FCCSS.

## 2.7 . Patients' characteristics

**Table 2.1:** Description of FCCSS patients

Covariate	Level	n
-----------	-------	---

Sex	Men	4200
	Women	3469
Age at CC diagnosis	<1 year old	1244
	1-4 years old	2482
	5-9 years old	1677
	10-14 years old	1623
	15+ years old	643
Treatment combination	Radiotherapy and chemotherapy	3105
	Radiotherapy alone	1088
	Chemotherapy alone	2574
	Nor radiotherapy nor chemotherapy	902
Anthracycline dose	0 mg/m <sup>2</sup>	506
	1-99 mg/m <sup>2</sup>	20
	100-249 mg/m <sup>2</sup>	125
	250+ mg/m <sup>2</sup>	114
Radiotherapy dose at the heart	None	3622
	0-4 Gy	2349
	5-34 Gy	1215
	35+ Gy	92
	Unknown	391
	02 - Lymphomas	1278

	03 - CNS tumor	1140
	04 - Peripheral nervous tumors	1034
	05 - Retinoblastomas	619
	06 - Renal tumors	1140
	07 - Hepatic tumors	79
	08 - Bone sarcomas	686
Type of childhood cancer	09 - Soft-tissue sarcomas	859
	10 - Germ cells and gonadal tumors	469
	11 - Other carcinomas	344
	12 - Other or unspecified tumors	11
	Unknown	10
	Yes	380
Cardiac Disease (grade $\geq 3$ )	No	7289
	Yes	828
Second Malignant Neoplasm	No	6841
	Yes	139
Diabetes	No	7530
	Yes	112
Chronic Renal Failure	No	7557

### **2.7.1 . Overall Description**

Table 2.1 details patients' characteristics. There is 3469 women and 4200 men. The age at childhood cancer diagnosis ranges from 0 to 20 years old, with 1 patient diagnosed before birth, 1243 diagnosed from birth to 1 year old, and 643 diagnosed at 15 years old or higher. The most common treatment combination was radiotherapy and chemotherapy (n=3105), followed by chemotherapy without radiotherapy (n=2574), and others split evenly between radiotherapy without chemotherapy (n=1088) and neither radiotherapy nor chemotherapy (n=902). Out of the 5679 treated with chemotherapy, 1144 of them were treated with cumulative anthracycline dose higher than 250 mg/m<sup>2</sup>, and 1272 with cumulative anthracycline dose ranging from 100 to 250 mg/m<sup>2</sup>. Out of the 4193 treated with radiotherapy, 391 had missing dose data after dose reconstruction because the information required for an estimation was missing. Those patients were included in all analysis, and multiple imputation was used to impute their doses. Out of the 3802 treated with radiotherapy for whom radiotherapy doses were available, 1307, 1354, 1266, and 1311 had a mean dose higher than 5 Gy at the heart, brain, left lung, and right lung respectively.

### **2.7.2 . Changes Across Years of Childhood Cancer Diagnosis**

The FCCSS follows patients over a wide time period during which childhood cancer treatments and survivorship changed a lot. We provide detailed figures of those changes in Appendix B.2, and highlights some important ones here. The age at childhood cancer is stable across year of childhood cancer diagnosis, although older patients were more frequent after 1983 (Figure B.1). The use of radiotherapy/chemotherapy differs across eras of treatment, with radiotherapy being always used before the arrival of chemotherapy in 1955. In the 1960s and 1970 radiotherapy chemotherapy were the most frequent combination, used ~60% of the time. In the 1980s, we see the rise of chemotherapy with chemotherapy, and in turn a lower use of chemotherapy with radiotherapy. (Figure B.2).

The mean radiotherapy dose at the heart, brain, left lung, and right lung are similar across eras of treatment. Use of chemotherapy comes with smaller doses at the brain, and higher doses at other organs for all decades of treatment. The cumulative dose of anthracyclines is similar when used with or without radiotherapy, but only started being used in the 1970s, and the use of doses higher than  $250 \text{ mg/m}^2$  was lower in the 1990s.

## **2.8 . Morbidities and Mortality**

The FCCSS recorded a total of 1496 deaths, 828 SMN, 379 CDs, 139 diabetes, and 112 CRF among 1276 unique patients.

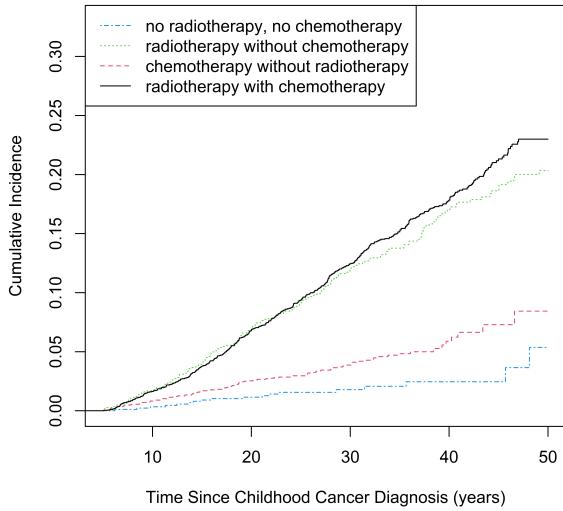
The cumulative incidence of death reaches 14.5% 30 years after diagnosis, and 36% 50 years after diagnosis. Patients treated with radiotherapy have a higher absolute risk of death at all times, regardless of the use of chemotherapy which shows little no effect (Figure 2.3).

Subsequent Malignant Neoplasm is the most frequent and severe iatrogenic event of this thesis, with a cumulative incidence of 9.1% 30 years after diagnosis, and 20.2% 50 years after diagnosis. Patients treated with radiotherapy developed more SMN than those treated without radiotherapy, regardless of chemotherapy use (Figure 2.2a).

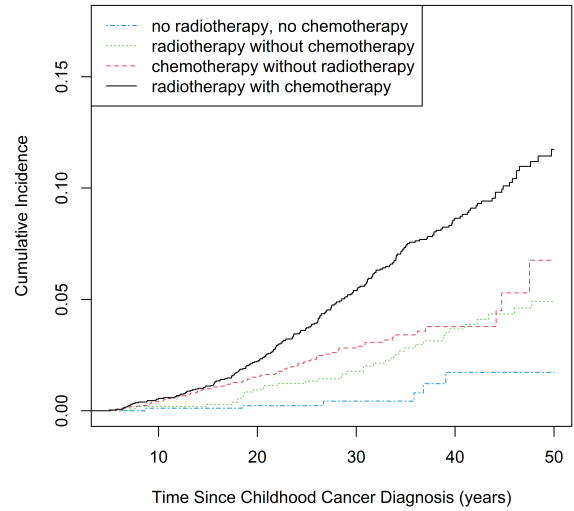
Cardiac diseases are also a frequent and severe iatrogenic event, with a cumulative incidence of of 3.8% 30 years after diagnosis, and 9.9% 50 years after diagnosis. Patients treated with both radiotherapy and chemotherapy experienced more CD than those treated with either one, who in turn experienced more CD than those treated with neither (Figure 2.2b).

The cumulative incidence of diabetes is of 1.2% 30 years after diagnosis, and 5.1% 50 years after diagnosis. Patients treated with radiotherapy developed more diabetes than those treated without radiotherapy (Figure 2.2c), regardless of chemotherapy status.

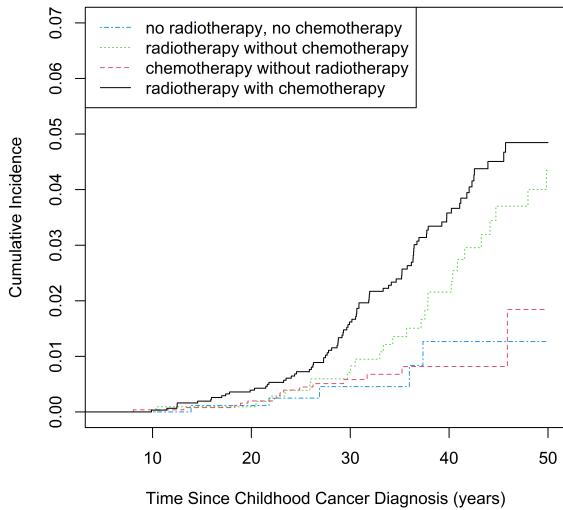
The cumulative incidence of CRF is of 1% 30 years after diagnosis, and 4% 50 years after diagnosis. Patients treated with both radiotherapy and chemotherapy appear to have experienced



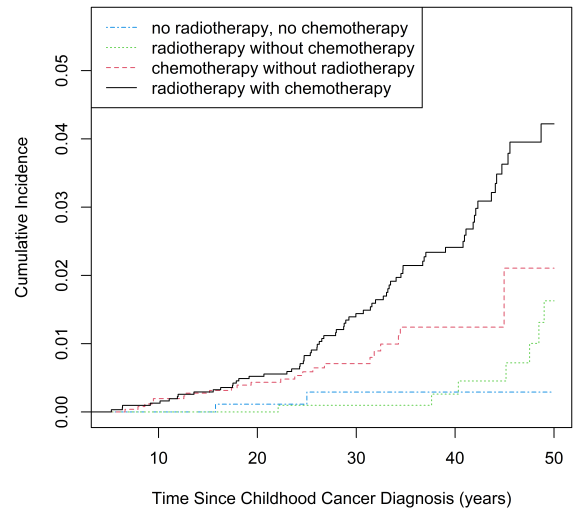
(a) Cumulative Incidence of SMN stratified on childhood cancer treatment.



(b) Cumulative Incidence of CD stratified on childhood cancer treatment.



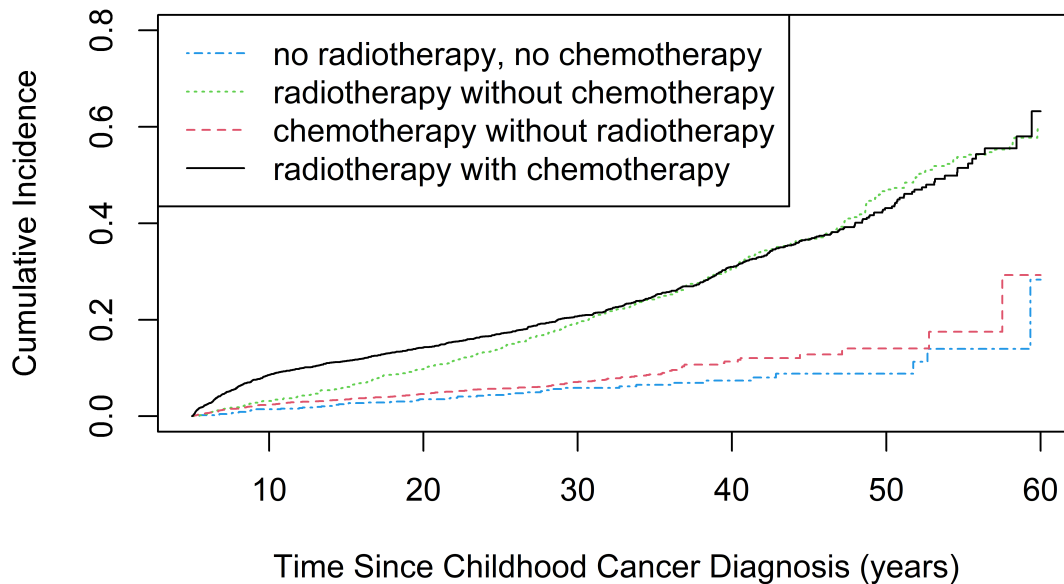
(c) Cumulative Incidence of diabetes stratified on childhood cancer treatment.



(d) Cumulative Incidence of CRF stratified on childhood cancer treatment.

**Figure 2.2:** Cumulative Incidence of iatrogenic events, stratified by the use of radiotherapy and/or chemotherapy during childhood cancer treatment. Death is included as a competing event.





**Figure 2.3:** Cumulative incidence of death, stratified by the use of radiotherapy and/or chemotherapy during childhood cancer treatment.

CRF earlier than those treated with only one or neither, but the cumulative incidence of CRF is low in all groups (<5% at 50 years after diagnosis) (Figure 2.2d).

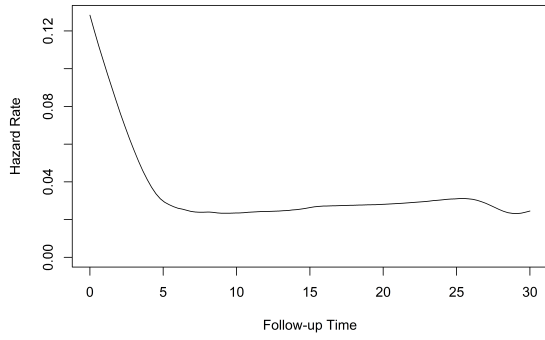
## 2.9 . Multi-Morbidity

### 2.9.1 . Multiple Events

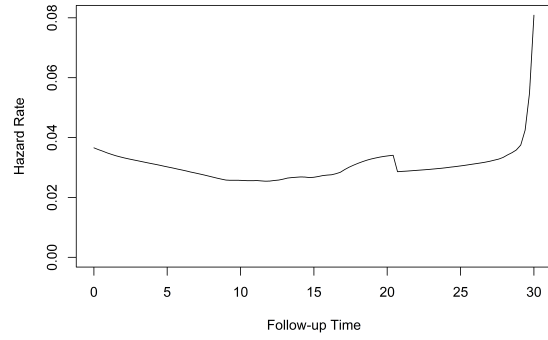
Out of the 828 patients who experienced a SMN, 203 experienced two or more subsequent malignant neoplasm. Out of the 379 patients who experienced a CD, 84 experienced two or more CD. A total of 83 patients experienced both a SMN and a CD.

### **2.9.2 . Hazard Functions**

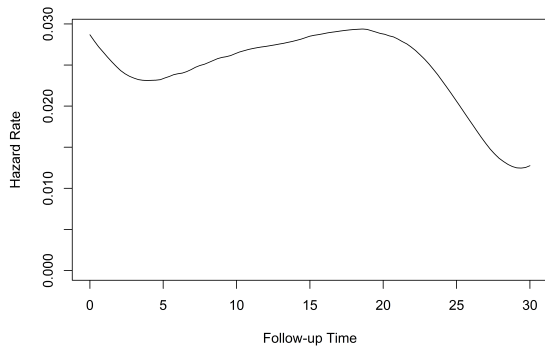
We looked at the smoothed hazard functions from CD to death/SMN/second CD and from SMN to death/third malignant neoplasm/CD (Figure 2.4) (Kenneth Hess and R. Gentleman 2021). We see that the instantaneous risk of death in the 5 years following the diagnosis of a SMN is very high and then reaches a plateau. The risk of third malignant neoplasm after a SMN is stable at all times, whereas the risk of first CD after SMN starts high, reaches a low point 5 years after SMN diagnosis, and then starts rising again 10 years after SMN. We see that the instantaneous risk of death remains constant after a CD. The risk of SMN after a first CD slowly lowers across time. The risk of second CD after first CD starts high and reaches a low point 5 years after CD diagnosis and remains constant.



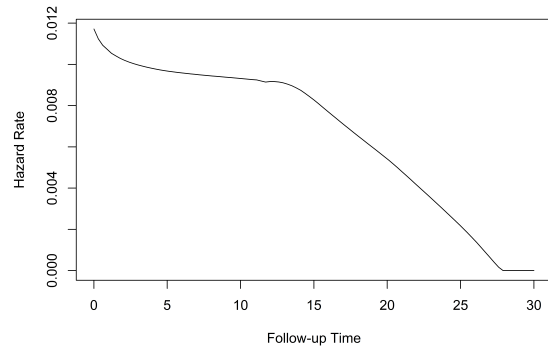
(a) Death after SMN



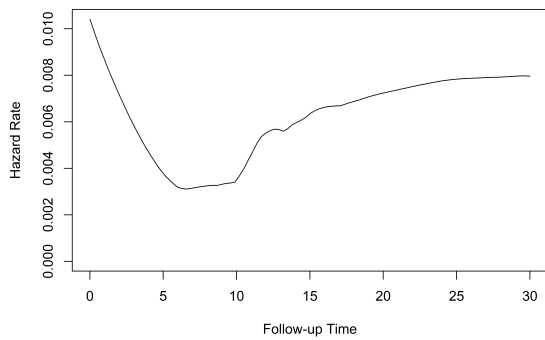
(b) Death after CD



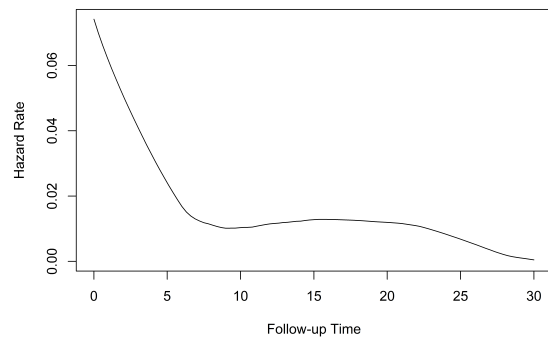
(c) Third malignant neoplasm after SMN



(d) SMN after CD



(e) CD after SMN



(f) Second CD after first CD

**Figure 2.4:** Hazards Functions estimated using kernel

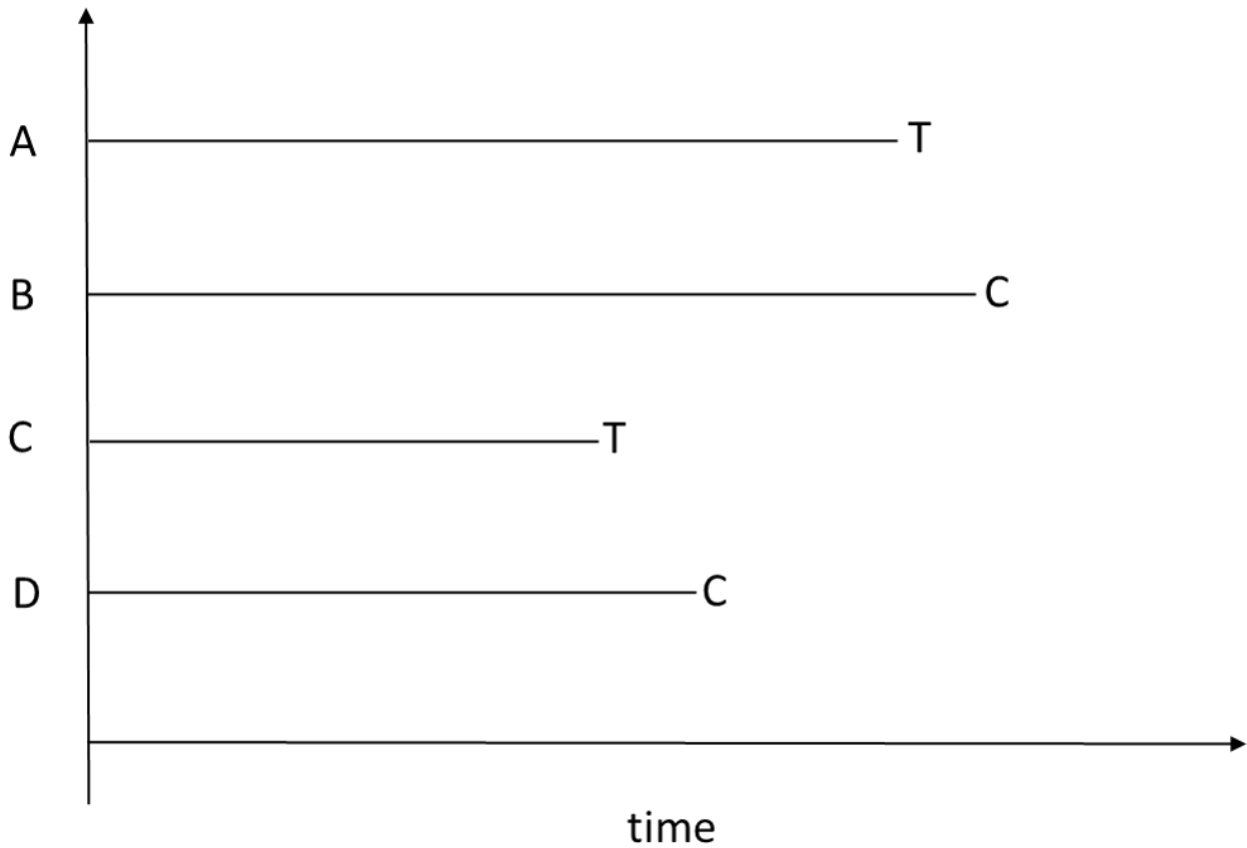
### **3 - Survival analysis and extensions to account for multiple type of events and time-dependent exposure**

The goals of this thesis requires to employ time-to-event history data (eg. the various adverse events encountered by the patients) and time-dependent exposure (the occurrence of a secondary event). This section details the formalism for handling such type of observations and exposures.

#### **3.1 . Event History Data**

Event history data is acquired by observing individuals over time, monitoring events times and events types. The specificity of such data, is that events times can be unobserved for some patients, due to mechanisms known as censoring and truncation.

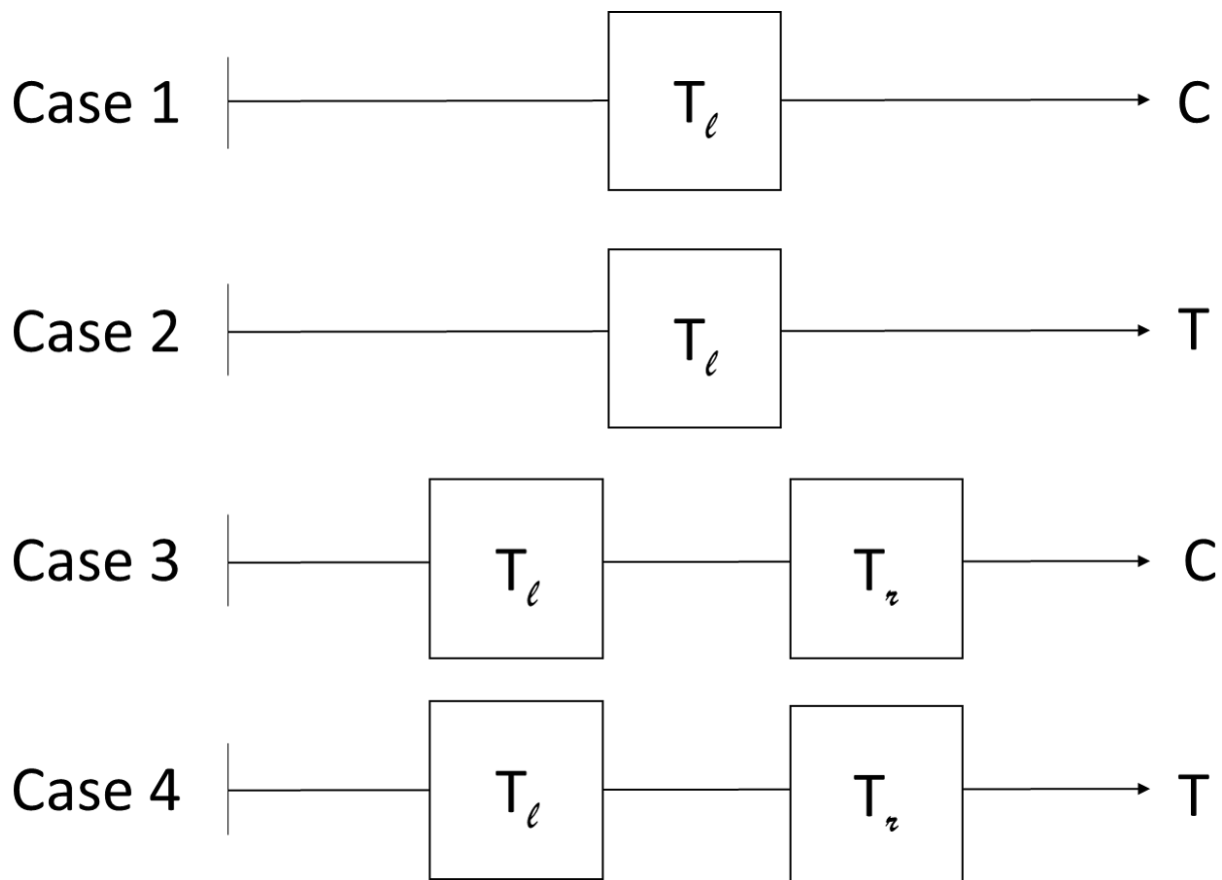
Censoring is classically classified into three categories: right censoring, left censoring, and interval censoring. In this thesis, only right censoring is of interest, which occurs when a subject's event is not observed because last follow-up time occurs before, as for patients B and D in Figure 3.1. This typically occurs when a study follow-up ends before all patients experience the event, or because a patient was lost to follow-up. Left censoring occurs when a subject's time-to-event is not observed but known to be less than a certain time. This typically occurs when data is recovered from administrative sources from the past, with missing information from early times. Interval censoring occurs when a subject's event time is known to be within an interval of time but exact time unknown. This typically occurs when an event needs to be diagnosed during a clinical visit, without means to know the exact start time of the condition (Figure 3.2) (Sabathé 2019). In this thesis, censoring time is assumed to be independent from time-to-event conditionally on covariates, and the case of dependent censoring is not discussed.



**Figure 3.1:** Representation of observed data in survival analysis, with  $T$  the time of occurrence of event, and  $C$  right censoring.

Truncation is another mechanism of unobservation. It occurs when only individuals whose time-to-event lies within a certain observational window  $(T_L, T_R)$  are observed (Klein and Moeschberger 2003). Individuals whose event time isn't in this interval aren't observed, and no information on them is available. This is in contrast to censoring where there is at least partial information on each individual. Because we are only aware of individuals with event time in the observation window, the inference is restricted to conditional estimation. When  $T_R$  is infinite, this is left truncation. Left truncation is the most common case, and also the only one discussed further.

In order to be enrolled in the FCCSS, patients need to be alive 5 years after their childhood

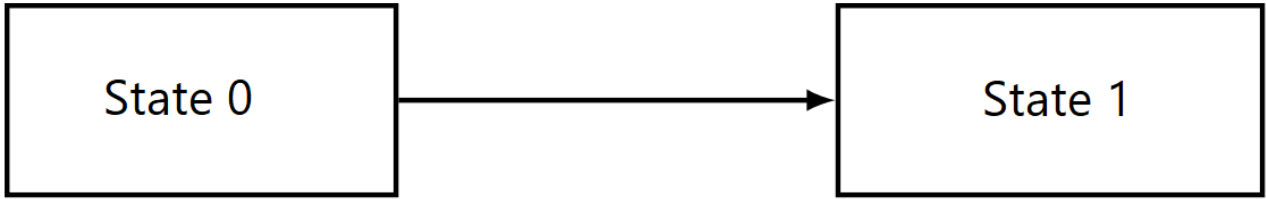


**Figure 3.2:** Representation of interval censored data, with  $T_l, T_r$  the time at which patients are observed,  $T$  the event and  $C$  the right censoring. (Sabathé, 2015)

cancer diagnosis, so on the age time scale patients are left truncated up to this time. FCCSS patients are right censored when lost to follow-up or last data update occurred before the event.

### 3.2 . Survival Analysis

The simplest case of event history data is survival data, when there is a single event occurring. All subjects start in State 0, and transition to State 1 upon occurrence of the event of interest (Figure 3.3). The observed failure time is  $\tilde{T} = \min(T, C)$ , with  $T$  the true event time and  $C$  the right censoring time. The indicator of event is denoted  $\delta$ , and defined as  $\delta = \mathbb{1}\{T \leq C\}$ , which equals 1 if the event of interest occurred, and 0 otherwise. Let  $T_L$  be the left truncation time.



**Figure 3.3:** Two states survival model.

The cumulative distribution function is defined as:

$$t \geq 0, \quad F(t) = \mathbb{P}(T \leq t) \quad (3.1)$$

Its complementary, the survival function, is:

$$S(t) = \mathbb{P}(T > t) = 1 - F(t) \quad (3.2)$$

The hazard function is defined as the instantaneous probability of event at time  $t$  given that the subject is still at risk at time  $t$ :

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{\mathbb{P}(t \leq T < t + dt | T \geq t)}{dt} \quad (3.3)$$

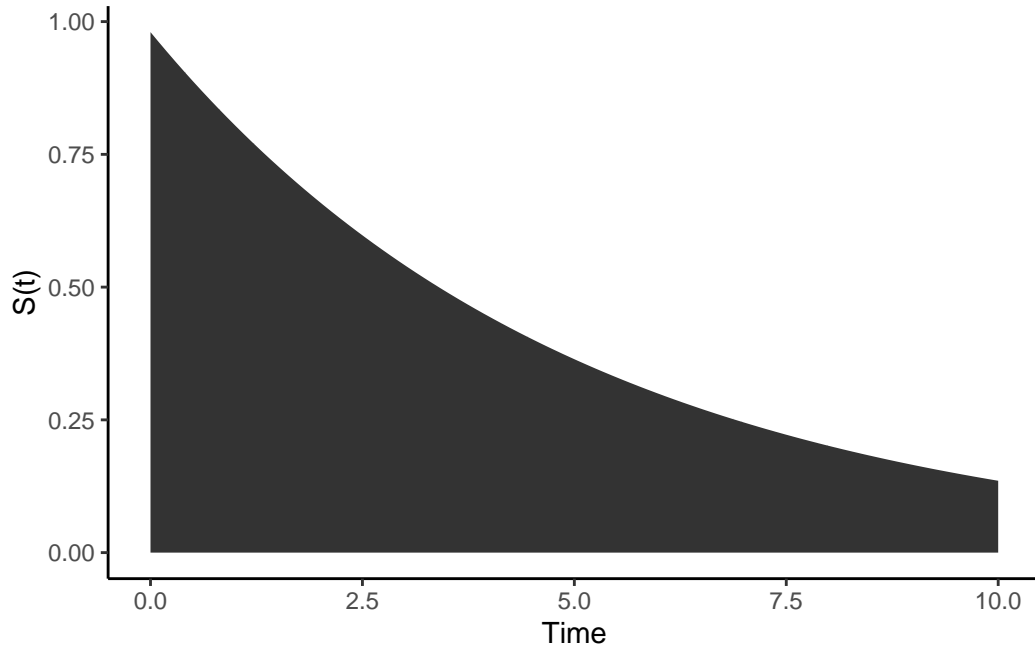
The cumulative hazard, which is the total accumulated hazard of experiencing the event before  $t$ , is defined as:

$$\Lambda(t) = \int_0^t \lambda(u) du \quad (3.4)$$

The survival function and the hazard function are linked such as:

$$S(t) = \exp(-\Lambda(t)) = \exp\left(-\int_0^t \lambda(u) du\right) \quad (3.5)$$

Those quantities are at the core of all event history data analysis.



**Figure 3.4:** RMST up to  $\tau = 10$

Another interesting measure, albeit less used, is the Restricted Mean Survival Time (RMST). RMST is defined using a pre-specified time-horizon  $\tau$ , as the area under the curve of the survival function up to time  $\tau$  (Figure 3.4), or formally:

$$\begin{aligned} \text{RMST}(0, \tau) &= \int_0^{\tau} S(t) dt \\ &= \mathbb{E}[\min(T, \tau)] \end{aligned} \tag{3.6}$$

corresponding to the average number of years spent alive before  $\tau$ . RMST is preferred over the Mean Survival Time (MST)  $= \int_0^{\infty} S(t) dt$ , also known as life expectancy at birth, because the MST is ill determined due to its strong dependence on the unobserved right-hand tail of the survival function. The choice of  $\tau$  is therefore restricted by the observed data. It can be further reduced for clinical relevance, when the focus is on the RMST within a short time span (Mia Klinten Grand and Putter 2016).

A quantity similar to the RMST is the Life Years Lost (LYL), which is often preferred for public



health communications:

$$\text{LYL}(0, \tau) = \tau - \text{RMST}(0, \tau) = \tau - \int_0^\tau S(t)dt \quad (3.7)$$

### 3.2.1 . Non Parametric Estimators for Survival Analysis

The described cumulative quantities are estimated using non-parametric estimators which describe these quantities in a crude way.

Non-parametric estimators of those quantities are built using the counting process theory. Let  $Y(t)$  be the number of subjects at risk at time  $t$ , and  $N(t)$  the number of observed events until  $t$ .

The Kaplan-Meier (KM) estimator of the survival function is (Kaplan and Meier 1958):

$$\hat{S}_{KM}(t) = \prod_{u \leq t} \left(1 - \frac{dN(u)}{Y(u)}\right) \quad (3.8)$$

The Nelson-Aalen estimator of the cumulative hazard function is:

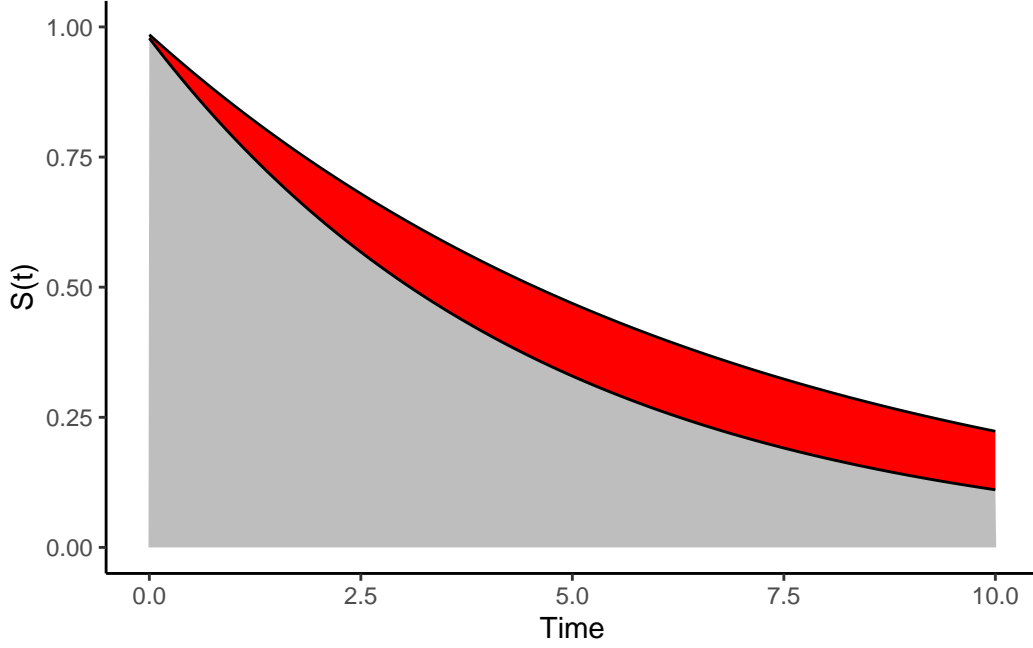
$$\hat{\Delta}_N(t) = \sum_{u \leq t} \frac{dN(u)}{Y(u)} \quad (3.9)$$

which can also be used to estimate the survival function using  $\hat{S}_N = \exp(-\hat{\Delta}_N(t))$ , and gives results asymptotically equivalent to the KM estimator (Borgan 2005).

A plug-in estimator of LYL can thus be obtained by:

$$\widehat{\text{LYL}}(0, \tau) = \tau - \int_0^\tau \hat{S}_{KM}(t)dt \quad (3.10)$$

### 3.2.2 . Life Years Lost by a Categorical Covariate



**Figure 3.5:** Life Years Lost up to 10 years by the binary covariate  $X$  is the red area.

LYL can also be defined as LYL by a covariate  $X$ . If  $X$  is binary, it is the difference in LYL among individuals with  $X = 1$  and those with  $X = 0$ , which is the area between the survival curves (Figure 3.5), or formally:

$$\begin{aligned} \text{LYL}(0, \tau|X) &= \text{LYL}(0, \tau|X = 1) - \text{LYL}(0, \tau|X = 0) \\ &= \mathbb{E}[\min(T, \tau)|X = 0] - \mathbb{E}[\min(T, \tau)|X = 1] \end{aligned} \quad (3.11)$$

The plug-in estimator is then obtained by:

$$\begin{aligned} \widehat{\text{LYL}}(0, \tau|X) &= \left( \tau - \int_0^\tau \hat{S}_{KM}(t|X = 1) dt \right) - \left( \tau - \int_0^\tau \hat{S}_{KM}(t|X = 0) dt \right) \\ &= \widehat{\text{LYL}}(0, \tau|X = 1) - \widehat{\text{LYL}}(0, \tau|X = 0) \end{aligned} \quad (3.12)$$

### 3.2.3 . Regression Model for the Hazard

In survival analysis, the interest is often to assess the effect of covariates on the hazard. This is most frequently done using a Proportional Hazard (PH) model. A PH model explains the hazard  $\lambda(t|X)$  using a baseline hazard  $\lambda_0(t|X)$ , a vector of regression parameters  $\beta$ , and a vector of covariates  $X$ , by assuming a multiplicative effect of  $X$  on  $\lambda_0(\cdot|X)$ :

$$\lambda(t|X) = \lambda_0(t|X) \exp(\beta X) \quad (3.13)$$

The interpretation is made simple by the PH assumption, with an increase of  $X$  by 1 multiplying the hazard by the Hazard Ratio (HR):  $\exp(\beta) = \frac{\lambda(t|X=1)}{\lambda(t|X=0)}$ .

Some authors choose to use a parametric PH model, in which  $\lambda_0$  is parametric. The distribution of  $\lambda_0(t|X)$  is often chosen to be exponential, Weibull, Gompertz, Gamma, piecewise constant functions, or spline functions (Royston and Parmar 2002). The choice of distribution for  $\lambda_0$  is made based on the observed shape of the hazard. However, the underline distribution is not always clear, and a misspecification of  $\lambda_0$  leads to biased estimates of the HR. Instead, a semi-parametric model leaving  $\lambda_0$  unspecified is often preferred – this is the Cox model.

The Cox Model solves the issue of  $\lambda_0$  misspecification, but remains very sensible to violation of the PH assumption, although this assumption is often fulfilled in practice. This is why verification of this assumption using the schoenfeld residuals is strongly recommended when fitting a Cox Model (Schoenfeld 1980). It is important to note that the Cox model is non collapsible, that is the introduction or removal of new covariates  $X_2$  will change the estimates of the old covariates  $X_1$ , regardless of the correlation between  $X_1$  and  $X_2$  (Sjölander, Dahlqwist, and Zetterqvist 2016). This can lead to challenging choices of covariates, when balancing data availability, clinical relevance, and the PH assumption.

The Cox model can also be written as a transformation model for the survival as:

$$\log(-\log(S(t|X))) = \log(\Lambda(t)) + \beta X \quad (3.14)$$

with a linear effect of  $X$ . This relationship is used to graphically test the hypothesis of proportional hazards (Grambsch and Therneau 1994). This formulation also makes explicit that  $\beta$  has an effect on the survival function. A  $\beta > 0$  increases the hazard and decreases the survival, whereas a  $\beta < 0$  decreases the hazard and increases the survival (Kalbfleisch and Prentice 2002).

### 3.2.4 . Pseudo Observations for Survival Analysis

The PH model is a widely and easily used regression model on the hazard. Other quantities (survival function, LYL) exist in survival analysis and a regression method on those is of interest, as they are quantities easily understood and communicated on because they are scaled as a probability or a time. Pseudo-observations is an approach used to estimate such regression models.

The original idea is that with complete data, the survival time would be observed for all individuals and standard statistical methods could be used on  $T$ . For the survival function  $S(t)$ , its expectation  $\mathbb{E}[S(t)] = \mathbb{E}[\mathbf{1}(T > t)]$  could then be estimated using  $\frac{1}{n} \sum_i \mathbf{1}(T > t)$  (Per Kragh Andersen and Pohar Perme 2010).

In practice, the data is incomplete, but the Kaplan-Meier is an unbiased estimator of  $\mathbb{E}[S(t)]$ . Using it, it is possible to define the pseudo-observation for individual  $i$  using a jackknife estimator, as the contribution of individual  $i$  to the estimator of  $\mathbb{E}[S(t)]$ :

$$\hat{\theta}_{it} = n \times \hat{S}_{KM}(t) - (n-1) \widehat{S}_{KM}^{(-i)}(t) \quad (3.15)$$

where  $n$  is the sample size, and  $\widehat{S}_{KM}^{(-i)}$  is the Kaplan-Meier estimator on the sample of size  $n-1$  where the individual  $i$  is removed. This new pseudo-observation data being complete, it is

possible to use standard statistical methods on it.

The use of the Kaplan-Meier and the survival function is for illustrative purpose, the pseudo-observations  $\hat{\theta}_{it}$  can be computed for any function  $G_t$  and any non-parametric unbiased estimator  $\hat{\theta}$  of its expectation  $\mathbb{E}[G_t]$ .

The two following results justify the use of pseudo-observations in regression models. When assuming that censoring is independent from event time conditional on the covariates (Graw, Gerds, and Schumacher 2009):

$$\mathbb{E}[G_t(T)|X] = g^{-1}(X\beta) \implies \mathbb{E}[\hat{\theta}_{it}|X] = g^{-1}(\beta_i X_i) + o_P(1) \quad (3.16)$$

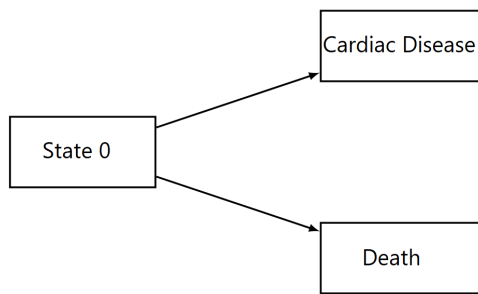
with  $g$  a link function of our choice and  $P$  the probability law of vector  $X_i$ . And in the pseudo-value approach, estimates of the regression coefficients are the solutions of the following generalized estimating equation (Graw, Gerds, and Schumacher 2009):

$$U_{(n)}(\beta_t) = \sum_{i=1}^n \left( \frac{\partial g^{-1}(\beta_t^T Z_k)}{\partial \beta_t} \right)^T V_k^{-1} \{ \hat{\theta}_{it} - g^{-1}(\beta_t^T Z_k) \} = 0 \quad (3.17)$$

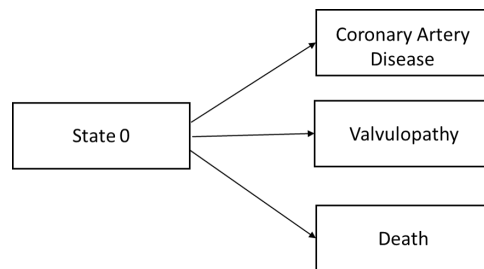
with  $V_k$  the working covariance matrix to account for the correlation structure of pseudo-observations.

Therefore, pseudo-observations can be used as response for a General Estimating Equation (GEE) with a link function  $g$  of our choice, to fit a model on  $G_t$ . The Kaplan-Meier estimator of the survival function and the RMST estimator are both valid choices of  $\hat{\theta}$ , therefore pseudo-observations can be used to perform direct regression on the survival function and LYL.

A single time point is enough to identify the regression coefficients for the life years lost. However, the survival function is a function of time, and therefore requires the use of several time points. By using several time points, it is possible to estimate the baseline risk function  $\theta(t|X = 0)$ , which can be of interest in itself. To do so, pseudo-observations are computed on a grid of time points. Using a grid of a dozen time points chosen as the deciles of the observed



(a) Illustrative example of a competing risk setting with Cardiac Disease and Death being the events of interest.



(b) Illustrative example of a competing risk setting with coronary artery disease and valvulopathy being the events of interest.

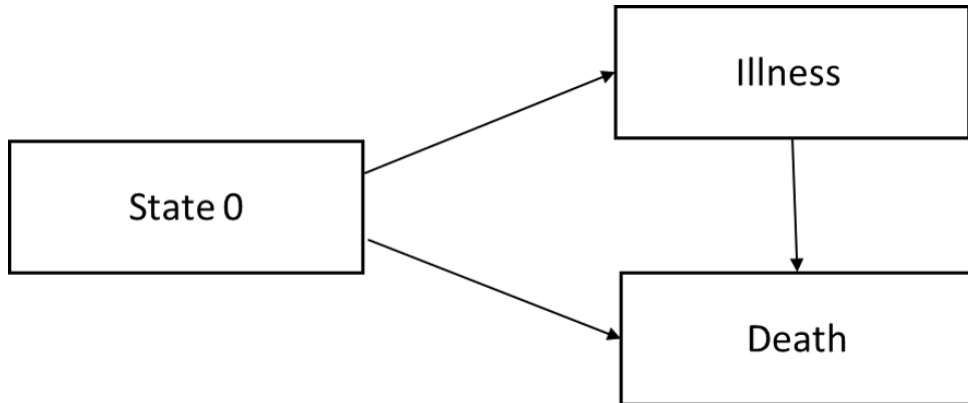
**Figure 3.6:** Examples of competing risk models

times is enough to provide an accurate estimation (Klein and Andersen 2005). Having multiple pseudo-observations for a given individual leads to correlated pseudo-observations which can be accounted for when specifying  $V_k$ . Nevertheless, the use of the identity matrix as a working covariance matrix is proven to provide good estimates and therefore often used for simplicity (Klein and Andersen 2005).

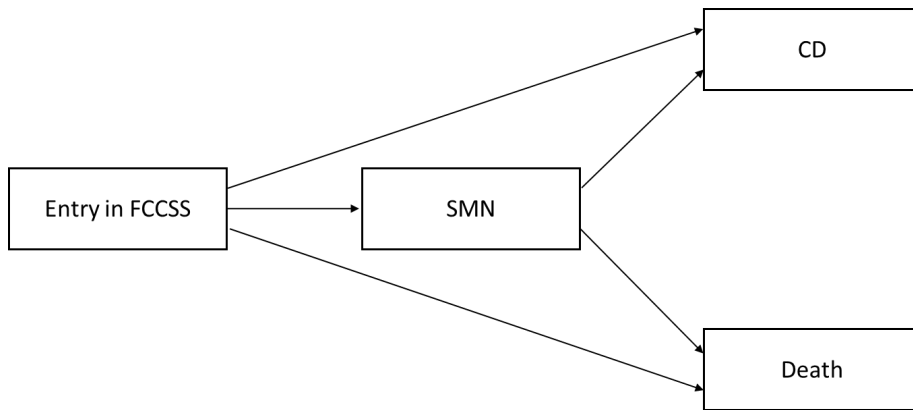
### 3.3 . Multi-State Models

Survival analysis studies only one event time and event type. When exclusives event types exist, competing risks models are used. Examples include the study of CD accounting for non-CD death (Figure 3.6a) and distinguishing CD types while accounting for non-CD death (Figure 3.6b). When events aren't exclusive, multi-states models are used. The original motivating example is the Illness-Death model, which allows for a transition from a healthy state to a diseased state, and from the healthy or diseased state to the death state (Figure 3.7). The motivating example for this thesis is Figure 3.8, to distinguish between a CD when healthy and a CD after a SMN while accounting for the competing risk of death.

Formally, states are defined as  $S = \{0, \dots, M\}$ , and the observed data is the state occupied at time  $t$ :  $E(t)$ , for  $t < C$  the censoring time. States are either absorbing states – when patients can never leave it, or transient states. In the survival analysis and competing risk settings, all



**Figure 3.7:** Illustrative example of the Illness-Death model.



**Figure 3.8:** Motivating example for the Phd of a multi-state setting with two absorbing states and one transient state.

individuals start in the same initial transient state and all others states are absorbing.

In the competing risk setting, it is often preferred to define the observed process using  $\tilde{T}$  the observed time, and  $e$  the type of event observed, with  $e = 0$  in case of censoring and  $e = k \in \{1, \dots, K\}$  the cause of failure otherwise.

The quantities used to study survival data are extended to the multi-state setting. Let  $Q_k(t)$  be the probability of being in state  $k$  at time  $t$ :

$$Q_k(t) = \mathbb{P}(E(t) = k)$$

In the survival data setting, we have that  $Q_0(t) = S(t)$ . In the competing risk setting, we define

the cause- $k$  cumulative incidence function as the probability to experience the event  $k$  before time  $t$ . It is equal to  $Q_k(t)$ .

$$CIF_k(t) = \mathbb{P}(T < t, e = k) \quad (3.18)$$

The hazard function equivalent is the transition intensity function (or cause-specific hazard), which is the instantaneous risk to move from state  $h$  to state  $k$  at time  $t$ , conditionally on being in state  $h$  at time  $t$  :

$$\alpha_{hk}(t) = \lim_{dt \rightarrow 0} \frac{\mathbb{P}(E(t+dt) = k | E(t) = h)}{dt} \quad (3.19)$$

As for the hazard function, the cumulative transition intensity from state  $h$  to state  $k$  is defined as:

$$\Lambda_{hk}(t) = \int_0^t \alpha_{hk}(u) du \quad (3.20)$$

The transition probability for an individual in state  $h$  at time  $s$  to be in state  $k$  at time  $t > s$  is:

$$P_{hk}(s, t) = \mathbb{P}(E(t) = k | E(s) = h) \quad (3.21)$$

In the competing risk setting, the cause- $k$  cumulative incidence is linked to the transition probability such as:

$$CIF_k(t) = P_{0k}(0, t) \quad (3.22)$$

LYL are hard to properly define in a multi-state setting, because in the absence of an absorbing state patients leave the risk set only when censored. Instead the Expected Length of Stay



(ELOS) is preferred (Mia Klinton Grand and Putter 2016). It is the average time spent in state  $k$ :

$$\text{ELOS}_k = \int_0^{\infty} \mathbf{1}(E(t) = k) dt \quad (3.23)$$

In the competing risk setting, LYL up to  $\tau$  can be defined as LYL up to  $\tau$  by all-cause

$$\text{LYL}(0, \tau) = \tau - \int_0^{\tau} P_{00}(0, t) dt \quad (3.24)$$

which can be decomposed as LYL up to  $\tau$  by cause  $k$  (P. K. Andersen 2013):

$$\begin{aligned} \text{LYL}(0, \tau) &= \sum_{k \neq 0} \text{LYL}_k(0, \tau) \\ &= \sum_{k \neq 0} \int_0^{\tau} P_{0k}(0, t) dt \end{aligned} \quad (3.25)$$

### 3.3.1 . Non Parametric Estimators for Multi-States

Let  $Y_h(t)$  be the number of subjects in state  $h$  at risk at time  $t$ , and  $N_{hk}(t)$  the number of observed direct transitions from state  $h$  to state  $k$  until  $t$ . In the following, the multi-state process is assumed to be Markovian and non-homogeneous. The Markovian assumption means that the previous states do not affect the transition intensities, and the non-homogeneous assumption means that individuals move from state to state. The cumulative transition intensities are estimated using the Nelson-Aalen estimator:

$$\widehat{\Delta}_{hk}^*(t) = \begin{cases} \int_0^t \frac{dN_{hk}(u)}{Y_h(u)} du, & h \neq k \\ -\sum_{h \neq k} \widehat{\Delta}_{hk}^*(t), & h = k \end{cases} \quad (3.26)$$

Let  $\widehat{\Delta}^*$  be the matrix of the cumulative intensities, with  $h, k$  elements being  $\widehat{\Delta}_{hk}^*$ . Using this matrix and the product-integral, the transition probabilities  $P_{hk}(s, t)$  can be estimated with

the Aalen-Johansen estimator (Aalen and Johansen 1978):

$$\widehat{P}_{hk}^*(s, t) = \prod_{(s,t]} (\mathbf{I} + d\widehat{\Delta}^*(u)) \quad (3.27)$$

with  $\mathbf{I}$  the identity matrix, and the product integral being a finite matrix product over the jump times of  $\widehat{\Delta}_{hk}^*(u)$  in  $(s, t]$ . An estimator for the probability to be in state  $k$  at time  $t$  can be easily deduced from it in absence of left-truncation:

$$\widehat{Q}_k(t) = \sum_{h \neq k} \widehat{Q}_h(0) \widehat{P}_{hk}^*(0, t) - \widehat{Q}_k(0) \widehat{P}_{kh}^*(0, t) \quad (3.28)$$

Because  $CIF_k(t) = Q_k(t)$  in the competing risk setting, this is also an estimator of the cumulative incidence function.

A plug-in estimator of the ELoS in state  $k$  can thus be obtained:

$$\widehat{\text{ELoS}}_k = \int_0^\tau \widehat{Q}_k(t) dt \quad (3.29)$$

In the competing risk setting, the LYL by cause  $k$  can be estimated similarly using:

$$\widehat{\text{LYL}}_k(0, \tau) = \int_0^\tau \widehat{Q}_k(t) dt \quad (3.30)$$

### 3.3.2 . Proportional Cause-Specific Hazard Models

The Cox Model can be extended to the multi-state settings, by modeling the cause-specific hazards:

$$\alpha_{hk}(t|X) = \alpha_{hk,0}(t) \exp(\beta_{h,k}X) \quad (3.31)$$

Following Klein (2006), in a competing risk settings, a desirable conclusion would be “The use of radiotherapy multiplies the cause-specific-hazard of each cause  $k$  by  $\beta_k$  and the all-cause

hazard by  $\beta_*$  (Gerds, Scheike, and Andersen 2012). However, such a result is unlikely to be true in practice, as the proportional assumption on each cause-specific hazard and on the all-cause hazard would lead to the equality:

$$\begin{aligned}\alpha_{0*}(t) &= \alpha_{0*,0}(t) \exp(\beta_* X) \\ \forall k, \alpha_{0k}(t) &= \alpha_{0k,0}(t) \exp(\beta_k X) \\ \implies \alpha_{0*,0}(t) \exp(\beta_* X) &= \sum_{k \in S} \alpha_{0k,0}(t) \exp(\beta_k X)\end{aligned}\tag{3.32}$$

where  $\alpha_{0*}$  is the all-causes hazard. This equality only holds in two special cases. If there exists a  $\beta$  such as for all cause  $k$ ,  $\beta = \beta_k$  in which case  $\beta_* = \beta$ , and if the baseline cause-specific hazards  $\alpha_{0k,0}(t)$  are all proportional. While both cases are mathematically possible, they are implausible in most competing risk problems. Indeed, covariates having the same effect on all competing risks is unlikely to ever happen, and in most competing risks settings those risks tend to be high at different time periods.

Another desirable conclusion would be “the probability of dying from cardiovascular disease during the next 5 years is  $\beta^*$  times as high for a patient treated with radiotherapy than for a patient treated without”. However, the cause-specific proportional hazard model doesn’t provide answers that allow such an interpretation. Instead, this model yields a conclusion such as “At all times during the next 5 years the instantaneous risk of dying from cardiovascular disease is  $\beta_k$  times as high for a patient treated with radiotherapy than for a patient treated without radiotherapy”. The reason this conclusion can’t be used to deduce a hypothetical  $\beta^*$  is that  $CIF_k$  depends on the cause-specific hazard of cause  $k$ , and of all other causes  $k' \neq k$ :

$$\begin{aligned}CIF_k(t) &= \exp\left(-\int_0^t \alpha_{0k}(u) P_{00}(u) du\right) \\ &= \exp\left(-\int_0^t \alpha_{0k,0}(u) \exp(\beta_k X) \left(\int_0^u \sum_{k' \neq k} \alpha_{0k',0}(v) \exp(\beta_{k'} X) dv\right) du\right)\end{aligned}\tag{3.33}$$

As a consequence, even a significant effect of  $\beta_k$  on the cause specific hazard can’t be generally

interpreted as an effect on the cause-k cumulative incidence function,  $CIF_k$  (Gray 1988).

To solve this interpretation problem in the competing risks setting, the Fine and Gray model formulates effects on the subdistribution hazard (Fine and Gray 1999). The subdistribution hazard is defined as:

$$\alpha_{0k}^{sub}(t) = \lim_{dt \rightarrow 0} \frac{\mathbb{P}(T \leq t + dt, e = k | T > t \vee (e \neq k \wedge e \neq 0))}{dt} \quad (3.34)$$

It is similar to the transition intensity  $\alpha_{0k}(\cdot)$  without removing patients who experienced a competing event from the risk set. If the new  $\beta_{FG}$  is positive, then we can draw conclusions such as “the probability of dying from cardiovascular disease during the next 5 years is significantly higher for individuals treated with radiotherapy than for those treated without radiotherapy”. The coefficients  $\beta_{FG}$  can’t be used to draw a conclusion on the magnitude of this increase, as they have value on the subdistribution hazard which makes little clinical sense.

All of this showed the need for use of other methods to perform regression with coefficients interpretable on the cumulative incidence scale in the competing risks setting.

### 3.3.3 . Pseudo Observations for Multi-State

Pseudo-observations can be used in the multi-state setting too, with all the estimator presented in Section 3.3.1 as  $\hat{\theta}$ . They have notably been used to study LYL by cause (Per Kragh Andersen, Canudas-Romo, and Keiding 2013; Aurélien Latouche et al. 2019), Expected Length of Stay (Mia Klinton Grand and Putter 2016), and cumulative incidence (Charrier et al. 2023) via direct regression models in competing risk and multi-states settings.

### 3.3.4 . Life Years Lost by Cause and Life Years Lost by Covariate

Let  $X$  be a binary covariate of interest. Let's assume that it is possible to distinguish whether or not the exposition to  $X$  is the cause of death, and let  $e$ , the cause of death be 2 in this case, and 1 otherwise. Then, an expected property of LYL would be that LYL by cause  $e = 2$  would be equal to LYL by  $X$ . However, this property doesn't hold. This section explains why.

The LYL by cause 2 can be simply written as:

$$LYL_2(0, \tau) = \int_0^\tau CIF_2(t) dt$$

The LYL by  $X$  can be written as :

$$LYL(0, \tau|X = 1) = \int_0^\tau S(t|X = 0) - S(t|X = 1) dt$$

Because  $S(t) = 1 - \sum_e CIF_e(t)$ , LYL by  $X$  can be written as:

$$\begin{aligned} LYL(0, \tau|X = 1) &= \int_0^\tau (1 - CIF_1(t|X = 0) - CIF_2(t|X = 1)) - (1 - CIF_1(t|X = 1) - CIF_2(t|X = 1)) dt \\ &= \int_0^\tau CIF_1(t|X = 1) + CIF_2(t|X = 1) - CIF_1(t|X = 0) dt \end{aligned}$$

Therefore, the equality  $LYL_2(0, \tau|X = 1) = LYL(0, \tau|X = 1)$  would hold only in the special case where the cumulative incidence of death by cause 1 are equal in populations with  $X = 0$  and  $X = 1$  at all time  $t$ .

However, this equality doesn't hold when  $CIF_2(t|X = 1) > 0$ . Intuitively, this can be thought of as death by cause 2 preventing death by cause 1, and therefore the cumulative incidence of death by cause 1 decreasing. Hence,  $CIF_1(t|X = 1) < CIF_2(t|X = 0)$  for all  $t$  such as  $CIF_2(t|X = 1) > 0$ , and  $LYL(0, \tau|X = 1) < LYL_2(0, \tau)$ .

### 3.4 . Time-Dependent Covariates

Time-dependent covariates occur frequently in biomedical and epidemiological research. These are covariates that change over time. In this thesis, the occurrence of a SMN is used as a time-dependent covariate in Section 4, by defining  $X_{SMN}(t) = \mathbf{1}(\text{a SMN occurred before } t)$  . Other examples include air pollution for the study of respiratory diseases, and blood pressure for diabetes.

Time-dependent covariates are classified as either external covariates and internal covariates (Kalbfleisch and Prentice 2002). External time-dependent covariates are *external* to the subject, can affect the failure process, but are not otherwise involved in the failure mechanism. Internal time-varying covariates are measured on the subject, can affect the failure process directly, and may also be impacted by the failure mechanism (Austin, Latouche, and Fine 2020). Formally, a time-dependent covariate is external if:

$$\forall (u, t) \in \mathbb{R}^2, u \leq t \implies \mathbb{P}(u \leq T < u + du | X(\mathbf{u}), T \geq u) = \mathbb{P}(u \leq T < u + du | X(\mathbf{t}), T \geq u) \quad (3.35)$$

External time-dependent covariates are typically out-patients information such as the weather or air pollution. All others time-dependent covariates are internal, examples include blood pressure and in our context the occurrence of a SMN.

#### 3.4.1 . Proportional Hazard Models

Time-dependent covariates can be included in proportional hazards models (T. M. Therneau and Grambsch 2000). In all cases, their effects is interpretable as having a multiplicative effect on the hazard (Austin and Fine 2017). However, to correctly interpret the results on the survival or cumulative incidence functions, external and internal time-dependent covariates must be distinguished, as well as the presence or absence of competing risks.

### 3.4.1.1 Time-Dependent Covariates without Competing Risk

In the absence of time-dependent covariates and competing risks, the coefficients of a Cox model can be interpreted as increasing or decreasing the survival function, as discussed in Section 3.2.3. With time-dependent covariates, this interpretation only holds when the time-dependent covariates are all external. In this case, the survival function can be written as:

$$\begin{aligned} S(t|X(s), s \leq t) &= \exp(-\Lambda(t|X(s), s \leq t)) \\ &= \int_0^t \lambda(s|X(s)) ds \\ &= \int_0^t \lambda_0(s) \exp(\beta X(s)) ds \end{aligned} \tag{3.36}$$

Because  $\exp(\beta X(s))$  is a function of the variable of integration, it can't be brought outside of the integral to give results similar to Equation 3.14. A consequence is that an effect of  $X(s)$  on the hazard doesn't imply an effect of the opposite direction on the survival function, as its effect depends on the entire history of  $X(s)$  over  $[0, t]$ .

Nevertheless, in the special case where  $X(s)$  is ordered, and its values for two individuals  $i, j$  are so that for all time  $s \leq t$ ,  $X_i(s) \leq X_j(s)$ , then  $\beta > 0$  implies that the survival function for individual  $j$  is lower than for individual  $i$  at all times  $s \leq t$ . In other cases where  $X(t)$  is unordered, then comparing the survival function of two individuals requires to look at  $\beta$ , the history of  $X(s)$  over  $[0, t]$ , and the baseline hazard function  $\lambda_0$  (Austin, Latouche, and Fine 2020).

If  $X(t)$  is an internal time-dependent covariate, a Cox model no longer yields results on the survival function. The core of the issue, is that  $X(t)$  is not known at all times regardless of the observation status. In particular, once the individual has experienced the event, the value of  $X(t)$  is unknown. Therefore, the integral in Equation 3.36 can't be evaluated, and the relationship between  $S(t|X(s), s \leq t)$  and  $\lambda(t|X(s), s \leq t)$  doesn't hold.

### 3.4.1.2 Time-Dependent Covariates with Competing Risk

With time-fixed covariates, we insisted on the point that in the presence of competing risks a PcsH model doesn't yield results directly interpretable on the cumulative incidence function (see Section 3.3.2).

In the presence of competing risks, the Fine-Gray model yields results interpretable on the cumulative incidence function without time-dependent covariates. When introducing time-dependent covariates, internal and external covariates must be distinguished too.

First, for external time-dependent covariates the usual relationship between the subdistribution hazard and the cumulative incidence holds:

$$CIF_k(t|X(s), s \leq t) = 1 - \exp\left(-\int_0^t \alpha_{0k,0}^{sub}(s) \exp(\beta X(s)) ds\right) \quad (3.37)$$

The integral above can be evaluated because  $X(s)$  is known at all time, regardless of the individual being observed or not. As previously,  $\exp(\beta X(s))$  can't be brought outside of the integral. Therefore, an effect of  $X(s)$  on the subdistribution hazard doesn't imply the same effect on the cumulative incidence, as its effect depends on the entire history of  $X(s)$  over  $[0, t]$  (Austin, Latouche, and Fine 2020).

In the special case where  $X(s)$  is ordered, the cumulative incidence of two individuals  $i, j$  can once more be compared using  $\beta$  if for all time  $s \leq t$ ,  $X_i(s) \leq X_j(s)$ . In any case, for a fixed realization of  $X(s)$ , individuals with different values of  $X(s)$  may be compared by calculating the cumulative incidence via either analytic or numerical integration. This is possible only because  $X(s)$  is known at all times.

Second, internal time-dependent covariates shouldn't be used in the Fine-Gray model to interpret results on the cumulative incidence (A. Latouche, Porcher, and Chevret 2005; Austin, Latouche, and Fine 2020). This is because internal covariates aren't observed at all times, therefore  $\int_0^t \alpha_{0k,0}^{sub}(s) \exp(\beta X(s)) ds$  can't be evaluated and the link between the subdistribution



hazard and the cumulative incidence doesn't hold.

Another issue occurs with internal covariates in the Fine-Gray model. When individuals experience a competing event they remain in the risk set for the subdistribution hazard. However, the occurrence of a competing event (eg. death) often precludes the observation of the internal covariate or makes its definition unclear. For this reason, the definition of an internal covariate after the occurrence of a competing event is at best unclear. The use of the last observation carried forward has been suggested, but isn't guaranteed to give valid nor interpretable results. Furthermore, this choice leads to a biased estimate, even in the simple case of a time jump process (Beyersmann and Schumacher 2008; A. Latouche, Porcher, and Chevret 2005).

### **3.4.2 . Landmark**

#### **3.4.2.1 The Concept of Landmark**

When studying a binary time-dependent covariate, a desirable conclusion would be "A change in  $X(t)$  increases the survival of individuals". For example, let  $X(t)$  be the patient's response to chemotherapy. A naive solution would be to stratify the population based on the observation of a response to chemotherapy in  $X(t)$  and to compare the survival functions of individuals who responded and those who didn't. However, this method is biased towards a higher survival for the responder group and can show a difference in survival despite  $X(t)$  not having any effect on survival. Intuitively, individuals who responded will be protected from death until they respond at time  $t_{immu} = \min(t|X(t) \neq X(0)) > 0$ , a protection that doesn't occur for individuals who didn't respond. This is called the immortality bias. For the same reasons, the interpretation that  $X(t)$  can be used to predict the future of individuals is also wrong. Another example of such a binary covariate, and the motivating example in this thesis, is the diagnosis of a SMN. A valid method, known as landmarking, is to select a fixed time  $l$  after the start of the study as a landmark for conducting the analysis (Anderson, Cain, and Gelber 1983) . Individuals still on study at the landmark time are separated into two categories according to  $X(l)$ , defined

as response/occurrence of event before that time (Cortese and Andersen 2010; Putter and Houwelingen 2017). Individuals are then followed forward in time to ascertain whether survival from the landmark depends on the patient’s covariate status at the landmark. Thus, inference is conditional on the status of patients at the landmark time. This method can handle continuous and/or categorical covariates.

In the presence of left truncation, individuals are included in the study conditionally on being at risk at the landmark time  $l$ . In this way, left truncation is discarded from the study and no further work is required for taking into account this unobservation mechanism.

When choosing this method, the choice of  $l$  can impact the estimation of the effect of  $X(t)$ . One solution is to choose a biologically or clinically meaningful landmark time point. In the context of studying multi-morbidity of childhood cancer survivors, the existence of biologically meaningful time point is unclear. This is because the different morbidities studied have different time of onset, and those times are also affected by the different treatment regimen used to treat childhood cancers. To make sure the choice of  $l$  doesn’t impact the results presented in this thesis, all landmark models were computed over a wide range of landmark times to investigate how the choice of  $l$  impacts the results. In the absence of a marked effect of landmark time point choice on the estimates of interest, results were shown at a subset of clinically meaningful landmark times.

### **3.4.2.2 Landmark for Time-Dependent Covariates with Competing Risks**

The landmark method was extended to competing risks and time-dependent covariates (Cortese and Andersen 2010). The idea is to use the last observation carried forward imputation to convert the time-dependent covariate  $X(t)$  into a time-fixed covariate  $X(l)$ , with  $l$  the landmark time.

This choices induces some bias. In a Cox model, let  $\beta_l$  be the estimated effect of  $X(l)$  on the hazard. Its value is constrained by  $\beta$ , the estimated effect of  $X(t)$  in a Cox model with time-

dependent covariate  $X(t)$ , such as  $|\beta_l| < |\beta|$  (Putter and Houwelingen 2017), and the difference between  $\beta$  and  $\beta_l$  increases with the frequency of changes in  $X(t), t > l$ . Intuitively, this is explained by a decrease in the validity of the value  $X(l)$  as a proxy for  $X(t)$ .

### 3.4.2.3 Landmark to Study Time-Varying Coefficients

Previously, only time-dependent covariate  $X(t)$  was discussed. Here, the possibility of a time-varying coefficient  $\beta(t)$  is discussed. Such coefficients can be specified in a Cox model, which becomes  $\lambda(t|X) = \lambda_0(t) \exp(\beta(t)X)$  for time-fixed covariates, and  $\lambda(t|X) = \lambda_0(t) \exp(\beta(t)X(t))$  for time-dependent covariates.

The landmark method can be used to fit a Cox model when the PH is not met, and the shape of  $\beta(t)$  unclear (Nicolaie et al. 2013b; Houwelingen and Putter 2015). The idea is to choose a set of landmark time points  $L = (l_j)_j$  such as the PH assumption is met on the intervals  $[l_j, l_{j+1}]$  and to fit models on those intervals by including patients conditionally on their survival at  $l_j$  and adding administrative censoring at  $l_{j+1}$  for all patients still at risk at that time.

The idea to split the time in intervals and fit a model on each interval can be used to estimate any time-varying effect of covariates. When doing so, a new  $\beta_l$  is estimated at each landmark time  $l$ . It is interesting to understand the link between all  $\beta_l$  to have a clear view of the time-varying effect of the covariates. This can be done using a **super model**, which fits the function  $\beta_{LM}(t)$  using  $(\beta_l)_{l \in L}$  and a known function  $h$  (Van Houwelingen 2007; Nicolaie et al. 2013a; Mia Klinten Grand and Putter 2016; Putter and Houwelingen 2022). In practice,  $h$  is chosen by plotting  $(\beta_l)_{l \in L}$ , and is often a polynomial function. The use of a super landmark models is valid in terms of prediction, but the estimates are different than when estimating  $\beta(t)$  directly (Van Houwelingen 2007). This is because  $\beta_{LM}(t)$  is estimated using information on the future, as  $\beta_l$  with  $l > t$  are used to estimate it. In practice, the difference between  $\beta(t)$  and  $\beta_{LM}(t)$  is small (Nicolaie et al. 2013a).

In the discussion above, super models were estimated using landmark models on  $[l_j, l_{j+1}]$ , as

the introductory example meant to solve the PH of the Cox model. In a general context, a super model can be estimated on any set of intervals  $([l_j, \tau_j])_j$ , where  $\tau_j$  is the time horizon for landmark time  $l_j$ . For example,  $\tau_j$  can be a fixed horizon  $\tau$  for all  $j$  to add information to landmark model at  $l_j$ , and reduce the variance of the estimators (Van Houwelingen 2007). Another example would be to use  $\tau_j = l_j + \omega$ , with  $\omega$  a fixed window time to study the effect on ELOS and interpret a decrease of the time-varying coefficient as a decrease over time of the effect of the covariate on the ELOS (Nicolaie et al. 2013a; Mia Klinten Grand and Putter 2016).

#### 3.4.2.4 Dynamic Pseudo Observations

Pseudo-observations can be combined with the landmark method to fit a wide range of model with time-dependent covariates and/or time-varying coefficients. This section describes in details how to estimate adapted pseudo-observations (Nicolaie et al. 2013a; Mia Klinten Grand and Putter 2016; Mia K. Grand et al. 2019; Aurélien Latouche et al. 2019).

##### 3.4.2.4.1 Pseudo-Observations with a Single Landmark Time Point

The simplest case consists of having a single landmark time point  $l$ . In this case, let  $I_l$  be the set of patients at risk at time  $l$ , and  $n(l)$  the number of patients at risk at time  $l$ . The new pseudo-observations are then computed only for the patients in  $I_l$ , using the new estimator  $\hat{\theta}^l$  computed using only those patients:

$$\forall i \in I_l, \hat{\theta}_{it}^l = n(l) \times \hat{\theta}^l(t) - (n(l) - 1) \times \widehat{\theta^{(-i),l}}(t) \quad (3.38)$$

The model is then fitted using a GEE and a grid of time points  $t > l$ , as for regular pseudo-observations.

### 3.4.2.4.2 Pseudo-Observations with Landmark Super Models

When the objective is to estimate time-varying coefficients, pseudo-observations can be used with a landmark super model. Pseudo-observations are then computed at a set of landmark time points  $L = (l_k)_k$ . Using the same notation as before:

$$\forall i \in I_{l_k}, \widehat{\theta}_{it}^{l_k} = n(l_k) \times \widehat{\theta}^{l_k}(t) - (n(l_k) - 1) \times \widehat{\theta}^{(-i), l_k}(t) \quad (3.39)$$

Here, the issue of choosing the grid of time points  $t$  arises. This is because for all  $l_k$ , it is required that  $l_k < t$ . First, if the grid is a single time point  $t_k$ , then two possibilities exist. One is to choose a shared  $\omega$  for all  $l_k$ , and to define  $t_k$  as  $l_k + \omega$ . In this way, the time-varying coefficient  $\beta(t)$  is an effect on a fixed measure, such as “cumulative incidence of CD within 5 years” with  $\omega = 5$ . This corresponds to the dynamic pseudo observations (Nicolai et al. 2013a).

A second option is to choose a common time point  $t$  for all  $l_k$ . An example would be to study the evolution of RMST up to 80 years old as patients age. Using a shared horizon is valuable to include the death of older patients at all landmark time points, to include the older and more frail patients at all landmark time points. The nature of the time-varying coefficient  $\beta(t)$  is then harder to interpret. Indeed, in this example at landmark at time  $l_1 = 20$  individuals will be at risk for longer than at landmark time  $l_2 = 60$ . This decreases in at-risk time induces a decreasing trend of the RMST as  $l_k$  increases, because RMST is bound within  $[0, 60]$  for  $l_1 = 20$ , and within  $[0, 20]$  for  $l_2 = 60$ . The same decreasing trend is expected for the estimated effect  $\beta(t)$  of a covariate  $X(t)$ .

The second case is when the grid is chosen with  $J > 1$  different time points  $(t_{kj})_{j \in (1, \dots, J)}$ . As when using a single time point, it would be intuitive to choose a range of  $(\omega_j)_j$  shared across landmark time points, and to define the grid of time points associated to  $l_k$  as  $(t_{kj} = l_k + \omega_j)_j$ . In

this way, it is possible to fit a function of “cumulative incidence of CD  $\omega_j$  years after landmark time point”. Using a landmark super model has the advantage to capture the time-varying trend of the cumulative incidence within  $\omega_j$  years as patients age. Similar results could be had by fitting  $J$  different landmark super models for each  $j \in (1, \dots, J)$  with the single time point  $t_{kj}$ . The advantage of using a grid of time points is an expected higher stability of the results (Klein and Andersen 2005).

The use of a shared horizon with a grid of  $J > 1$  different time points is less obvious. If all the grid of time points  $t_{.j}$  is greater than the maximum of  $l_k$ , then it can be used as with a grid of one time point. This could be used to answer questions such as “What is the cumulative incidence of death on the 60-80 years old range based on the age of cancer onset, for patients experiencing cancer before 60 years old”. This type of investigation is rarely done, because of its interpretation challenges.

### 3.4.2.4.3 Pseudo Observations with Left Truncation

The dynamic pseudo-observations can be linked to the question of pseudo-observations with left truncation. The original issue with pseudo-observations in the presence of left truncation, is that for individual  $i$  and time  $t$  such as the left truncation time of individual  $i$  is greater than time  $t$ , the pseudo observation is

$$\begin{aligned}
 \widehat{\theta}_{it} &= n \times \widehat{\theta}(t) - (n - 1)\widehat{\theta}^{(-i)}(t) \\
 &= n \times \widehat{\theta}(t) - (n - 1)\widehat{\theta}(t) \\
 &= \widehat{\theta}(t)
 \end{aligned}
 \tag{3.40}$$

because individual  $i$  doesn’t contribute to  $\widehat{\theta}(t)$ . A first suggested solution is to use the landmark method, and to compute the pseudo observations of Equation 3.38. This is called strict pseudo-observations (Mia Klinten Grand and Putter 2016).

Strict pseudo-observations can result in a high loss of power, especially in earlier times when few patients are included because of left truncation. The idea to only create pseudo-observations for individuals that actually contribute to the estimator was suggested (Mia K. Grand et al. 2019). The new pseudo-observations are

$$\widehat{\theta}_{it} = n(t) \times \widehat{\theta}(t) - (n(t) - 1)\widehat{\theta}^{(-i)}(t)$$

where  $i \in \{i | T_{L_i} < t\}$  and  $n(t)$  the number of such subjects. The new subset of individuals correspond to individuals who contributed to the pseudo-observations. Those new pseudo-observations are called stopped pseudo-observations. Simulation studies have shown great results of this method, with little bias over multiple scenarios of censoring and truncation (Mia K. Grand et al. 2019).

### 3.4.3 . Modelling a Transient State with a Time-Dependent Covariate

The motivating example for this thesis is the multi-state model Figure 3.8, where SMN is a transient state. A transient state could also be modeled as a time-dependent covariate, where  $X(t) = \mathbf{1}(t \leq t_{SMN})$  (Beyersmann and Schumacher 2008). In this way, the association between a SMN and the risk of CD and non-cardiac death can be estimated using a regression model and the landmark strategy described above. In Section 4, this strategy is used to investigate the association of a SMN occurrence on the cumulative incidence of CD at various landmark time points. Because SMN is considered a time-dependent covariate, it is intuitive to fit a regression model appropriately chosen to summarize the association of SMN and CD with a single coefficient  $\beta$ . Summarizing this association with a single number is useful to communicate with clinicians. This is especially true considering that SMN is an internal time-dependent covariate and death is included as a competing event, which is out of the scope of validity of the Cox model and Fine-Gray model for interpreting an effect on the cumulative incidence. Converting the SMN state into a time-dependent covariate also facilitates the investigation of

a possible time-varying effect of SMN on the risk of CD by using a super model. Indeed, it is plausible that the effect of a SMN on the occurrence of CD depends on the individual's age, as older patients are at increased risk of both CD and non-cardiac death.

The use of the multi-state model has its own advantage. Notably, by using the ELOS estimators it is possible to investigate the average time spent in the healthy state and the average time spent with a SMN without further complications. This is especially useful to investigate the temporality of occurrence of SMN complications.





## 4 - Cardiac Disease Risk after a Second Malignant Neoplasm

The goal of this chapter is to study the impact of a second malignant neoplasm on the risk of severe cardiac disease. Some of the results are already published, along with the code (Charrier et al. (2023), github).

### 4.1 . Context

Advances in cancer treatment have significantly improved childhood cancer survival, with the 5-year survival rate exceeding 80% in Europe today (Gatta et al. 2014). Most of those 5-year survivors live long after, and experience many iatrogenic events, with SMN and CDs being among the most severe and life threatening frequently experienced (Robison and Hudson 2014; Oeffinger et al. 2006; Mertens et al. 2008; Armstrong et al. 2009; Reulen et al. 2010; Tukenova et al. 2010). Previous studies identified several main risk factors for both of them, as discussed in Section 1.4.

The starting point of this thesis on multi-morbidity among childhood cancer survivors has been to study the impact of SMN on CD risk. The idea is twofold.

First, to validate the hypothesis that childhood cancer survivors' iatrogenic events occurrence increase the risk of further iatrogenic events occurring. To do so, well defined late effects with plausible causal mechanism and sufficient number of events in the cohort are required. SMN and CD meet all those requirements. Indeed, they are validated by health care professionals according to international guidelines, and the plausible causal mechanism is the well established cardio-toxicity of cancer treatments, such as anthracyclines and radiotherapy at the heart.

Second, to provide clinical insight on a possibly important risk factor of CD among childhood

cancer survivors. Guidelines already recommend to take into account exposition to cardiotoxic treatment of all cancers when evaluating patients' risk of CD. However, evidence of the cumulative effect of such treatments for childhood cancer survivors is still lacking. Furthermore, despite SMN likely increasing the hazard of CD, previous studies estimating the impact of childhood cancer treatment on CD have ignored it. This could have lead to an over-estimation of childhood cancer treatment effect on CD (Pearl and Mackenzie 2018).

## 4.2 . Methods

The primary event of interest was the first occurrence of severe cardiac disease, or death resulting from cardiac disease, with death from other cause as a competing event. SMN was considered as a time-dependent exposure. This corresponds to the multi-state model of Figure 3.8. We defined  $T$  as the time of the first occurrence of cardiac disease or death resulting from cardiac disease (event of interest), death from other cause (competing event), or last follow-up (censor)

### 4.2.1 . Data

Data used comes from the FCCSS which is detailed in Section 2. In 7,670 CCS over a median follow-up of 30 years (Inter Quartile Range: 22-38 years), there were 828 cases of second malignant neoplasms and 379 cases of cardiac disease, of which 49 patients experienced a second malignant neoplasm before.

### 4.2.2 . Metrics of Interest

The goal of this work was to provide an estimation of the effect of SMN on the cause-specific-hazard of CD and on the cumulative incidence of CD.

#### 4.2.2.1 Cumulative Incidence

To estimate the effect of a SMN on the cumulative incidence of cardiac disease, we used the landmark method described in Section 3.4.2.2, along with the strict pseudo-observations of Section 3.4.2.4. That is, we chose a set of  $K$  landmark time points  $L = (l_k)_{k=1,\dots,K}$ , computed pseudo-observations for each of those landmark times using the Aalen-Johansen estimator, and estimated the parameters of the regression model using general estimating equations (Aalen and Johansen 1978). We chose to use an additive regression model to provide easily interpretable results. Therefore, we assumed the following form of the cumulative incidence of CD for each landmark time  $l_k$ :

$$CIF_{CD}(t|X(l_k), T \geq l_k) = CIF_{0,CD}(t|X(l_k), T \geq l_k) + \beta_{l_k} X(l_k) \quad (4.1)$$

with  $CIF_{0,CD}(t|X(l_k), T \geq l_k)$  the baseline cumulative incidence function. The estimated  $\beta_{l_k}$  has the interpretation that “Compared to patients who didn’t experience a SMN,  $\beta_{l_k}\%$  more patients who experienced a SMN are going to experience a CD while accounting for death of any cause”.

We chose to analyze the cumulative incidence on two different time scales, the time since diagnosis and the patient’s age. The time since diagnosis time scale is the standard to study iatrogenic events among childhood cancer survivors, as the the focus in on the effect of childhood cancer treatments. We used this time scale to give results interpretable alongside previous studies. Nonetheless, cardiac disease occurrence and mortality are strongly influenced by patient’s age. We find it valuable to investigate whether the effect of SMN is mediated by age.

Indeed, providing results conditional on patient’s attained age provides information about their risk once they reached a life milestone.

For the time since diagnosis time scale, we used landmark time points from 15 to 35 years since diagnosis. For the patient’s age time scale, we used landmark time points from 20 to 35 years old. For each landmark time points  $l_k$ , we computed pseudo-observations on the grid of time points  $(l_k + i)_{i \in (2,4,\dots,20)}$ .

#### 4.2.2.2 Cause Specific Hazard

To estimate the effect of a SMN on the cause-specific-hazard of cardiac disease we used a proportional cause specific hazard regression model (Cox 1972; Prentice et al. 1978):

$$\alpha_{0CD}(t|X(t)) = \alpha_{0CD,0}(t) \exp(\beta X(t))$$

We use this method in complement to the regression on the cumulative incidence, as it provides information on the effect of SMN on the instantaneous risk of CD, but not on the cumulative incidence of CD. This is discussed in details in Section 3.4.1.2.

We used time since diagnosis as a time scale, and excluded the 31 patients who had an event within 5 years of diagnosis, because patients aren’t at risk of death during those 5 years. This *immunity* comes from the cohort inclusion criteria of surviving 5 years after childhood cancer diagnosis.

#### 4.2.3 . Childhood Cancer Information

Childhood cancer treatment is known to increase the absolute and instantaneous risk of CD, SMN, and death (Lipshultz et al. 2013; Haddy et al. 2016; Chow et al. 2015, 2018). We adjusted for it by including this information in the regression models as time-fixed covariates. In clinical practice, it is rare to have access to detailed information about childhood cancer treatment, notably radiotherapy and chemotherapy doses. To account for this disparity of

information available to clinicians, we used different combinations of covariates. Thus we were available to provide both clinically relevant results and an estimation of SMN effect on CD with little confounding.

The information included in all models is age at childhood cancer, sex, and year of childhood cancer diagnosis. The smallest model includes radiotherapy (yes/no) and chemotherapy (yes/no) use status. The largest model includes cumulative radiotherapy dose at the heart, cumulative anthracycline dose, cumulative alkylating agent dose, and cumulative platinum agent dose. We also used an univariable model. Table 4.1 details all the covariates combinations.

**Table 4.1:** Covariates combinations used to study SMN effect on CD.

	SMN	Age at diagnosis	Sex	Year of diagnosis	Radiotherapy	Chemotherapy	Dose of RT to the heart	Dose of RT to the brain	Exposition of the thyroid to RT	Dose of anthracyclines	Dose of alkylating agent	Use of platinum agent
Univariable	X											
Model adjusted for RT (Yes/No) and CT (Yes/No)	X	X	X	X	X	X						
Model adjusted for cumulative doses for RT (Yes/No) and CT (mg/m <sup>2</sup> )	X	X	X	X	X					X	X	X
Model adjusted for cumulative doses for RT (Gy) and CT (Yes/No)	X	X	X	X		X	X	X	X			

Model adjusted for											
cumulative doses for RT (Gy) and CT (mg/m <sup>2</sup> )	X	X	X	X		X	X	X	X	X	X

#### 4.2.4 . SMN Type

This work is motivated by the cardio-toxicity of SMN treatments. Therefore, it is pertinent to look into the details of SMN treatments and their association with CD risk. However, we did not have access to this data. Nonetheless, we did have access to SMN types, which are strongly correlated with the treatments used and their cardio-toxicity. With the help of a clinician we defined three categories of SMN based on their associated cardio-toxicity: “Breast cancer among women” (radiotherapy at the heart), “Sarcoma, bone, and soft tissue cancer” (anthracyclines) and “Others or unknown” (no specific cardio-toxicity). We used those categories to perform an additional analysis in which SMN is a 4-level factor covariate.

#### 4.2.5 . Time Dependent Effect of SMN

We showed in Section 2.9.2 that death and CD instantaneous risk is high after a SMN, and plummets down shortly after. With the current method, we would therefore expect a lower estimate at later time, due to patients labelled “with SMN” being a mix of recently diagnosed and long term SMN survivors, who already survived the peak of CD and death risk. We decided to do an additional analysis splitting “with SMN” patients in two categories, “with SMN diagnosed less than 5 years ago” and “with SMN diagnosed more than 5 years ago”. The threshold of 5 years was used by looking at the transition intensities shown in Section 2.9.2, the amount of patients in each category for each landmark time point, and because 5 years is a commonly used threshold in cancer survivorship studies.

## **4.3 . Results**

### **4.3.1 . Population at Landmark Time**

A summary of our population can be found in Table 2.1. Table 4.2 and Table 4.3 show the differences across the principal landmark time points. We see that the baseline characteristics are stables across landmark times. Most importantly, we have between 125 and 266 SMN and between 9 and 21 CDs after a SMN at all landmark times. We also observe a shift in time since SMN as landmark time increases, with the mean ranging from 5.04 to 12.13 years.



**Table 4.2:** Description of the population at landmark times on the patient's age scale.

covariate	level	20	25	30	35
n		6420	6272	5382	4336
SMN		150	200	222	266
CD	without SMN	250	228	183	123
	after SMN	10	17	20	21
Time from SMN to landmark		5.05 [1.88-6.99]	7.12 [3.04-9.93]	9.05 [3.06-13.94]	9.52 [3.28-15.05]
Type of SMN	No SMN	5734	5582	4766	3799
	Breast: women	72	94	91	84
	Sarcoma, bone, soft tissue	78	66	53	30
	Others or unknown	536	530	472	423
	<1 year old	1134	872	659	489
Age at cancer	1-4 years old	2224	1897	1491	1150
	5-9 years old	1507	1443	1257	954
	10-14 years old	1555	1469	1405	1198
	15+ years old	0	591	570	545

Sex	Men	3523	3465	2971	2402
	Women	2897	2807	2411	1934
Decade of diagnosis	1940s	10	9	9	8
	1950s	124	126	124	122
	1960s	565	544	517	497
	1970s	1483	1469	1423	1354
	1980s	2060	2188	2105	1670
	1990s	2178	1936	1204	685

**Table 4.3:** Description of the population at landmark times on the time since diagnosis scale.

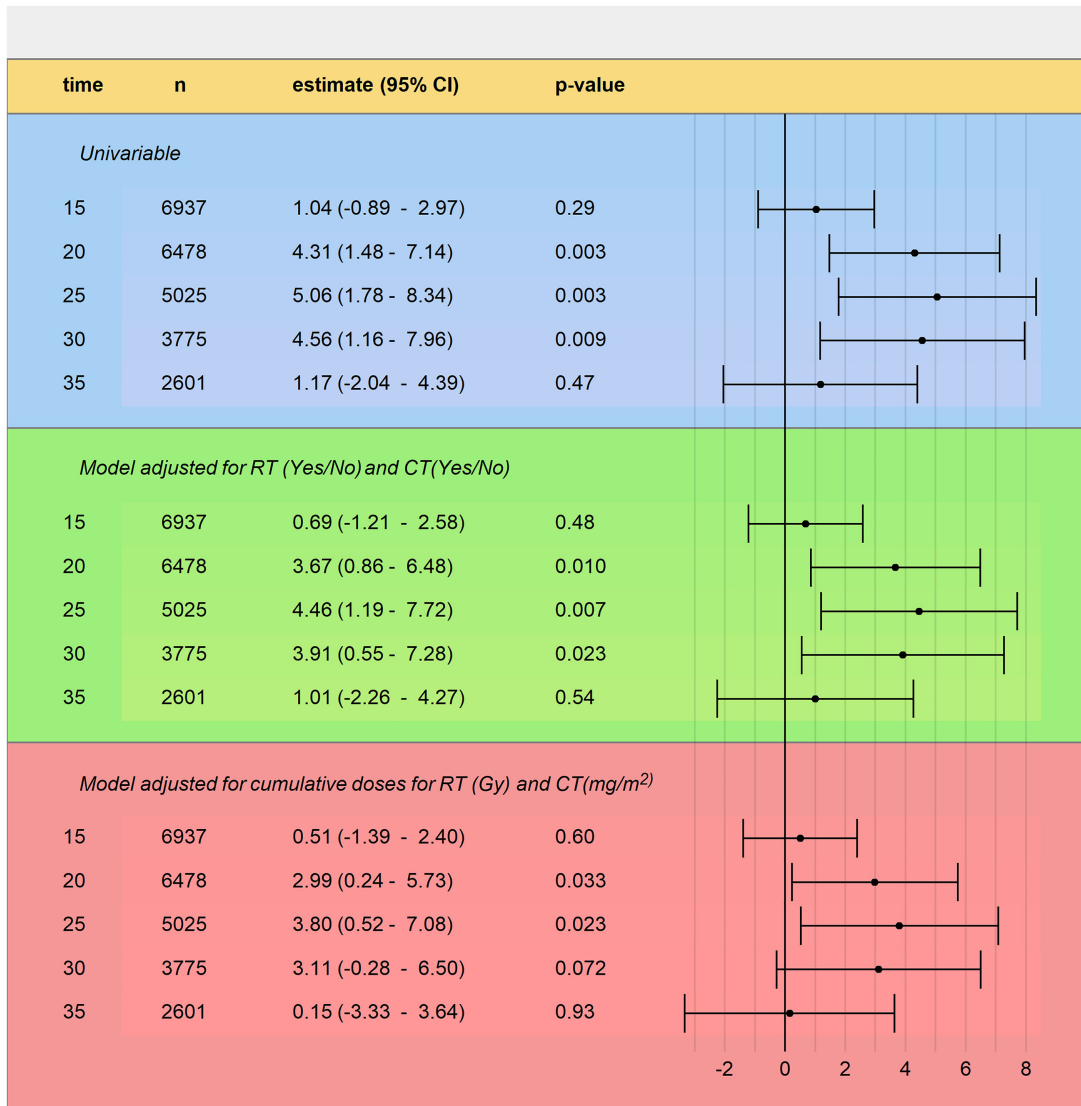
covariate	level	15	20	25	30	35
n		4929	5956	5023	3776	2602
SMN		125	231	260	262	225
CD	without SMN	178	189	147	95	55
	after SMN	9	19	21	20	9
Time from SMN to landmark		5.04 [1.34-7.87]	6.77 [2.40-10.81]	9.10 [3.92-13.04]	10.51 [4.04-15.75]	12.13 [5.19-18.73]
Type of SMN	No SMN	4438	5332	4458	3316	2255
	Breast: women	33	70	75	55	44
	Sarcoma, bone, soft tissue	67	63	48	31	18
	Others or unknown	391	491	442	374	285
	<1 year old	1183	1104	846	643	477
1-4 years old	2264	2119	1683	1312	959	
5-9 years old	1482	1394	1121	844	583	
10-14 years old	0	1339	1034	771	493	

## Age at cancer

	15+ years old	0	0	339	206	90
Sex	Men	2727	3271	2792	2091	1434
	Women	2202	2685	2231	1685	1168
Decade of diagnosis	1940s	6	9	9	9	6
	1950s	102	121	123	122	122
	1960s	456	531	514	489	450
	1970s	1162	1401	1391	1326	1246
	1980s	1557	1983	2085	1827	778
	1990s	1646	1911	901	3	0

### 4.3.2 . Cumulative Incidence

#### 4.3.2.1 Effect of any SMN occurrence



**Figure 4.1:** Effect of a SMN on the cumulative incidence of CD.

The cumulative incidence of cardiac disease was higher when a SMN occurred for landmark times 20, 25, and 30 years after diagnosis, as well as 30 and 35 years of attained age. The maximum estimated excess was 3.8% (95% CI: 0.5% - 7.1%) for the landmark time of 25 years after diagnosis Figure 4.1. Note that, here, estimates correspond to whether and how much the

probability of experiencing cardiac disease changes between SMN survivors and baseline CCS. Figure 4.1 summarizes the estimated excess of cumulative incidence of cardiac disease after a SMN for each configuration. As detailed in Table 4.3, the 3.8% excess for landmark time 25 years after diagnosis translated into approximately 10 cardiac disease cases (260 patients with a SMN at this time,  $260 * 3.8 / 100 = 9.8$ ) out of the 30 expected after a SMN (260 patients with a SMN at this time, 11.5% cumulative incidence of cardiac disease,  $260 * 0.115 = 29.8$ ) in this category.

For the landmark times 15 and 35 years after diagnosis, as well as 20 and 25 years of attained age, no significant excess of cardiac disease incidence due to a SMN was observed. In the univariable analysis for patients who survived 25 years after diagnosis Figure 4.1, those experiencing a SMN had a cumulative incidence 5.1% (95% CI: 1.8% - 8.3%) higher than those who did not have a SMN. When adjusting for age at first diagnosis, radiotherapy (yes/no), and chemotherapy (yes/no), the cumulative incidence of cardiac disease was 4.5% (95% CI: 1.2% - 7.7%) higher when a SMN occurred in the first 25 years after childhood cancer diagnosis. When including all treatment information (age at diagnosis, sex, year of diagnosis, average radiotherapy dose at the heart and brain, exposure of the neck to radiotherapy, cumulative anthracycline dose, use of alkylating agent), the cumulative incidence of cardiac disease was 3.8% (95% CI: 0.5% - 7.1%) higher when a SMN occurred in the first 25 years after childhood cancer.

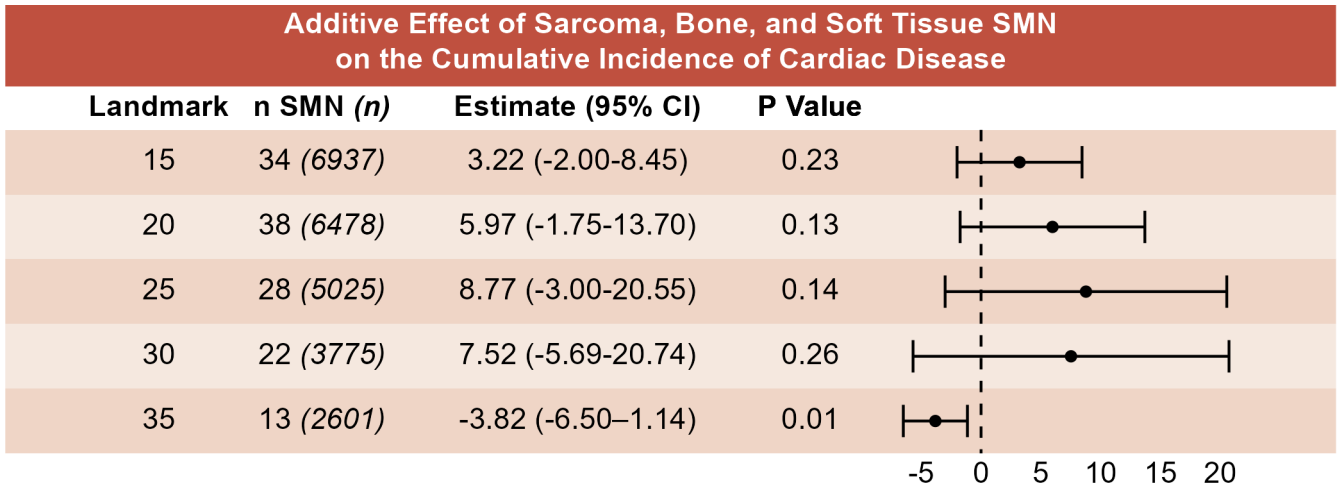
Results on the attained age time scale were similar. In the univariable analysis for patients who survived  $\geq 35$  years of age, those experiencing a SMN had a cumulative incidence 4.4% (95% CI: 1.1% - 7.7%) higher than those who did not have a SMN. When adjusting for age at first diagnosis, radiotherapy (yes/no), and chemotherapy (yes/no), the cumulative incidence of cardiac disease was 4.0% (95% CI: 0.8% - 7.3%) higher when a SMN occurred before 35 years of age. When including all treatment information (age at diagnosis, sex, year of diagnosis, average radiotherapy dose at the heart and brain, exposure of the neck to radiotherapy, cumulative anthracycline dose, use of alkylating agent), the cumulative incidence of cardiac disease was

3.3% (95% CI: 0.0% - 6.5%) higher when a SMN occurred before 35 years of age.

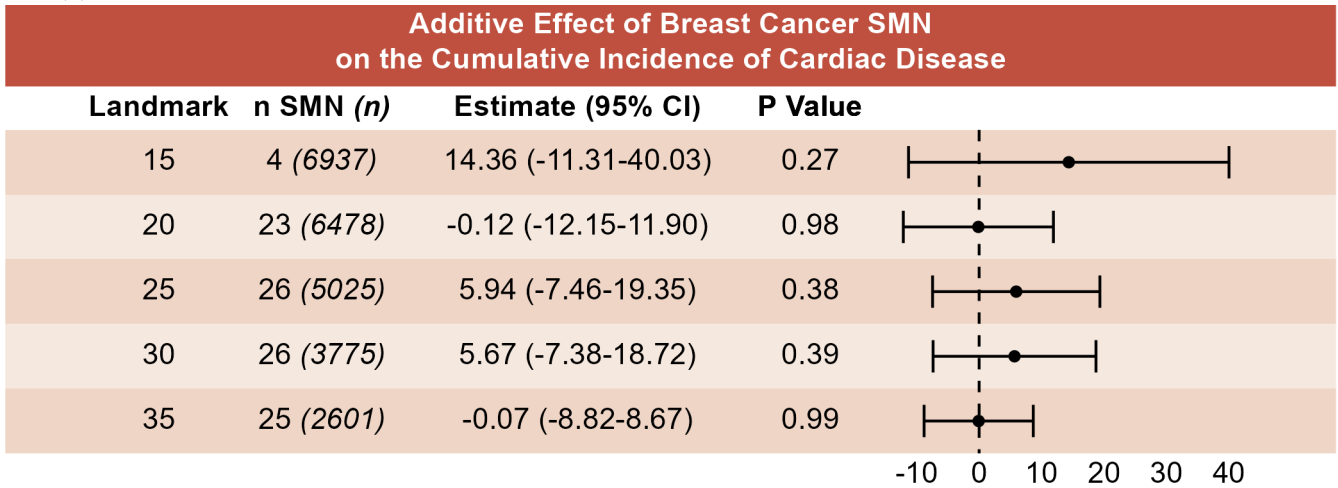
#### 4.3.2.2 Effect by Type of SMN

When categorizing the SMN status into 4 types, we did not observe any statistically significant effect, due to high standard errors. However, we did observe tendencies in the pointwise estimates (Figure 4.2, Figure 4.3) among the cancer types. When adjusting for all treatment information, we observed that breast cancer (10.9%; 95% CI: -12.6% to 34.5%); sarcoma, bone, or soft tissue cancer (7.7%; 95% CI: -1.2% to 16.6%); and other SMN types (2.0%; 95% CI: -1.0% to 5.0%) showed an increase in cardiac disease cumulative incidence when a SMN occurred before the patient was 30 years of age, though all weren't statistically significant.

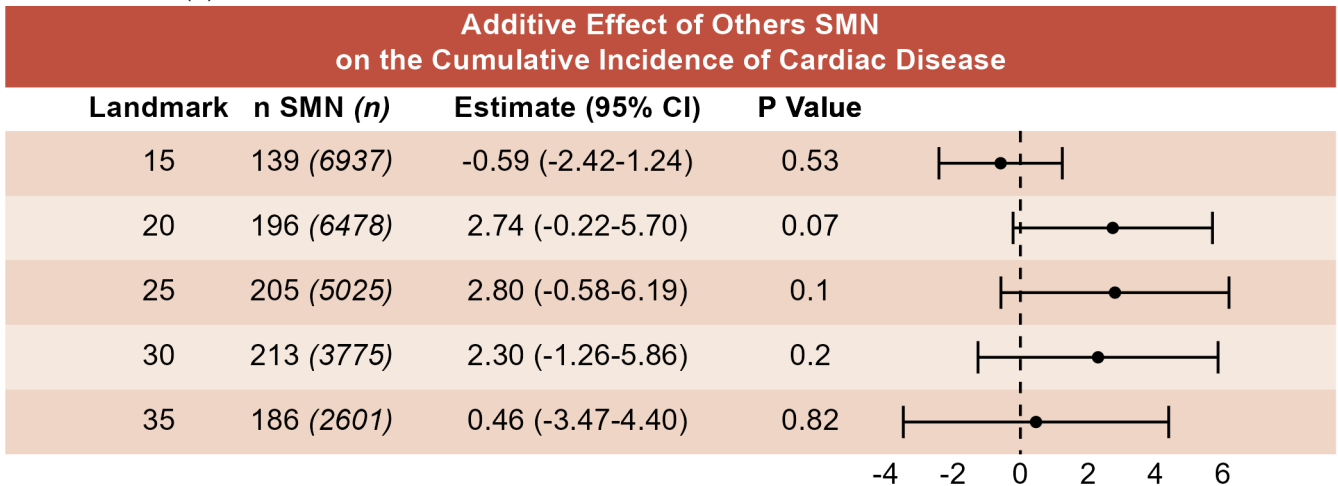
When adjusting for all treatment information, we found the effect of Sarcoma, bone, and soft tissue cancers on the cumulative incidence of death to be higher than for Other or unknown cancers at all landmark times (Figure 4.4, Figure 4.5). The effect of sarcoma, bone, and soft tissue cancer on the cumulative incidence of death range from 30.5% (95% CI: 16.8% - 44.2%) at landmark time 20 years old to 7.8% (95% CI: -12.0% - 27.6%) at landmark time 35 years since diagnosis. For breast cancer, the same effect range from 44.4% (95% CI: 17.3% - 71.5%) at landmark time 32 years old to -5.0% (95% CI: -8.6% - -1.4%) at landmark time 22 years old. For others and unknown cancer, the effect ranges from 13.9% (95% CI: 7.2% - 20.6%) at landmark time 21 years old to 5.5% (95% CI: 0% - 12.8%) at landmark time 35 years since diagnosis.



(a) Additive effect of Sarcoma, Bone, and Soft Tissue Cancer SMN on the cumulative incidence of Cardiac Disease.



(b) Additive effect of Breast Cancer SMN on the cumulative incidence of Cardiac Disease.

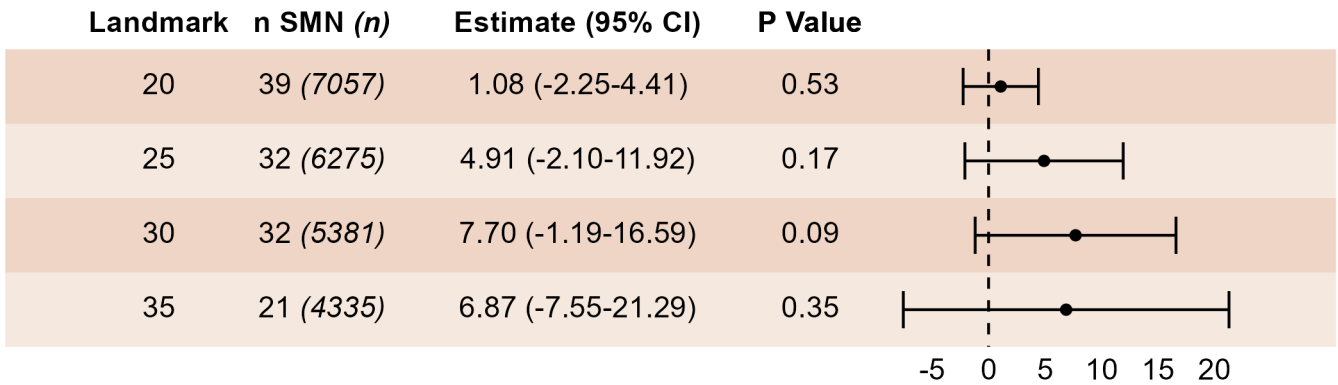


(c) Additive effect of Others SMN on the cumulative incidence of Cardiac Disease.

**Figure 4.2:** Additive effect of each type of SMN on the cumulative incidence of Cardiac Disease. All landmark times are years since childhood cancer diagnosis. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).

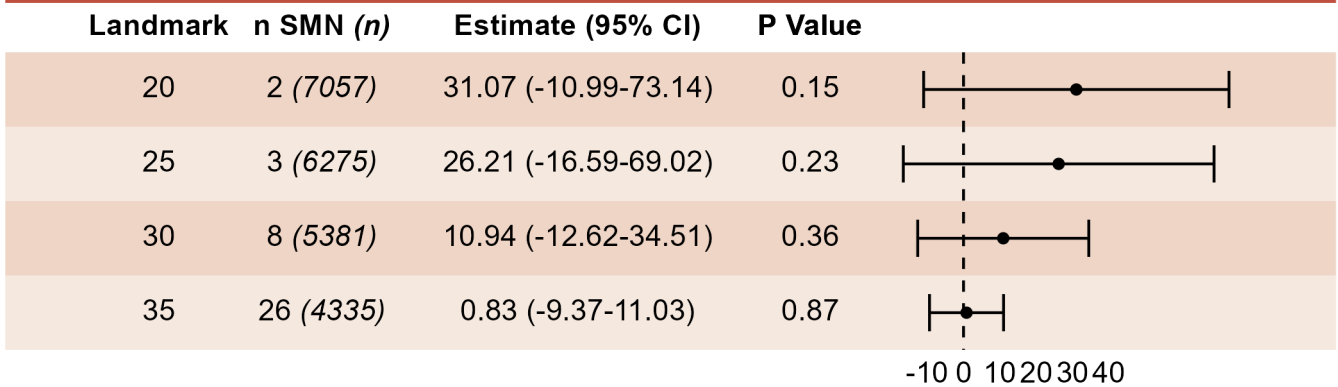


**Additive Effect of Sarcoma, Bone, and Soft Tissue SMN  
on the Cumulative Incidence of Cardiac Disease**



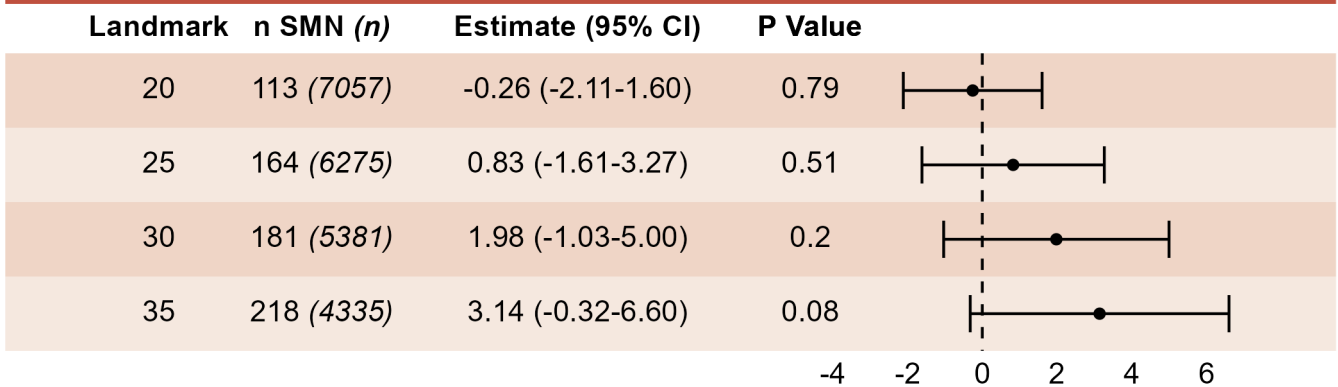
(a) Additive effect of Sarcoma, Bone, and Soft Tissue Cancer SMN on the cumulative incidence of Cardiac Disease.

**Additive Effect of Breast Cancer SMN  
on the Cumulative Incidence of Cardiac Disease**



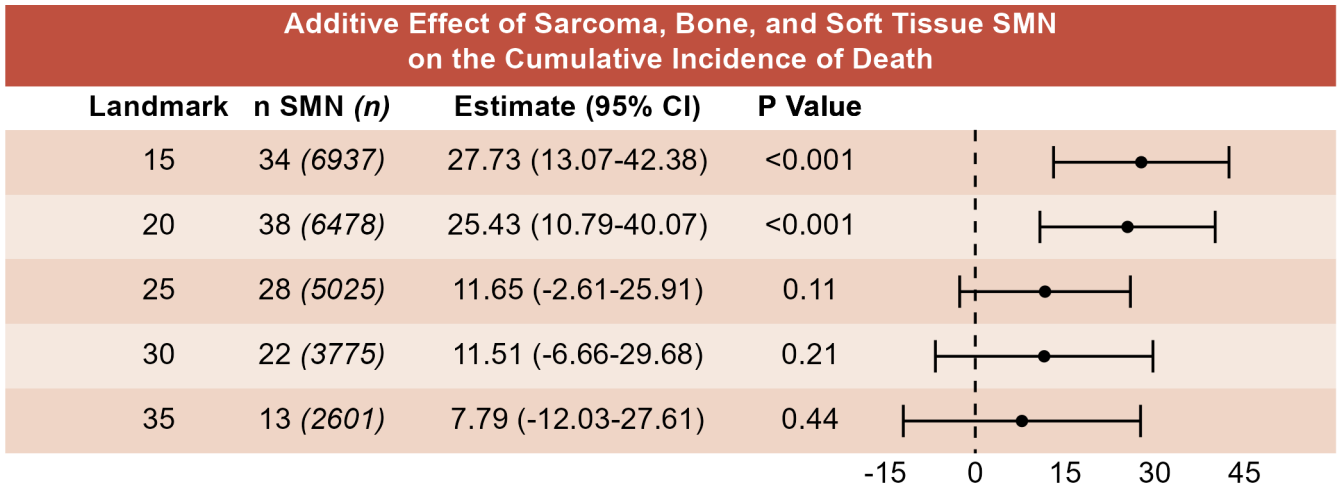
(b) Additive effect of Breast Cancer SMN on the cumulative incidence of Cardiac Disease.

**Additive Effect of Others SMN  
on the Cumulative Incidence of Cardiac Disease**

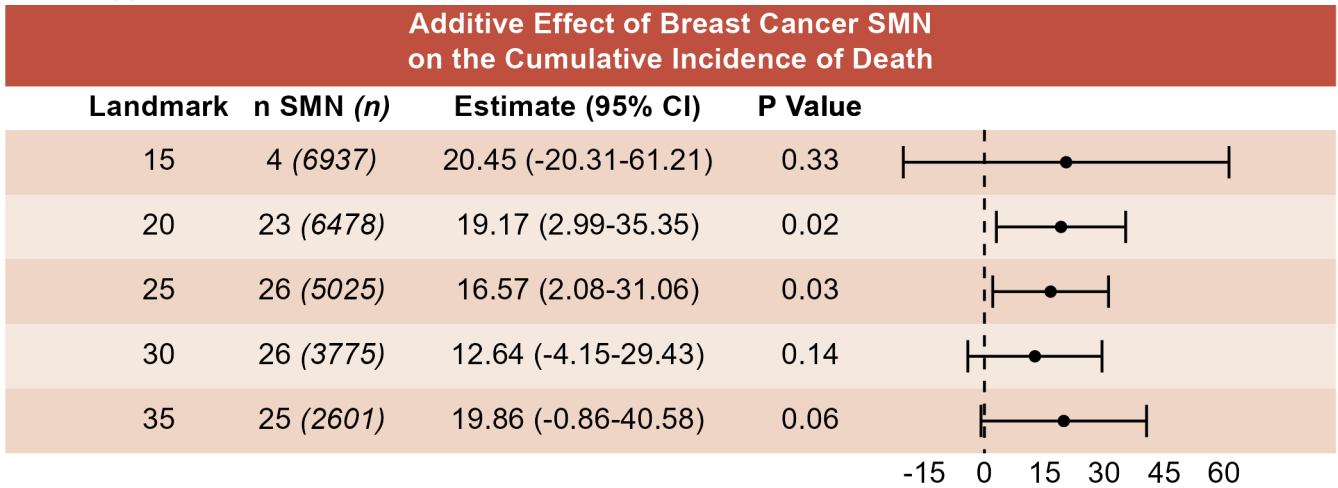


(c) Additive effect of Others SMN on the cumulative incidence of Cardiac Disease.

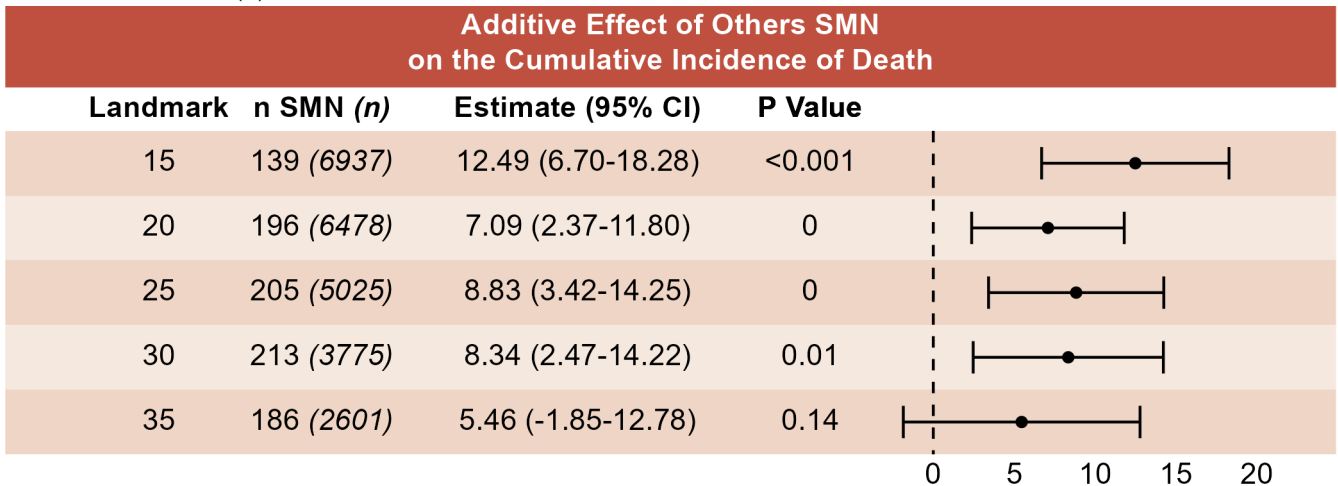
**Figure 4.3:** Additive effect of each type of SMN on the cumulative incidence of Cardiac Disease. All landmark times are patients' age in years. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).



(a) Additive effect of Sarcoma, Bone, and Soft Tissue Cancer SMN on the cumulative incidence of death.



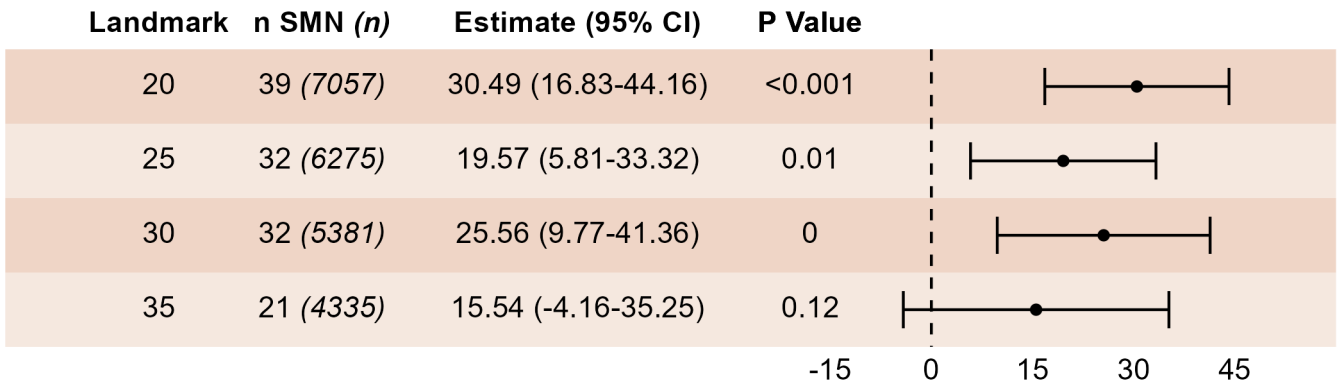
(b) Additive effect of Breast Cancer SMN on the cumulative incidence of death.



(c) Additive effect of Others SMN on the cumulative incidence of death.

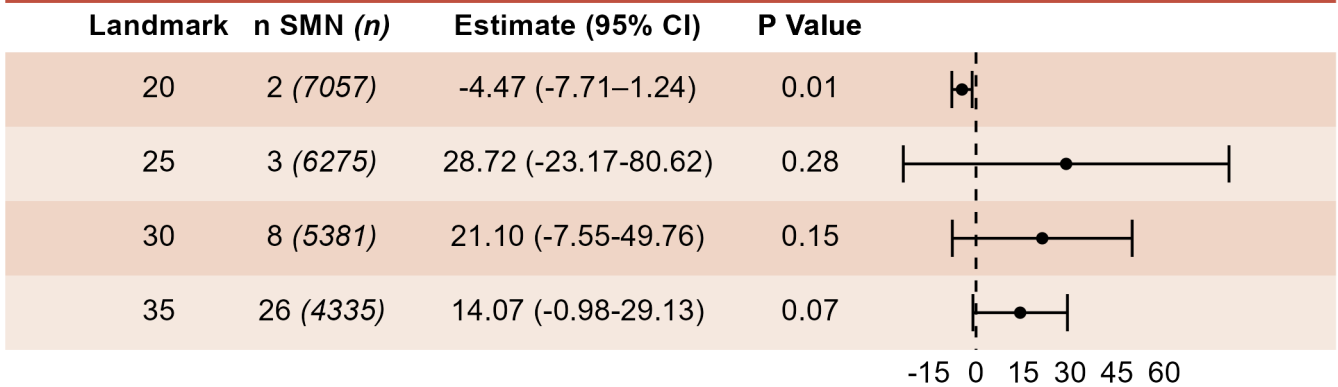
**Figure 4.4:** Additive effect of each type of SMN on the cumulative incidence of death. All landmark times are years since childhood cancer diagnosis. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).

**Additive Effect of Sarcoma, Bone, and Soft Tissue SMN  
on the Cumulative Incidence of Death**



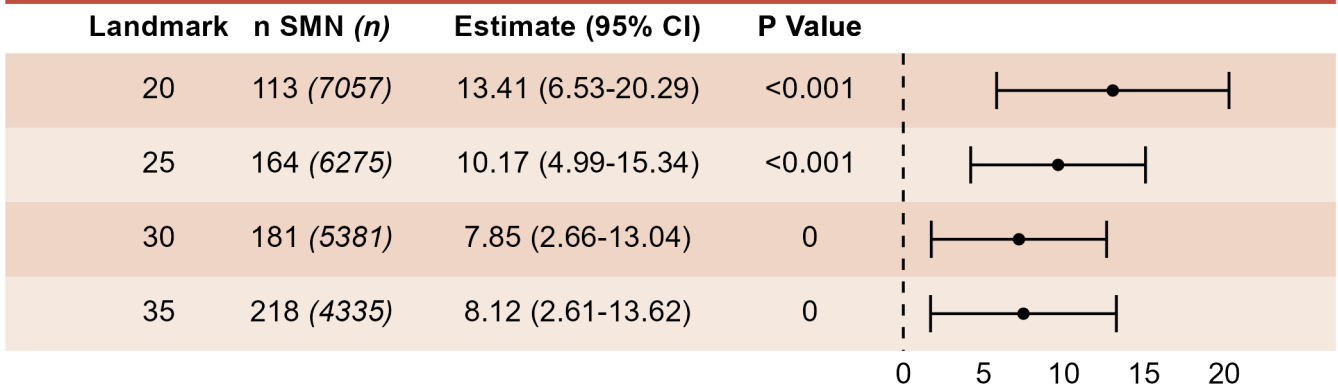
(a) Additive effect of Sarcoma, Bone, and Soft Tissue Cancer SMN on the cumulative incidence of death.

**Additive Effect of Breast Cancer SMN  
on the Cumulative Incidence of Death**



(b) Additive effect of Breast Cancer SMN on the cumulative incidence of death.

**Additive Effect of Others SMN  
on the Cumulative Incidence of Death**



(c) Additive effect of Others SMN on the cumulative incidence of death.

**Figure 4.5:** Additive effect of each type of SMN on the cumulative incidence of death. All landmark times are patients' age in years. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).

### **4.3.2.3 Effect of SMN within 5 Years and Later**

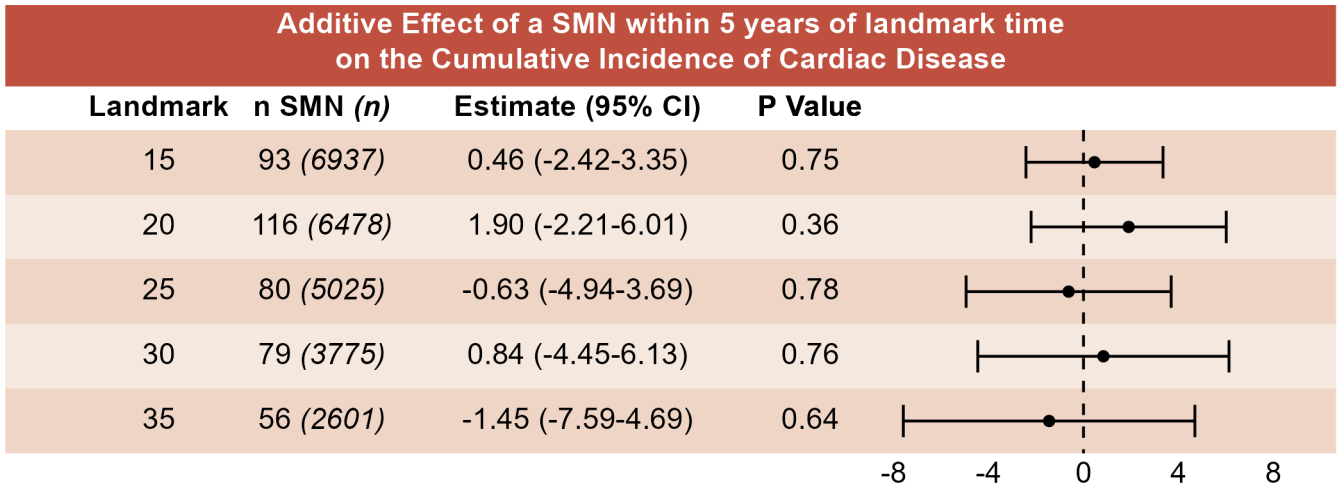
When distinguishing SMN based on its occurrence less than five years before landmark time, we observe a higher effect on the cumulative incidence of SMN diagnosed more than 5 years ago compared to the scenario where time since SMN diagnosis isn't considered (Figure 4.6, Figure 4.7). When adjusting for all treatment information we find that a SMN diagnosed more than 5 years ago increases the cumulative incidence of CD by 5.74% (95 %CI : 1.47 - 10.01) at landmark time 25 years after diagnosis, compared to the 3.80% (0.52 - 7.08) reported above. However, we don't observe any effect of a SMN diagnosed less than 5 years before landmark time on the cumulative incidence of cardiac disease.

When looking at the cumulative incidence of death, we find on the contrary that a SMN diagnosed less than 5 years before landmark time has a stronger impact on death than one diagnosed more than 5 years ago (Figure 4.8, Figure 4.9). For instance, at landmark time 25 years after diagnosis, and after adjusting for all treatment information, we find an increase of 13.92% (95% CI: 4.73 - 23.11) for a recent SMN, and of 8.14% (95% CI: 2.62 - 13.66) for a SMN more than 5 years before landmark time.

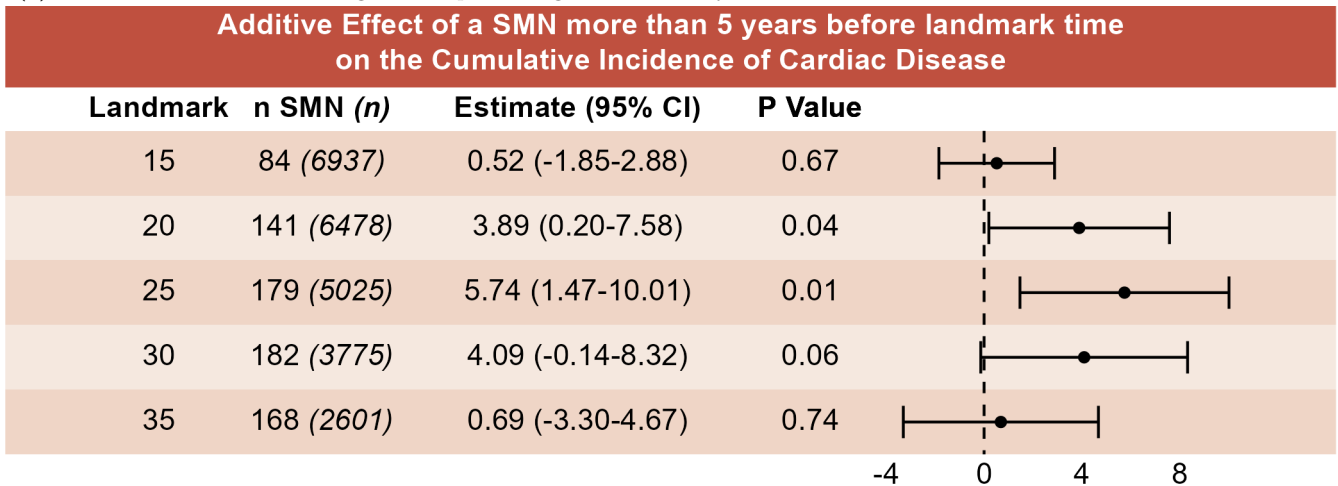
### **4.3.3 . Cox Model**

#### **4.3.3.1 Effect of any SMN occurrence**

For all combinations of treatment-related risk factors, the occurrence of a SMN had a deleterious effect on the csH of cardiac disease (Figure 4.10). In univariable analysis, the occurrence of a SMN resulted in a 3-fold increase in the hazard of cardiac disease (csHR: 2.7; 95% CI: 2.0 - 3.7). When adjusting for age at diagnosis, radiotherapy (yes/no), and chemotherapy (yes/no), the occurrence of a SMN caused a 2-fold increase in the hazard of cardiac disease (csHR: 2.3; 95% CI: 1.6 - 3.1). When including all treatment information (age at diagnosis, sex, year of diagnosis, average radiotherapy dose at the heart and brain, exposure of the neck to

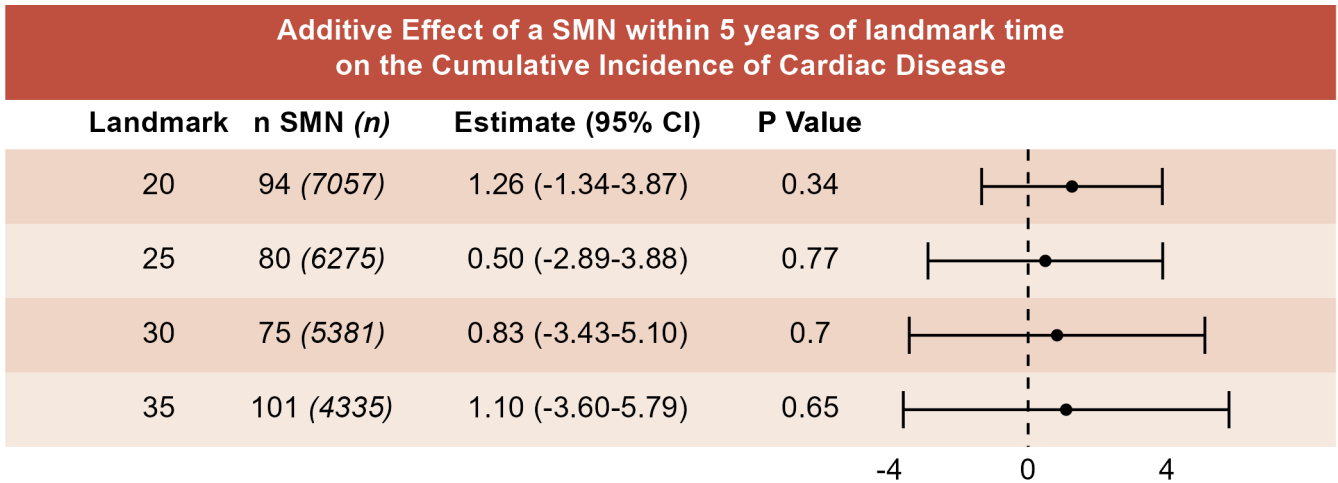


(a) Additive effect of Second Malignant Neoplasm diagnosed within 5 years on the cumulative incidence of Cardiac Disease.

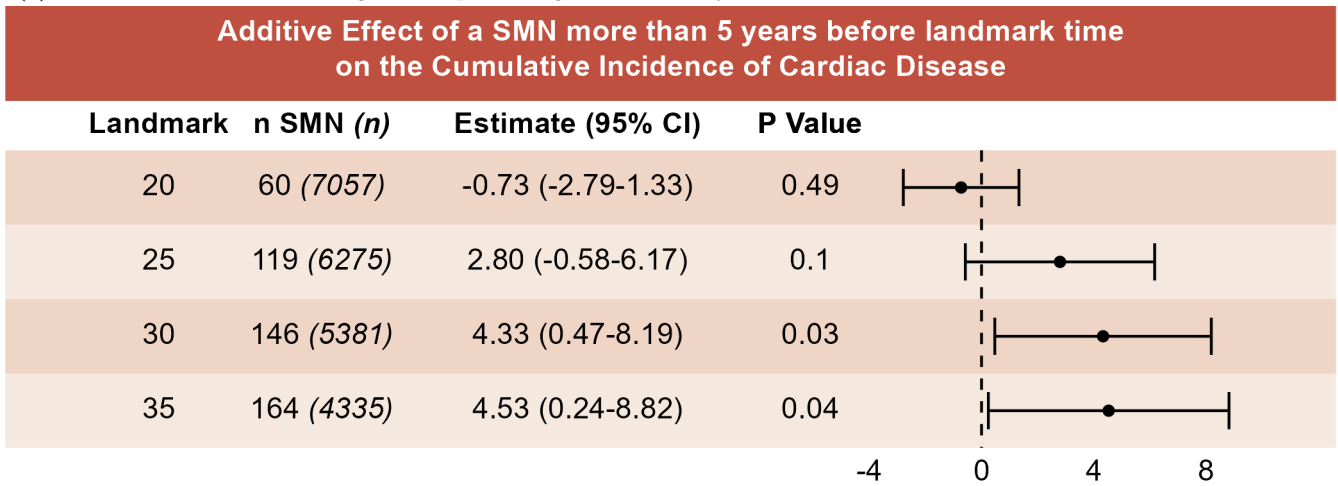


(b) Additive effect of Malignant Neoplasm diagnosed more than 5 years ago on the cumulative incidence of Cardiac Disease.

**Figure 4.6:** Additive effect of Second Malignant Neoplasm based on time since its diagnosis on the cumulative incidence of Cardiac Disease. All landmark times are years since childhood cancer diagnosis. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).

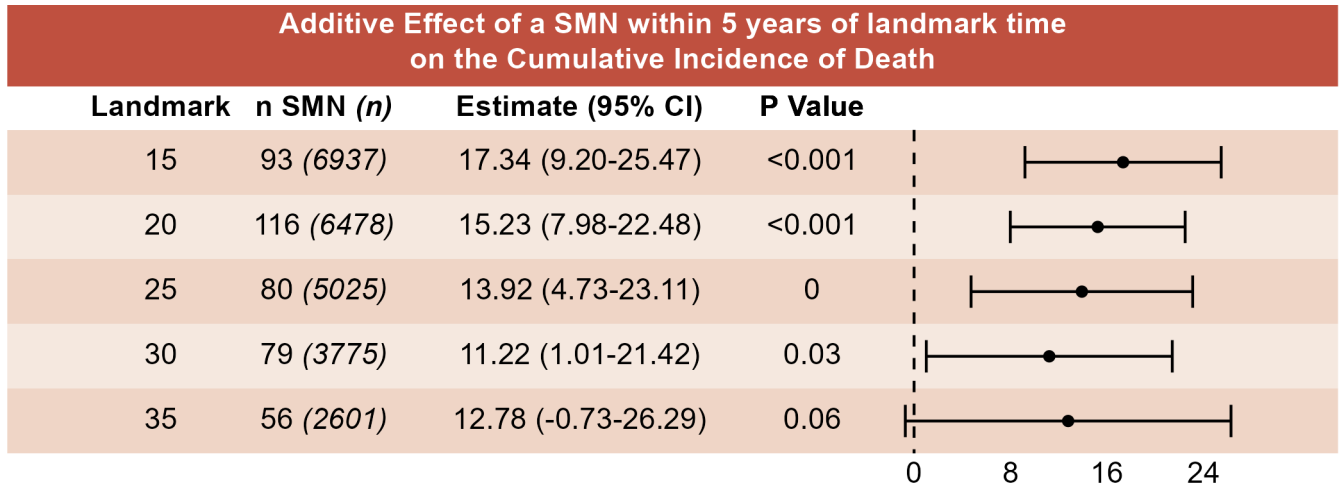


(a) Additive effect of Second Malignant Neoplasm diagnosed within 5 years on the cumulative incidence of Cardiac Disease.

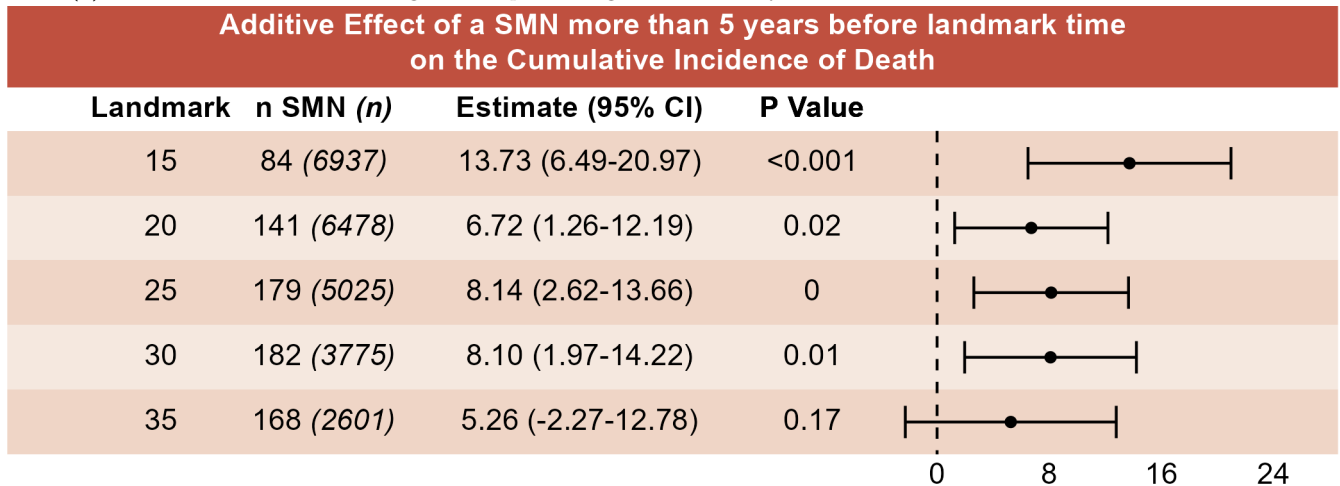


(b) Additive effect of Malignant Neoplasm diagnosed more than 5 years ago on the cumulative incidence of Cardiac Disease.

**Figure 4.7:** Additive effect of Second Malignant Neoplasm based on time since its diagnosis on the cumulative incidence of Cardiac Disease. All landmark times are patients' age in years. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).

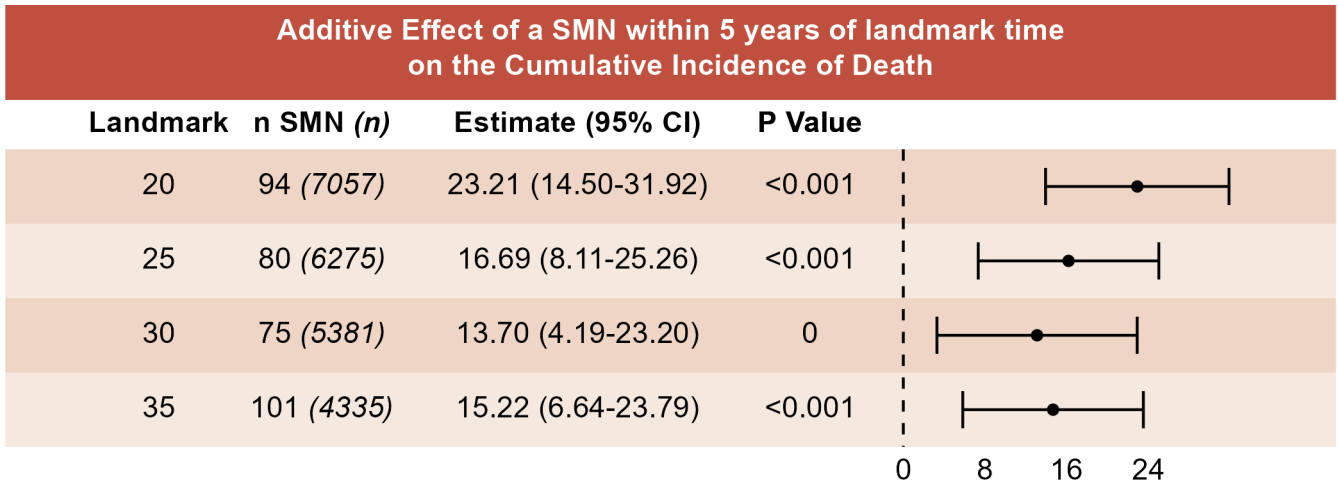


(a) Additive effect of Second Malignant Neoplasm diagnosed within 5 years on the cumulative incidence of death.

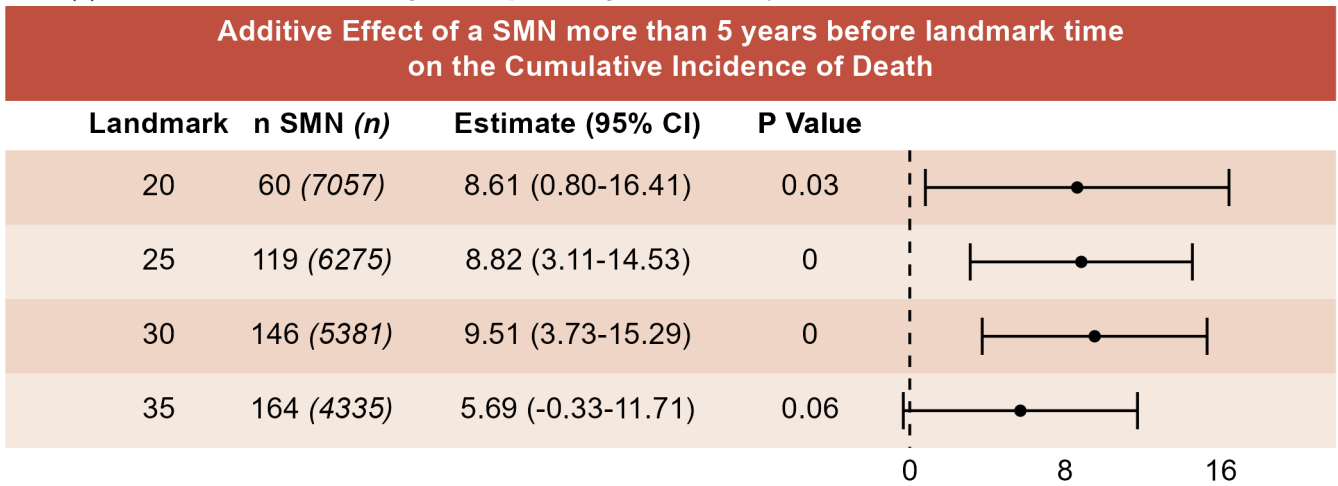


(b) Additive effect of Malignant Neoplasm diagnosed more than 5 years ago on the cumulative incidence of death.

**Figure 4.8:** Additive effect of Second Malignant Neoplasm based on time since its diagnosis on the cumulative incidence of death. All landmark times are years since childhood cancer diagnosis. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).



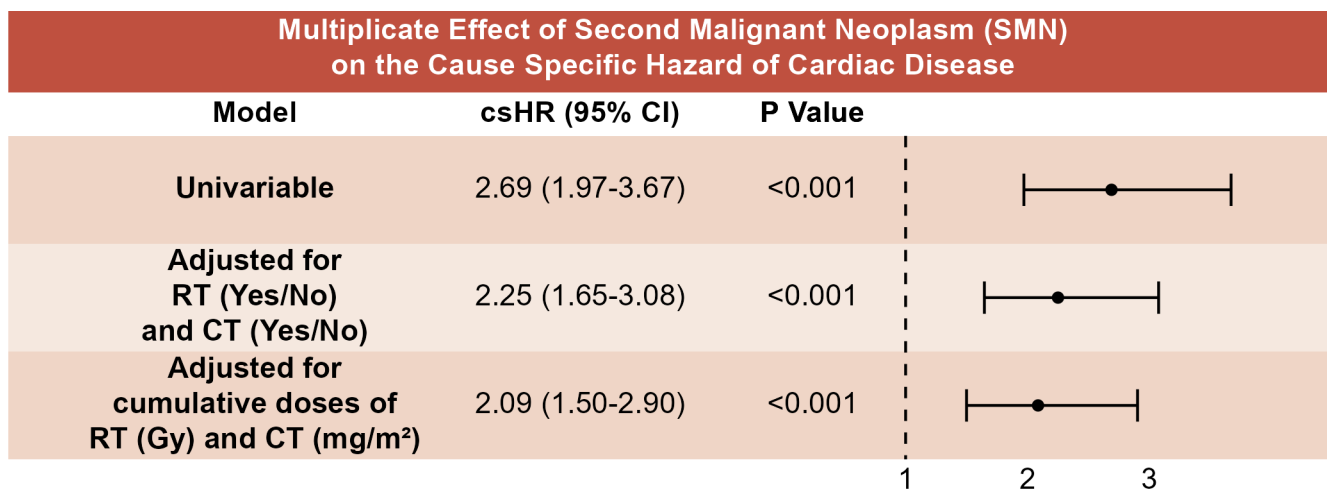
(a) Additive effect of Second Malignant Neoplasm diagnosed within 5 years on the cumulative incidence of death.



(b) Additive effect of Malignant Neoplasm diagnosed more than 5 years ago on the cumulative incidence of death.

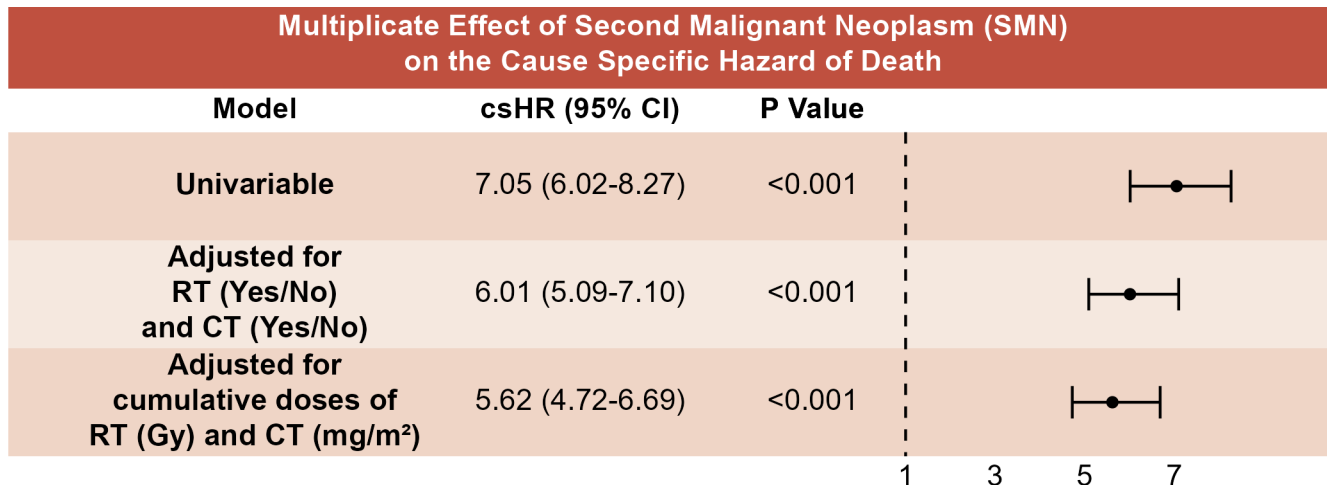
**Figure 4.9:** Additive effect of Second Malignant Neoplasm based on time since its diagnosis on the cumulative incidence of death. All landmark times are patients' age in years. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).





**Figure 4.10:** Multiplicative effect of SMN on the cause specific hazard of CD. The errorbar is the 95% confidence interval.

radiotherapy, cumulative anthracycline dose, use of alkylating agent, use of platinum agent), the occurrence of a SMN increased the likelihood of cardiac disease (csHR: 2.1; 95% CI: 1.5 - 2.9).



**Figure 4.11:** Multiplicative effect of SMN on the cause specific hazard of death. The errorbar is the 95% confidence interval.

#### **4.3.3.2 Effect by Type of SMN**

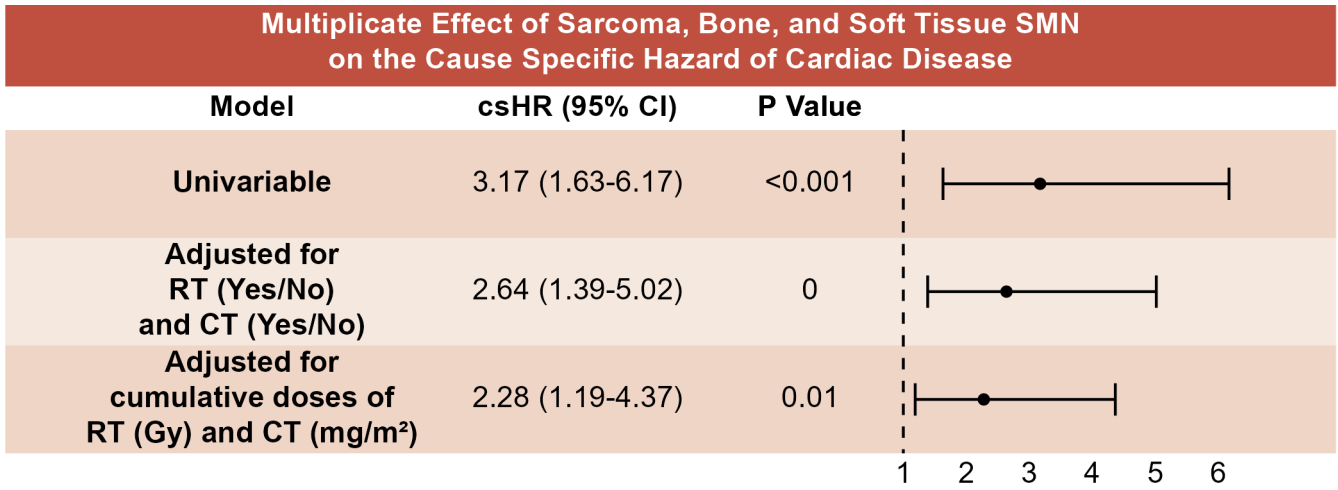
When categorizing the SMN status into 4 types, we observed some differences (Figure 4.12). In univariable analysis, the occurrence of breast cancer (csHR: 3.5; 95% CI: 2.1 - 5.8); sarcoma, bone, or soft tissue cancer (csHR: 3.2; 95% CI: 1.6 - 6.2); and other SMN (csHR: 1.9; 95% CI: 1.5 - 2.6) was associated with an increased risk of cardiac disease. When adjusting for all treatment information (same as previously), the occurrence of breast cancer (csHR: 1.7; 95% CI: 1.0 - 3.0); sarcoma, bone, or soft tissue cancer (csHR: 2.3; 95% CI: 1.2 - 4.4); and other SMN (csHR: 1.4; 95% CI: 1.1 - 1.9) also was associated with an increased hazard of cardiac disease.

We observed stark differences on the csHR of death between each SMN types (Figure 4.13). When adjusting for all treatment information (same as previously), the occurrence of sarcoma, bone, or soft tissue cancer resulted in a 6-fold increase in the hazard of death (csHR: 5.8; 95% CI: 4.4 - 7.7), while breast cancer didn't increase it (csHR: 1.2; 95% CI: 0.8 - 1.7), and other SMN resulted in a 1.6 fold increase (csHR: 1.6; 95% CI: 1.3 - 1.8).

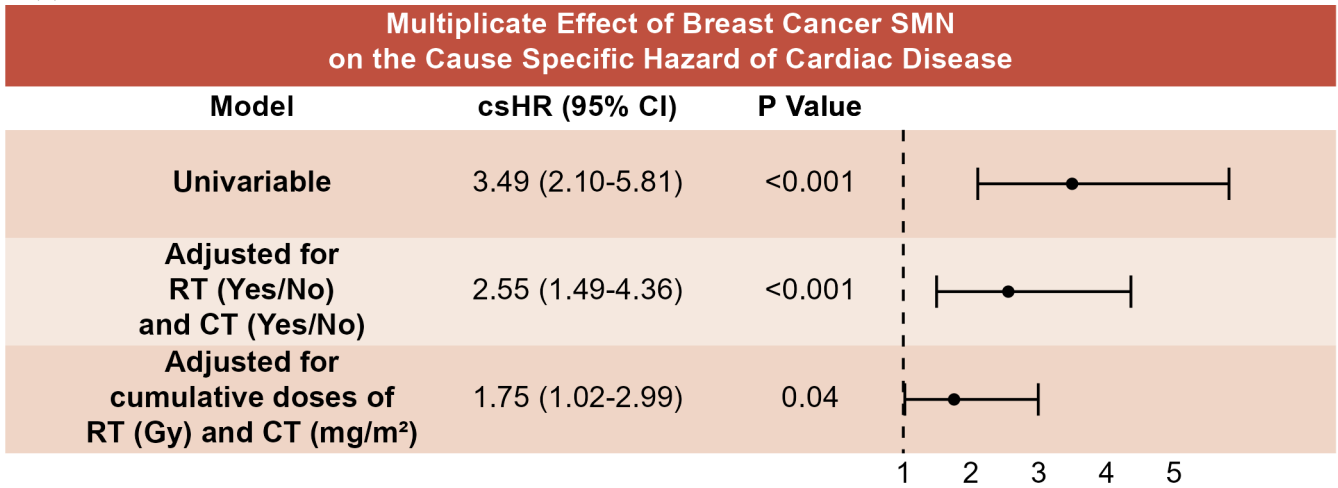
#### **4.3.3.3 Effect of SMN within 5 Years and Later**

When distinguishing SMN based on its occurrence less than five years before landmark time, we did not observe any statistically difference on the CSH of CD (Figure 4.14). When adjusting for all treatment information (same as previous), the occurrence of SMN within 5 years resulted in a 2.5-fold increase of the CSH of CD (csHR: 2.5; 95% CI: 1.5 - 4.1), and its occurrence more than 5 years ago resulted in a 1.9-fold increase (csHR: 1.9; 95% CI: 1.3 - 2.8).

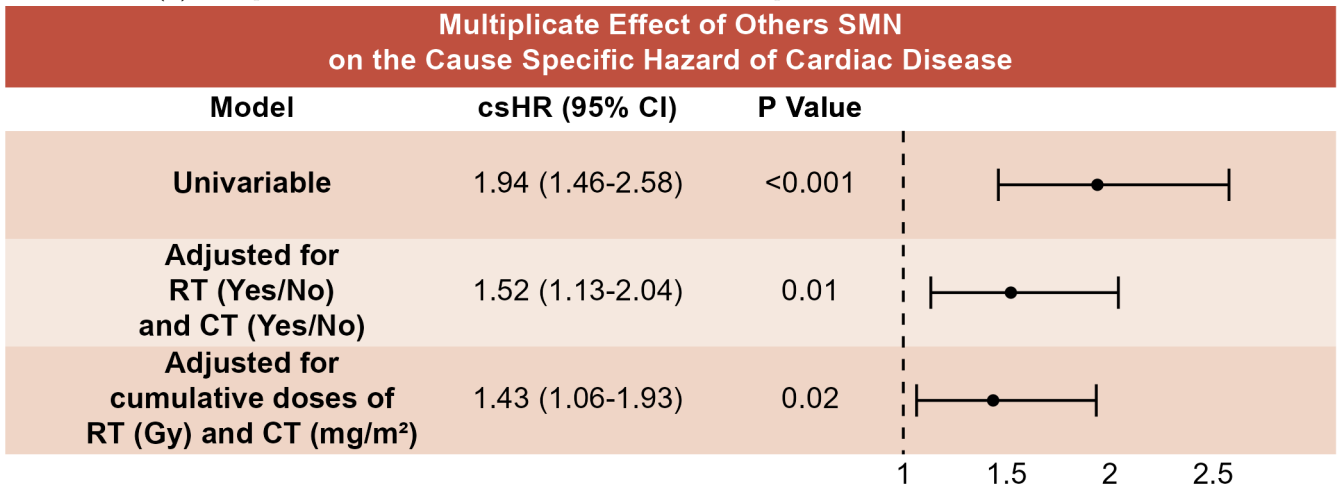
We did observe a difference on the CSH of death, with a SMN within 5 years having a stronger effect than a SMN more than 5 years ago (Figure 4.15). When adjusting for all treatment information, the occurrence of SMN within 5 years resulted in a 10-fold increase of the CSH of death (csHR: 10.5; 95% CI: 8.8 - 12.5), and its occurrence more than 5 years ago resulted in a



(a) Multiplicative effect of Sarcoma, Bone, and Soft Tissue Cancer SMN on the cause specific hazard of Cardiac Disease.

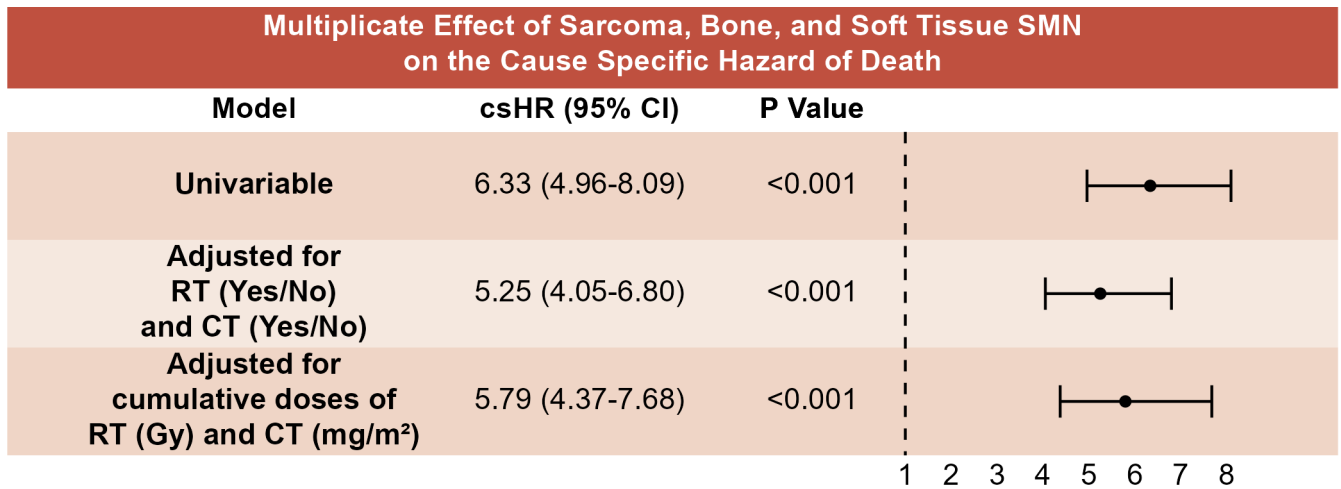


(b) Multiplicative effect of Breast Cancer SMN on the cause specific hazard of Cardiac Disease.

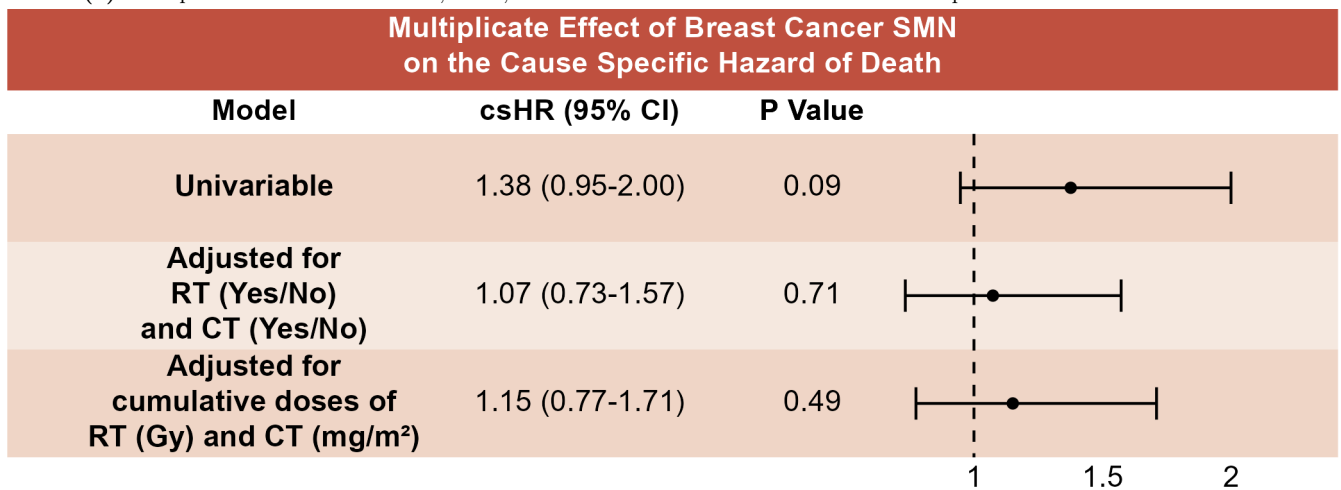


(c) Multiplicative effect of Others SMN on the cause specific hazard of Cardiac Disease.

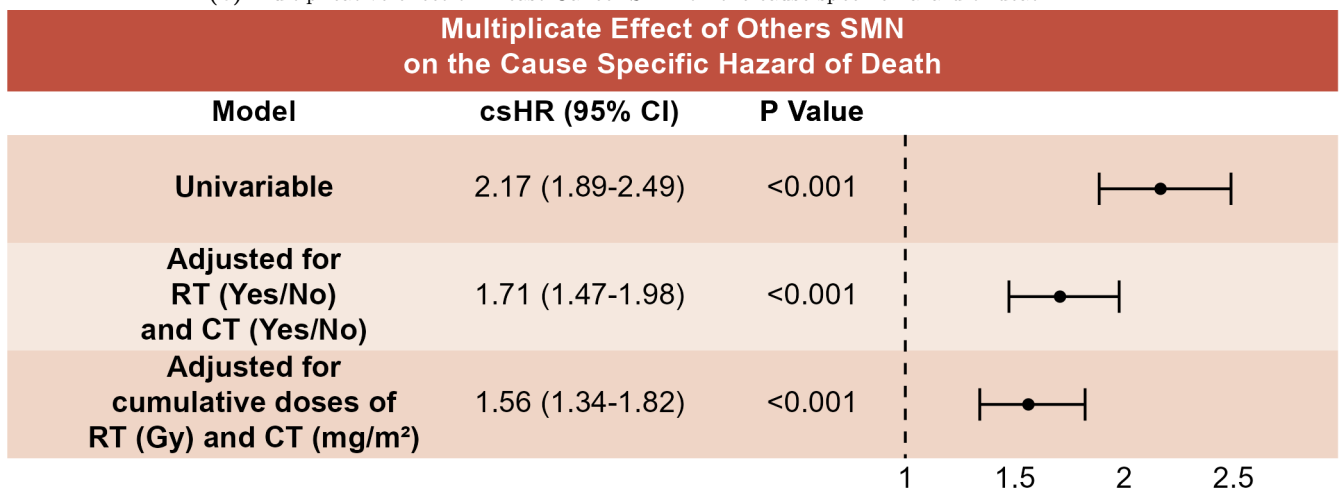
**Figure 4.12:** Multiplicative effect of each type of Second Malignant Neoplasm on the cause specific hazard of Cardiac Disease. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).



(a) Multiplicative effect of Sarcoma, Bone, and Soft Tissue Cancer SMN on the cause specific hazard of death.



(b) Multiplicative effect of Breast Cancer SMN on the cause specific hazard of death.



(c) Multiplicative effect of Others SMN on the cause specific hazard of death.

**Figure 4.13:** Multiplicative effect of each type of Second Malignant Neoplasm on the cause specific hazard of death. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).

3-fold increase (csHR: 2.9; 95% CI: 2.3 - 3.6).

#### **4.3.4 . Effects of Childhood Cancer Treatments**

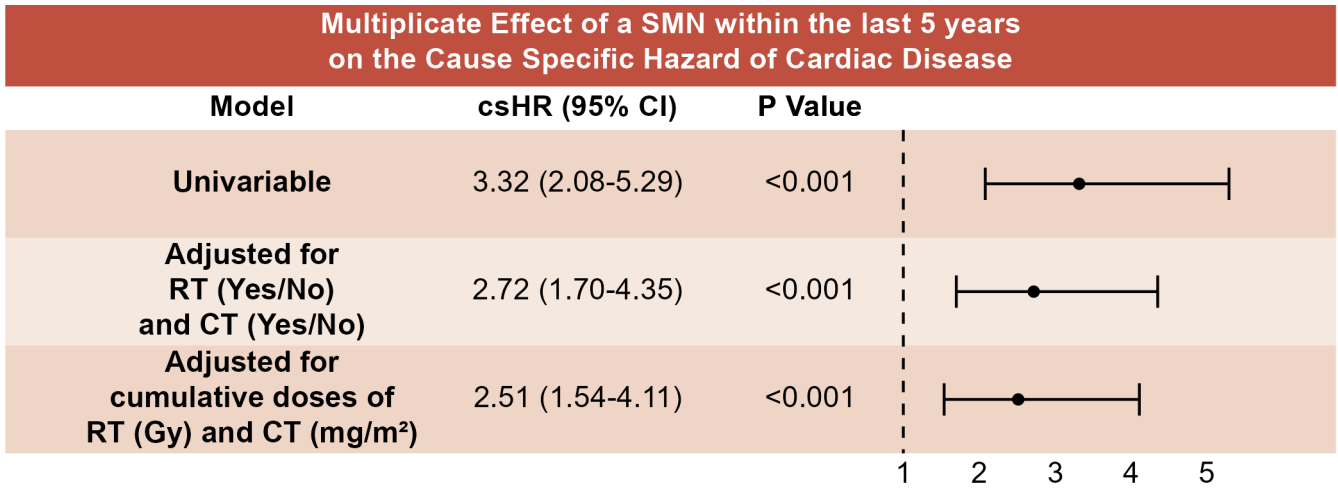
Although this study focused on estimating the effect of a SMN and did not aim to estimate treatment effects, we did obtain csHR estimates for specific covariates while estimating the effect of a SMN. Among the risk factors analyzed, the most important ones were a  $> 5$  Gy average heart dose,  $> 250$  mg/ m<sup>2</sup> cumulative dose of anthracycline, and the use of alkylating agents, which align with findings of previous studies. As previously determined, we did not find an increase of cardiac disease risk for  $< 100$  mg/ m<sup>2</sup> cumulative anthracycline dose.

### **4.4 . Discussion**

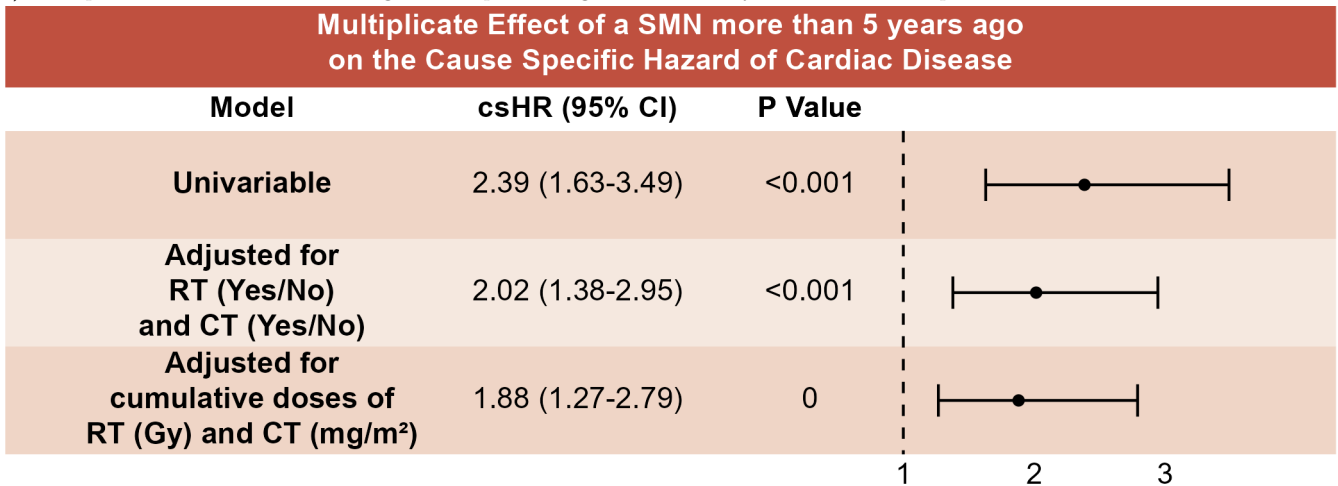
#### **4.4.1 . Increased Risk of Cardiac Disease After a Second Malignant Neoplasm**

We conducted this study on a large, well-defined population of childhood cancer survivors over an extended treatment period (1945-2000). Our data collection involved a comprehensive approach, including self-reported questionnaires, hospital-based databases/registries, and clinically assessed data from survivors participating in long-term clinical follow-ups. We observed a deleterious effect of SMNs on the cause-specific-hazard scale and the cumulative incidence of cardiac disease among childhood cancer survivors. We conclude that CCS who experienced a SMN had higher instantaneous and absolute risks of severe cardiac disease compared with those who did not experience a SMN. To the best of our knowledge, this is the first study to explore severe cardiac disease occurrence after experiencing a SMN among CCS.

This study was a first attempt to understand in depth the relationship between the iatrogenic events of CCS, specifically between SMN and severe cardiac disease, two of the most frequent and important life-threatening adverse events. Those results further argue for the existence of high risk of multi-morbidity among CCS, with iatrogenic events increasing the risk of each other. In our population of 5-year CCS, the cumulative incidence of severe cardiac disease

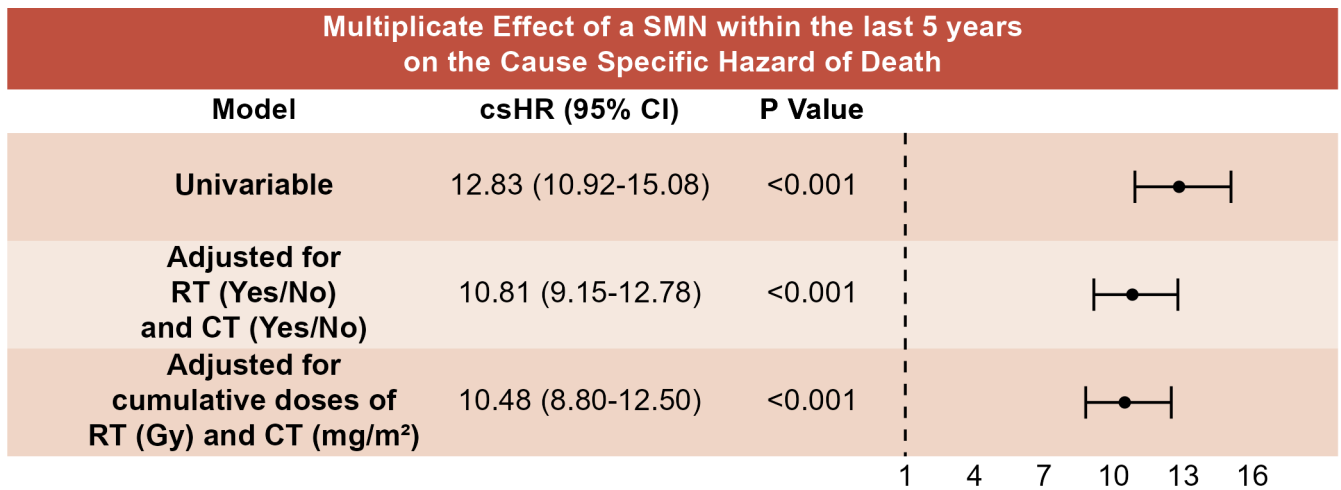


(a) Multiplicative effect of Second Malignant Neoplasm diagnosed within 5 years on the cause specific hazard of Cardiac Disease.

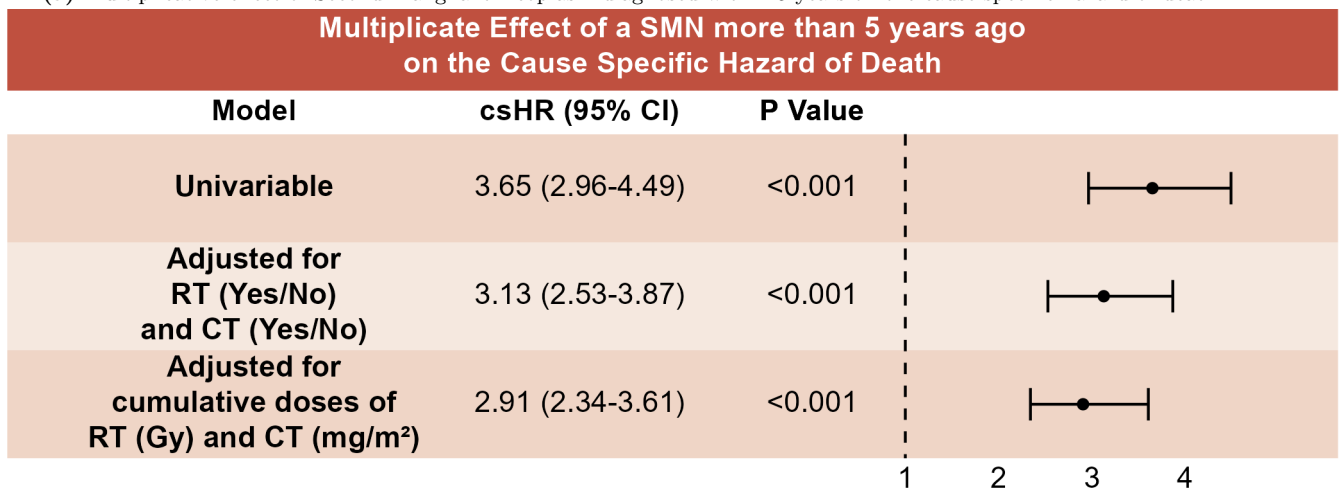


(b) Multiplicative effect of Malignant Neoplasm diagnosed more than 5 years ago on the cause specific hazard of Cardiac Disease.

**Figure 4.14:** Multiplicative effect of Second Malignant Neoplasm based on time since its diagnosis on the cause specific hazard of Cardiac Disease. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).



(a) Multiplicative effect of Second Malignant Neoplasm diagnosed within 5 years on the cause specific hazard of death.



(b) Multiplicative effect of Malignant Neoplasm diagnosed more than 5 years ago on the cause specific hazard of death.

**Figure 4.15:** Multiplicative effect of Second Malignant Neoplasm based on time since its diagnosis on the cause specific hazard of death. Estimates correspond to the model adjusted for cumulative doses of RT (Gy) and CT (mg/m<sup>2</sup>).

was 3.9% (95% CI: 3.4% - 4.3%) and 8.4% (95% CI: 7.3% - 9.5%) at 30 and 50 years of age, respectively. These proportions closely align with the 30-year cumulative incidence of first severe cardiac disease in the Netherlands Cancer Institute study (4.2%; 95% CI: 2.8% - 5.6%) and the Childhood Cancer Survivor Study across 27 participating institutions in the United States and Canada (4.8%; 95% CI: 4.3% - 5.2%) (Pal et al. 2012; Bates et al. 2019).

#### **4.4.2 . The confounding of childhood cancer treatments**

It is well known that both heart radiation and anthracycline-containing chemotherapy can increase the risk of severe cardiac disease in CCS (Haddy et al. 2016; Bates et al. 2019). Our findings concur with these previous studies and support current surveillance guidelines that take into account cardiotoxicity from all cancer treatments for risk stratification. Our results also suggest a possible increase in cardiac risk after a SMN. A major strength of this study is the access to a cohort of CCS with long-term follow-up and detailed clinical history, including information on both anthracycline doses and radiotherapy doses at the heart. This allowed us to adjust for multiple configurations of childhood cancer treatment and explore the evolution of the excess of severe cardiac disease after a SMN over time.

Many cardiotoxic childhood cancer treatments are also known to increase the risk of a SMN. We have detailed records available, allowing us to adjust on important factors, such as radiotherapy, chemotherapy, and age at childhood cancer. We observed the expected effect of these factors on cardiac risk, and the estimated effect of a SMN decreased when adjusting for them, which is consistent with previous results indicating shared risk factors for both a SMN and cardiac disease.



#### **4.4.3 . Changes of SMN effect on CD risk over time.**

We observed an increase in the likelihood of cardiac disease at all times, while the effect on the cumulative incidence was time-dependent. This difference may be explained by the increased risk of death of older patients after a SMN, which can compensate for the increased risk of cardiac disease. Indeed, if the effect on the risk of cardiac disease remains stable but the instantaneous risk of death increases, the effect on the cumulative incidence will decrease, possibly disappearing or even reversing. This does not mean that CCS who experience a SMN at an older age have a better prognosis than those who are diagnosed earlier; rather, their death rate becomes so high that cardiac diseases become less relevant. The evolution of the effect of a SMN on the cumulative incidence of cardiac disease may also be influenced by the overall increase of cardiac disease at a later age, possibly reducing the excess of cardiac disease caused by a SMN, as patients may have had a cardiac disease with or without a SMN.

Finally, it is possible that an effect may exist but we may have not detected it due to the small number of patients followed for an extended period and the small number of observed events. In the latter case, larger cohorts could induce clearer results. The difference observed for younger patients could be explained by their more recent diagnosis of a SMN, resulting in a higher death rate compared with patients with longer survival periods.

#### **4.4.4 . The effect of cardio-toxic treatments of SMN.**

We didn't have access to SMN treatment data, precluding us to perform a robust analysis of the association of SMN treatments with the risk of cardiac disease. However, we did have access to the type of SMN and used it as a proxy for SMN treatments. Due to a small number of events, we cannot draw conclusive effects of each SMN type on the cumulative incidence and the cause-specific-hazard, although point estimates suggest there may be differences. Those differences are likely induced by the use of radiotherapy at the heart to treat breast cancer, and the use of anthracyclines to treat sarcoma, bone, and soft tissue cancer.

#### 4.4.5 . Limitations

Some limitations of this study merit discussion.

A limitation that affected all analysis is the small number of events, with a total of 49 observed cardiac diseases after a second malignant neoplasm. This small number of event was further reduced by the use of the landmark method, with a number of event ranging from 9 to 21 for landmark time points. This small number of event lead to imprecise estimates (wide confidence intervals), which in turn rendered imprecise the estimates by cancer types and the estimates of time-varying coefficients.

Due to the lack of SMN treatment data, we were precluded from performing a robust analysis of the association of SMN cardio-toxic treatments with cardiac disease onset. We did use SMN type as a proxy, and our results suggest differences exist based on SMN type.

Aside from the cardio-toxicity of SMN treatments, unobserved shared risk factors of a SMN and cardiac disease may exist, and they could explain the increased cardiac risk observed for the category “Others or unknown” SMN. The effects of these factors are uncertain pending data on SMN treatment doses, which are required to conduct a mediation analysis. Another explanation could be a potential effect of having a SMN, apart from cardio-toxic SMN therapy, which may be related to an overall increased frailty of the patients.

Finally, some detection bias is likely to occur due to increased care provided to SMN patients during and after their SMN treatment. To minimize this bias, we considered only cardiac events of grade 3 or higher, but this may not have fully removed the bias. We suggest that the excess of cardiac disease after a SMN is partially due to cardiotoxicity introduced by the SMN treatment, because the mentioned biases are unlikely to account for the full differences between cancer types.

#### **4.4.6 . Perspectives**

This first exploration of cardiac disease risk after a second malignant neoplasm could be further extended. First, bigger cohorts of childhood cancer survivors exist (Pancare, Childhood Cancer Survivors Study) and could be used to get more precise estimates (smaller confidence intervals) (Hjorth et al. 2015; Robison et al. 2002). Next, an analysis including SMN treatment data would be able to conclude on the effect of cardio-toxic treatment of SMN and on the existence of a baseline effect of SMN on the risk of cardiac disease. Finally, the inclusion of life long covariates such as the smoking status would also help to isolate the effect of SMN and reduce confounding.

#### **4.5 . Summary**

This study provides new insight on the multi-morbidity experienced by childhood cancer survivors and its causes, by showing that a second malignant neoplasm increases the cumulative incidence of cardiac disease among childhood cancer survivors, despite an increased risk of death. Our results also suggest that cardio-toxic treatments of second malignant neoplasms contribute to much of this increase. Those results can potentially impact the long-term follow-up recommendations of childhood cancer survivors.

## 5 - Life Years Lost by Iatrogenic Events

The goal of this chapter is to study the life years lost by childhood cancer treatments and iatrogenic events among childhood cancer survivors. Some of the results are already published (Charrier et al. (2024)).

### 5.1 . Context

Despite childhood cancer survivors living long after the 5-year threshold, they still are at increased risk of death compared to the general population (Moskalewicz et al. 2024). Furthermore, many of the iatrogenic events they experience are further increasing the risk of death – examples include subsequent malignant neoplasms and cardiac diseases.

Previous works have summarized the magnitude, causes, and temporal patterns of childhood cancer late mortality using Standardized Mortality Ratios, Absolute Excess Risk, Poisson models, and Cox models (Moskalewicz et al. 2024). They showed three important facets of mortality : (i) excess mortality magnitude varies by cancer type and childhood cancer treatment (Armstrong et al. 2016), (ii) excess mortality is important at all ages and does not seem to waver at later ages (Dixon et al. 2023), (iii) death by cancer recurrence/progression is the leading cause of death during the first years after diagnosis while death by subsequent malignant neoplasm, cardiac disease, and pulmonary disease is the leading cause of death starting 15-25 years after diagnosis (Armstrong et al. 2016).

Two studies have investigated life years lost by childhood cancer treatment and iatrogenic events. A first one provided LYL estimates using hazard-based models for stratified treatment effects, and cause-specific LYL (Yeh et al. 2010). Because those estimates are model-based they are susceptible to give results different than a direct regression on LYL, which avoids

unnecessary uncertainties in intermediate estimates (Per Kragh Andersen 2017). A second one investigated LYL due to iatrogenic events at various ages among children, teenagers, and young adult cancer survivors, but didn't link them to cancer treatments, however a necessary step to accurately guide harm reduction of childhood cancer treatments (Chang et al. 2022).

We aimed to build upon those previous studies and investigate concomitantly the association of LYL with childhood cancer treatments and iatrogenic events. Furthermore, we also investigated the time-varying effects of those factors. To achieve those goals, we used regression models directly on life years lost. To investigate the time-varying effects of covariates on LYL, we fitted those models across a sequence of landmark times, using age as a time scale.

All the previous studies on the mortality of childhood cancer survivors have focused on the childhood cancer treatments, specific iatrogenic events, or multiple chronic health conditions without distinction of types. Multi-morbidity among childhood cancer survivors is frequent, and we showed in Section 4 that morbidities can increase the risk of each other. Building on those previous works, we've investigated the existence of a two-way interaction of iatrogenic events on the life years lost of childhood cancer survivors.

## **5.2 . Methods**

### **5.2.1 . Data**

Data used comes from the FCCSS which is detailed in Section 2, with 1496 deaths (1458 of which occurred before patients reached 60 years old), 828 SMN, 379 CDs of grade  $\geq 3$ , 130 diabetes, and 112 Chronic Renal Failures among 1276 unique patients. We have respectively 5371, 5110, and 1488 patients included for analysis at landmark times 16, 32, and 48 years of attained age.

### 5.2.2 . Life Years Lost

The estimand of interest is LYL by all-cause of death, up to  $\tau$  (Equation 3.7). We considered two types of analysis, one with  $\tau = 60$  and one with a 10 years window after the landmark time point.

#### 5.2.2.1 Covariates

Our interest lies in both the LYL by childhood cancer treatments and the LYL by iatrogenic events (SMN, CD, diabetes, and chronic renal failure). We used a linear regression model to analyze the LYL by childhood cancer treatments and iatrogenic events. We also used it to adjust for childhood cancer treatments when estimating the LYL by iatrogenic events, as those are correlated. The model is then

$$\text{LYL}(\tau|X) = \beta_0 + \beta X$$

where  $\beta_0$  is the expected number of life years lost for patients with  $X = 0$ .

We chose the childhood cancer treatments and iatrogenic events of interest using literature and clinical relevance (Chang et al. 2022; Tukenova et al. 2010; Wong et al. 2016; Ehrhardt et al. 2023). We retained the following childhood cancer treatment covariates: use of radiotherapy, chemotherapy, or both radiotherapy and chemotherapy, sex, and age at childhood cancer diagnosis. We dichotomized exposure to radiotherapy and chemotherapy because it is unclear how to include doses in a model for a population with a wide range of childhood cancer diagnosis. This range of cancer diagnosis induces big differences regarding treatment regimens and body area treated. Including all types of doses is not possible because of the high collinearity of radiotherapy doses.

We estimated models including (i) only childhood cancer treatments, (ii) only one iatrogenic event, (iii) childhood cancer treatments and one iatrogenic event, (iv) childhood cancer treat-

ments and two iatrogenic events, (v) and childhood cancer treatments, two iatrogenic events, and the two-way interaction of those events. This is visually summarized in Table 5.1

**Table 5.1:** Covariates combinations used to study life years lost among childhood cancer survivors

	i	ii	iii	iv	v
Childhood Cancer Treatments	X		X	X	X
Iatrogenic Event #1		X	X	X	X
Iatrogenic Event #2				X	X
Two ways interaction of Iatrogenic Events #1 and #2					X

### 5.2.2.2 Life Years Lost with Left Truncation and Right Censoring

We used patient’s age as the time scale to interpret results as life years lost. As a consequence, patients were left truncated (to 5 years after childhood cancer diagnosis) and right censored (lost to follow-up, ...). Because we are interested in internal time-dependent covariates, we chose to use the landmark method again, which is detailed in section 3.4.2.2. The new model is therefore:

$$\text{LYL}(\tau|l_k, X(l_k)) = \beta_{0,l_k} + \beta_{l_k}X(l_k)$$

We chose landmark time points ( $l_k$ ) ranging from 16 to 48 to include a wide range of age and time at iatrogenic events. We used  $\tau = 60$  by considering the number of patients at risk at all times. Of note, life years lost up to  $\tau$  aren’t directly comparable, as the maximum life years lived is  $\tau - l_k$ , which decreases as  $l_k$  increases.

We did another analysis using  $\tau_k = l_k + 10$  to investigate time-varying coefficients. We used the same  $l_k$  as above. The goal of this second analysis is to detect time-varying trends of estimates, for which it is better to fix the length of the window in order to have a shared maximum time

at risk.

### 5.2.2.3 Pseudo-Observations to Fit the Coefficients

We estimated the coefficients using strict pseudo-observations (Section 3.4.2.4), computed using the integral under the curve of the Kaplan-Meier estimator (Equation 3.10). Strict pseudo-observations are computed using only patients at risk at landmark time  $l$ , using:

$$\hat{\theta}_i^l = n(l) \times \widehat{LYL}(\tau|l) - (n(l) - 1) \times \widehat{LYL}^{(-i)}(\tau|l)$$

where  $n(l)$  is the number of patients at risk at time  $l$ ,  $\widehat{LYL}(\tau|l)$  is the non-parametric estimator of LYL up to  $\tau$  conditionally on being at risk at time  $l$  (Equation 3.10), and  $\widehat{LYL}^{(-i)}(\tau|l)$  this estimator on the sample with patient  $i$  removed.

### 5.2.3 . Life Years Lost by Cardiac Death

We didn't have the data required to move from LYL by all-cause death to LYL by cause of death. Nonetheless, we were able to identify death by CD from the grade of observed CDs. We used cause-specific life years lost to perform another analysis on cardiac death. The methodological details are similar to the analysis on all-cause death, with the exception of the pseudo-observations which were computed using the non-parametric estimator Equation 3.30. Because cardiac death is specific to a small region of the body, we also present results adjusted for radiotherapy doses at the heart and brain. The small to non-existent number of life years lost by chemotherapy has also raised the question of a drug-specific effect on mortality, so we adjusted the same model for cumulative dose of anthracyclines, a drug class known for its cardio-toxicity.



## 5.3 . Results

### 5.3.1 . Life Years Lost up to 60 Years Old

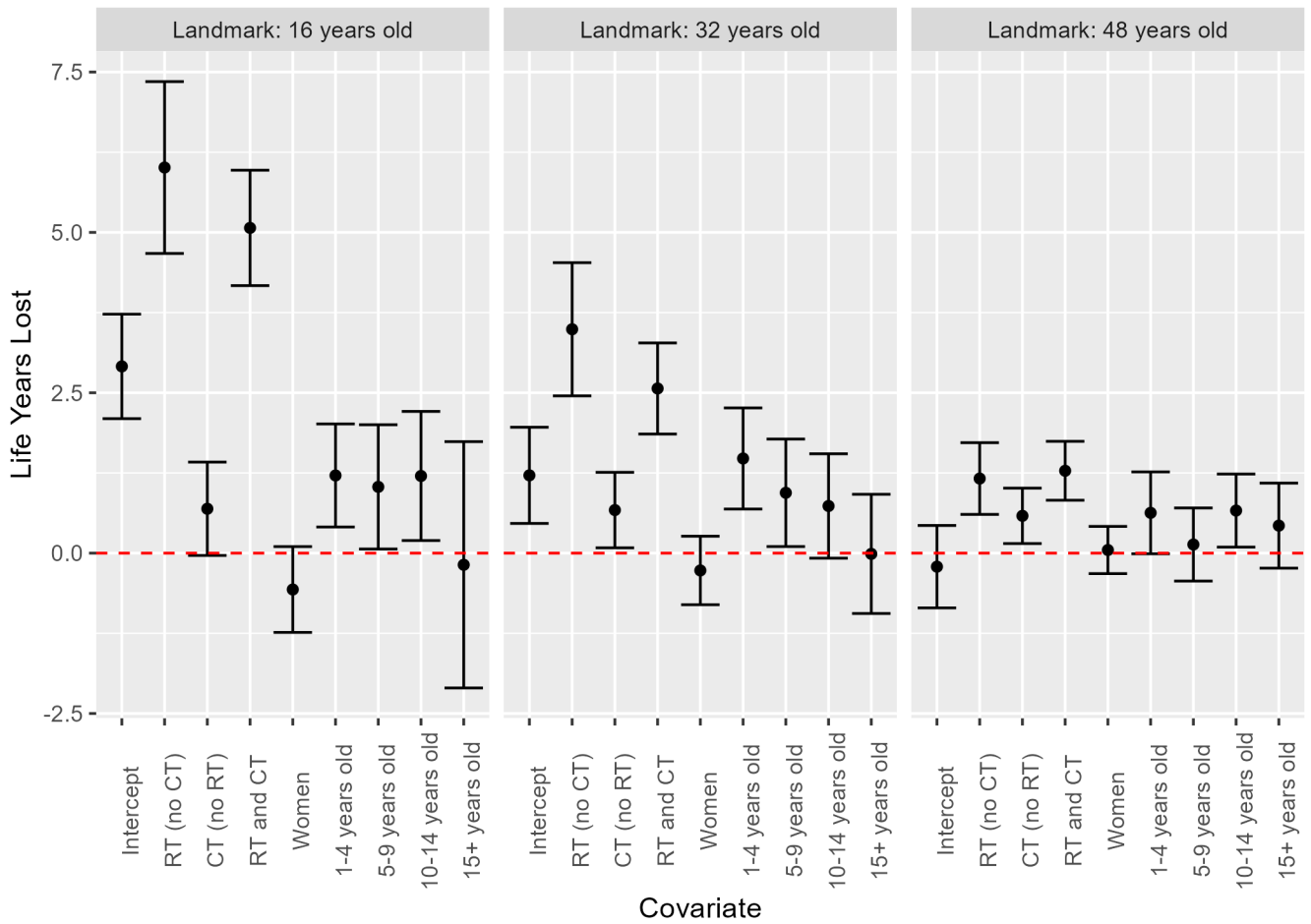
Life Years Lost up to 60 years by childhood cancer treatment and iatrogenic events are summarized in Figure 5.1, Figure 5.2, and Figure 5.3.

Results are to be interpreted as “conditional on survival at time points 16, 32, and 48 years, exposition to covariate is associated with a reduction of restricted life expectancy of  $\beta$  years.” Restricted life expectancy is the life expectancy up to time  $\tau$  (here  $\tau = 60$  years old).

#### 5.3.1.1 Life Years Lost by Childhood Cancer Treatment

When adjusting for childhood cancer treatment only, we found that treating patients with radiotherapy and no chemotherapy was the treatment regiment contributing the most to Life Years Lost, with 6.0 (95% CI: 4.7-7.3), 3.5 (95% CI: 2.5-4.5), 1.2 (95% CI: 0.6-1.7) Life Years Lost at landmark times 16 years old, 32 years old, and 48 years old, respectively. Use of both radiotherapy and chemotherapy was also associated with Life Years Lost, with 5.1 (95% CI: 4.2-6.0), 2.6 (95% CI: 1.9-3.3), 1.3 (95% CI: 0.8-1.7) Life Years Lost at landmark times 16 years old, 32 years old, and 48 years old, respectively. Use of chemotherapy and no radiotherapy was associated with Life Years Lost, but contributed a lot less to Life Years Lost than other treatment regimen, with 0.7 (95% CI: -0.0-1.4), 0.7 (95% CI: 0.1-1.3), 0.6 (95% CI: 0.1-1.0) Life Years Lost at landmark times 16 years old, 32 years old, and 48 years old, respectively.

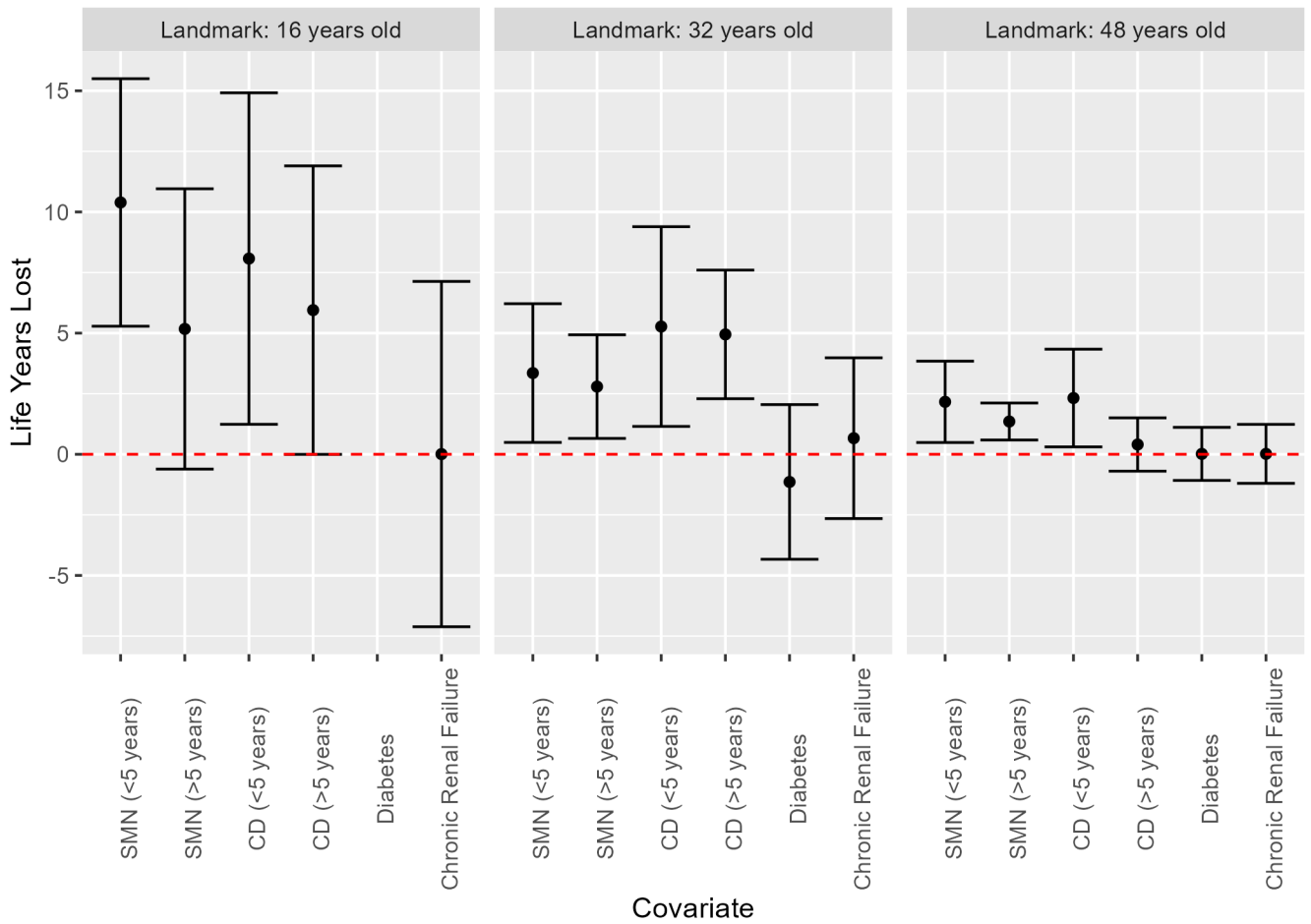
We found no effect of sex, and only a small effect of age at diagnosis, with patients diagnosed at 1-5 years old faring the worst, with 1.2 (95% CI: 0.4-2.0), 1.5 (95% CI: 0.7-2.3), 0.6 (95% CI: -0.0-1.3) Life Years Lost at landmark times 16 years old, 32 years old, and 48 years old respectively, using patients diagnosed at before 1 year old as reference.



**Figure 5.1:** Life Years Lost by all cause death up to 60 years old, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, and age at childhood cancer diagnosis.

### 5.3.1.2 Life Years Lost by Iatrogenic Events

SMN and CD were both associated with Life Years Lost. When adjusting for childhood cancer treatment, SMN point estimates were higher than CD point estimates, which is consistent with the usual hierarchy of SMN contributing more to mortality than CD. SMN diagnosed less than 5 years before landmark time point was associated with 10.5 (95% CI: 5.4-15.7), 4.3 (95% CI: 1.8-6.7), 2.0 (95% CI: 0.7-3.4) Life Years Lost at landmark times 16 years old, 32 years old, and 48 years old respectively. CD diagnosed less than 5 years before landmark time point was



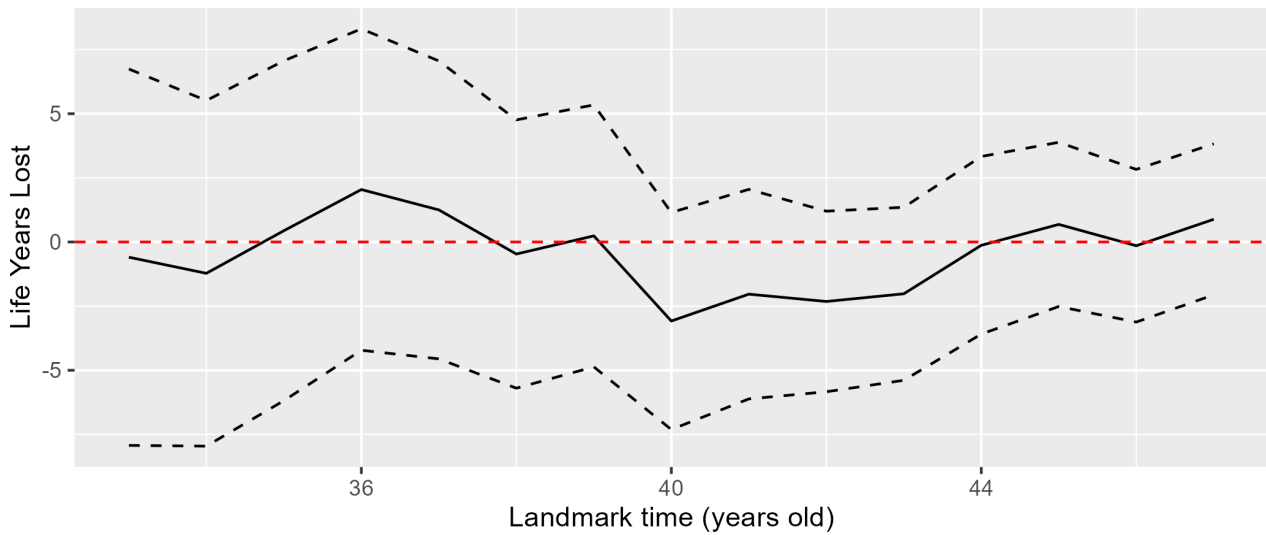
**Figure 5.2:** Life Years Lost by all cause death up to 60 years old, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, age at childhood cancer diagnosis, and the studied iatrogenic event.

associated with 7.7 (95% CI: 1.1-14.3), 6.4 (95% CI: 2.5-10.3), 2.2 (95% CI: 0.4-4.0) Life Years Lost at landmark times 16 years old, 32 years old, and 48 years old respectively.

Univariable models showed Life Years Lost by SMN and CD. At landmark time 32 years old, we found respectively 3.8 (95% CI: 0.9 - 6.7) Life Years Lost by SMN, 5.9 (95% CI: 1.6 - 10.0) by CD. We found no Life Years Lost by diabetes and chronic renal failure, even in the case of univariable models (at landmark time 32 years old, 95% CI for diabetes: -4.0 - 3.4, and 95% CI for chronic renal failure: -2.5 - 4.1). The results for diabetes at landmark time 16 years old

isn't displayed, because only one patients was diagnosed with diabetes before this time.

### 5.3.1.3 Life Years Lost by Multi-Morbidity



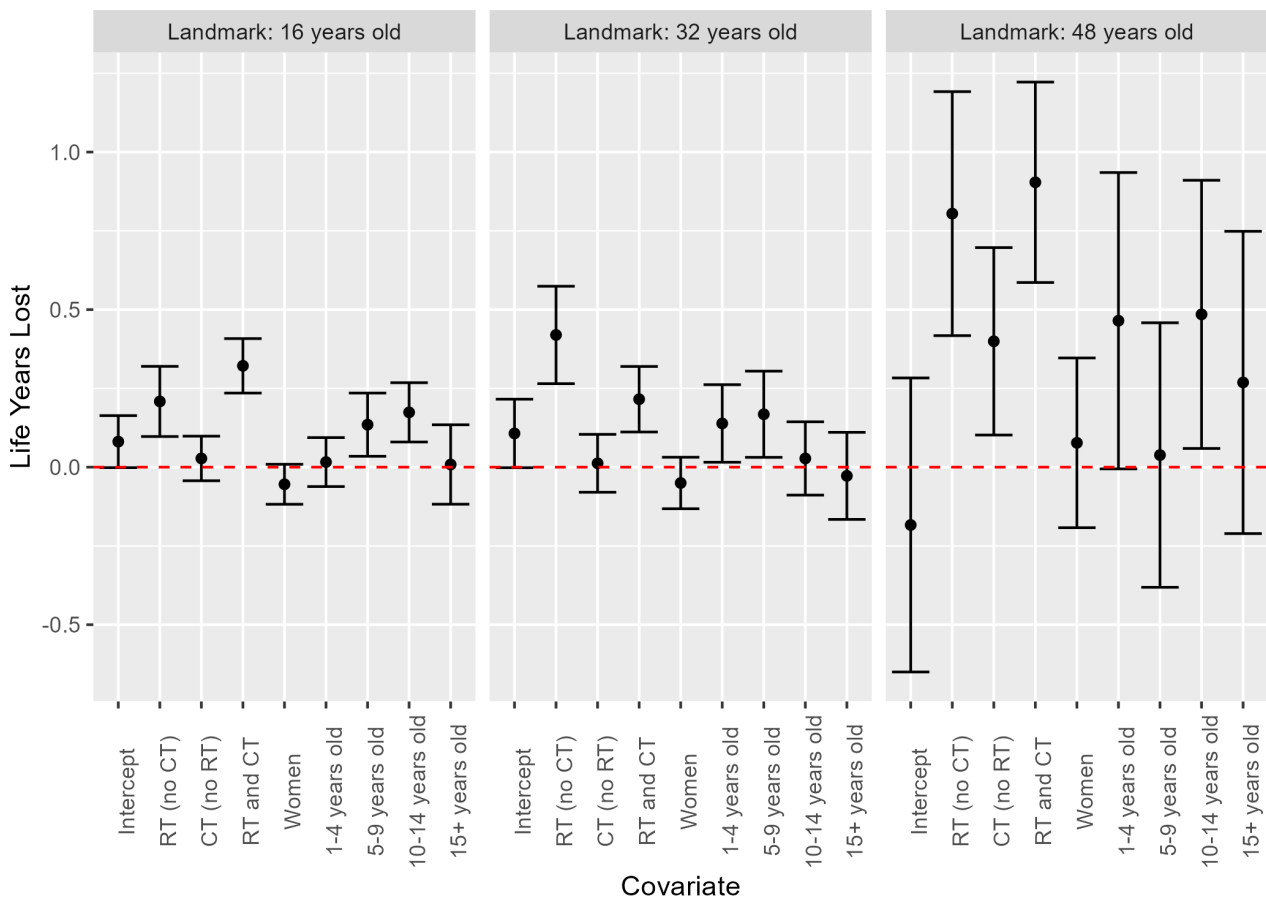
**Figure 5.3:** All cause death Life Years Lost up to 60 years old by two-ways interaction of subsequent malignant neoplasm and cardiac disease, conditional on survival at each landmark time. Black dotted lines represent the 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, age at childhood cancer diagnosis, subsequent malignant neoplasm, and cardiac disease.

Only a small number of patients experienced two specific iatrogenic events before landmark times. Therefore, we only display the results for the two-ways interaction of subsequent malignant neoplasm and cardiac disease, as they are both the most frequent iatrogenic events and the only ones contributing to life years lost by childhood cancer survivors. When adjusting for childhood cancer treatment and the occurrence of two iatrogenic events, we found no effect of two-ways interaction at all landmark times (Figure 5.3).

### 5.3.2 . Life Years Lost up to 10 Years after Landmark Time

Life Years Lost up to 10 years after landmark time by childhood cancer treatment and iatrogenic events are summarized in Figure 5.4, Figure 5.5, and Figure 5.6.

#### 5.3.2.1 Life Years Lost by Childhood Cancer Treatment



**Figure 5.4:** Life Years Lost within 10 years of landmark time, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, and age at childhood cancer diagnosis.

We found an increasing effect of radiotherapy as patients get older, with 0.21 LYL (95% CI: 0.10 - 0.32) at landmark time 16 years, 0.42 LYL (95% CI: 0.26 - 0.57) at 32 years, and 0.08 LYL (95% CI: 0.42 - 1.19) at 48 years. We found a similar pattern for those treated with

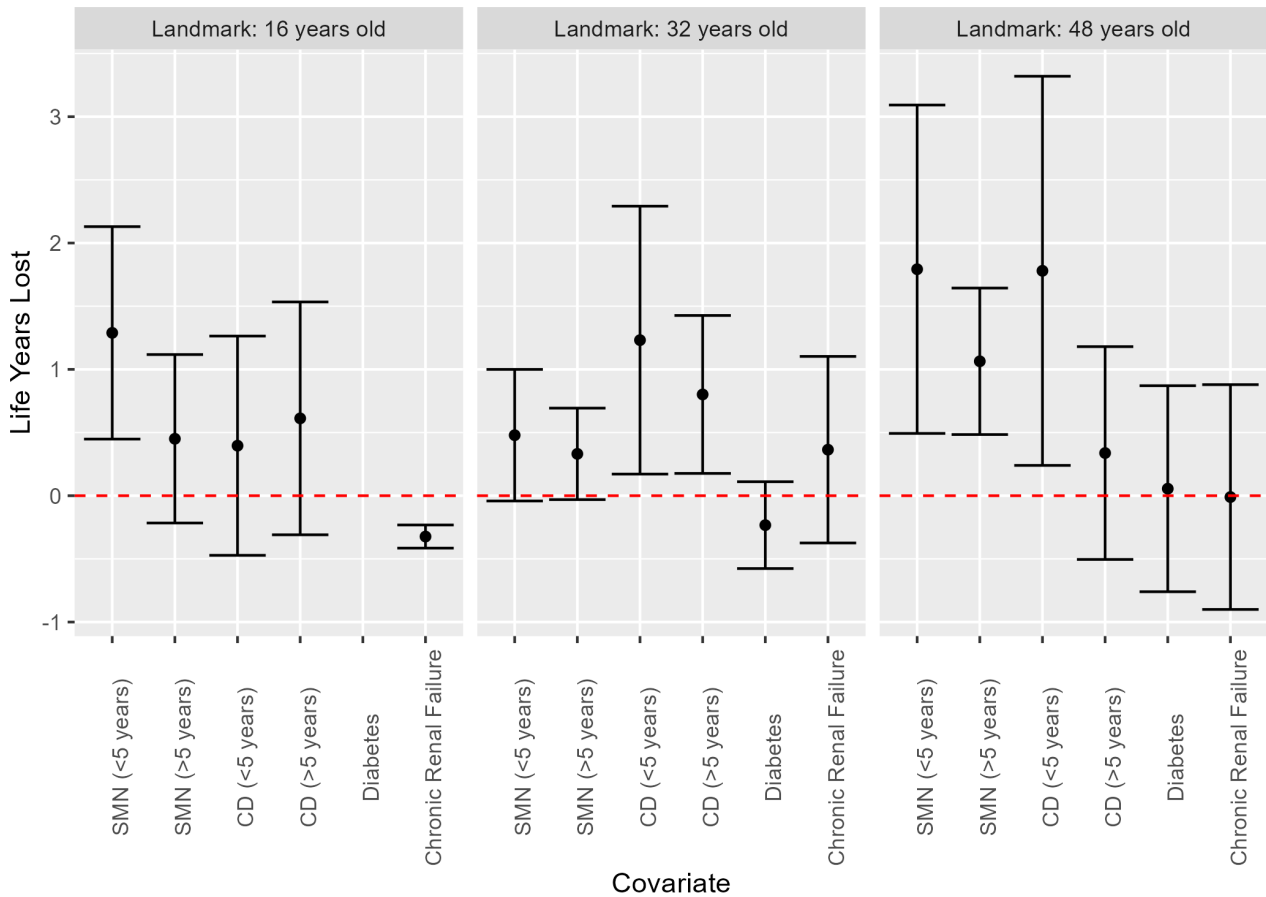
both radiotherapy and chemotherapy, with 0.32 LYL (95% CI: 0.23 - 0.41) at landmark time 16 years, 0.22 LYL (95% CI: 0.11 - 0.32) at 32 years, 0.90 LYL (95% CI: 0.59 - 1.22) at 48 years. We found no difference between males and females, and only small differences by age at childhood cancer diagnosis. Those differences don't appear to change over time, with 0.13 LYL (95% CI: 0.03 - 0.24) of patients diagnosed between 5 and 9 years old at landmark time 16 years, 0.17 LYL (95% CI: 0.03 - 0.30) at 32 years, and 0.04 LYL (95% CI: -0.38 - 0.45) at 48 years.

### **5.3.2.2 Life Years Lost by Iatrogenic Events**

We found a higher impact of a SMN diagnosis within the 5 years before landmark time on LYL at early and late age compared to middle age — 1.3 (95% CI: 0.45 - 2.13) at 16 years, 0.48 (95% CI: -0.04 - 1.00) at 32 years, and 1.80 (95% CI: 0.49 - 3.09) at 48 years. We found an increase in LYL-10 by a recent CD as patients get older, with 0.40 LYL (95% CI: -0.47 - 1.26) at landmark time 16 years, 1.23 (95% CI: 0.17 - 2.29) at 32 years, and 1.78 (95% CI: 0.24 - 3.32) at 48 years.

### **5.3.2.3 Life Years Lost by Multi-Morbidity**

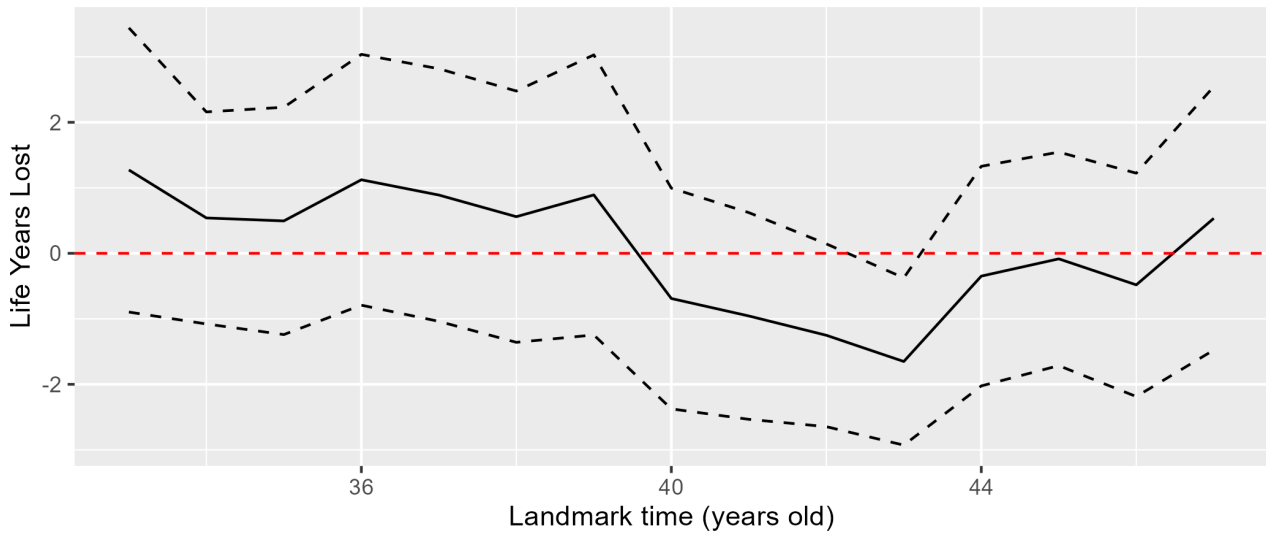
Once more, we only show results for the two-ways interaction of subsequent malignant neoplasm and cardiac disease because of the small number of patients for other combinations. When adjusting for childhood cancer treatment and the occurrence of two iatrogenic events, we found no effect of two-ways interaction at all landmark times but the 43 years old one, where a small protective effect was observed (Figure 5.3).



**Figure 5.5:** Life Years Lost within 10 years of landmark time, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, age at childhood cancer diagnosis, and the studied iatrogenic event.

### 5.3.3 . Life Years By Cardiac Death

Using the whole cohort and  $\tau = 60$ , we found a total of 0.3 Life Years Lost by cardiac death, out of the 8 Life Years Lost by all-cause death. The decomposition of Life Years Lost by cardiac death up to 60 years is shown in Figure 5.7 and Figure 5.8. When adjusting for radiotherapy and chemotherapy use, sex, and age at diagnosis, we found some Life Years Lost by cardiac death among patients treated with both radiotherapy and chemotherapy (0.34 [95% CI: 0.15 - 0.53] at 16 years, 0.20 [95% CI: 0.08 - 0.33] at 32 years, and 0.02 [95% CI: -0.02 - 0.06] at 48 years), and none for patients treated with chemotherapy and no radiotherapy (0.60 [95% CI:



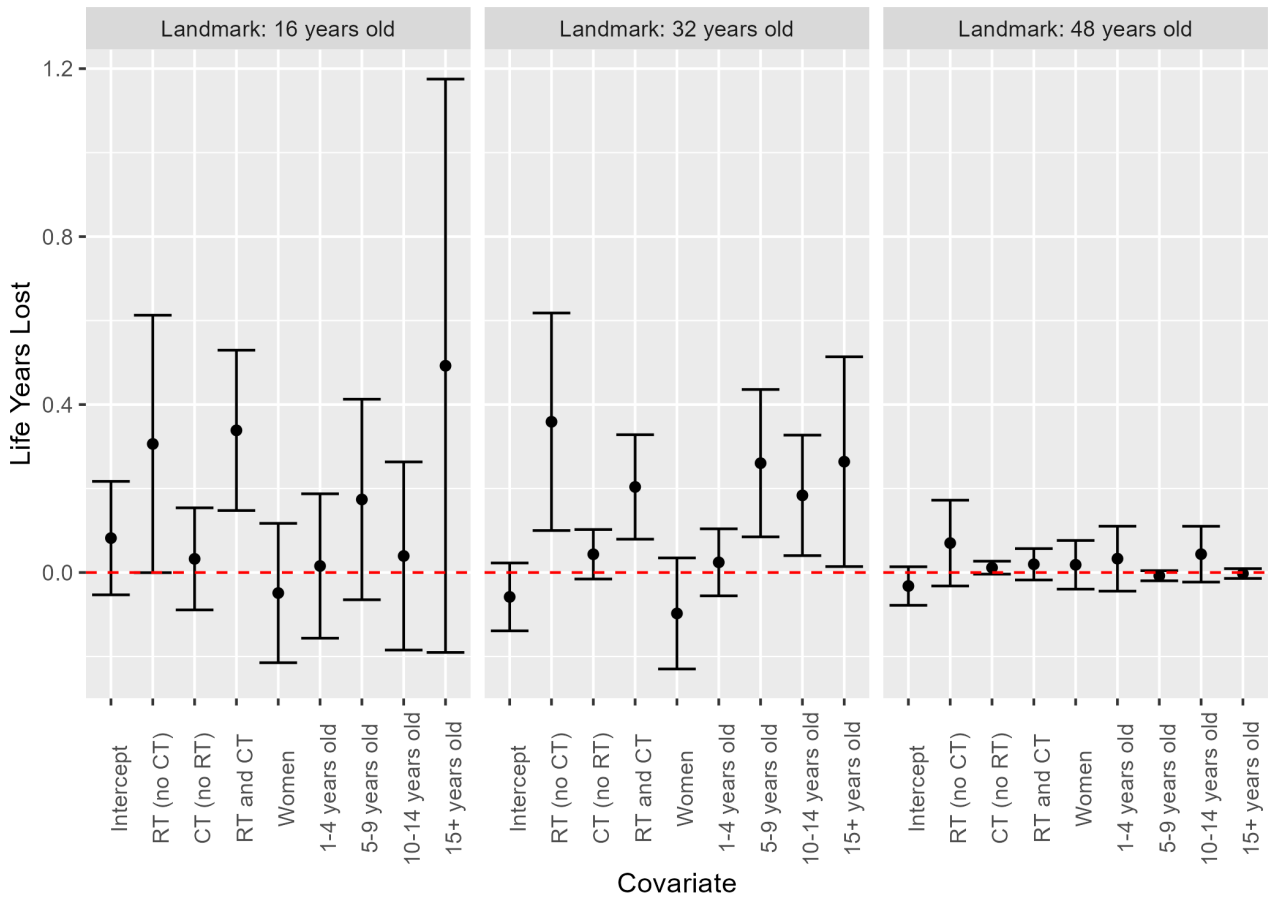
**Figure 5.6:** All cause death Life Years Lost within 10 years of landmark time old by two-ways interaction of subsequent malignant neoplasm and cardiac disease. Black dotted lines represent the 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, age at childhood cancer diagnosis, subsequent malignant neoplasm, and cardiac disease.

-0.01 - 0.14] at 16 years, 0.0 [95% CI: -0.0 - 0.1] at 32 years, 0.01 [95% CI: -0.00 - 0.03] at 48 years (Figure 5.7).

When adjusting for radiotherapy doses at the heart and brain, and cumulative dose of anthracyclines, we found radiotherapy at the heart superior to 20 Gy to be a major contributor, with 2.59 (95% CI: 1.53 - 3.6) Life Years Years Lost by cardiac death at landmark time 16 years old, and 1.82 (95% CI: 1.00 - 2.64) at 32 years old. This effect had disappeared at landmark time 48 years old (0.78 [95% CI: -0.14 - 0.29]). Dependent on landmark time, we found either no effect or a small protective effect of cumulative dose of anthracycline higher than 250 mg/m<sup>2</sup>, with a reference of cumulative dose inferior to 100 mg/m<sup>2</sup> — 0.15 (95% CI: -0.20 - 0.50) at landmark time 16 years old, -0.23 (95% CI: -0.42 - -0.05) at 32 years old, and 0.00 (95% CI: -0.02 - 0.03) at 48 years old (Figure 5.8).

We do not show any results of Life Years Lost by iatrogenic events, because of the small number of cardiac death observed in the FCCSS.



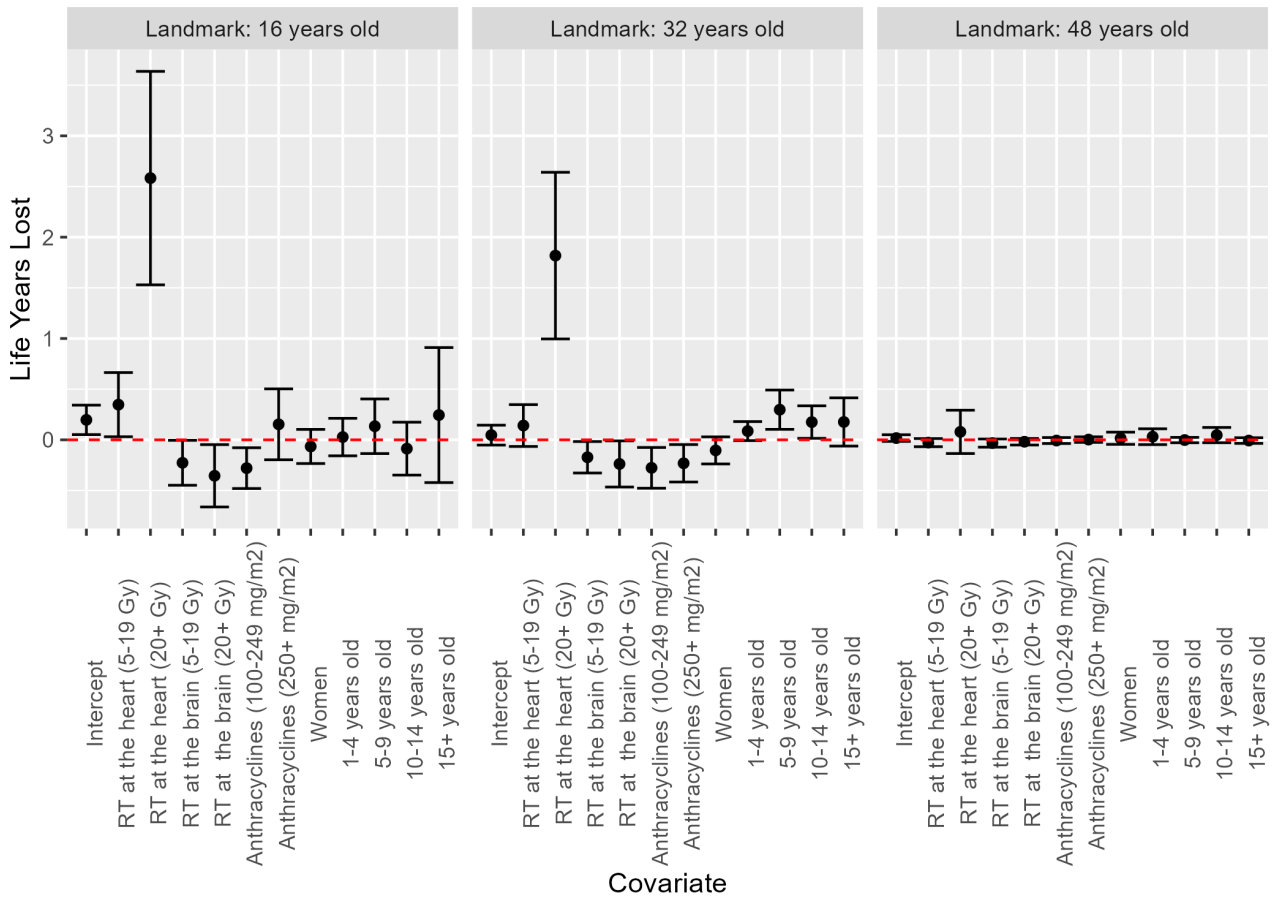


**Figure 5.7:** Life Years Lost by cardiac death up to 60 years old, conditional on survival at each landmark time. Error bars represent 95% confidence interval. The red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, sex, and age at childhood cancer diagnosis.

## 5.4 . Discussion

### 5.4.1 . Important Contributors to Life Years Lost

In this retrospective with a prospective follow-up cohort study of childhood cancer survivors, we examined: i) the Life Years Lost by childhood cancer treatments, ii) the Life Years Lost by iatrogenic events, iii) the Life Years Lost by multi-morbidity, iv) and time-varying aspect of those life years lost. We found that radiotherapy was the childhood cancer treatment most



**Figure 5.8:** Life Years Lost by all cause death up to 60 years old, conditional on survival at each landmark time. Error bars represent 95% confidence interval. Red dotted line corresponds to the null. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, Sex, and age at diagnosis.

strongly associated with Life Years Lost and that other childhood cancer treatments' associations with Life Years Lost were very small in comparison. We also showed that both subsequent malignant neoplasms and cardiac diseases were important contributors to the Life Years Lost by childhood cancer survivors. We found no effect of multi-morbidity on Life Years Lost, other than the respective effects of each iatrogenic events.

Our results on Life Years Lost by iatrogenic events are coherent with those reported by Chang et al. (2022). Similarly to us, they assessed iatrogenic events associations with Life Years Lost depending on the age at the health condition onset. They found Cardiac Disease to be associated with 10.13 (95% CI: 7.13 - 14.30) Life Years Lost at 32.5 years old, and -0.29 (95% CI:

-0.29 - -0.29) Life Years Lost at 45 years old. They also found Subsequent Malignant Neoplasm to be associated with 11.67 (95% CI: 9.29 - 15.27) Life Years Lost at 32.5 years old, and 1.06 (95% CI: 0.00 - 2.98) Life Years Lost at 45 years old. The differences with our results can be attributed to differences in population included, methodology used regarding Life Years Lost computation, time since iatrogenic event onset, and inclusion of childhood cancer treatment covariates.

#### **5.4.2 . The Advantage of the Life Years Lost Quantity**

Most of the previous works on the mortality of childhood cancer survivors used epidemiological quantities such as Standardized Mortality Ratios, Absolute Excessive Risk, and Proportional Hazard Ratios (PHR). They found Standardized Mortality Ratios ranging from 5.6 to 17.2 and Absolute Excess Risk of death ranging from 34.1 to 70.1 excess deaths per 10,000 person years (Moskowitz et al. 2019).

These quantities are useful for clinicians and public health decision-makers, but require extra work and care to be interpreted and communicated to patients who may lack the knowledge to interpret relative metrics. Furthermore, Life Years Lost results translate into life expectancy, a quantity used to communicate with patients. It is also useful to evaluate how childhood cancer treatment evolved to reduce long term mortality (Yeh et al. 2020).

#### **5.4.3 . Cause Specific Life Years Lost**

Life Years Lost are often combined with causes of death and studied as Cause Specific LYL. This allows for a simple visualization ranking causes of death by the magnitude of Life Years Lost. This also enables the study of risk factors associated with each cause of death, such as investigating if anthracyclines contribute to non-cardiac death. Causes of death were not usable in our case, therefore, we studied all-cause death. Nonetheless, we still provided valuable results on all-cause mortality and did a side analysis on cardiac death. We found that radiotherapy at the heart was the biggest contributor to life years lost cardiac death, and found no effect

of anthracyclines. The absence of effect of anthracyclines on LYL by cardiac death, despite its known association with cardiac disease risk could be induced by a higher surveillance of patients treated with anthracyclines leading to care being provided before the cardiac disease becomes deadly.

#### 5.4.4 . Limitations

A limitation of the study is the lack of access to alcohol consumption, smoking status, and socio-economic data. Those factors are known to be associated with mortality, and previous works have shown an association between those and childhood cancer type (and therefore treatment regimen and iatrogenic events) (A. Dumas et al. 2015; Ernst et al. 2023; Langeveld et al. 2002). This may bias the results towards a higher association of iatrogenic events and death. Nonetheless, we believe the marginal associations of secondary cancer and cardiac disease with LYL to be of interest. We also suggest that using a landmark strategy and looking at the effect of SMN/CD already diagnosed at that time provide estimates less susceptible to those biases than Cause Specific LYL. The design of the study (analysis at various landmark time points) shifts away from survival bias to results conditional on survival. The interpretation of results still must be made by keeping in mind that patients surviving up to a landmark time point are survivors of survivors, and therefore a special case of initial survivors. This is especially true regarding iatrogenic events, which often occurred months before each landmark time point. Regarding the null results on the Life Years Lost by multi-morbidity, it is possible that those are false negative. Indeed, SMN effect on mortality wavers with time since SMN diagnosis, but we only used a simple dichotomization, based on the diagnosis of the SMN within 5 years of landmark time or more. Patients with multi-morbidity have a higher time between first iatrogenic event and landmark time, therefore reducing their life years lost by first iatrogenic event. This loss of life years lost is then part of the estimated life years lost by multi-morbidity, resulting in a bias towards  $-\infty$ .

#### **5.4.5 . Perspectives**

Further work could be done to provide an fuller view of life years lost by childhood cancer survivors. First, the inclusion of socio-economic data could be valuable, as it is known that childhood cancer survivors socio-economic status is affected by the care they received and the risk of iatrogenic events they perceive to be at risk of. Second, accounting for time since iatrogenic event would be valuable. It would both complete the insight on life years lost by iatrogenic event and reassess the null result of life years lost by multi-morbidity.

#### **5.5 . Summary**

This study showed that radiotherapy induces a lot of Life Years Lost among childhood cancer survivors, while chemotherapy has little to no effect. Iatrogenic events also contribute to Life Years Lost among childhood cancer survivors, with Subsequent Malignant Neoplasms and Cardiac Diseases being the most important factors. We found no effect of multi-morbidity per se, although it might be induced by the method used.

## 6 - Landmark Estimation of Life Years Lost with Time-Dependent Covariates, Time-Varying Effect, and their Interaction

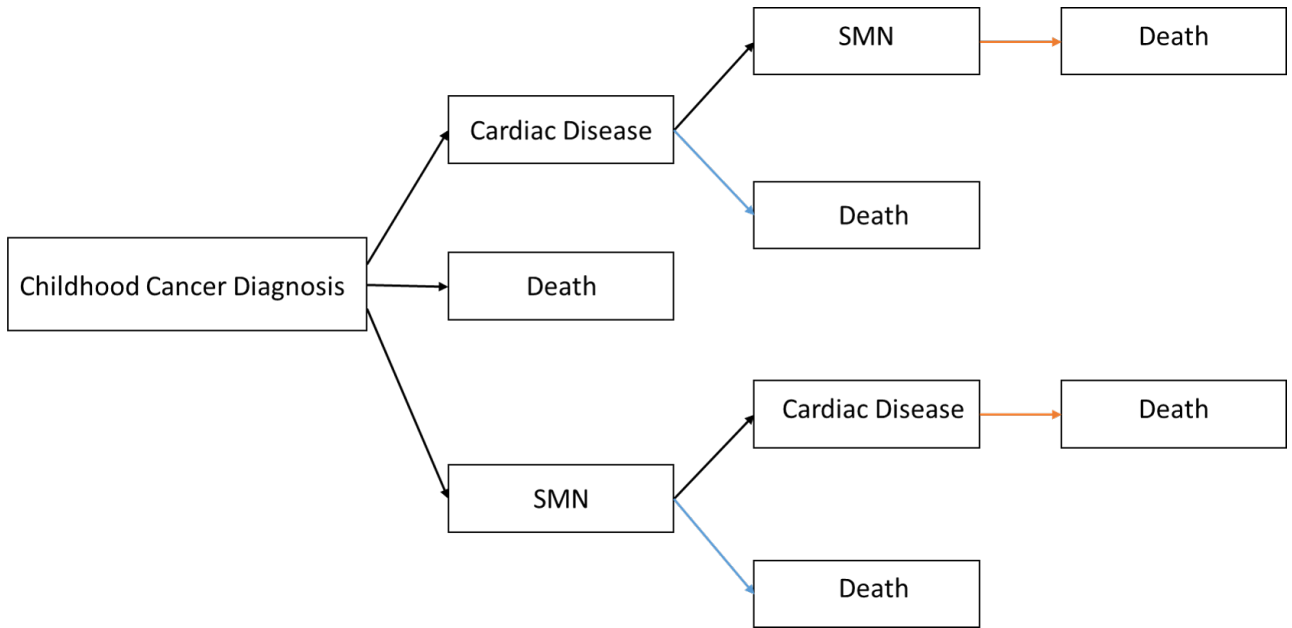
This chapter describes the third contribution of this thesis, which is still undergoing.

### 6.1 . Introduction

The landmark method is suitable in survival analysis to study time-dependent covariates and time-varying effects (Cortese and Andersen 2010; Putter and Houwelingen 2017). We notably employed it in Section 4 to study the time-varying effect of the internal time-dependent covariate  $\mathbb{1}(\text{a SMN occurred before } t)$  on the cumulative incidence of cardiac disease while accounting for the competing event death. We also used it in Section 5 to study the life years lost by iatrogenic events.

One of the iatrogenic event of interest for LYL was Second Malignant Neoplasm. We found (a) that the SMN effect on the absolute and instantaneous risk of death depends on the time since SMN. Thus, our estimate of LYL by SMN is affected by the time between SMN occurrence and landmark time. This means that our estimate of LYL by SMN **is not** a valid description of LYL by SMN occurrence **at** landmark time, but **it is** a valid description of LYL by SMN **before** landmark time conditional on survival at landmark time.

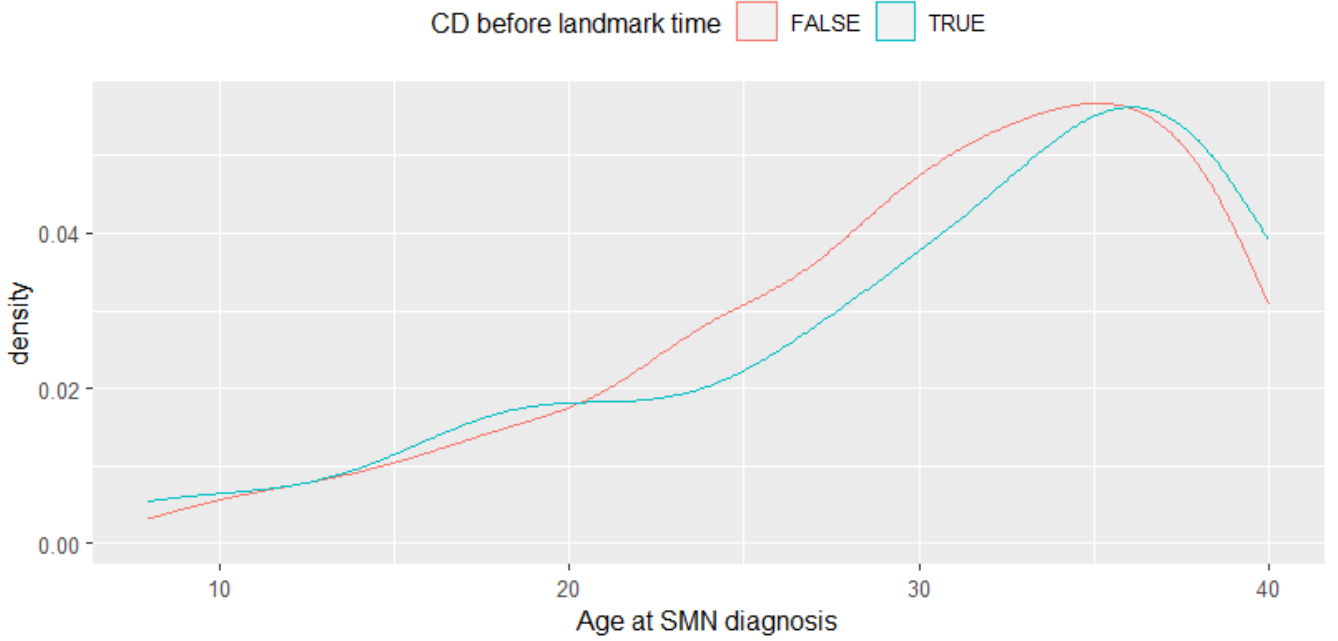
When investigating the LYL by co-occurrence of SMN and CD (further named here LYL by multi-morbidity), we found no evidence of such LYL. Which means our LYL from the orange transitions in Figure 6.1 can be summarized using the LYL from the blue transitions. We delved into potential underlying mechanisms for this result and noticed (b) a difference in time



**Figure 6.1:** Multi-State model showing the transitions to death after multi-morbidity in orange, and the transitions to death after a single morbidity in blue.

between SMN occurrence and landmark time between the individuals who experienced only a SMN before landmark time, and the individuals who experienced both a SMN and CD before landmark time (Figure 6.2).

As a consequence of (a) and (b), we have that even in the absence of an effect of CD on the risk of death, for landmark time  $l$ ,  $S(t|T > l, T_{SMN} < l, T_{CD} < l) \neq S(t|T > l, T_{SMN} < l, T_{CD} > l)$ , because the extra risk attributable to SMN is lower in the group with the earliest SMN occurrence. (c) This difference in the survival function translates into a difference in the LYL. Therefore, if assuming no effect of CD for simplicity and denoting  $LYL^{SMN\&CD}(l, \tau)$  to be the LYL by multi-morbidity after accounting for LYL by SMN, we have the following decomposition:



**Figure 6.2:** Density of SMN times, conditional on survival up to 40 years old and diagnosis of SMN before 40 years old. Lines are stratified based on diagnosis of a CD before 40 years old.

$$\begin{aligned}
 LYL(l, \tau | T > l, T_{SMN} < l, T_{CD} < l) &= LYL(l, \tau | T > l, T_{SMN} < l, T_{CD} > l) \\
 &+ \underbrace{LYL^{SMN \& CD}(l, \tau) + \text{Extra } LYL(l, \tau)}_{LYL^{MM}(l, \tau)}
 \end{aligned}$$

where  $ExtraLYL(l, \tau)$  is the difference in LYL induced by (c).

Using an additive model on the LYL as done in Section 5, we obtain an estimator of  $LYL^{MM}(l, \tau)$ . However, this result isn't properly adjusted for LYL by SMN when using the indicator of SMN before  $t$ , and can't be interpreted solely as LYL by multi-morbidity. Instead, it is the sum of LYL by multi-morbidity, and some residual unadjusted LYL by SMN (namely  $ExtraLYL(l, \tau)$ ). In our case, we expect  $ExtraLYL(l, \tau)$  to be positive, because of the shape of (a) and (b) (decreased time since SMN in the multi-morbidity group, and decreasing effect of SMN on survival as time since SMN increases). This property remains with an effect of CD on the survival func-



tion. We don't have any evidence of a time-varying effect of CD on survival, hence why our explanation focuses on the effect of SMN.

We conducted a simulation study to investigate how our estimator of LYL by multi-morbidity behaves under various scenarios of correlation between SMN and CD, and time-varying effect of SMN, CD, and multi-morbidity on the survival function. In Section 4 we showed an effect of SMN on the instantaneous and absolute risk of CD, but for simplicity, we didn't include this relation in our analysis, and restricted ourselves to shared causes of SMN and CD to simulate their correlation.

## 6.2 . Simulation Scenarios

### 6.2.1 . Simulating SMN and CD

We simulated times of SMN ( $T_{SMN}$ ), and CD ( $T_{CD}$ ) using a gompertz distribution for them to be distributed similarly to those observed in our cohort, but multiplied their hazard to increase the number of events observed and increase the power of our analysis.

Our initial idea requires that among patients who survival up to  $l$  and experienced a SMN before  $l$  ( $T_{SMN} < l$ ), the distribution of  $T_{SMN}$  is different conditional on occurrence of CD before  $l$ . This requires SMN and CD occurrence to not be independent. Therefore, we conducted scenarios accounting for various correlation of SMN and CD, starting with an independent scenario to confirm the estimates acquired to be correct. We then simulated a time-fix covariate  $z \in \{0, 1\}$ , and applied a HR of 2, 3, or 4 to the hazards of SMN and CD.

### 6.2.2 . Simulating Death

We plan on analyzing data using a landmark strategy. Therefore, we are at risk of the bias induced by the last observation carried forward principle, as some individuals with no SMN at landmark time  $l$  will experience a SMN later, and their survival will be affected by it. In order to neutralize this bias which doesn't interest us, we choose to simulate data according to the landmark time  $l$  at which they will be studied. That is, we choose  $l \in \{20, 30, 40, 50\}$  and set times of SMN and CD higher than  $l$  to be equal to  $\infty$ . In this way, the individuals who didn't experience a SMN or CD before  $l$  will never experience one.

Once  $T_{SMN}, T_{CD}$  simulated and landmark time  $l$  chosen, we simulated times of death ( $T_D$ ). We used a baseline survival function  $S_0(t)$  close to the one observed in the FCCSS. We modeled the effect of SMN, CD, and multi-morbidity (ie. interaction of SMN and CD) on the survival function using an additive model on the survival function:

$$\begin{aligned} S(t|T_{SMN}, T_{CD}) = & S_0(t) \\ & + \beta_{SMN}^S(t - T_{SMN})\mathbf{1}(T_{SMN} < t) \\ & + \beta_{CD}^S(t - T_{CD})\mathbf{1}(T_{CD} < t) \\ & + \beta_{SMN\&CD}^S(t - \max(T_{SMN}, T_{CD}))\mathbf{1}(T_{SMN} < t, T_{CD} < t) \end{aligned}$$

The aim of this parametrization is to control the LYL by  $SMN$ ,  $CD$ , and their interaction easily.

We studied three different scenarios, where  $\beta_{SMN}^S$ , and  $\beta_{CD}^S$  and  $\beta_{SMN\&CD}$  are either:

$$\beta_{SMN}^S(t) = \begin{cases} 0, & \text{(Null effect)} \\ \mathbf{1}(t > 0) \times a_{SMN}^c t, & \text{(Constant Effect)} \\ \mathbf{1}(t > 5) \times (a_{SMN}^d t + 5b_{SMN}^d) + \mathbf{1}(t \in ]0, 5]) \times (a_{SMN}^d + b_{SMN}^d)t, & \\ & \text{(Decreasing Effect)} \end{cases}$$

Those effects respectively correspond to a **null effect** of SMN on the survival, a **constant effect** on the survival ( $a_{SMN}^c$  shift in survival every year), or a **decreasing with time since diagnosis effect** on the survival ( $a_{SMN}^d + b_{SMN}^d$  shift in survival every year during the first 5 years after SMN diagnosis, and  $a_{SMN}^d$  afterwards). The decreasing scenario is modeled as a spline for simplicity of visualization and true life years lost computation. We used the same configurations for CD and multi-morbidity, but changed  $a_{SMN}^c, a_{SMN}^d, b_{SMN}^d$ . We used the following values for  $a_*^*$  and  $b_*^d$ :

$$\begin{cases} a_{SMN}^c = a_{SMN}^d = b_{SMN}^d = -0.01 \\ a_{CD}^c = a_{CD}^d = b_{CD}^d = -0.005 \\ a_{SMN\&CD}^c = a_{SMN\&CD}^d = b_{SMN\&CD}^d = -0.004 \end{cases}$$

In an earlier analysis, we found no difference under different censoring scenarios. Hence, we made the choice to apply no censoring in this analysis to lower the required computational cost. Under each scenario, we simulated 1000 data set of 10000 individuals.

### 6.2.3 . Summary of simulation procedure

Here is a quick summary of the steps of the simulation procedure:

1. Simulation of binary covariate  $z$ .
2. Simulation of times of SMN and CD.

3. Choice of landmark time  $l$ .
4. Deletion of all SMN and CD times bigger than  $l$ .
5. Computation of survival times.

#### 6.2.4 . Estimated Models

Once the data simulated, we estimated models from Section 5 using pseudo-observations and the landmark method. Let  $k$  be the identifier of time-dependent covariates ( $SMN$ ,  $CD$ ,  $SMN\&CD$ ). We fitted models with and without multi-morbidity effect. We fitted all models using a constant  $\beta$ , but defined  $X_k(l)$  as the binary covariate  $\mathbb{1}(\text{event } k \text{ occurred before } l)$ :

$$LYL(\tau|X_{SMN}(l), X_{CD}(l), T \geq l) = \beta_0^l + \beta_{SMN}^l X_{SMN}(l) + \beta_{CD}^l X_{CD}(l) + \beta_{SMN\&CD}^l X_{SMN}(l) X_{CD}(l)$$

### 6.3 . Simulation Results

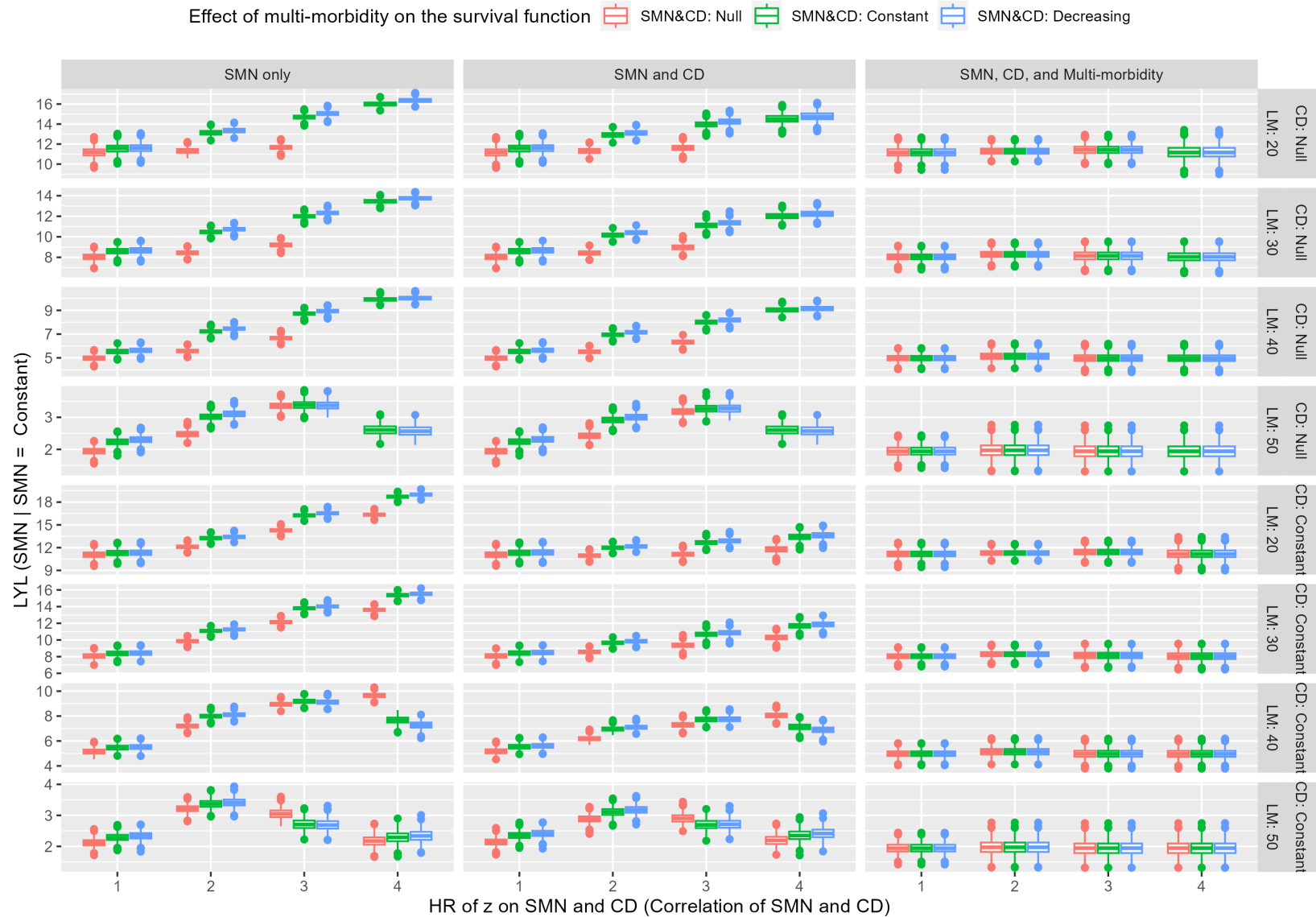
Some results are currently missing from figures, as they are still being computed.

#### 6.3.1 . Estimates of Life Years Lost by SMN

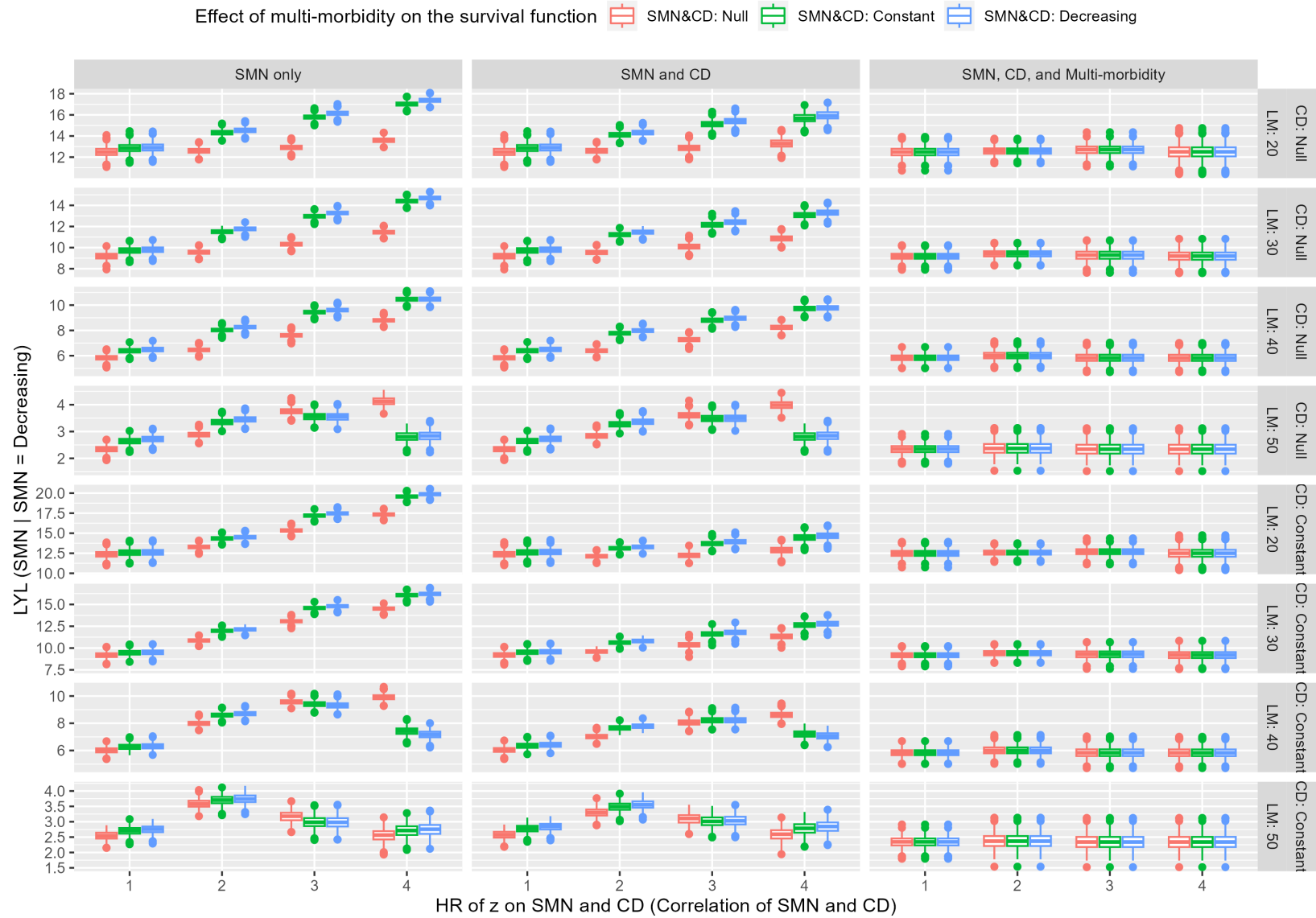
We start by discussing our estimated LYL by SMN under different scenarios. Figure 6.3 and Figure 6.4 show the estimated LYL by SMN when the true effect of SMN on survival is respectively constant and decreasing. Both figures are very similar, suggesting that the patterns observed here are unaffected by the time-varying coefficient of SMN on survival.

When adjusting for all three covariates (SMN, CD, and multi-morbidity), the estimated LYL by SMN are unaffected by changes in the effect of CD on survival, the effect of multi-morbidity on survival, and the strength of correlation between SMN and CD.

When we don't adjust for multi-morbidity, we observe differences when the effect of multi-morbidity on survival changes, and when the correlation between SMN and CD changes. Al-



**Figure 6.3:** Life Years Lost by SMN, when the true effect of SMN on the survival function is constant. Each column corresponds to a different configuration of covariates adjusted on (either only SMN, both SMN and CD, or SMN and CD and their interaction). Each row corresponds to a combination of the true effect of CD on the survival function (either null or constant) and the landmark time at which estimation is done (20, 30, 40).



**Figure 6.4:** Life Years Lost by SMN, when the true effect of SMN on the survival function is decreasing. Each column corresponds to a different configuration of covariates adjusted on (either only SMN, both SMN and CD, or SMN and CD and their interaction). Each row corresponds to a combination of the true effect of CD on the survival function (either null or constant) and the landmark time at which estimation is done (20, 30, 40).

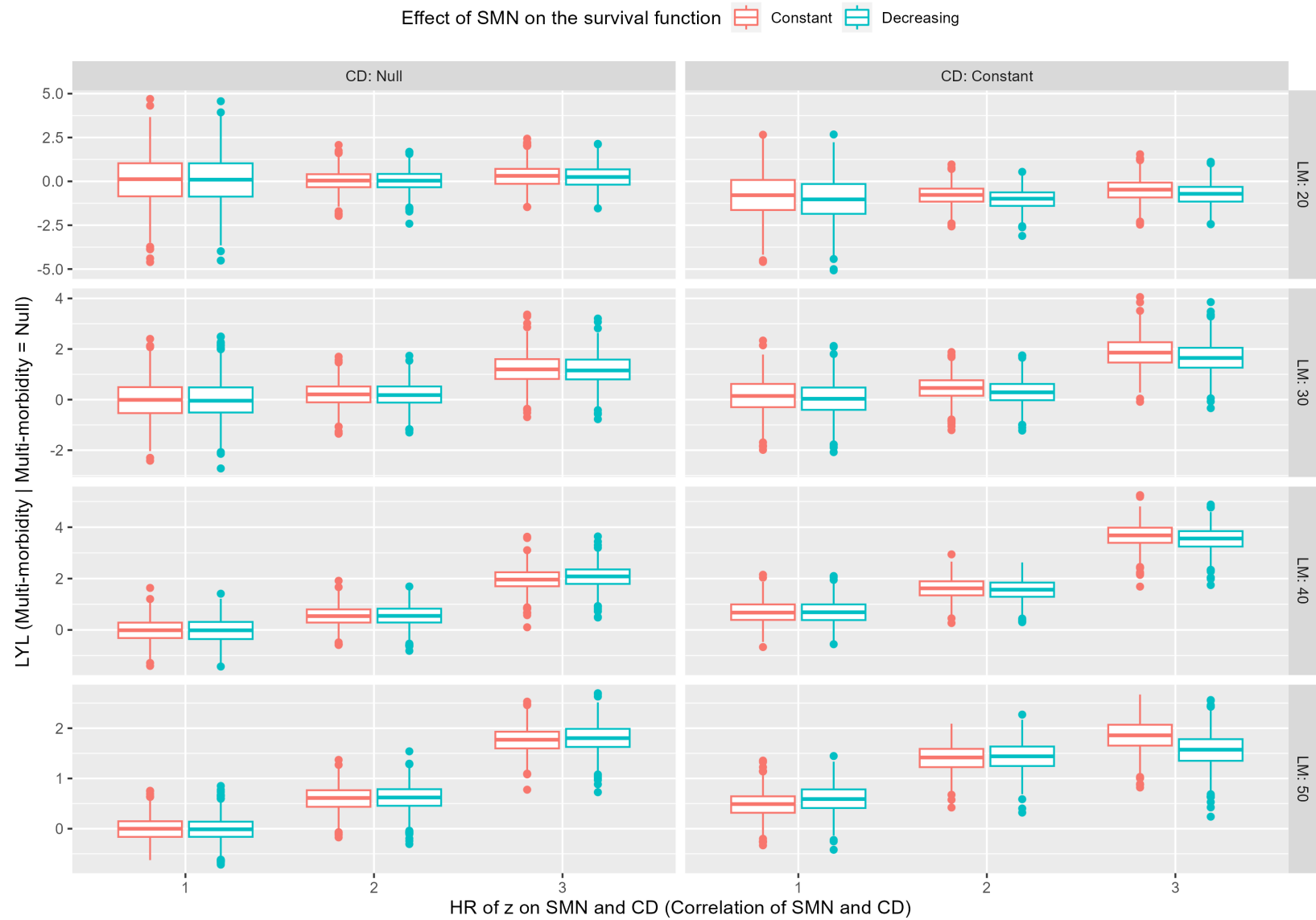
though in most scenarios the LYL by SMN increase as the correlation between SMN and CD increases, the results at the highest landmark time ( $l = 50$ ) don't conform to this and instead show a bell shape, increasing at first and then decreasing. In most scenarios the LYL by SMN are increased by the inclusion of a multi-morbidity effect on survival, but this result notably doesn't hold at landmark time  $l = 40$  when the correlation between SMN and CD is due to a hazard ratio of 4.

### **6.3.2 . Estimates of Life Years Lost by Multi-morbidity**

We summarize our estimated life years lost by multi-morbidity in Figure 6.5 and Figure 6.6 which correspond to the scenarios where the true effect of multi-morbidity on survival is respectively null and constant. We see the shape of the results are similar for both configurations. The values are different, as expected by the addition of a multi-morbidity effect on the survival function.

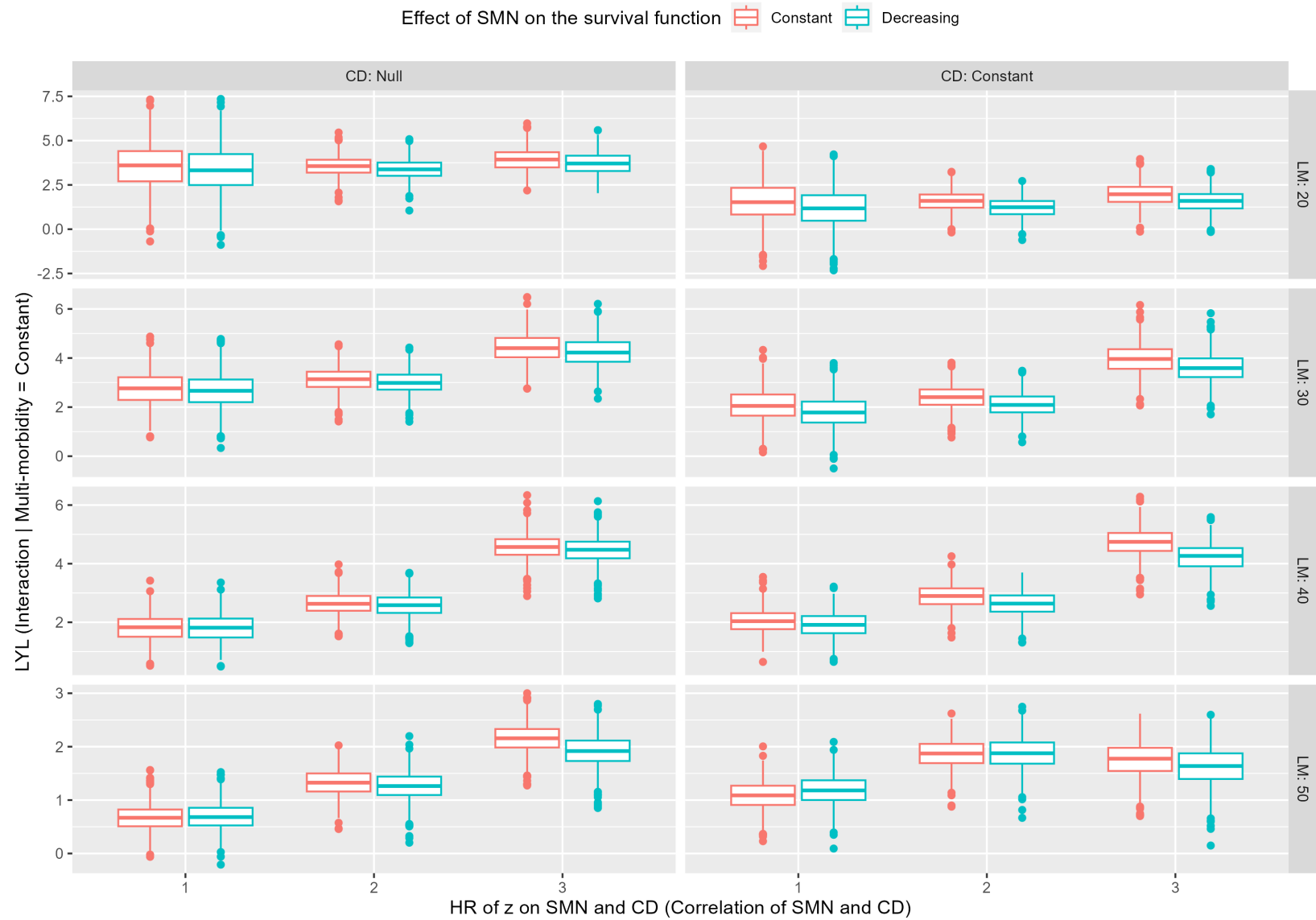
Under a null effect of CD and multi-morbidity on the survival, and independence of SMN and CD, we estimate 0 LYL by multi-morbidity at all landmark times. We note that under independence of SMN and CD and null effect of multi-morbidity on the survival function but non-null effect of CD on the survival, the LYL by multi-morbidity are greater than 0 for landmark times 40, 50 and smaller than 0 for landmark time 20. We observe an increase in the LYL by multi-morbidity as the strength of the correlation between SMN and CD increases.

Under all scenarios, we observe a small difference in LYL by multi-morbidity between the scenarios where SMN effect on survival is constant and the ones where SMN effect on survival is decreasing. This difference appears to slightly increase as the strength of the correlation between CD and SMN increases.



**Figure 6.5:** Life Years Lost by interaction of SMN and CD, when adjusting for SMN and CD, when the true effect of this interaction on survival is null. Columns correspond to the true effect of CD on survival. Rows correspond to the landmark time at which estimation is done (20, 30, 40, 50).





**Figure 6.6:** Life Years Lost by interaction of SMN and CD, when adjusting for SMN and CD, when the true effect of this interaction on survival is constant. Columns correspond to the true effect of CD on survival. Rows correspond to the landmark time at which estimation is done (20, 30, 40, 50).

## 6.4 . Expected coefficients

### 6.4.1 . Landmark time = 0

We start by using time-fixed covariates to observe how our estimator behaves at landmark time  $l = 0$ . Let  $a, b$  be binary time-fixed covariates, and  $a : b$  their interaction. We are interested in the LYL by  $a : b$  up to  $\tau$ . In order to have 0 LYL by  $a : b$  for all  $\tau$ , we must have for all time  $t \in [0, \tau]$  that

$$\begin{aligned} S(t|a = 0, b = 0) - S(t|a = 1, b = 1) = & (S(t|a = 0, b = 0) - S(t|a = 1, b = 0)) + \\ & (S(t|a = 0, b = 0) - S(t|a = 0, b = 1)) \end{aligned}$$

Let  $S_0(t) = S(t|a = 0, b = 0)$ ,  $\beta_a(t) = S_0(t) - S(t|a = 1, b = 0)$  and  $\beta_b(t) = S_0(t) - S(t|a = 0, b = 1)$ .

Then, we have 0 LYL by  $a : b$  for all  $\tau$  if:

$$S(t|a = 1, b = 1) = S_0(t) - \beta_a(t) - \beta_b(t)$$

Therefore, once the independent effects of  $a$  and  $b$  on  $S(t)$  are set, there is only solution such as  $LYL^\tau(a : b) = 0$ .

### 6.4.2 . Landmark time > 0

Using the survival functions above, we show that when using a landmark time  $l > 0$ , the equality  $LYL_l^\tau(a : b) = 0$  holds only under a small restrictive sets of assumptions. Once more, the equality holds for all  $\tau$  if and only if for all  $t \in [l, \tau]$ , we have

$$S_0(t|T > l) - S(t|T > l, a = 1, b = 1) = S_0(t|T > l) - S(t|T > l, a = 1, b = 0) + \\ S_0(t|T > l) - S(t|T > l, a = 0, b = 0)$$

Under our scenario with constant effect of  $a, b$  (ie.  $\beta_a(t) = \gamma_a t, \beta_b(t) = \gamma_b t$ ), this equality holds only if one of the following equality holds:

$$\left\{ \begin{array}{l} \gamma_a = 0 \\ \gamma_b = 0 \\ l = 0 \\ S_0(t) = \frac{t}{l} S_0(l) \\ S_0(l) = \frac{l}{2} (\gamma_a + \gamma_b) \end{array} \right.$$

Those equalities are either unlikely, too restrictive for our purpose, or impossible to achieve:

1. The first two ones correspond to the case where one of the covariates has no effect on the survival function.
2. The third case to landmark time  $l = 0$  discussed above.
3. The fourth one requires an increasing survival function and is therefore impossible.

4. The fifth one implies an increase of  $S_0(l)$  as  $l$  increases, and can therefore only be achieved at a single landmark time point.

Those results explain why our estimated LYL by multi-morbidity is non-null when the effect of multi-morbidity on survival is null but SMN and CD have a non-null effect on survival in Figure 6.6.

## **6.5 . Discussion**

### **6.5.1 . Preliminary Results**

In this undergoing work, our preliminary results show that our estimated life years lost by multi-morbidity are affected by the correlation between SMN and CD, and by the respective effects of SMN and CD on the survival function. The changes in LYL by multi-morbidity by the introduction of a time-varying coefficient of SMN and/or CD are small to non-existent.

### **6.5.2 . Design of the Simulations**

We simulated data using an additive model on the survival function. We did so to control the life years lost by each covariate. We showed in Section 6.4 that this control doesn't hold once we use a landmark method. We also showed that this design shifts the LYL by multi-morbidity in an unexpected way when both SMN and CD have an effect on the survival function. This shift could be partially or totally responsible for the differences observed between the scenarios where SMN effect on survival is constant and the ones where it is decreasing. Therefore, our design seems inappropriate to isolate the effect of interest, but we showed that its importance is small relatively to other factors.

Of note, we designed the scenarios such as using the last observation carried forward at landmark time  $l$  yields the correct value of our time-dependent covariates at all times  $t > l$ . In this way, we neutralized a source of bias in our estimation to be able to limit the number of out of scope effects in our results.

### **6.5.3 . Work in Progress**

As is stated at the start of this section, this work is actively undergoing. Therefore, we state here further simulation scenarios and analysis we believe to be of interest for further work.

The central hypothesis justifying this work is the difference in time since SMN and landmark time between the population exposed to CD and the one unexposed. Our scenarios induce such a difference under correlation between SMN and CD, but further control of this could lead to settings closer to the one observed in our data. Furthermore, including an effect of SMN on the risk of CD would be of interest to create those scenarios.

Our scenarios of simulation imply an additive effect on the survival, which implies an increasing effect on the hazard as time increases. Therefore, our scenario of decreasing effect on the survival isn't a scenario of decreasing effect on the hazard as time since SMN/CD diagnosis increases. Framing our simulation model on the hazard could be valuable to be closer to real data behavior.

## **6.6 . Summary**

In conclusion, our preliminary results confirmed that a correlation between two time-dependent covariates, changes the estimated life years lost by their interaction when using a landmark strategy. We showed a small effect of a time-varying coefficient of SMN on the life years lost by multi-morbidity, but this effect is very small when compared to the one induced by the correlation between SMN and CD.

## 7 - General Discussion and Conclusion

This chapters summarizes the main findings of this thesis, their strengths, limitations, clinical implications, and finally proposes future research.

### 7.1 . Main Findings

Childhood Cancer Survivors are at long-term risk of many iatrogenic events of childhood cancer treatment. They had also been identified as at high risk of developing multiple iatrogenic events. In this thesis, we investigated the consequences of a first iatrogenic event occurrence on the health of patients, comparing them to morbidity-free childhood cancer survivors.

We presented three studies conducted using the FCCSS cohort, which includes 7,670 5-year french childhood cancer survivors treated between 1945 and 2000. In those studies, (i) we quantified the shift in absolute risk of cardiac disease after a second malignant neoplasm occurrence; (ii) we quantified the life years lost by childhood cancer treatment, second malignant neoplasm, cardiac disease, diabetes, chronic renal failure, and two ways interaction of those iatrogenic events; and (iii) we investigated how our choice to use the landmark method and to dichotomize the effect of iatrogenic event could be responsible for us finding no life years lost by two ways interaction of iatrogenic events.

Our main findings can be summarized as the following 4 key points:

1. *Second Malignant Neoplasm Occurrence Increases the Absolute Risk of Cardiac Disease.*

In Section 4 we found that the occurrence of a Second Malignant Neoplasm multiplies the instantaneous risk of Cardiac Disease by 2-fold (95% CI: 1.5 - 2.9), and increases its cumulative incidence by 3.8% (95% CI: 0.5 - 7.1) conditional on survival 25 years after diagnosis. We found those effects to depend on the type of Second Malignant Neoplasm,

with those usually treated using cardio-toxic treatments leading to a bigger increase – 2.3 fold (95% CI: 1.2 - 4.4) for breast cancer, 1.7 fold (95% CI: 1.0 - 3.0) for sarcoma, bone, and soft tissue cancer, and 1.4 fold (95% CI: 1.1 - 1.9) for other cancers.

2. *Life Years Lost by Childhood Cancer Treatments Can Mostly Be Attributed to Radiotherapy.* In Section 5 we found 6.0 (95% CI: 4.7 - 7.3) life years lost by radiotherapy at 16 years old, and only 0.7 (95% CI: -0.0 - 1.4) life years lost by chemotherapy. We found some differences based on age at childhood cancer diagnosis, with survivors diagnosed between 1 and 5 years old faring the worst, with 1.2 (95% CI: 0.4 - 2.0) life years lost at 16 years old, compared to survivors diagnosed before 1 year old.
3. *Second Malignant Neoplasms and Cardiac Diseases Contribute to Many Life Years Lost.* In Section 5 we found respectively 10.5 (95% CI: 5.4 - 15.7) and 7.7 (95% CI: 1.1 - 14.3) life years lost by Second Malignant Neoplasms and Cardiac Diseases before 16 years old after adjusting for childhood cancer treatments. Therefore, we conclude that a relevant part of the early mortality experienced by childhood cancer survivors can be attributed to the occurrence of those iatrogenic events. We found no life years lost by diabetes and chronic renal failure, and therefore conclude that mortality by iatrogenic event must be studied going into the details of the type of iatrogenic event.
4. *Occurrence of Both Second Malignant Neoplasm and Cardiac Disease Doesn't Reduce Life Expectancy Further.* In Section 5 we found no life years lost by two-ways interaction of Second Malignant Neoplasm and Cardiac Disease. Therefore, we conclude that childhood cancer survivors life years lost by SMN and CD can be summarized as the sum of life years lost by each one alone.

## 7.2 . Strengths and Limitations

A major strength of this thesis is the access to the FCCSS, a well-defined cohort of childhood cancer survivors with both long-term follow-up of patients and detailed information regarding childhood cancer treatment. This allowed us to investigate morbidities occurrence over a long period of time and to adjust for confounding by childhood cancer treatment.

Another strength is the use of regression models on probability and time scales. The use of those scales, as opposed to the used of the cause specific hazard, allowed us to answer questions relevant in a clinical and public health setting. For instance, we were able to answer the question “Is the proportion of Cardiac Diseases observed higher among childhood cancer survivors who experienced a Second Malignant Neoplasm compared to those who didn’t ?” by the affirmative. Whereas results on the cause-specific hazard scale can’t be used to answer this question. Thus, we provided results interpretable for both public health and patient care.

Some limitations of this thesis merit discussion. The main limitation of the contributions are related to data available.

First, we didn’t study the treatments of second malignant neoplasms. Although we did study the types of second malignant neoplasms, which is a proxy for the treatments. Therefore, our results are limited to a general effect of second malignant neoplasms, despite having showed differences based on the type of cancer. In practice, the treatments used for adult cancer are often known and used to adapt patients follow-up, hence why investigations on the consequences of those treatments would have provided valuable information for clinical care.

Second, we didn’t adjust for smoking status and socio-economic status which are both associated with occurrence of second malignant neoplasms and cardiac diseases, because data was missing for most patients. This confounding is likely to have caused some overestimation of our estimates, and could be the reason why we found a cardio-toxic effect of the second malignant neoplasms types not associated with cardio-toxic treatments. Conducting a mediation analy-



sis would have been valuable to assess the importance of smoking and socio-economic status and conclude on the existence of a baseline cardio-toxicity of subsequent malignant neoplasm. However, we didn't have access to the data required for doing it.

Third, we conducted our analysis of life years lost using iatrogenic events occurrence before a pre-specified landmark time. This leads to an underestimation of the life years lost by iatrogenic events, especially by second malignant neoplasm whose mortality occurs shortly after diagnosis. Those results are also different than life years lost by cause, which is often used in public health settings because it can be used to decompose the mortality of patients. Hence, those results are not directly comparable to the preferred life years lost by cause.

Finally, we limited ourselves to the study of multi-morbidity as the occurrence of two iatrogenic events. Therefore, we only provided a succinct glimpse into the multi-morbidity patterns experienced by childhood cancer survivors. Furthermore, this investigation was restricted to very severe consequences, and didn't look into less severe iatrogenic events whose accumulation could severely affect the quality of life of patients. We did so to have enough statistical power. Even with the large cohort size, the number of some types of events were low and including them would have compromised the statistical power of our analysis.

### **7.3 . Clinical Implications**

The results of this thesis provide new evidence that childhood cancer survivors need appropriate long-term follow-up that is re-updated on occurrence of subsequent malignant neoplasm. This refinement needs to include treatment data for both childhood cancer and adult cancers, as we showed that both significantly contribute to cardiac risk.

Our results also imply that occurrence of a subsequent malignant neoplasm and cardiac disease have a substantial impact on survivors' life expectancy. We also showed that those patients remain at increased risk more than 5 years since the iatrogenic event diagnosis, implying the

need for greater supervision of this frail population.

#### **7.4 . Future Research Perspectives**

Further work could try to overcome those limitations. First the acquisition of second malignant neoplasm treatment data would be valuable, and could be used to expand both the first and second contribution for valuable clinical and public health insight about the causality of the mechanisms at play. Second, the use of bigger cohorts such as Pancare and the FCCSS to run the analysis for both reproducibility and to have a higher number of events could be done to achieve a better precision in our estimates.

#### **7.5 . Conclusions**

In conclusion, we have promoted additive regression models, albeit rarely used, to a well-defined cohort of childhood cancer survivors, and provided a first glance into the multi-morbidity of those patients and its consequences. We also investigated the property of the landmark method in our scenario with time-varying coefficients and time-dependent covariates. We are confident this work can be used both clinically and to initiate further research on the topic of multi-morbidity among childhood cancer survivors.



## References

- Aalen, Odd O., and Søren Johansen. 1978. “An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations.” *Scandinavian Journal of Statistics* 5 (3): 141–50. <https://doi.org/10.2307/4615704>.
- Aba, N., S. Belhechmi, B. Fresneau, C. El-Fayech, C. Rubino, R. Allodji, E. Morel, et al. 2023. “CO12.1 - A Case-Control Study to Identify Potential Genetic Biomarkers Related to Cardiac Diseases Occurrence in Childhood Cancer Survivors.” *Revue d'Épidémiologie Et de Santé Publique*, EPICLIN 202317e Conférence francophone d'épidémiologie Clinique30e Journées des statisticiens des Centres de Lutte contre le Cancer, 71 (May): 101634. <https://doi.org/10.1016/j.respe.2023.101634>.
- Allodji, Rodrigue S., Nadia Haddy, Giao Vu-Bezin, Agnès Dumas, Brice Fresneau, Imene Mansouri, Charlotte Demoor-Goldschmidt, et al. 2019. “Risk of Subsequent Colorectal Cancers After a Solid Tumor in Childhood: Effects of Radiation Therapy and Chemotherapy.” *Pediatric Blood & Cancer* 66 (2): e27495. <https://doi.org/10.1002/pbc.27495>.
- Allodji, Rodrigue S., Boris Schwartz, Cristina Veres, Nadia Haddy, Carole Rubino, Marie-Cécile Le Deley, Martine Labbé, et al. 2015. “Risk of Subsequent Leukemia After a Solid Tumor in Childhood: Impact of Bone Marrow Radiation Therapy and Chemotherapy.” *International Journal of Radiation Oncology\*Biophysics* 93 (3): 658–67. <https://doi.org/10.1016/j.ijrobp.2015.07.2270>.
- Ambrose, John A., and Rajat S. Barua. 2004. “The Pathophysiology of Cigarette Smoking and Cardiovascular Disease: An Update.” *Journal of the American College of Cardiology* 43 (10): 1731–37. <https://doi.org/10.1016/j.jacc.2003.12.047>.
- Andersen, P. K. 2013. “Decomposition of Number of Life Years Lost According to Causes of Death.” *Statistics in Medicine* 32 (30): 5278–85. <https://doi.org/10.1002/sim.5903>.

- Andersen, Per Kragh. 2017. "Life Years Lost Among Patients with a Given Disease." *Statistics in Medicine* 36 (22): 3573–82. <https://doi.org/10.1002/sim.7357>.
- Andersen, Per Kragh, Vladimir Canudas-Romo, and Niels Keiding. 2013. "Cause-Specific Measures of Life Years Lost." *Demographic Research* 29 (41): 1127–52. <https://doi.org/10.4054/DemRes.2013.29.41>.
- Andersen, Per Kragh, John P. Klein, and Susanne Rosthøj. 2003. "Generalised Linear Models for Correlated Pseudo-observations, with Applications to Multi-state Models." *Biometrika* 90 (1): 15–27. <https://doi.org/10.1093/biomet/90.1.15>.
- Andersen, Per Kragh, and Maja Pohar Perme. 2010. "Pseudo-Observations in Survival Analysis." *Statistical Methods in Medical Research* 19 (1): 71–99. <https://doi.org/10.1177/0962280209105020>.
- Anderson, J. R., K. C. Cain, and R. D. Gelber. 1983. "Analysis of Survival by Tumor Response." *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 1 (11): 710–19. <https://doi.org/10.1200/JCO.1983.1.11.710>.
- Armstrong, Gregory T., Yan Chen, Yutaka Yasui, Wendy Leisenring, Todd M. Gibson, Ann C. Mertens, Marilyn Stovall, et al. 2016. "Reduction in Late Mortality Among 5-Year Survivors of Childhood Cancer." *The New England Journal of Medicine* 374 (9): 833–42. <https://doi.org/10.1056/NEJMoa1510795>.
- Armstrong, Gregory T., Qi Liu, Yutaka Yasui, Joseph P. Neglia, Wendy Leisenring, Leslie L. Robison, and Ann C. Mertens. 2009. "Late Mortality Among 5-Year Survivors of Childhood Cancer: A Summary from the Childhood Cancer Survivor Study." *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 27 (14): 2328–38. <https://doi.org/10.1200/JCO.2008.21.1425>.
- Austin, Peter C., and Jason P. Fine. 2017. "Practical Recommendations for Reporting Fine-Gray Model Analyses for Competing Risk Data." *Statistics in Medicine* 36 (27): 4391–4400. <https://doi.org/10.1002/sim.7501>.

- Austin, Peter C., Aurélien Latouche, and Jason P. Fine. 2020. “A Review of the Use of Time-Varying Covariates in the Fine-Gray Subdistribution Hazard Competing Risk Regression Model.” *Statistics in Medicine* 39 (2): 103–13. <https://doi.org/10.1002/sim.8399>.
- Bates, James E., Rebecca M. Howell, Qi Liu, Yutaka Yasui, Daniel A. Mulrooney, Sughosh Dhakal, Susan A. Smith, et al. 2019. “Therapy-Related Cardiac Risk in Childhood Cancer Survivors: An Analysis of the Childhood Cancer Survivor Study.” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 37 (13): 1090–1101. <https://doi.org/10.1200/JCO.18.01764>.
- Bejarano-Quisoboni, Daniel, Henri Panjo, Brice Fresneau, Chiraz El-Fayech, François Doz, Aurore Surun, Florent de Vathaire, and Nathalie Pelletier-Fleury. 2023. “Excess Healthcare Expenditure in Adults Treated for Solid Cancer in Childhood: A Cohort Study in France.” *The European Journal of Health Economics*, June. <https://doi.org/10.1007/s10198-023-01606-6>.
- Bejarano-Quisoboni, Daniel, Nathalie Pelletier-Fleury, Rodrigue S. Allodji, Brice Fresneau, Majorie Boussac, Hélène Pacquement, François Doz, et al. 2022. “Long-Term Hospitalisations in Survivors of Paediatric Solid Tumours in France.” *Scientific Reports* 12 (1): 18068. <https://doi.org/10.1038/s41598-022-22689-w>.
- Beyersmann, Jan, and Martin Schumacher. 2008. “Time-Dependent Covariates in the Proportional Subdistribution Hazards Model for Competing Risks.” *Biostatistics* 9 (4): 765–76. <https://doi.org/10.1093/biostatistics/kxn009>.
- Bhakta, Nickhill, Qi Liu, Kirsten K Ness, Malek Baassiri, Hesham Eissa, Frederick Yeo, Wassim Chemaitilly, et al. 2017. “The Cumulative Burden of Surviving Childhood Cancer: An Initial Report from the St Jude Lifetime Cohort Study (SJLIFE).” *The Lancet* 390 (10112): 2569–82. [https://doi.org/10.1016/S0140-6736\(17\)31610-0](https://doi.org/10.1016/S0140-6736(17)31610-0).
- Borgan, Ørnulf. 2005. “Kaplan–Meier Estimator.” In *Encyclopedia of Biostatistics*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/0470011815.b2a11042>.

- Botta, Laura, Gemma Gatta, Riccardo Capocaccia, Charles Stiller, Adela Cañete, Luigino Dal Maso, Kaire Innos, et al. 2022. “Long-Term Survival and Cure Fraction Estimates for Childhood Cancer in Europe (EUROCORE-6): Results from a Population-Based Study.” *The Lancet Oncology*, November. [https://doi.org/10.1016/S1470-2045\(22\)00637-4](https://doi.org/10.1016/S1470-2045(22)00637-4).
- Boudoulas, Konstantinos Dean, Filippou Triposkiadis, Richard Gumina, Daniel Addison, Cezar Iliescu, and Harisios Boudoulas. 2022. “Cardiovascular Disease, Cancer and Multimorbidity Interactions: Clinical Implications.” *Cardiology*, January. <https://doi.org/10.1159/000521680>.
- Chang, Wai Hoong, Michail Katsoulis, Yen Yi Tan, Stefanie H. Mueller, Katherine Green, and Alvina G. Lai. 2022. “Late Effects of Cancer in Children, Teenagers and Young Adults: Population-Based Study on the Burden of 183 Conditions, in-Patient and Critical Care Admissions and Years of Life Lost.” *The Lancet Regional Health - Europe* 12 (January): 100248. <https://doi.org/10.1016/j.lanepe.2021.100248>.
- Charrier, Thibaud, Nadia Haddy, Brice Fresneau, Boris Schwartz, Neige Journy, Charlotte Demoor-Goldschmidt, Ibrahima Diallo, et al. 2024. “Life Years Lost by Childhood Cancer Treatment and Health Related Late Effects Among Childhood Cancer Survivors.” *Cancer Epidemiology* 93 (December): 102692. <https://doi.org/10.1016/j.canep.2024.102692>.
- Charrier, Thibaud, Nadia Haddy, Boris Schwartz, Neige Journy, Brice Fresneau, Charlotte Demoor-Goldschmidt, Ibrahima Diallo, et al. 2023. “Increased Cardiac Risk After a Second Malignant Neoplasm Among Childhood Cancer Survivors, a FCCSS Study.” *JACC: CardioOncology*, October. <https://doi.org/10.1016/j.jaccao.2023.07.008>.
- Chounta, Stefania, Rodrigue Allodji, Maria Vakalopoulou, Mahmoud Bentriou, Duyen Thi Do, Florent De Vathaire, Ibrahima Diallo, et al. 2023. “Dosiomics-Based Prediction of Radiation-Induced Valvulopathy After Childhood Cancer.” *Cancers* 15 (12): 3107. <https://doi.org/10.3390/cancers15123107>.
- Chounta, Stefania, Sarah Lemler, Nadia Haddy, Brice Fresneau, Imene Mansouri, Mahmoud

- Bentriou, Charlotte Demoor-Goldschmidt, et al. 2023. “The Risk of Valvular Heart Disease in the French Childhood Cancer Survivors’ Study: Contribution of Dose-Volume Histogram Parameters.” *Radiotherapy and Oncology*, January, 109479. <https://doi.org/10.1016/j.radonc.2023.109479>.
- Chow, Eric J., Yan Chen, Melissa M. Hudson, Elizabeth A. M. Feijen, Leontien C. Kremer, William L. Border, Daniel M. Green, et al. 2018. “Prediction of Ischemic Heart Disease and Stroke in Survivors of Childhood Cancer.” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 36 (1): 44–52. <https://doi.org/10.1200/JCO.2017.74.8673>.
- Chow, Eric J., Yan Chen, Leontien C. Kremer, Norman E. Breslow, Melissa M. Hudson, Gregory T. Armstrong, William L. Border, et al. 2015. “Individual Prediction of Heart Failure Among Childhood Cancer Survivors.” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 33 (5): 394–402. <https://doi.org/10.1200/JCO.2014.56.1373>.
- Cortese, Giuliana, and Per K. Andersen. 2010. “Competing Risks and Time-Dependent Covariates.” *Individual Prediction of Heart Failure Among Childhood Cancer Survivor* 52 (1): 138–58. <https://doi.org/10.1002/bimj.200900076>.
- Cox, D. R. 1972. “Regression Models and Life-Tables.” *Journal of the Royal Statistical Society: Series B (Methodological)* 34 (2): 187–202. <https://doi.org/10.1111/j.2517-6161.1972.tb00899.x>.
- Davies, Laurie E., Stewart W. Mercer, Katie Brittain, Carol Jagger, Louise Robinson, and Andrew Kingston. 2022. “The Association Between Multimorbidity and Mobility Disability-Free Life Expectancy in Adults Aged 85 Years and over: A Modelling Study in the Newcastle 85+ Cohort.” *PLOS Medicine* 19 (11): e1004130. <https://doi.org/10.1371/journal.pmed.1004130>.
- Dixon, Stephanie B., Qi Liu, Eric J. Chow, Kevin C. Oeffinger, Paul C. Nathan, Rebecca



- M. Howell, Wendy M. Leisenring, et al. 2023. “Specific Causes of Excess Late Mortality and Association with Modifiable Risk Factors Among Survivors of Childhood Cancer: A Report from the Childhood Cancer Survivor Study Cohort.” *The Lancet* 0 (0). [https://doi.org/10.1016/S0140-6736\(22\)02471-0](https://doi.org/10.1016/S0140-6736(22)02471-0).
- Dong, Huiru, Leslie L. Robison, Wendy M. Leisenring, Leah J. Martin, Gregory T. Armstrong, and Yutaka Yasui. 2015. “Estimating the Burden of Recurrent Events in the Presence of Competing Risks: The Method of Mean Cumulative Count.” *American Journal of Epidemiology* 181 (7): 532–40. <https://doi.org/10.1093/aje/kwu289>.
- Ducos, C., F. Brice, F. Rosselli, G. Vu-Bezin, B. Schwartz, R. Allodji, G. Marenne, et al. 2023. “Identification of Rare Variants Involved in the Risk of Second Cancer Following Radiotherapy and Chemotherapy Treatment of Pediatric Cancers.” *Revue d’Épidémiologie Et de Santé Publique*, EPICLIN 202317e Conférence francophone d’épidémiologie Clinique30e Journées des statisticiens des Centres de Lutte contre le Cancer, 71 (May): 101799. <https://doi.org/10.1016/j.respe.2023.101799>.
- Dumas, A., I. Cailbault, C. Perrey, O. Oberlin, F. De Vathaire, and P. Amiel. 2015. “Educational Trajectories After Childhood Cancer: When Illness Experience Matters.” *Social Science & Medicine* 135 (June): 67–74. <https://doi.org/10.1016/j.socscimed.2015.04.031>.
- Dumas, Agnès, Florent De Vathaire, and Gilles Vassal. 2016. “Access to Loan-Related Insurance for French Cancer Survivors.” *The Lancet Oncology* 17 (10): 1354–56. [https://doi.org/10.1016/S1470-2045\(16\)30452-1](https://doi.org/10.1016/S1470-2045(16)30452-1).
- Ehrhardt, Matthew J., Kevin R. Krull, Nickhill Bhakta, Qi Liu, Yutaka Yasui, Leslie L. Robison, and Melissa M. Hudson. 2023. “Improving Quality and Quantity of Life for Childhood Cancer Survivors Globally in the Twenty-First Century.” *Nature Reviews Clinical Oncology*, July, 1–19. <https://doi.org/10.1038/s41571-023-00802-w>.
- Ernst, Mareike, Andreas Hinz, Elmar Brähler, Hiltrud Merzenich, Jörg Faber, Philipp S. Wild,

- and Manfred E. Beutel. 2023. “Quality of Life After Pediatric Cancer: Comparison of Long-Term Childhood Cancer Survivors’ Quality of Life with a Representative General Population Sample and Associations with Physical Health and Risk Indicators.” *Health and Quality of Life Outcomes* 21 (1): 65. <https://doi.org/10.1186/s12955-023-02153-7>.
- Feijen, Elizabeth A. M., Wendy M. Leisenring, Kayla L. Stratton, Kirsten K. Ness, Helena J. H. van der Pal, Elvira C. van Dalen, Gregory T. Armstrong, et al. 2019. “Derivation of Anthracycline and Anthraquinone Equivalence Ratios to Doxorubicin for Late-Onset Cardiotoxicity.” *JAMA Oncology* 5 (6): 864–71. <https://doi.org/10.1001/jamaoncol.2018.6634>.
- Feller, William. 1991. “An Introduction to Probability Theory and Its Applications, Volume 2, 2nd Edition | Wiley.” *Wiley.com*. <https://www.wiley.com/en-us/An+Introduction+to+Probability+Theory+and+Its+Applications%2C+Volume+2%2C+2nd+Edition-p-9780471257097>.
- Fine, Jason P., and Robert J. Gray. 1999. “A Proportional Hazards Model for the Subdistribution of a Competing Risk.” *Journal of the American Statistical Association* 94 (446): 496–509. <https://doi.org/10.1080/01621459.1999.10474144>.
- Gatta, Gemma, Laura Botta, Silvia Rossi, Tiiu Aareleid, Magdalena Bielska-Lasota, Jacqueline Clavel, Nadya Dimitrova, et al. 2014. “Childhood Cancer Survival in Europe 1999–2007: Results of EURO CARE-5—a Population-Based Study.” *The Lancet. Oncology* 15 (1): 35–47. [https://doi.org/10.1016/S1470-2045\(13\)70548-5](https://doi.org/10.1016/S1470-2045(13)70548-5).
- Gerds, Thomas A., Thomas H. Scheike, and Per K. Andersen. 2012. “Absolute Risk Regression for Competing Risks: Interpretation, Link Functions, and Prediction.” *Statistics in Medicine* 31 (29): 3921–30. <https://doi.org/10.1002/sim.5459>.
- Gibson, Todd M, Sogol Mostoufi-Moab, Kayla L Stratton, Wendy M Leisenring, Dana Barnea, Eric J Chow, Sarah S Donaldson, et al. 2018. “Temporal Patterns in the Risk of Chronic Health Conditions in Survivors of Childhood Cancer Diagnosed 1970–99: A Report from the Childhood Cancer Survivor Study Cohort.” *The Lancet Oncology* 19 (12): 1590–1601.

[https://doi.org/10.1016/S1470-2045\(18\)30537-0](https://doi.org/10.1016/S1470-2045(18)30537-0).

- Grambsch, Patricia M., and Terry M. Therneau. 1994. "Proportional Hazards Tests and Diagnostics Based on Weighted Residuals." *Biometrika* 81 (3): 515–26. <https://doi.org/10.1093/biomet/81.3.515>.
- Grand, Mia Klinten, and Hein Putter. 2016. "Regression Models for Expected Length of Stay." *Statistics in Medicine* 35 (7): 1178–92. <https://doi.org/10.1002/sim.6771>.
- Grand, Mia K., Hein Putter, Arthur Allignol, and Per K. Andersen. 2019. "A Note on Pseudo-Observations and Left-Truncation." *Biometrical Journal* 61 (2): 290–98. <https://doi.org/10.1002/bimj.201700274>.
- Graw, Frederik, Thomas A. Gerds, and Martin Schumacher. 2009. "On Pseudo-Values for Regression Analysis in Competing Risks Models." *Lifetime Data Analysis* 15 (2): 241–55. <https://doi.org/10.1007/s10985-008-9107-z>.
- Gray, Robert J. 1988. "A Class of  $K$ -Sample Tests for Comparing the Cumulative Incidence of a Competing Risk." *The Annals of Statistics* 16 (3): 1141–54. <https://doi.org/10.1214/aos/1176350951>.
- Groot, Patricia de, and Reginald F. Munden. 2012. "Lung Cancer Epidemiology, Risk Factors, and Prevention." *Radiologic Clinics of North America* 50 (5): 863–76. <https://doi.org/10.1016/j.rcl.2012.06.006>.
- Haddy, Nadia, Stéphanie Diallo, Chiraz El-Fayech, Boris Schwartz, François Pein, Mike Hawkins, Cristina Veres, et al. 2016. "Cardiac Diseases Following Childhood Cancer Treatment: Cohort Study." *Circulation* 133 (1): 31–38. <https://doi.org/10.1161/CIRCULATIONAHA.115.016686>.
- Heckman, James J. 2001. "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture." *Journal of Political Economy* 109 (4): 673–748. <https://doi.org/10.1086/322086>.
- Heckman, James J., and George J. Borjas. 1980. "Does Unemployment Cause Future Unem-

- ployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence.” *Economica* 47 (187): 247–83. <https://doi.org/10.2307/2553150>.
- Hjorth, Lars, Riccardo Haupt, Roderick Skinner, Desiree Grabow, Julianne Byrne, Sabine Karner, Gill Levitt, et al. 2015. “Survivorship After Childhood Cancer: PanCare: A European Network to Promote Optimal Long-Term Care.” *European Journal of Cancer* 51 (10): 1203–11. <https://doi.org/10.1016/j.ejca.2015.04.002>.
- Houwelingen, Hans C. van, and Hein Putter. 2015. “Comparison of Stopped Cox Regression with Direct Methods Such as Pseudo-Values and Binomial Regression.” *Lifetime Data Analysis* 21 (2): 180–96. <https://doi.org/10.1007/s10985-014-9299-3>.
- Hudson, Melissa M., Smita Bhatia, Jacqueline Casillas, and Wendy Landier. 2021. “Long-Term Follow-up Care for Childhood, Adolescent, and Young Adult Cancer Survivors.” *Pediatrics* 148 (3): e2021053127. <https://doi.org/10.1542/peds.2021-053127>.
- Jaworski, Catherine, Justin A. Mariani, Greg Wheeler, and David M. Kaye. 2013. “Cardiac Complications of Thoracic Irradiation.” *Journal of the American College of Cardiology* 61 (23): 2319–28. <https://doi.org/10.1016/j.jacc.2013.01.090>.
- Johnston, W. T., Friederike Erdmann, Robert Newton, Eva Steliarova-Foucher, Joachim Schüz, and Eve Roman. 2021. “Childhood Cancer: Estimating Regional and Global Incidence.” *Cancer Epidemiology*, *Childhood Cancer: A Global Perspective*, 71 (April): 101662. <https://doi.org/10.1016/j.canep.2019.101662>.
- Kalbfleisch, John D., and Ross L. Prentice. 2002. *The Statistical Analysis of Failure Time Data*. 1st ed. Wiley Series in Probability and Statistics. Wiley. <https://doi.org/10.1002/9781118032985>.
- Kaplan, E. L., and Paul Meier. 1958. “Nonparametric Estimation from Incomplete Observations.” *Journal of the American Statistical Association* 53 (282): 457–81. <https://doi.org/10.1080/01621459.1958.10501452>.

- Kenneth Hess, S original by, and R port by R. Gentleman. 2021. *Muhaz: Hazard Function Estimation in Survival Analysis*. <https://CRAN.R-project.org/package=muhaz>.
- Kessing, L. V., E. W. Olsen, and P. K. Andersen. 1999. “Recurrence in Affective Disorder: Analyses with Frailty Models.” *American Journal of Epidemiology* 149 (5): 404–11. <https://doi.org/10.1093/oxfordjournals.aje.a009827>.
- Klein, John P. 2006. “Modelling Competing Risks in Cancer Studies.” *Statistics in Medicine* 25 (6): 1015–34. <https://doi.org/10.1002/sim.2246>.
- Klein, John P., and Per Kragh Andersen. 2005. “Regression Modeling of Competing Risks Data Based on Pseudovalues of the Cumulative Incidence Function.” *Biometrics* 61 (1): 223–29. <https://doi.org/10.1111/j.0006-341X.2005.031209.x>.
- Klein, John P., and Melvin L. Moeschberger. 2003. *Survival Analysis: Techniques for Censored and Truncated Data*. 2nd ed. Statistics for Biology and Health. New York: Springer.
- Langeveld, N., H. Stam, M. Grootenhuis, and B. Last. 2002. “Quality of Life in Young Adult Survivors of Childhood Cancer.” *Supportive Care in Cancer* 10 (8): 579–600. <https://doi.org/10.1007/s00520-002-0388-6>.
- Latouche, A., R. Porcher, and S. Chevret. 2005. “A Note on Including Time-Dependent Covariate in Regression Model for Competing Risks Data.” *Biometrical Journal* 47 (6): 807–14. <https://doi.org/10.1002/bimj.200410152>.
- Latouche, Aurelien, Arthur Allignol, Jan Beyersmann, Myriam Labopin, and Jason P. Fine. 2013. “A Competing Risks Analysis Should Report Results on All Cause-Specific Hazards and Cumulative Incidence Functions.” *Journal of Clinical Epidemiology* 66 (6): 648–53. <https://doi.org/10.1016/j.jclinepi.2012.09.017>.
- Latouche, Aurélien, Per Kragh Andersen, Grégoire Rey, and Margarita Moreno-Betancur. 2019. “A Note on the Measurement of Socioeconomic Inequalities in Life Years Lost by Cause of Death.” *Epidemiology* 30 (4): 569–72. <https://doi.org/10.1097/EDE.0000000000001022>.

- Lipshultz, Steven E., M. Jacob Adams, Steven D. Colan, Louis S. Constine, Eugene H. Herman, Daphne T. Hsu, Melissa M. Hudson, et al. 2013. “Long-Term Cardiovascular Toxicity in Children, Adolescents, and Young Adults Who Receive Cancer Therapy: Pathophysiology, Course, Monitoring, Management, Prevention, and Research Directions.” *Circulation* 128 (17): 1927–95. <https://doi.org/10.1161/CIR.0b013e3182a88099>.
- Mertens, A. C., Q. Liu, J. P. Neglia, K. Wasilewski, W. Leisenring, G. T. Armstrong, L. L. Robison, and Y. Yasui. 2008. “Cause-Specific Late Mortality Among 5-Year Survivors of Childhood Cancer: The Childhood Cancer Survivor Study.” *JNCI Journal of the National Cancer Institute* 100 (19): 1368–79. <https://doi.org/10.1093/jnci/djn310>.
- Michel, Gisela, Renée L. Mulder, Helena J. H. van der Pal, Roderick Skinner, Edit Bárdi, Morven C. Brown, Janine Vetsch, et al. 2019. “Evidence-Based Recommendations for the Organization of Long-Term Follow-up Care for Childhood and Adolescent Cancer Survivors: A Report from the PanCareSurFup Guidelines Working Group.” *Journal of Cancer Survivorship* 13 (5): 759–72. <https://doi.org/10.1007/s11764-019-00795-5>.
- Moskalewicz, Alexandra, Benjamin Martinez, Elizabeth M. Uleryk, Petros Pechlivanoglou, Sumit Gupta, and Paul C. Nathan. 2024. “Late Mortality Among 5-Year Survivors of Childhood Cancer: A Systematic Review and Meta-Analysis.” *Cancer* n/a (n/a). <https://doi.org/10.1002/cncr.35213>.
- Moskowitz, Chaya S., Joanne F. Chou, Joseph P. Neglia, Ann H. Partridge, Rebecca M. Howell, Lisa R. Diller, Danielle Novetsky Friedman, et al. 2019. “Mortality After Breast Cancer Among Survivors of Childhood Cancer: A Report From the Childhood Cancer Survivor Study.” *Journal of Clinical Oncology* 37 (24): 2120–30. <https://doi.org/10.1200/JCO.18.02219>.
- Nicolaie, M. A., J. C. van Houwelingen, T. M. de Witte, and H. Putter. 2013a. “Dynamic Pseudo-Observations: A Robust Approach to Dynamic Prediction in Competing Risks.” *Biometrics* 69 (4): 1043–52. <https://doi.org/10.1111/biom.12061>.

- . 2013b. “Dynamic Prediction by Landmarking in Competing Risks.” *Statistics in Medicine* 32 (12): 2031–47. <https://doi.org/10.1002/sim.5665>.
- Oeffinger, Kevin C, Toana Kawashima, Debra L Friedman, Nina S Kadan-Lottick, and Leslie L Robison. 2006. “Chronic Health Conditions in Adult Survivors of Childhood Cancer.” *N Engl J Med*, 11. <https://doi.org/10.1056/NEJMsa060185>.
- Pal, Helena J. van der, Elvira C. van Dalen, Evelien van Delden, Irma W. van Dijk, Wouter E. Kok, Ronald B. Geskus, Elske Sieswerda, et al. 2012. “High Risk of Symptomatic Cardiac Events in Childhood Cancer Survivors.” *Journal of Clinical Oncology* 30 (13): 1429–37. <https://doi.org/10.1200/JCO.2010.33.4730>.
- Pearl, Judea, and Dana Mackenzie. 2018. *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.
- Pénichoux, Juliette. 2012. “Modèles Multiplicatifs Du Risque Pour Des Événements Successifs En Présence d’hétérogénéité.” PhD thesis, Université Paris Sud - Paris XI. <https://tel.archives-ouvertes.fr/tel-00997551>.
- Peters, Elisabeth, Laura Mendoza Schulz, and Monika Reuss-Borst. 2016. “Quality of Life After Cancer-How the Extent of Impairment Is Influenced by Patient Characteristics.” *BMC Cancer* 16 (1): 787. <https://doi.org/10.1186/s12885-016-2822-z>.
- Prentice, R. L., J. D. Kalbfleisch, A. V. Peterson, N. Flournoy, V. T. Farewell, and N. E. Breslow. 1978. “The Analysis of Failure Times in the Presence of Competing Risks.” *Biometrics* 34 (4): 541–54. <https://doi.org/10.2307/2530374>.
- Putter, Hein, and Hans C. van Houwelingen. 2017. “Understanding Landmarking and Its Relation with Time-Dependent Cox Regression.” *Statistics in Biosciences* 9 (2): 489–503. <https://doi.org/10.1007/s12561-016-9157-9>.
- . 2022. “Landmarking 2.0: Bridging the Gap Between Joint Models and Landmarking.” *Statistics in Medicine* 41 (11): 1901–17. <https://doi.org/10.1002/sim.9336>.
- Reulen, Raoul C., David L. Winter, Clare Frobisher, Emma R. Lancashire, Charles A. Stiller,

- Meriel E. Jenney, Roderick Skinner, Michael C. Stevens, Michael M. Hawkins, and British Childhood Cancer Survivor Study Steering Group. 2010. “Long-Term Cause-Specific Mortality Among Survivors of Childhood Cancer.” *JAMA* 304 (2): 172–79. <https://doi.org/10.1001/jama.2010.923>.
- “RNCE, Registre National Des Cancers de l’enfant.” 2024. <https://rnce.inserm.fr/rnce/les-chiffres/>. <https://rnce.inserm.fr/rnce/les-chiffres/>.
- Robison, Leslie L., and Melissa M. Hudson. 2014. “Survivors of Childhood and Adolescent Cancer: Life-Long Risks and Responsibilities.” *Nature Reviews Cancer* 14 (1): 61–70. <https://doi.org/10.1038/nrc3634>.
- Robison, Leslie L., Ann C. Mertens, John D. Boice, Norman E. Breslow, Sarah S. Donaldson, Daniel M. Green, Frederic P. Li, et al. 2002. “Study Design and Cohort Characteristics of the Childhood Cancer Survivor Study: A Multi-Institutional Collaborative Project.” *Medical and Pediatric Oncology* 38 (4): 229–39. <https://doi.org/10.1002/mpo.1316>.
- Royston, Patrick, and Mahesh K. B. Parmar. 2002. “Flexible Parametric Proportional-Hazards and Proportional-Odds Models for Censored Survival Data, with Application to Prognostic Modelling and Estimation of Treatment Effects.” *Statistics in Medicine* 21 (15): 2175–97. <https://doi.org/10.1002/sim.1203>.
- Sabathé, Camille. 2019. “Modélisation de l’effet de Facteurs de Risque Sur La Probabilité de Devenir Dément Et d’autres Indicateurs de Santé.” These de doctorat, Bordeaux. <https://www.theses.fr/2019BORD0224>.
- Schoenfeld, David. 1980. “Chi-Squared Goodness-of-Fit Tests for the Proportional Hazards Regression Model.” *Biometrika* 67 (1): 145–53. <https://doi.org/10.1093/biomet/67.1.145>.
- Schwartz, B., A. Thiébaud, A. di Meglio, A. Dumas, K. Berkane, A. Bourmaud, G. Vu-Bezin, et al. 2024. “Utilisation d’opioïdes à Long-Terme Après Traitement d’un Cancer Dans l’enfance : Résultats de La Cohorte Française FCCSS-« French Childhood Cancer Survivor



- Study ».” *Journal of Epidemiology and Population Health*, Congrès éMOIS 2024, 72 (March): 202210. <https://doi.org/10.1016/j.jep.h.2024.202210>.
- Sjölander, Arvid, Elisabeth Dahlqvist, and Johan Zetterqvist. 2016. “A Note on the Noncollapsibility of Rate Differences and Rate Ratios.” *Epidemiology (Cambridge, Mass.)* 27 (3): 356–59. <https://doi.org/10.1097/EDE.0000000000000433>.
- Ssenyonga, Naomi, Charles Stiller, Kayo Nakata, Jaime Shalkow, Sheilagh Redmond, Jean-Luc Bulliard, Fabio Girardi, et al. 2022. “Worldwide Trends in Population-Based Survival for Children, Adolescents, and Young Adults Diagnosed with Leukaemia, by Subtype, During 2000-14 (CONCORD-3): Analysis of Individual Data from 258 Cancer Registries in 61 Countries.” *The Lancet. Child & Adolescent Health* 6 (6): 409–31. [https://doi.org/10.1016/S2352-4642\(22\)00095-5](https://doi.org/10.1016/S2352-4642(22)00095-5).
- Steliarova-Foucher, Eva, Charles Stiller, Peter Kaatsch, Franco Berrino, Jan-Willem Coebergh, Brigitte Lacour, and Max Parkin. 2004. “Geographical Patterns and Time Trends of Cancer Incidence and Survival Among Children and Adolescents in Europe Since the 1970s (the ACCISproject): An Epidemiological Study.” *Lancet (London, England)* 364 (9451): 2097–2105. [https://doi.org/10.1016/S0140-6736\(04\)17550-8](https://doi.org/10.1016/S0140-6736(04)17550-8).
- Steliarova-Foucher, Eva, Charles Stiller, Brigitte Lacour, and Peter Kaatsch. 2005. “International Classification of Childhood Cancer, Third Edition.” *Cancer* 103 (7): 1457–67. <https://doi.org/10.1002/cncr.20910>.
- Therneau, Terry M., and Patricia M. Grambsch. 2000. *Modeling Survival Data: Extending the Cox Model*. Edited by K. Dietz, M. Gail, K. Krickeberg, J. Samet, and A. Tsiatis. Statistics for Biology and Health. New York, NY: Springer New York. <https://doi.org/10.1007/978-1-4757-3294-8>.
- Therneau, Terry, Cynthia Crowson, and Elizabeth Atkinson. 2024. “Multi-State Models and Competing Risks,” 29.
- Tukenova, Markhaba, Catherine Guibout, Odile Oberlin, Françoise Doyon, Abdeddahir Mou-

- sannif, Nadia Haddy, Sylvie Guérin, et al. 2010. “Role of Cancer Treatment in Long-Term Overall and Cardiovascular Mortality After Childhood Cancer.” *Journal of Clinical Oncology* 28 (8): 1308–15. <https://doi.org/10.1200/JCO.2008.20.2267>.
- Van Houwelingen, Hans C. 2007. “Dynamic Prediction by Landmarking in Event History Analysis.” *Scandinavian Journal of Statistics* 34 (1): 70–85. <https://doi.org/10.1111/j.1467-9469.2006.00529.x>.
- Veres, Cristina, Rodrigue S. Allodji, Damien Llanas, Jérémie Vu Bezin, Jean Chavaudra, Jean Pierre Mège, Dimitri Lefkopoulos, et al. 2014. “Retrospective Reconstructions of Active Bone Marrow Dose-Volume Histograms.” *International Journal of Radiation Oncology\*Biography\*Physics* 90 (5): 1216–24. <https://doi.org/10.1016/j.ijrobp.2014.08.335>.
- Whelan, Kimberly, and Elizabeth Alva. 2018. “Chapter 1 - Epidemiology of Childhood Cancer.” In *Pediatric Cancer Genetics*, edited by Nathaniel H. Robin and Meagan B. Farmer, 1–20. Elsevier. <https://doi.org/10.1016/B978-0-323-48555-5.00001-6>.
- Wohlfahrt, Peter, Dominik Jenča, Josef Stehlik, Vojtěch Melenovský, Jolana Mrázková, Vladimír Staněk, Jiří Kettner, et al. 2023. “Heart Failure-Related Quality-of-Life Impairment After Myocardial Infarction.” *Clinical Research in Cardiology: Official Journal of the German Cardiac Society* 112 (1): 39–48. <https://doi.org/10.1007/s00392-022-02008-z>.
- Wong, Kwok F., Raoul C. Reulen, David L. Winter, Joyeeta Guha, Miranda M. Fidler, Julie Kelly, Emma R. Lancashire, Kathryn Pritchard-Jones, Helen C. Jenkinson, and Elaine Sugden. 2016. “Risk of Adverse Health and Social Outcomes up to 50 Years After Wilms Tumor: The British Childhood Cancer Survivor Study.” *Journal of Clinical Oncology* 34 (15): 1772–79. <https://doi.org/10.1200/JCO.2015.64.4344>.
- Yeh, Jennifer M., Larissa Nekhlyudov, Sue J. Goldie, Ann C. Mertens, and Lisa Diller. 2010. “A Model-Based Estimate of Cumulative Excess Mortality in Survivors of Childhood Cancer.” *Annals of Internal Medicine* 152 (7): 409–17. <https://doi.org/10.7326/0003-4819-152-7-20100400>.

- Yeh, Jennifer M., Zachary J. Ward, Aeysha Chaudhry, Qi Liu, Yutaka Yasui, Gregory T. Armstrong, Todd M. Gibson, et al. 2020. "Life Expectancy of Adult Survivors of Childhood Cancer Over 3 Decades." *JAMA Oncology* 6 (3): 350–57. <https://doi.org/10.1001/jamaoncol.2019.5582>.
- Yu, Ruby, Derek Lai, Grace Leung, Lok-Yan Tam, Clara Cheng, Sara Kong, Cecilia Tong, and Jean Woo. 2024. "Transitions in Intrinsic Capacity Among Community-Dwelling Older People and Their Associated Factors: A Multistate Modelling Analysis." *The Journal of Nutrition, Health and Aging* 28 (7): 100273. <https://doi.org/10.1016/j.jnha.2024.100273>.

## A - Anthracyclines Doses Equivalent

We used doxorubicin equivalents in order to account for the disparity in cardio-toxicity by  $\text{mg}/\text{m}^2$  of each drug (Feijen et al. 2019). We used the coefficients of Table A.1.

**Table A.1:** Ponderations used to compute cumulative dose of anthracyclines. Doses are in  $\text{mg}/\text{m}^2$

Drug	Range	Coefficient
Daunorubicine	< 150	0.8
	150 - 300	0.6
	> 300	0.5
Epirubicine	< 150	1.3
	150 - 300	0.6
	> 300	0.5
Novantrone	< 300	4.2
	> 300	48.3
Rubidazone	> 0	0.05625
Epiadriamycine	< 150	1.3
	150 - 300	0.6
	> 300	0.5



## B - Details of FCCSS data

### B.1 . Radiotherapy by Childhood Cancer Type

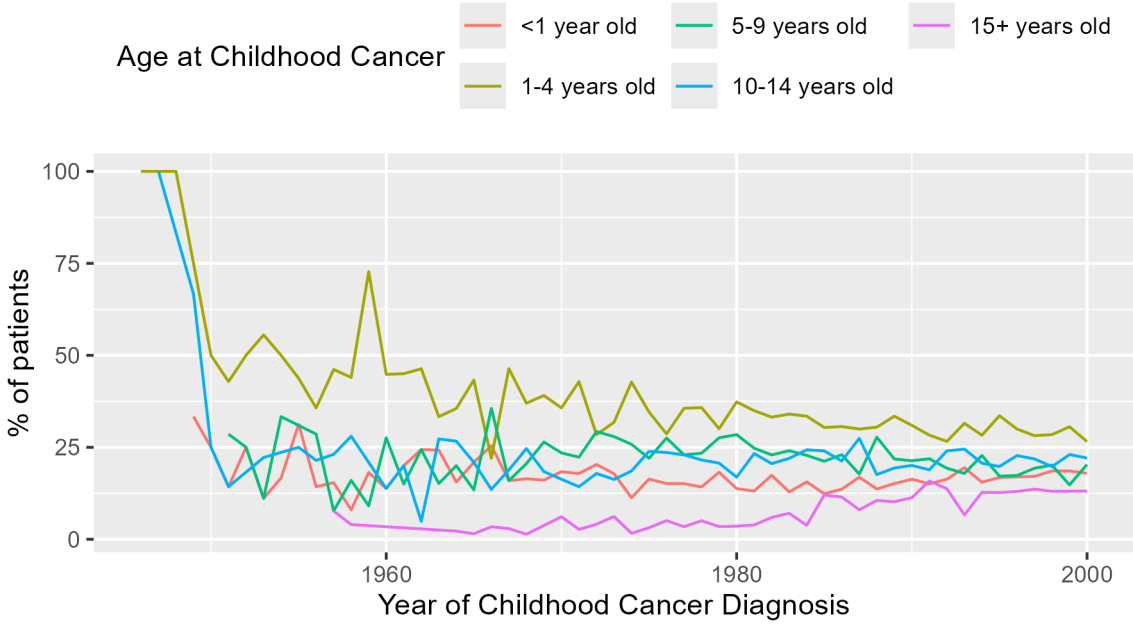
**Table B.1:** Percent of FCCSS patient treated with radiotherapy by type of childhood cancer.

Type of Childhood Cancer	Number of patients in FCCSS	Percent treated of radiotherapy
Unknown	10	20.0
01 -Leukemias	1	100.0
02 -Lymphomas	1278	55.7
03 -CNS tumor	1140	86.1
04 -Peripheral nervouus tumors	1034	37.9
05 -Retinnoblastomas	619	54.4
06 -Renal tumors	1140	56.1
07 -Hepatic tumors	79	20.3
08 -Bone sarcomas	686	43.3
09 -Soft-tissue sarcomas	859	57.4
10 -Germ cells and gonadal tumors	469	30.1
11 -Other carcinomas	344	50.9
12 -Other or unspecified tumors	11	63.6

**Table B.2:** Percent of FCCSS patients treated with high doses of radiotherapy at the brain and heart, by type of childhood cancer.

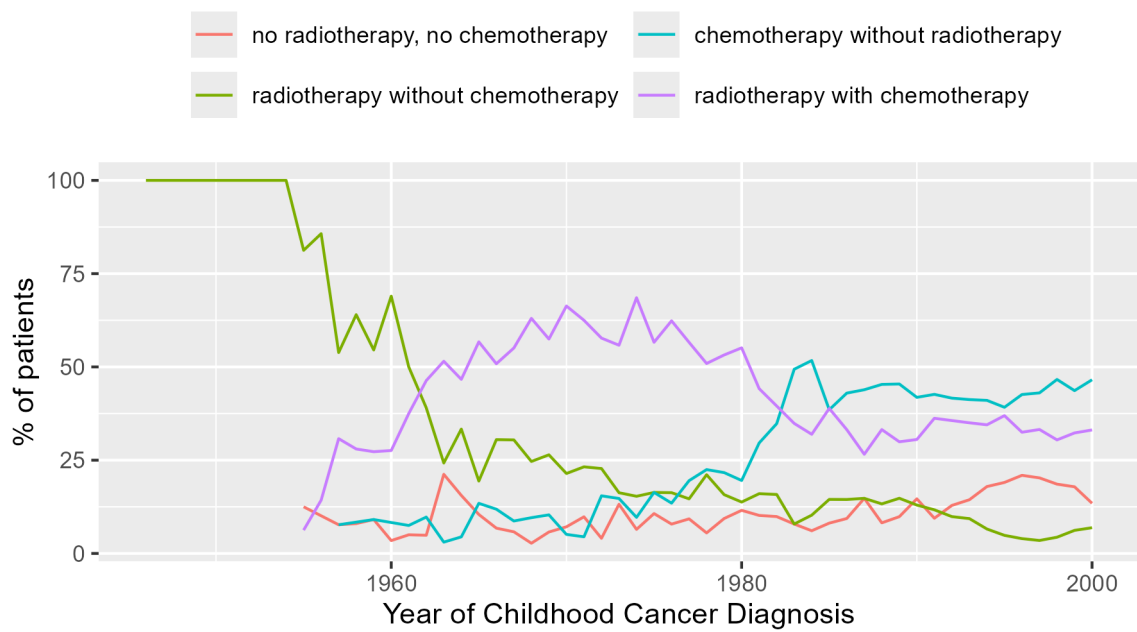
Type of Childhood Cancer	Number of patients in FCCSS	% brain RT > 20 Gy	% brain RT > 30 Gy	% heart RT > 20 Gy	% heart RT > 30 Gy
Unknown	10	0.0	0.0	0.0	0.0
01 -Leukemias	1	0.0	0.0	0.0	0.0
02 -Lymphomas	1278	8.5	1.3	16.4	10.6
03 -CNS tumor	1140	50.4	32.3	4.6	0.9
04 -Peripheral nervouus tumors	1034	1.5	0.9	3.1	0.9
05 -Retinoblastomas	619	0.6	0.0	0.2	0.0
06 -Renal tumors	1140	0.1	0.0	4.1	0.7
07 -Hepatic tumors	79	0.0	0.0	0.0	0.0
08 -Bone sarcomas	686	1.2	0.4	3.9	1.3
09 -Soft-tissue sarcomas	859	1.9	0.5	2.0	0.8
10 -Germ cells and gonadal tumors	469	5.5	3.0	1.3	0.2
11 -Other carcinomas	344	1.7	0.6	1.5	0.3
12 -Other or unspecified tumors	11	0.0	0.0	0.0	0.0

**B.2 . Changes in Patients Characteristics by Year of Childhood Cancer Diagnosis**



**Figure B.1:** Age at Childhood Cancer Diagnosis among FCCSS patients by Year of Childhood Cancer Diagnosis





**Figure B.2:** Use of Radiotherapy and Chemotherapy among FCCSS patients by Year of Childhood Cancer Diagnosis

## C - Published Works

ORIGINAL RESEARCH

# Increased Cardiac Risk After a Second Malignant Neoplasm Among Childhood Cancer Survivors



## A FCCSS Study

Thibaud Charrier, PhD,<sup>a,b,c,d</sup> Nadia Haddy, PhD,<sup>a,b,c</sup> Boris Schwartz, MSc,<sup>a,b,c</sup> Neige Journy, PhD,<sup>a,b,c</sup> Brice Fresneau, MD, PhD,<sup>a,e</sup> Charlotte Demoor-Goldschmidt, MD, PhD,<sup>b,c,f,g,h</sup> Ibrahima Diallo, PhD,<sup>i,j</sup> Aurore Surun, MD,<sup>k</sup> Isabelle Aerts, MD,<sup>k</sup> François Doz, MD,<sup>k,l</sup> Vincent Souchard, MSc,<sup>a,b,c</sup> Giao Vu-Bezin, MSc,<sup>a,b,c</sup> Anne Laprie, MD,<sup>m</sup> Sarah Lemler, PhD,<sup>n</sup> Véronique Letort, PhD,<sup>n</sup> Carole Rubino, MD, PhD,<sup>a,b,c</sup> Stéfania Chounta, PhD,<sup>a,b,c,n</sup> Florent de Vathaire, PhD,<sup>a,b,c</sup> Aurélien Latouche, PhD,<sup>d,o</sup> Rodrigue S. Allodji, PhD<sup>a,b,c,p</sup>

### ABSTRACT

**BACKGROUND** Childhood cancer survivors (CCS) are at an elevated risk of developing both a second malignant neoplasm (SMN) and cardiac disease.

**OBJECTIVES** This study sought to assess the excess of occurrence of cardiac disease after a SMN among CCS.

**METHODS** Analyses included 7,670 CCS from the French Childhood Cancer Survivors Study cohort diagnosed between 1945 and 2000. To account for the time dependence of the occurrence of a SMN, we employed a landmark approach, considering an additive regression model for the cumulative incidence of cardiac disease. We estimated the effect of a SMN on the instantaneous risk of cardiac disease using a proportional cause-specific hazard model, considering a SMN as a time-dependent exposure. In both models, we adjusted for demographic and treatment information and considered death as a competing event.

**RESULTS** In 7,670 CCS over a median follow-up of 30 years (IQR: 22-38 years), there were 378 cases of cardiac disease identified, of which 49 patients experienced a SMN. Patients who survived 25 years after their childhood cancer diagnosis and had a SMN in that time frame had a significantly increased cumulative incidence of cardiac disease, which was 3.8% (95% CI: 0.5% to 7.1%) higher compared with those without a SMN during this period. No SMN-induced excess of cardiac disease was observed at subsequent landmark times. SMNs were associated with a 2-fold increase (cause-specific HR: 2.0; 95% CI: 1.4-2.8) of cardiac disease.

**CONCLUSIONS** The occurrence of a SMN among CCS is associated with an increased risk of cardiac disease occurrence and risk at younger ages. (J Am Coll Cardiol CardioOnc 2023;5:792-803) © 2023 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

From the <sup>a</sup>Cancer and Radiation Team, Centre de Recherche en Epidemiologie et Santé des Populations, Université Paris-Saclay-Université Paris-Sud-Université de Versailles Saint-Quentin-en-Yvelines, Villejuif, France; <sup>b</sup>Cancer and Radiation Team, Centre de Recherche en Epidemiologie et Santé des Populations, Institut National de la Santé et de la Recherche Médicale, Villejuif, France; <sup>c</sup>Cancer and Radiation Team, Department of Clinical Research, Gustave Roussy, Villejuif, France; <sup>d</sup>U900, Institut National de la Santé et de la Recherche Médicale, PSL Research University, Institut Curie, Saint-Cloud, France; <sup>e</sup>Department of Pediatric Oncology, Gustave Roussy, Université Paris-Saclay, Villejuif, France; <sup>f</sup>Department of Pediatric Oncology, Centre Hospitalier Universitaire, Angers, France; <sup>g</sup>Department of Radiotherapy, Centre François Baclesse, Caen, France; <sup>h</sup>Department of Supportive Care, Centre François Baclesse, Caen, France; <sup>i</sup>Department of Radiation Oncology, Gustave Roussy, Villejuif, France; <sup>j</sup>Radiothérapie Moléculaire et Innovation Thérapeutique, Institut National de la Santé et de la Recherche Médicale, Gustave Roussy, Paris-Saclay University, Villejuif, France; <sup>k</sup>SIREDO Oncology Center, Institut Curie, Paris, France; <sup>l</sup>Université Paris Cité, Paris, France; <sup>m</sup>Department of Radiation Oncology, Centre Antoine-Lacassagne, Nice, France; <sup>n</sup>Mathématiques et Informatique pour la Complexité et les Systèmes, CentraleSupélec, Université Paris-Saclay, Gif-sur-Yvette, France; <sup>o</sup>Conservatoire National des Arts et Métiers, Paris, France; and the <sup>p</sup>Polytechnic School of Abomey-Calavi, University of Abomey-Calavi, Cotonou, Benin.

Advances in cancer treatment have significantly improved childhood cancer survival, with the 5-year survival rate exceeding 80% in most European and North American countries today.<sup>1</sup> The population of childhood cancer survivors (CCS) in Europe now exceeds 300,000 people, with approximately 1 in every 1,000 individuals being a CCS.<sup>2</sup> This increased population has also led to a higher rate of health issues related to late effects<sup>3</sup> compared with the general population,<sup>4</sup> with cardiac disease and the occurrence of a second malignant neoplasm (SMN) being among the most severe and life-threatening conditions experienced by CCS.<sup>5-8</sup>

Previous studies have identified several main risk factors for cardiac disease,<sup>9,10</sup> with the most important ones being cumulative doses of anthracyclines, chest radiotherapy (including heart radiation as low as 5 Gy), alkylating agents administration, and treatment at a young age. Cardiac disease has been found to contribute to excess mortality of CCS who survived breast cancer, as reported by Moskowitz et al.<sup>11</sup>

Despite this knowledge, the effects of a SMN on cardiac disease remain unknown. We hypothesized that the risk of cardiac disease may be influenced by both common risk factors of a SMN and cardiac disease, as well as treatments for a SMN, but whether this increased risk would translate to a higher cumulative incidence of cardiac disease among CCS is unknown. Indeed, patients who experienced a SMN have been found to have a higher mortality rate than those who did not, which could counterbalance the increased risk of cardiac disease. Therefore, in the present study, we aim to quantify both the increased risk and increased cumulative incidence of cardiac disease after a SMN among CCS.

## METHODS

**STUDY POPULATION.** The FCCSS (French Childhood Cancer Survivors Study)<sup>12</sup> cohort follows 7,670 5-year survivors treated between 1945 and 2000 for solid cancer in 5 cancer centers in France. The FCCSS received approval from a National Committee on Ethics and the French National Agency Regulating Data Protection (agreement nos. 902287 and 12038829). Written informed consent was obtained from all patients, parents, or guardians in accordance with national research ethics requirements. The present analysis included all 7,670 survivors, but to

address potential bias introduced by missing data, we used multiple imputation as recommended (see the [Supplemental Appendix](#) for more details).<sup>13</sup>

**CANCER THERAPY EXPOSURES.** Chemotherapy, surgery, and radiotherapy information for childhood cancer treatment were abstracted from medical records. For this study, chemotherapy exposures were defined as follows: 1) receipt (or not) of any chemotherapy agent, anthracyclines, alkylating agents, or platinum agents; and 2) cumulative doses for each chemotherapy class. For each survivor, cumulative doses of each drug received per m<sup>2</sup> were computed by summing chemotherapy doses across cycles. Anthracycline cumulative dose was computed using doxorubicin equivalents.<sup>14</sup>

For external beam radiotherapy and/or brachytherapy, radiation dose distributions to the heart were retrospectively reconstructed on patient-specific voxel phantoms, considering individual patient treatment information. This information included treatment machine, type of radiation, beam energy, irradiation technique, field size and shape, gantry and collimator angles, use of accessories, target volume location, and total delivered dose. This retrospective reconstruction was necessary, because computed tomography scans were not used on many patients. More details on the methodology and dosimetry software package used have already been published.<sup>15</sup> The use of voxel phantoms, in which the heart was carefully delineated, allowed the computation of the mean radiation dose.

**OUTCOME DEFINITIONS: IDENTIFICATION AND VALIDATION OF DEATHS, SMN, AND CARDIAC DISEASE.** Vital status of all patients and causes of death for deceased patients were obtained from C epiDC.<sup>16</sup> Clinical and epidemiological follow-up, including self-administered questionnaires and cohort linkage with the French Hospital Database and Health Insurance Information System,<sup>17</sup> was performed to identify the occurrence of iatrogenic effects. For patients treated at Gustave Roussy, clinic long-term follow-up was also performed. SMNs and cardiac diseases were identified through these different sources and subsequently validated by a trained and experienced clinical research associate. Validation was based on medical, pathology, or radiological reports obtained from the

## ABBREVIATIONS AND ACRONYMS

**CCS** = childhood cancer survivor(s)

**CCSS** = Childhood Cancer Survivor Study

**csH** = cause-specific hazard

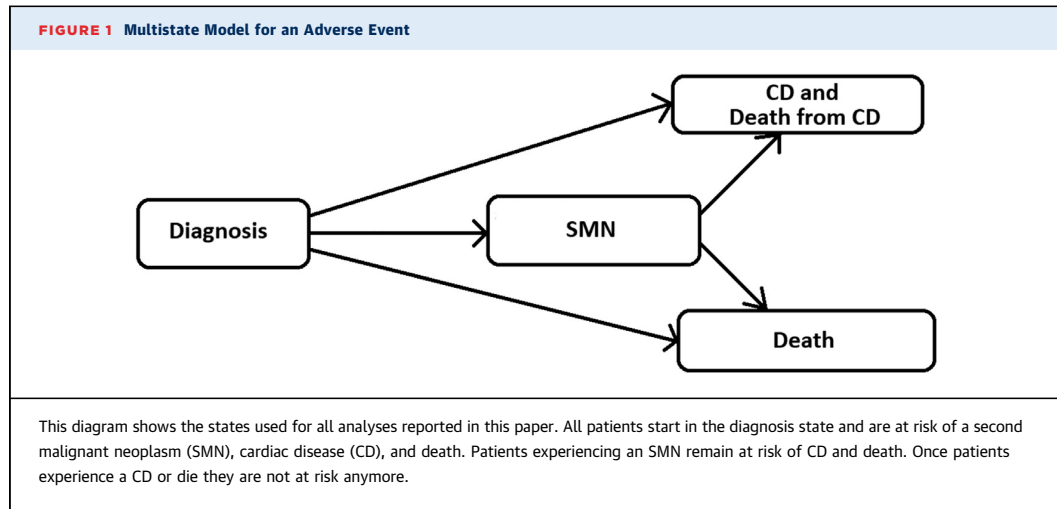
**csHR** = cause-specific HR

**FCCSS** = French Childhood Cancer Survivors Study

**SMN** = second malignant neoplasm

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the [Author Center](#).

Manuscript received December 1, 2022; revised manuscript received July 18, 2023; accepted July 19, 2023.



treating centers or from referring doctors, regardless of the data source used for first identification. Cardiac disease events were graded according to the Common Terminology Criteria for Adverse Events (version 4.03). In this study, we included only severe cardiac disease events (grade  $\geq 3$ ), based on the consideration that nonsevere cardiac disease events may be self-reported and could cause reporting bias in the data.

**STATISTICAL ANALYSIS.** For the study, our primary event of interest was the first occurrence of severe cardiac disease or death resulting from cardiac disease, with death from any cause considered to be a competing event. SMN was considered to be a time-dependent exposure, rather than a competing event. We defined  $T$  as the time of the first occurrence of cardiac disease or death resulting from cardiac disease (event of interest), death from other cause (competing event), or last follow-up (censor) (Figure 1). Patients ( $n = 35$ ) whose first event occurred within 5 years of their childhood cancer diagnosis were excluded from the analysis.

**CUMULATIVE INCIDENCE OF CARDIAC DISEASE.** To estimate the effect of a SMN on cardiac disease incidence, we chose to quantify the excess in the occurrence of cardiac disease observed after a SMN, that is, the percentage of the population experiencing a cardiac disease after a SMN, not explained by non-SMN-related information. We employed a landmark analysis<sup>18</sup> at 9 different time points. We included patients conditional on their survival at  $s \in \{15, 20, 25, 30, 35\}$  years from diagnosis and at  $s \in \{20, 25, 30, 35\}$  years of age. We chose those landmark times by looking at the number of events observed after each landmark time (Supplemental Table 1) and setting 5 years between

each time, which is reasonable and clinically relevant for a prediction range. Following Cortese and Andersen's suggestion,<sup>19</sup> SMN status was defined as the binary time-dependent indicator  $X_{SMN}(s) = 1_{\{SMN \text{ before } s\}}$  for each landmark time. We then included  $X_{SMN}(s)$  as a covariate in an additive model for the cumulative incidence of cardiac disease (Equation 1),

$$P(T \leq t, e = \text{cardiac disease} | T > s, S_{MN}, X) = F_{s, \text{cardiac disease}}(t) + X_{SMN}(s)\beta_{s, SMN}^T + X\beta_s^T \quad (1)$$

where  $F_{s, CD}(t) = \mathbb{P}(T \leq t, e = \text{cardiac disease} | T > s)$  is the baseline cumulative incidence function for cardiac disease,  $\beta_s$  is the vector of effects, and  $X$  is the vector of covariates containing treatment information. The motivation for employing an additive model for the cumulative incidence of cardiac disease was due to the simplicity of interpreting the model coefficients. This model allows us to estimate the proportion of patients who would experience a cardiac disease given they experienced a SMN.

To account for the disparity of information available to clinicians when evaluating patients' risks, we used different combinations of covariates. The most detailed combination included, among others, both radiotherapy doses to the heart and cumulative anthracyclines doses. These groupings were chosen based on previous works identifying major risk factors<sup>9,10</sup> and developing risk prediction tools.<sup>20,21</sup> Supplemental Table 2 summarizes the combinations of covariates considered. All treatment information used pertain to childhood cancer.

The previous regression model was estimated using pseudovalues<sup>22</sup> computed using the Aalen-

Johansen<sup>23</sup> estimator and generalized estimating equations with identity as the link function. For regression analysis, pseudovalues are computed at a grid of time points. A small subset of time points is enough to have reliable estimates,<sup>22</sup> but there must be enough events before the first grid time point and after the last. So, we have chosen the grid {2, 4, ..., 20} years after landmark time.

Furthermore, we decided to perform a sensitivity analysis to determine whether the estimates of the above-mentioned method can be separated into a baseline SMN effect and a treatment of SMN effect. We performed the same analysis changing SMN information from a binary time-dependent indicator to a categorical time-dependent covariate, named  $X_{SMN.type}(t)$ . The levels of  $X_{SMN.type}(t)$  were chosen with a clinician based on their knowledge of standard cancer treatments and their known cardiotoxicity. These levels are “no SMN,” “Breast cancer (women),” “Sarcoma, bone, soft tissue cancer,” and “Others or unknown SMN.” The levels of  $X_{SMN}(t)$  are “no SMN” and “SMN.” SMN status was still determined for each landmark time.

**CAUSE-SPECIFIC HAZARD OF CARDIAC DISEASE.** To estimate the effect of an SMN on cardiac disease instantaneous risk, we chose to use an extension of the Cox model<sup>24,25</sup> with considerations for competing risks. This measure has the same interpretation as the HR of the Cox model on the instantaneous risk. Of note, the effect of an SMN on the cause-specific hazard (csH) of cardiac disease cannot be directly converted into an effect of an SMN on the cumulative incidence of cardiac disease.<sup>19</sup>

We considered patients who survived at least 5 years after childhood cancer diagnosis and conducted the analysis using a proportional csH model, including SMN status as a time-dependent covariate  $X_{SMN}(t) = 1_{(SMN \text{ before } t)}$ .

We used 5 years since diagnosis of first cancer as the time when patients start being at risk of cardiac disease and death. Similar to the previous analysis on the additive scale for the cumulative incidence, we considered various sets of covariates to adjust our estimation of the csH, using treatment information for the childhood cancer. Those sets of covariates are detailed in [Supplemental Table 2](#).

We performed a sensitivity analysis by categorizing SMN into 4 types. This categorization was consistent with the one used in the cumulative incidence sensitivity analysis, where we used  $X_{SMN.type}(t)$  as a time-dependent covariate instead of  $X_{SMN}(t)$ .

All analyses were conducted using R software (version 4.2.0, R Foundation for Statistical

Computing) and packages survival,<sup>26,27</sup> geepack,<sup>28-30</sup> and pseudo.<sup>31</sup> To allow interested readers to implement the methods used, we have made the R scripts available online.<sup>32</sup>

## RESULTS

Continuous data are presented as median (IQR). Cumulative incidence estimates are presented as a percentage with 95% CI. csH model results are presented as the cause-specific HR (csHR) with 95% CI.

**DESCRIPTIVE ANALYSIS OF THE COHORT.** The demographic and treatment characteristics of 5-year CCS from the FCCSS cohort are presented in [Table 1](#). Among the 7,670 survivors of cancer, 4,201 (54.8%) were female, almost half (48.6%) were younger than 5 years of age at their first childhood cancer diagnosis, and more than 80% of the cohort was still alive as of the last follow-up contact. Approximately 40% of all survivors received a combination of chemotherapy and radiation. Specifically, 31% of the population was treated with cumulative doses of anthracyclines higher than 100 mg/m<sup>2</sup>, and more than 49% of the cohort had been exposed to heart radiation during radiotherapy.

Over a median follow-up (*T*) of 30 years (IQR: 22-38 years), 795 individuals developed an SMN, 329 developed a cardiac disease without an SMN, and 49 developed a cardiac disease after an SMN. The median time to a cardiac disease was 23 years (IQR: 15-32 years), with a median age at event of 32 years (IQR: 21-40 years). The median time to an SMN was 20 years (IQR: 13-29 years), with a median age at event of 29 years (IQR: 19-38 years). The majority of the population with an SMN was treated more recently than 2000: only 6% (*n* = 45) were diagnosed before 1980, 11% (*n* = 90) between 1980 and 1989, and 22% (*n* = 177) between 1990 and 1999, while 61% (*n* = 483) were diagnosed after 2000. The number of patients included at each landmark time and the number of observed events are listed in [Supplemental Table 1](#).

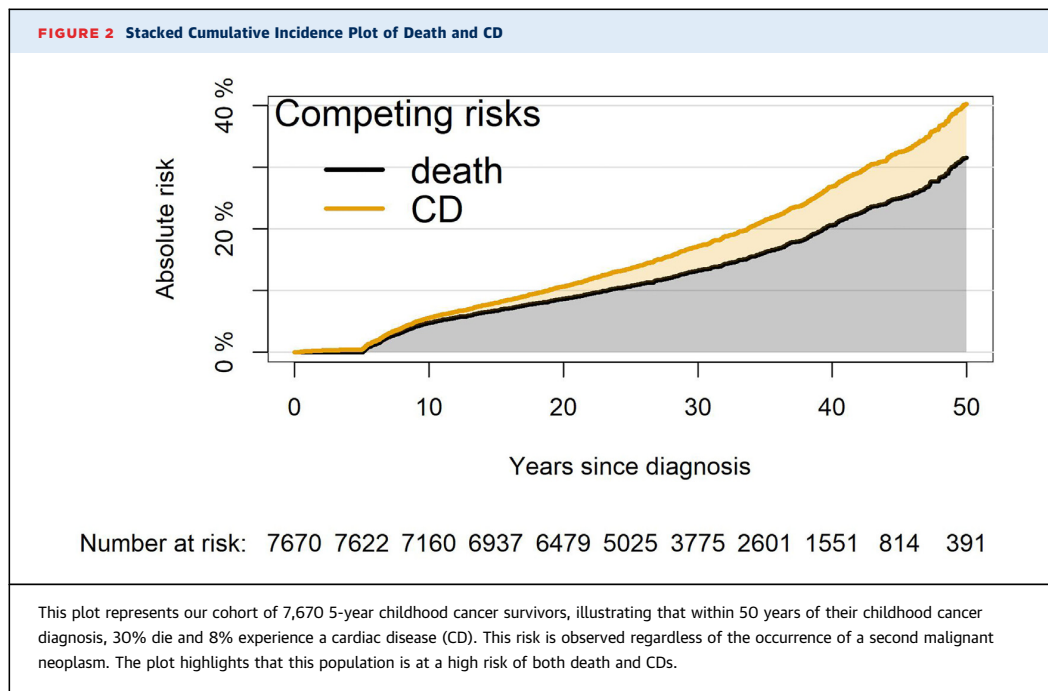
The identified cardiac diseases among those with and without an SMN included heart failure (*n* = 194 [51.3%]), valvular heart disease (*n* = 55 [14.6%]), arrhythmia (*n* = 46 [12.2%]), pericardial disease (*n* = 28 [7.4%]), ischemic heart disease (*n* = 28 [7.4%]), and other heart diseases (*n* = 27 [7.1%]). For SMN, all malignant neoplasms were included. Of the 795 SMNs, the most common ones identified were breast cancer (*n* = 90 [11.3%]), thyroid cancer (*n* = 82 [10.3%]), bone cancer (*n* = 70 [8.8%]), and skin epitheliomas and carcinoma (*n* = 68 [8.5%]). A fourth of post-SMN cardiac diseases occurred after breast cancer ([Supplemental Table 3](#)).

<b>TABLE 1 Cohort Description</b>	
<b>Sex</b>	
Female	4,201 (54.8%)
Male	3,469 (45.2%)
<b>Age at diagnosis</b>	
<5 y	3,726 (48.6%)
5-10 y	1,678 (21.9%)
10-15 y	1,623 (21.2%)
>15 y	643 (8.4%)
<b>Deceased</b>	
No	6,173 (80.5%)
Yes	1,497 (19.5%)
<b>Treatment combination</b>	
Nor radiotherapy nor chemotherapy	902 (11.8%)
Radiotherapy alone	1,088 (14.2%)
Chemotherapy alone	2,574 (33.6%)
Radiotherapy and chemotherapy	3,106 (40.5%)
<b>Anthracyclines doses</b>	
0 mg/m <sup>2</sup>	5,060 (66.0%)
0-100 mg/m <sup>2</sup>	209 (2.7%)
100-250 mg/m <sup>2</sup>	1,257 (16.4%)
>250 mg/m <sup>2</sup>	1,144 (14.9%)
<b>Alkylating agent</b>	
No	4,060 (52.9%)
Yes	3,610 (47.1%)
<b>Platinum agent</b>	
No	6,045 (78.8%)
Yes	1,625 (21.2%)
<b>Mean heart RT dose</b>	
0 Gy	3,622 (49.8%)
0-5 Gy	2,349 (32.3%)
5-15 Gy	573 (7.9%)
15-35 Gy	643 (8.8%)
>35 Gy	92 (1.3%)
<b>Mean brain RT dose</b>	
0 Gy	3,567 (49.0%)
0-20 Gy	2,954 (40.6%)
20-30 Gy	342 (4.7%)
30-50 Gy	409 (5.6%)
>50 Gy	7 (0.1%)
<b>Neck RT</b>	
No	3,609 (49.6%)
Yes	3,670 (50.4%)
<b>Type of childhood cancer</b>	
Unknown	10 (0.1%)
01 - Leukemias	1 (0.0%)
02 - Lymphomas	1,278 (16.7%)
03 - CNS tumor	1,140 (14.9%)
04 - Peripheral nervous tumors	1,034 (13.5%)
05 - Retinoblastomas	619 (8.1%)
06 - Renal tumors	1,140 (14.9%)
07 - Hepatic tumors	79 (1.0%)
08 - Bone sarcomas	686 (8.9%)
09 - Soft-tissue sarcomas	859 (11.2%)
10 - Germ cells and gonadal tumors	469 (6.1%)
11 - Other carcinomas	344 (4.5%)
12 - Other or unspecified tumors	11 (0.1%)
Values are n (%).	
RT = radiotherapy.	

The cumulative incidence of cardiac disease was 3.9% (95% CI: 3.4% to 4.3%) at 30 years after diagnosis and 8.4% (95% CI: 7.3% to 9.5%) at 50 years after diagnosis for the whole cohort, presented in **Figure 2**. For patients treated with both radiotherapy (any dose) and chemotherapy (any agent), the cumulative incidence of cardiac disease was 5.8% (95% CI: 4.9% to 6.7%) at 30 years after diagnosis and 12.0% (95% CI: 10.2% to 13.8%) at 50 years after diagnosis (**Supplemental Figure 1**). In comparison, for patients treated with neither radiotherapy nor chemotherapy, the cumulative incidence was 1.7% (95% CI: 0.2% to 3.3%) at 50 years after diagnosis.

**SMN EFFECT ON THE CUMULATIVE INCIDENCE OF CARDIAC DISEASE: ADDITIVE REGRESSION MODEL.** The cumulative incidence of cardiac disease was higher when an SMN occurred for landmark times 20, 25, and 30 years after diagnosis, as well as 30 and 35 years of attained age. The maximum estimated excess was 3.8% (95% CI: 0.5% to 7.1%) for the landmark time of 25 years after diagnosis (**Figure 3**). Note that, here, estimates correspond to whether and how much the probability of experiencing cardiac disease changes between SMN survivors and baseline CCS. The **Central Illustration**, **Figure 3**, and **Table 2** summarize the estimated excess of cumulative incidence of cardiac disease after an SMN for each configuration. As detailed in **Supplemental Table 1**, the 3.8% excess for landmark time 25 years after diagnosis translated into approximately 10 cardiac diseases (259 patients with an SMN at this time,  $259 \cdot 3.8 / 100 = 9.8$ ) out of the 30 expected after an SMN (259 patients with an SMN at this time, 11.5% cumulative incidence of cardiac disease,  $259 \cdot 0.115 = 29.8$ ) in this category.

For the landmark times 15 and 35 years after diagnosis, as well as 20 and 25 years of attained age, no significant excess of cardiac disease incidence due to an SMN was observed. In the univariable analysis for patients who survived  $\geq 25$  years after diagnosis (**Figure 3, Central Illustration**), those experiencing an SMN had a cumulative incidence 5.1% (95% CI: 1.8% to 8.3%) higher than those who did not have an SMN. When adjusting for age at first diagnosis, radiotherapy (yes/no), and chemotherapy (yes/no), the cumulative incidence of cardiac disease was 4.5% (95% CI: 1.2% to 7.7%) higher when an SMN occurred in the first 25 years after childhood cancer diagnosis. When including all treatment information (age at diagnosis, sex, year of diagnosis, average radiotherapy dose at the heart and brain, exposure of the neck to radiotherapy, cumulative anthracycline dose, use of alkylating agent), the cumulative incidence of



cardiac disease was 3.8% (95% CI: 0.5% to 7.1%) higher when an SMN occurred in the first 25 years after childhood cancer.

Results on the attained age time scale were similar. In the univariable analysis for patients who survived  $\geq 35$  years of age (Table 2), those experiencing an SMN had a cumulative incidence 4.4% (95% CI: 1.1% to 7.7%) higher than those who did not have an SMN. When adjusting for age at first diagnosis, radiotherapy (yes/no), and chemotherapy (yes/no), the cumulative incidence of cardiac disease was 4.0% (95% CI: 0.8% to 7.3%) higher when an SMN occurred before 35 years of age. When including all treatment information (age at diagnosis, sex, year of diagnosis, average radiotherapy dose at the heart and brain, exposure of the neck to radiotherapy, cumulative anthracycline dose, use of alkylating agent), the cumulative incidence of cardiac disease was 3.3% (95% CI: 0.0% to 6.5%) higher when an SMN occurred before 35 years of age.

When categorizing the SMN status into 4 types, we did not observe any statistically significant effect, due to high standard errors. However, we did observe tendencies in the pointwise estimates (Supplemental Figure 2) among the cancer types. When adjusting for all treatment information, we observed that breast cancer (10.9%; 95% CI: -12.7% to 34.5%); sarcoma, bone, or soft tissue cancer (7.7%; 95% CI: -1.2% to

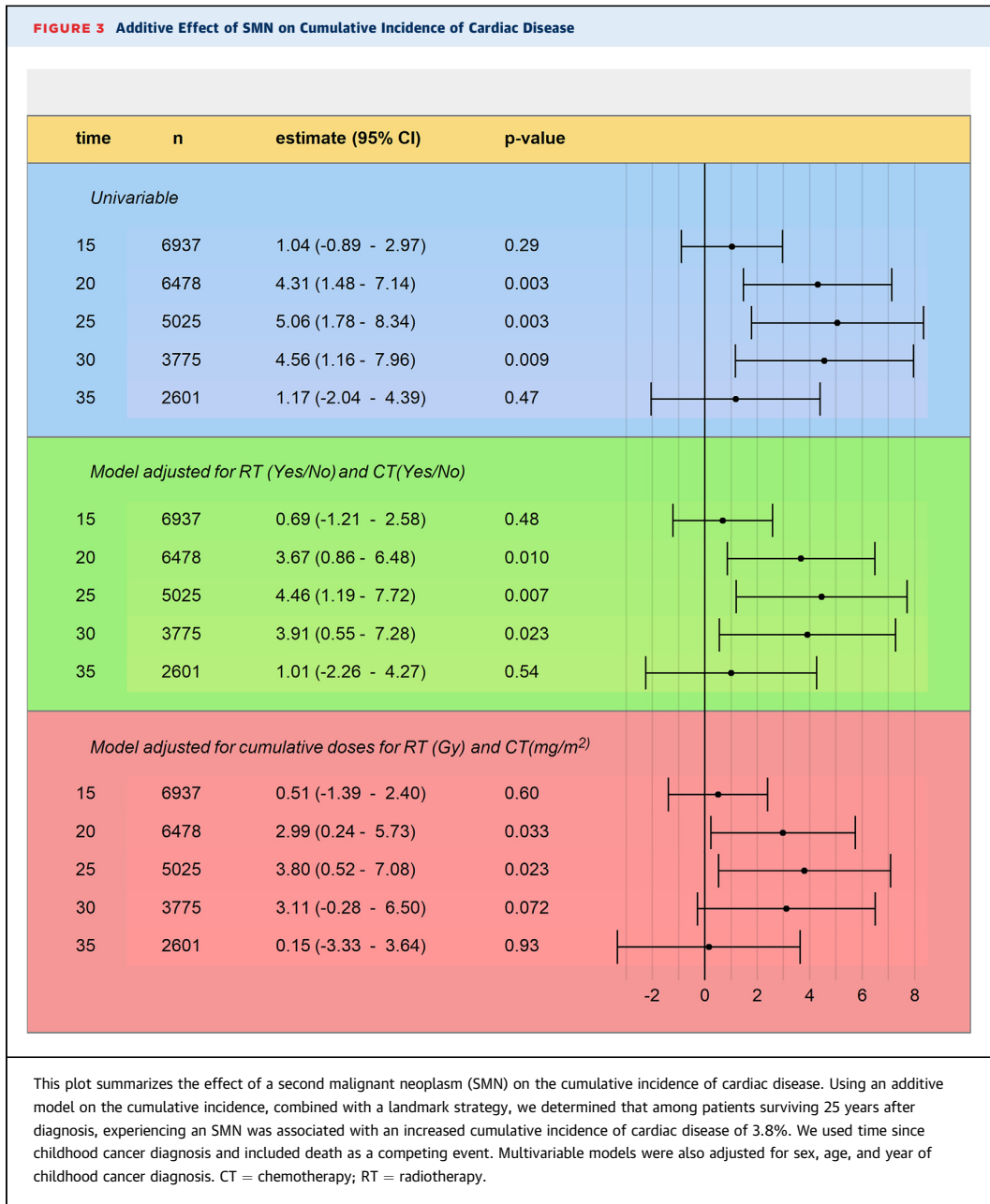
16.6%); and other SMN types (2.0%; 95% CI: -1.0% to 5.0%) showed an increase in cardiac disease cumulative incidence when an SMN occurred before the patient was 30 years of age, though not all were statistically significant.

To better understand the importance of those quantities, we provided the cumulative incidence curves for each landmark time in Supplemental Figure 3.

Following recommendations for the reporting of competing risks results, the effects of an SMN on cumulative incidences of death are provided in Supplemental Figures 4 and 5, and cumulative incidences conditional on survival time are provided in Supplemental Figure 6. The effect of an SMN on the cumulative incidence of death varies with time and by cancer type, with earlier times having the highest increase (17.5%; 95% CI: 11.3% to 23.7%) for patients who survived up to 20 years of age. The cumulative incidences of death conditional on survival at 15, ..., 35 years after childhood cancer diagnosis (Supplemental Figure 6) show that older patients have a death rate slightly higher than younger patients, while remaining comparable.

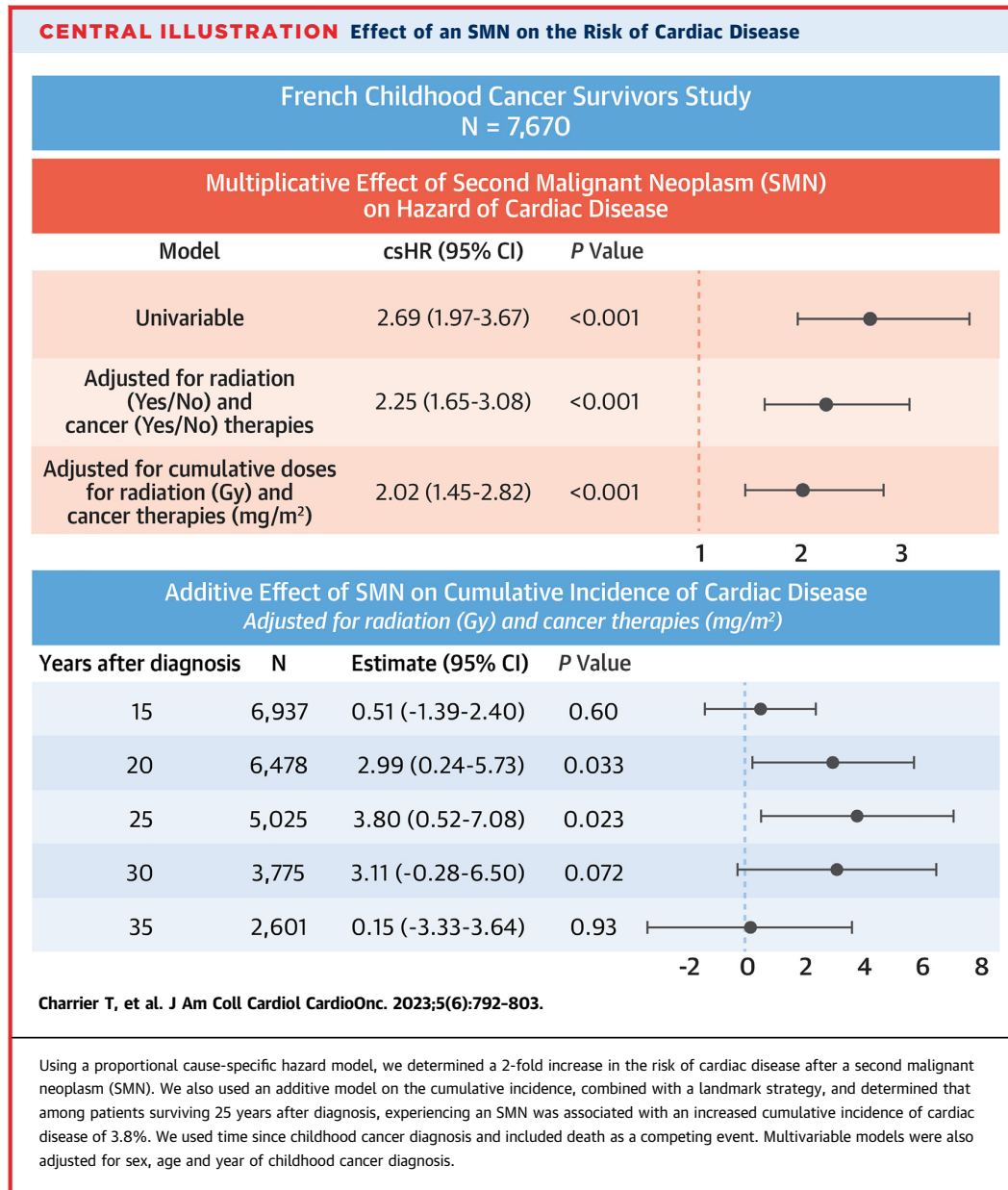
**SMN EFFECT ON THE csH OF CARDIAC DISEASE.** For all combinations of treatment-related risk factors, the occurrence of an SMN had a deleterious effect on the csH of cardiac disease (Central Illustration,





Supplemental Figure 7). In univariable analysis, the occurrence of an SMN resulted in a 3-fold increase in the hazard of cardiac disease (csHR: 2.7; 95% CI: 2.0 to 3.7). When adjusting for age at diagnosis, radiotherapy (yes/no), and chemotherapy (yes/no), the occurrence of an SMN caused a 2-fold increase in the hazard of cardiac disease (csHR: 2.3; 95% CI: 1.6 to 3.1). When including all treatment information (age at

diagnosis, sex, year of diagnosis, average radiotherapy dose at the heart and brain, exposure of the neck to radiotherapy, cumulative anthracycline dose, use of alkylating agent, use of platinum agent), the occurrence of an SMN increased the likelihood of cardiac disease (csHR: 2.0; 95% CI: 1.5 to 2.8). Using detailed doses for chemotherapy and radiotherapy doses instead of binary information did not modify



the estimation of the effect of an SMN on the csH of cardiac disease.

When categorizing the SMN status into 4 types, we observed some differences (Supplemental Figure 8). In univariable analysis, the occurrence of breast cancer (csHR: 3.5; 95% CI: 2.1 to 5.8); sarcoma, bone, or soft tissue cancer (csHR: 3.2; 95% CI: 1.6 to 6.2); and other SMN (csHR: 1.9; 95% CI: 1.5 to 2.6) was associated with an increased risk of cardiac disease. When

adjusting for all treatment information (same as previous), the occurrence of breast cancer (csHR: 1.9; 95% CI: 1.1 to 3.2); sarcoma, bone, or soft tissue cancer (csHR: 2.6; 95% CI: 1.3 to 5.0); and other SMN (csHR: 1.4; 95% CI: 1.1 to 2.0) also was associated with an increased hazard of cardiac disease. HR estimates for the competing risk of death overall and by SMN type are provided in Supplemental Figures 9 and 10, respectively.

**TABLE 2 Estimation of the Additive Effect of an SMN Occurrence on the Cumulative Incidence of Cardiac Disease**

Attained Age (y)	Covariate	$\beta_{SMN}$ (95% CI)	P Value
20	Univariable	0.62 (–1.23 to 2.46)	0.51
	Model adjusted for RT (yes/no) and CT (yes/no) <sup>a</sup>	0.29 (–1.55 to 2.13)	0.76
	Model adjusted for cumulative doses for RT (Gy) and CT (mg/m <sup>2</sup> ) <sup>b</sup>	0.50 (–1.30 to 2.29)	0.59
25	Univariable	2.33 (–0.19 to 4.84)	0.070
	Model adjusted for RT (yes/no) and CT (yes/no) <sup>a</sup>	2.04 (–0.46 to 4.55)	0.11
	Model adjusted for cumulative doses for RT (Gy) and CT (mg/m <sup>2</sup> ) <sup>b</sup>	1.88 (–0.56 to 4.33)	0.14
30	Univariable	3.87 (0.86 to 6.88)	0.012
	Model adjusted for RT (yes/no) and CT (yes/no) <sup>a</sup>	3.49 (0.52 to 6.47)	0.021
	Model adjusted for cumulative doses for RT (Gy) and CT (mg/m <sup>2</sup> ) <sup>b</sup>	3.17 (0.22 to 6.11)	0.035
35	Univariable	4.39 (1.09 to 7.68)	0.009
	Model adjusted for RT (yes/no) and CT (yes/no) <sup>a</sup>	4.01 (0.76 to 7.26)	0.016
	Model adjusted for cumulative doses for RT (Gy) and CT (mg/m <sup>2</sup> ) <sup>b</sup>	3.25 (0.01 to 6.49)	0.049

Death is included as a competing event. The time scale is attained patient age. For patients who survived  $\geq 35$  years of age, the univariable analysis shows that patients experiencing an SMN had a cumulative incidence 4.4% (95% CI: 1.1%-7.7%) higher than those who did not have an SMN. When adjusting for age at first diagnosis, RT (yes/no), and chemotherapy (yes/no), the cumulative incidence of cardiac disease was 4.0% (95% CI: 0.8%-7.3%) higher when an SMN occurred before 35 years of age. When including all treatment information (age at diagnosis, sex, year of diagnosis, average RT dose at the heart and brain, exposure of the neck to RT, cumulative anthracycline dose, use of alkylating agent), the cumulative incidence of cardiac disease was 3.3% (95% CI: 0.0%-6.5%) higher when an SMN occurred before 35 years of age. <sup>a</sup>Adjusted on sex, age at diagnosis, cumulative anthracyclines doses, alkylating agent (yes/no), mean RT dose at the heart, mean RT dose at the brain, RT at the neck (yes/no), and the year of first cancer diagnosis (before/after 1980). <sup>b</sup>Adjusted on RT (yes/no), chemotherapy (yes/no), sex, age at diagnosis, and the year of first cancer diagnosis (before/after 1980).  
CT = chemotherapy; RT = radiotherapy; SMN = second malignant neoplasm.

**TREATMENT EFFECTS.** Although this study focused on estimating the effect of an SMN and did not aim to estimate treatment effects, we did obtain csHR estimates for specific covariates while estimating the effect of an SMN (Supplemental Table 4). Among the risk factors analyzed, the most important ones were a  $>5$  Gy average heart dose,  $>250$  mg/m<sup>2</sup> cumulative dose of anthracycline, and the use of alkylating agents, which align with findings of previous studies.<sup>9,10</sup> As previously determined,<sup>33</sup> we did not find an increase of cardiac disease risk for  $<100$  mg/m<sup>2</sup> cumulative anthracycline dose.

## DISCUSSION

**CARDIAC DISEASES AMONG CCS.** In this study, we followed a well-defined population of CCS (FCCSS) over an extended treatment period spanning from 1946 to 2000. Our data collection involved a comprehensive approach, including self-reported questionnaires, hospital-based databases/registries, and clinically assessed data from survivors participating in long-term clinical follow-ups. We observed a deleterious effect of SMNs on the csH scale and the cumulative incidence of cardiac disease among CCS. We conclude that CCS who experienced an SMN had higher instantaneous and absolute risks of severe cardiac disease compared with those who did not experience an SMN. To the best of our knowledge,

this is the first study to explore severe cardiac disease occurrence after experiencing an SMN among CCS. Previous research has identified that CCS are at high risk of multimorbidity,<sup>4</sup> which is consistent with our results. This study was a first attempt to understand in depth the relationship between the multimorbidity among CCS, specifically an SMN and severe cardiac disease, 2 of the most frequent and important life-threatening adverse events. In our population of 5-year CCS, the cumulative incidence of severe cardiac disease was 3.9% (95% CI: 3.4% to 4.3%) and 8.4% (95% CI: 7.3% to 9.5%) at 30 and 50 years of age, respectively. These proportions closely align with the 30-year cumulative incidence of first severe cardiac disease in the Netherlands Cancer Institute study (4.2%; 95% CI: 2.8% to 5.6%)<sup>34</sup> and the Childhood Cancer Survivor Study (CCSS) across 27 participating institutions in the United States and Canada (4.8%; 95% CI: 4.3% to 5.2%).<sup>35</sup>

It is well known that both heart radiation and anthracycline-containing chemotherapy can increase the risk of severe cardiac disease in CCS.<sup>10,35</sup> Our findings concur with these previous studies and support current surveillance guidelines that take into account cardiotoxicity from all cancer treatments for risk stratification. Our results also suggest a possible increase in cardiac risk after an SMN. A major strength of this study is the access to a cohort of CCS with long-term follow-up and detailed clinical history,

including information on both anthracycline doses and radiotherapy doses at the heart. This allowed us to adjust for multiple configurations of childhood cancer treatment and explore the evolution of the excess of severe cardiac disease after an SMN over time. However, the small number of severe cardiac diseases after an SMN somewhat limits this strength. It will be useful to conduct the proposed analysis on larger cohorts, such as Pancare<sup>36</sup> or the CCSS.<sup>37</sup>

Many cardiotoxic childhood cancer treatments are also known to increase the risk of an SMN. We have detailed records available, allowing us to adjust on important factors, such as radiotherapy, chemotherapy, and age at childhood cancer. We observed the expected effect of these factors on cardiac risk, and the estimated effect of an SMN decreased when adjusting for them, which is consistent with previous results indicating shared risk factors for both an SMN and cardiac disease.

We observed an increase in the likelihood of cardiac disease at all times, while the effect on the cumulative incidence was time-dependent (**Figure 3, Central Illustration**). This difference may be explained by the increased risk of death of older patients after an SMN (**Supplemental Figure 6**), which can compensate for the increased risk of cardiac disease. Indeed, if the effect on the risk of cardiac disease remains stable but the instantaneous risk of death increases, the effect on the cumulative incidence will decrease, possibly disappearing or even reversing. This does not mean that CCS who experience an SMN at an older age have a better prognosis than those who are diagnosed earlier; rather, their death rate becomes so high that cardiac diseases become less relevant. The evolution of the effect of an SMN on the cumulative incidence of cardiac disease may also be influenced by the overall increase of cardiac disease at a later age, possibly reducing the excess of cardiac disease caused by an SMN, as patients may have had a cardiac disease with or without an SMN.

Finally, it is possible that an effect may exist but we may have not detected it due to the small number of patients followed for an extended period and the small number of observed events. In the latter case, larger cohorts could induce clearer results. The difference observed for younger patients could be explained by their more recent diagnosis of an SMN (**Supplemental Figure 11**), resulting in a higher death rate (**Supplemental Figure 4**) compared with patients with longer survival periods.

**STUDY LIMITATIONS.** The lack of SMN treatment data precluded a more robust analysis of the associations of SMN treatments and other factors with the

risk of cardiac disease. This information was unavailable for most of our cohort, as patients were treated at different hospitals for their SMN than for their childhood cancer. However, we did have access to the type of SMN, which was used as a proxy for SMN treatments. Due to a small number of events, we cannot draw conclusive effects of each SMN type on the cumulative incidence (**Supplemental Figure 2**), although point estimates suggest there may be differences. Results on the csH scale (**Supplemental Figure 8**) are more conclusive and show that sarcoma, bone, and soft tissue SMN induce the highest increase. This suggests that cardiotoxic SMN treatments are contributing to the increase of cardiac risk after an SMN. Aside from the cardiotoxicity of those treatments, unobserved shared risk factors of an SMN and cardiac disease may exist, and they could explain the increased cardiac risk observed for the category “Others or unknown” SMN. The effects of these factors are uncertain pending data on SMN treatment doses, which are required to conduct a mediation analysis. Another explanation could be a potential effect of having an SMN, apart from cardiotoxic SMN therapy, which may be related to an overall increased frailty of the patients. Finally, some detection bias is likely to occur due to increased care provided to SMN patients during and after their SMN treatment. To minimize this bias, we considered only cardiac events of grade 3 or higher, but this may not have fully removed the bias. We suggest that the excess of cardiac disease after an SMN is partially due to cardiotoxicity introduced by the SMN treatment, because the mentioned biases are unlikely to account for the full differences between cancer types.

The estimates by cancer type may be imprecise (large CI) due to the low number of observed events. Therefore, similar analyses in a larger cohort (Pancare, CCSS)<sup>36,37</sup> may help detect differences by cancer type. Further work, including the collection of SMN treatment dose data, would be necessary to better quantify the extent to which the increase is due to SMN treatment, identify the treatments that contribute the most to the increased occurrence of cardiac disease, and investigate if the occurrence of an SMN increases cardiac risk in a way unrelated to the cardiotoxicity of treatments.

## CONCLUSIONS

Our study shows that CCS diagnosed with an SMN have both a higher csH of cardiac disease and cumulative incidence of cardiac disease, likely related to the cardiotoxicity of SMN treatments. This finding is important because, although childhood cancer

survival rates increased through the improvement of modern therapies, this success carries an increased risk of late effects, including SMNs and cardiac disease. Therefore, our findings provide important insights for clinical practice guidelines concerning SMNs and cardiac disease post-therapy surveillance and risk-reducing strategies.

#### FUNDING SUPPORT AND AUTHOR DISCLOSURES

This work was supported and funded by the Gustave Roussy Foundation (Pediatric Program "Guérir le Cancer de l'Enfant"), the ITMO (Instituts thématiques multiorganismes) Cancer d'Aviesan Program (RadioPrediTool project no. 20CM112-00), the INCa (Institut national du cancer)/ARC (Foundation ARC for Cancer Research) foundation (CHART project), the Foundation ARC for Cancer Research (grant no. Pop-HaRC 201401208), the "START" PAIR Research Program (grant no. INCa-Fondation ARC-LNCC 11902), and the Ligue Nationale Contre le Cancer association. These funding agencies had no role in the design and conduct of the study, in the collection, management, analysis, and interpretation of the data, or in the preparation, review, and approval of the manuscript. The authors have reported that they have no relationships relevant to the contents of this paper to disclose.

**ADDRESS FOR CORRESPONDENCE:** Dr Thibaud Charrier, B2M, 114, rue Édouard Vaillant 94805 Villejuif, France. E-mail: [thibaud.charrier@gustaveroussy.fr](mailto:thibaud.charrier@gustaveroussy.fr). OR Dr Rodrigue Allodji, B2M, 114, rue Édouard Vaillant 94805 Villejuif, France. E-mail: [rodrigue.allodji@gustaveroussy.fr](mailto:rodrigue.allodji@gustaveroussy.fr).

#### PERSPECTIVES

**COMPETENCY IN MEDICAL KNOWLEDGE:** CCS experiencing an SMN are at increased risk of cardiac disease, with sarcoma, bone, soft tissue, and breast cancer survivors being at a higher risk.

**TRANSLATIONAL OUTLOOK:** Further research is needed to assess SMN treatment effects and whether or not there is a baseline effect unrelated to treatments.

#### REFERENCES

- Gatta G, Botta L, Rossi S, et al. Childhood cancer survival in Europe 1999-2007: Results of EURO-CARE-5-a population-based study. *Lancet Oncol*. 2014;15(1):35-47.
- Whelan K, Alva E. Chapter 1 - epidemiology of childhood cancer. In: Robin NH, Farmer MB, eds. *Pediatric Cancer Genetics*. Elsevier; 2018:1-20.
- Robison LL, Hudson MM. Survivors of childhood and adolescent cancer: life-long risks and responsibilities. *Nat Rev Cancer*. 2014;14(1):61-70.
- Oeffinger KC, Mertens AC, Sklar CA, et al. Chronic health conditions in adult survivors of childhood cancer. *N Engl J Med*. 2006;355:1572-1582.
- Mertens AC, Liu Q, Neglia JP, et al. Cause-specific late mortality among 5-year survivors of childhood cancer: the childhood cancer survivor study. *J Natl Cancer Inst*. 2008;100(19):1368-1379.
- Armstrong GT, Liu Q, Yasui Y, et al. Late mortality among 5-year survivors of childhood cancer: a summary from the childhood cancer survivor study. *J Clin Oncol*. 2009;27(14):2328-2338.
- Reulen RC, Winter DL, Frobisher C, et al. Long-term cause-specific mortality among survivors of childhood cancer. *JAMA*. 2010;304(2):172-179.
- Tukenova M, Guibout C, Oberlin O, et al. Role of cancer treatment in long-term overall and cardiovascular mortality after childhood cancer. *J Clin Oncol*. 2010;28(8):1308-1315.
- Lipshultz SE, Adams MJ, Colan SD, et al. Long-term cardiovascular toxicity in children, adolescents, and young adults who receive cancer therapy: pathophysiology, course, monitoring, management, prevention, and research directions. *Circulation*. 2013;128(17):1927-1995.
- Haddy N, Diallo S, El-Fayech C, et al. Cardiac diseases following childhood cancer treatment: cohort study. *Circulation*. 2016;133(1):31-38.
- Moskowitz CS, Chou JF, Neglia JP, et al. Mortality after breast cancer among survivors of childhood cancer: a report from the childhood cancer survivor study. *J Clin Oncol*. 2019;37(24):2120-2130.
- Mansouri I, Allodji RS, Hill C, et al. The role of irradiated heart and left ventricular volumes in heart failure occurrence after childhood cancer. *Eur J Heart Fail*. 2019;21(4):509-518.
- van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45(3):1-67.
- Feijen EAM, Leisenring WM, Stratton KL, et al. Derivation of anthracycline and anthraquinone equivalence ratios to doxorubicin for late-onset cardiotoxicity. *JAMA Oncol*. 2019;5(6):864-871.
- Veres C, Allodji RS, Llanas D, et al. Retrospective reconstructions of active bone marrow dose-volume histograms. *Int J Radiat Oncol Biol Phys*. 2014;90(5):1216-1224.
- Center of Epidemiology on Medical Causes of Death. Accessed August 2, 2022. <https://www.cepidc.inserm.fr/>
- Système National des Données de Santé. Accessed August 2, 2022. <https://www.snds.gov.fr/SNDS/Composantes-du-SNDS>
- Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. *J Clin Oncol*. 1983;1(11):710-719.
- Cortese G, Andersen PK. Competing risks and time-dependent covariates. *Biom J*. 2010;52(1):138-158.
- Chow EJ, Chen Y, Kremer LC, et al. Individual prediction of heart failure among childhood cancer survivors. *J Clin Oncol*. 2015;33(5):394-402.
- Chow EJ, Chen Y, Hudson MM, et al. Prediction of ischemic heart disease and stroke in survivors of childhood cancer. *J Clin Oncol*. 2018;36(1):44-52.
- Klein JP, Andersen PK. Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. *Biometrics*. 2005;61(1):223-229.
- Aalen OO, Johansen S. An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *J R Stat Soc Series B Stat Methodol*. 1978;5(3):141-150.
- Cox DR. Regression models and life-tables. *J R Stat Soc Series B Stat Methodol*. 1972;34(2):187-202.
- Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. *Biometrics*. 1978;34(4):541-554.
- Therneau T. A package for survival analysis in R. 2022. Accessed September 5, 2022. <https://CRAN.R-project.org/package=survival>
- Therneau TT, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. Springer; 2000.
- Yan J. Geepack: yet another package for generalized estimating equations. *R J*. 2002;2/3:12-14.
- Yan J, Fine JP. Estimating equations for association structures. *Stat Med*. 2004;23:859-880.
- Halekoh U, Højsgaard S, Yan J. The r package geepack for generalized estimating equations. *J Stat Softw*. 2006;15(2):1-11.

31. Perme MP, Gerster M. *Pseudo: computes pseudo-observations for modeling*. 2017. Accessed September 5, 2022. <https://CRAN.R-project.org/package=pseudo>
32. Charrier T. Code used to conduct the analysis: github. Accessed March 31, 2023. <https://github.com/T-Charrier/Effect-of-SMN-on-CD-risk>
33. Chen Y, Chow EJ, Oeffinger KC, et al. Traditional cardiovascular risk factors and individual prediction of cardiovascular events in childhood cancer survivors. *J Natl Cancer Inst*. 2019;112(3):256-265.
34. van der Pal HJ, van Dalen EC, van Delden E, et al. High risk of symptomatic cardiac events in

childhood cancer survivors. *J Clin Oncol*. 2012;30(13):1429-1437.

35. Bates JE, Howell RM, Liu Q, et al. Therapy-related cardiac risk in childhood cancer survivors: an analysis of the childhood cancer survivor study. *J Clin Oncol*. 2019;37(13):1090-1101.

36. Hjorth L, Haupt R, Skinner R, et al. Survivorship after childhood cancer: PanCare: a European network to promote optimal long-term care. *Eur J Cancer*. 2015;51(10):1203-1211.

37. Robison LL, Mertens AC, Boice JD, et al. Study design and cohort characteristics of the childhood cancer survivor study: a multi-institutional

collaborative project. *Med Pediatr Oncol*. 2002;38(4):229-239.

---

**KEY WORDS** additive model, anthracycline chemotherapy, cardiac disease, cardio-oncology, cumulative incidence, late effect, radiation

---

**APPENDIX** For an expanded Methods section and supplemental tables and figures, please see the online version of this paper.



Contents lists available at ScienceDirect

Cancer Epidemiology

journal homepage: [www.elsevier.com/locate/canep](http://www.elsevier.com/locate/canep)

## Life years lost by childhood cancer treatment and health related late effects among childhood cancer survivors

Thibaud Charrier<sup>a,b,c,\*</sup>, Nadia Haddy<sup>a,b,d</sup>, Brice Fresneau<sup>a,e</sup>, Boris Schwartz<sup>a,b,d</sup>, Neige Journy<sup>a,b,d</sup>, Charlotte Demoor-Goldschmidt<sup>b,d,f,g,h</sup>, Ibrahima Diallo<sup>i,j</sup>, Isabelle Aerts<sup>k</sup>, François Doz<sup>k,l</sup>, Vincent Souchard<sup>a,b,d</sup>, Giao Vu-Bezin<sup>a,b,d</sup>, Anne Laprie<sup>m</sup>, Sarah Lemler<sup>n</sup>, Véronique Letort<sup>n</sup>, Carole Rubino<sup>a,b,d</sup>, Kaniav Kamary<sup>n</sup>, Naïla Myriam Aba<sup>a,b</sup>, Claire Ducos<sup>a,b</sup>, Médéa Locquet<sup>a,b,d</sup>, Florent de Vathaire<sup>a,b,d</sup>, Rodrigue S. Allodji<sup>a,b,d,o</sup>, Aurélien Latouche<sup>c,p</sup>

<sup>a</sup> Université Paris-Saclay, UVSQ, Inserm, CESP, Villejuif 94807, France

<sup>b</sup> CESP, INSERM U1018, Université Paris-Saclay, UVSQ, Villejuif, France

<sup>c</sup> INSERM, U900, Institut Curie, PSL Research University, Saint-Cloud, France

<sup>d</sup> Gustave Roussy, Department of Clinical Research, Cancer and Radiation Team, Villejuif F-94805, France

<sup>e</sup> Gustave Roussy, Department of Children and Adolescent Oncology, Villejuif F-94805, France

<sup>f</sup> Department of Pediatric Oncology, Centre Hospitalier Universitaire, 4 rue Larrey, Angers 49000, France

<sup>g</sup> Department of radiotherapy, Centre François Baclesse, 3 av du Général Harris, Caen 14000, France

<sup>h</sup> Department of supportive care, Centre François Baclesse, 3 av du Général Harris, Caen 14000, France

<sup>i</sup> Department of Radiation Oncology, Gustave Roussy, Villejuif, Paris, France

<sup>j</sup> Gustave Roussy, Inserm, Radiothérapie Moléculaire et Innovation Thérapeutique, Paris-Saclay University, Villejuif, Île-de-France, France

<sup>k</sup> SIREDO Oncology Center (Care, Innovation and Research for Children, Adolescents and Young Adults with Cancer), Institut Curie, Paris, France

<sup>l</sup> Université Paris Cité, Paris, France

<sup>m</sup> Department of Radiation Oncology, Centre Antoine-Lacassagne, Nice, France

<sup>n</sup> Université Paris-Saclay, CentraleSupélec, Mathématiques et Informatique pour la Complexité et les Systèmes, Gif-sur-Yvette 91190, France

<sup>o</sup> Polytechnic School of Abomey-Calavi (EPAC), University of Abomey-Calavi, 01 P.O. Box 2009, Cotonou, Benin

<sup>p</sup> Conservatoire National des Arts et Métiers, Paris, France

### ARTICLE INFO

#### Keywords:

Childhood cancer survivors  
Late mortality  
Iatrogenic effect  
Life years lost  
Years of life lost  
Pseudo observations  
Landmark

### ABSTRACT

**Background:** Identifying risk factors contributing the most to mortality of childhood cancer survivors is essential to guide harm reduction efforts in childhood cancer treatments, and long-term follow-up of childhood cancer survivors.

**Methods:** We assessed Life Years Lost from childhood cancer treatments and their health-related late effects among the French Childhood Cancer Survivors Study, a cohort of 7670 5-year childhood cancer survivors. Using a landmark strategy, we also assessed time-varying effects of risk factors, and how the multi-morbidity affects life years lost.

**Results:** We found subsequent malignant neoplasm (9.0 years [95 %CI: 4.3–13.7]), severe cardiac disease (8.0 years [95 %CI: 1.2–14.9]), and the use of radiotherapy (6.0 years [95 %CI: 4.7–7.3]) to be the highest contributors to Life Years Lost among childhood cancer survivors. We found no interaction impact on life years lost between health related late effects considered.

**Conclusions:** Those findings suggest that radiotherapy is the root cause of early mortality among childhood cancer survivors. Moreover patients experiencing a subsequent malignant neoplasm or a cardiac disease should be monitored closely after the event, as comorbidity is common and causes premature deaths.

\* Corresponding author at: Université Paris-Saclay, UVSQ, Inserm, CESP, Villejuif 94807, France.

E-mail address: [thibaud.charrier@gustaveroussy.fr](mailto:thibaud.charrier@gustaveroussy.fr) (T. Charrier).

<https://doi.org/10.1016/j.canep.2024.102692>

Received 5 August 2024; Received in revised form 17 October 2024; Accepted 25 October 2024

Available online 7 November 2024

1877-7821/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



## 1. Introduction

Clinical progress has greatly improved 5-year survival of childhood cancer, reaching  $\sim 82\%$  in Western Europe and the USA nowadays [1]. Most 5-year survivors become very long-term survivors, but they are at increased risk for a wide range of treatment-related health-related late effects, up to and including death [1–3]. We defined mortality 5 years after childhood cancer diagnosis as late mortality, a time point at which patients are rarely at direct risk of death due to the childhood cancer itself. Previous works have summarized the magnitude, causes, and temporal patterns of childhood cancer late mortality using Standardized Mortality Ratios (SMR), Absolute Excess Risk (AER), Poisson models, and Cox models [4]. They showed three important facets of mortality: (i) excess mortality magnitude varies by cancer type and childhood cancer treatment [5], (ii) excess mortality is important at all ages and does not seem to waver at later ages [6], (iii) death by cancer recurrence/progression is the leading cause of death during the first years after diagnosis while death by Subsequent Malignant Neoplasm (SMN), Cardiac disease (CD), and Pulmonary disease is the leading cause of death starting 15–25 years after diagnosis [5].

Other studies have investigated Life Years Lost (LYL) among childhood cancer survivors, giving valuable insight into survivors' life expectancy. Yeh et al. provided LYL estimates using hazard-based models for stratified treatment effects, and cause-specific LYL [7]. Because those estimates are model-based they are susceptible to give results different than a direct regression on LYL, which avoids unnecessary uncertainties in intermediate estimates [8]. Chang et al. have studied LYL due to health-related late effects at various ages among children, teenagers, and young adult cancer survivors, but didn't link them to cancer treatments, however a necessary step to accurately guide harm reduction of childhood cancer treatments [9]. Furthermore, previous works showed that multi-morbidity is important among Childhood Cancer Survivors, with over a third of survivors who experience multiple health conditions; and is likely to be associated with both poor quality of life and premature death [10].

We aimed to build upon those previous studies and investigate concomitantly the association of LYL with childhood cancer treatments and cancer treatment-related late effects. Furthermore, we also investigated the time-varying effects of those factors. To achieve those goals, we used regression models directly on life years lost. To investigate the time-varying effects of covariates on LYL, we fitted those models across a sequence of landmark times, using age as a time scale. Finally, we investigated the presence of a two-way interaction between health-related late effects on LYL, to better the understanding of multi-morbidity consequences among childhood cancer survivors.

## 2. Material and methods

### 2.1. Data

#### 2.1.1. Population

The French Childhood Cancer Survivors Study (FCCSS [11]) cohort follows 7670 5-year survivors treated between 1945 and 2000 for solid cancer in five cancer centers in France. Included survivors aren't representative of the general french childhood cancer survivors population, as they were treated for more severe and complex cases due to the centers used for recruitment. The FCCSS received approval from a National committee on ethics and the French national agency regulating data protection (Commission Nationale Informatique et Liberté, agreements no. 902287 and no. 12038829). Written informed consent was obtained from all patients, parents, or guardians, following national research ethics requirements. The present analysis included all 7670 survivors.

#### 2.1.2. Cancer treatments exposures

Chemotherapy and radiotherapy information for childhood cancer

treatment were abstracted from medical records. For this study, chemotherapy exposures were defined as receipt (or not) of any chemotherapy agent, anthracyclines, alkylating agents, or platinum compounds.

#### 2.1.3. Outcome definitions: identification and validation of deaths and late effects

The vital status of all patients were obtained from CépiDC, which registers deaths in France [12]. Causes of deaths from CépiDC were available only for the most recent deaths, and were therefore not used in this analysis. Clinical and epidemiological follow-up, including self-administered questionnaires and cohort linkage with the French Hospital Database and Health Insurance Information System (SNDS [13]), were performed to identify the occurrence of iatrogenic effects. Long-term clinic follow-up was also performed for patients treated at Gustave Roussy Institute. SMNs and CDs were identified through these different sources and subsequently validated by a clinical research associate. Validation was based on medical, pathology, or radiological reports from the treating centers or referring doctors, regardless of the data source used for first identification. All events were graded according to the Common Terminology Criteria for Adverse Events (CTCAE version 4.03) (Fig. 1).

We included late effects occurring within the first five years after childhood cancer diagnosis. For all results including time since late effect occurrence, we defined it as the time since the last occurrence of the late effect.

### 2.2. Statistical analysis

For each subject  $i = 1, \dots, n$ , let  $T_i, C_i, \tilde{T}_i := \min(T_i, C_i)$  respectively denote the time of death by any cause, the right censoring time which is the time of last follow-up or death, Let  $\mathbf{X}$  be a time-fixed covariate, and  $\mathbf{Z}(t)$  be a time-dependent covariate at time  $t$ . We used attained age as the time scale.

#### 2.2.1. Life years lost

LYL up to a pre-specified time horizon  $\tau$  is defined as  $\tau$  - the area under the survival curve [14]. With  $S$  the Kaplan-Meier estimator,  $LYL(\tau) = \tau - \int_0^\tau S(t)dt$ . We have to use a pre-specified time horizon  $\tau$ , because few patients are observed at later times due to right censoring, therefore it would be difficult to correctly estimate LYL without  $\tau$ . Here, we set  $\tau = 60$  years old. In the following sections, whenever we talk about LYL, we're actually talking about LYL up to  $\tau$ .

LYL by a binary covariate  $\tilde{X}$  is the difference of LYL for population  $\tilde{X} = 1$  and  $\tilde{X} = 0$ ,  $LYL(\tau, \tilde{X}) = LYL(\tau|\tilde{X} = 1) - LYL(\tau|\tilde{X} = 0)$ . In a more general case we used a regression model using pseudo observations [14, 15].

#### 2.2.2. Pseudo observations for LYL

Let  $\hat{\theta} := \tau - \int_0^\tau \hat{S}(t)dt$  be a non parametric estimator of LYL, where  $S$  is the Kaplan-Meier estimator. Then for each patient  $i$ , we define the pseudo observation

$$\hat{\theta}_i = n\hat{\theta} - (n-1)\hat{\theta}^{(-i)} \quad (1)$$

where  $\hat{\theta}$  is the estimate based on the entire data set, and  $\hat{\theta}^{(-i)}$  is the estimate where subject  $i$  was discarded. Using a linear model on  $\hat{\theta}_i$ , we can estimate a linear model on LYL [15].

For more details see Appendix A.

#### 2.2.3. Regression with left truncation and right censoring

Patients were included conditionally on their survival 5 years after their childhood cancer diagnosis. Therefore failure times are subject to left truncation and right censoring. We modeled the occurrence of



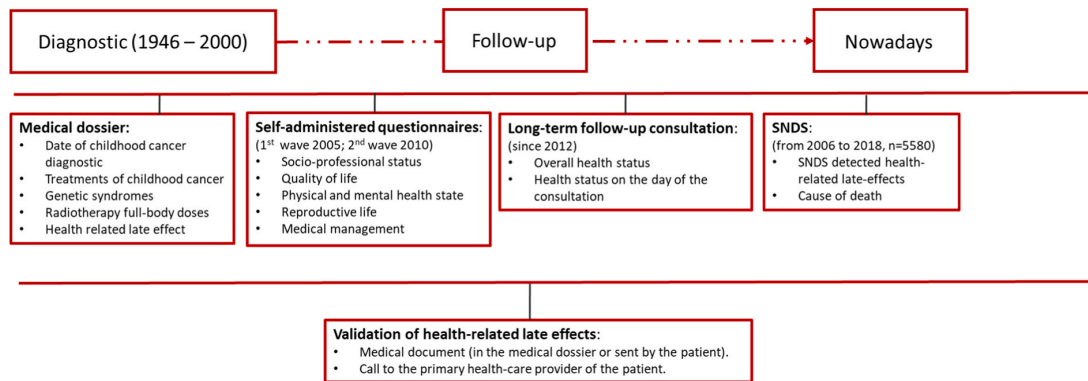


Fig. 1. Flowchart of data acquisition process in the FCCSS.

health-related late effects (e.g. second cancer) by time dependent covariate  $Z(\cdot)$ . Following [16,17] we used the landmark method to account for left truncation and time-dependent covariates. That is, we chose  $K$  landmark time points  $(l_k)_{k \in \{1, \dots, K\}}$  and performed the analysis at each  $l_k \in (l_k)_k$  using only patients at risk at  $l_k$  and substituting the time dependent covariate  $Z(\cdot)$  for its time fixed equivalent  $Z(l_k)$ . Our new pseudo observations are the *strict* pseudo observations of Grand [16].

We chose  $(l_k) = (16, 32, 48)$ , corresponding to various stages of aging while being close to the quartiles of times of death and  $l_1$  being late enough for late effects of interest to have occurred. We did not use a consensus on biologically meaningful times [18] because it did not exist, and chose equidistant time points because the use of the Aalen additive model for all-cause hazards [19] showed an overall linear trend. We also fitted our models at  $(l_k) = (6, \dots, 50)$  to verify that our choice of landmark time points didn't condition our results. To choose  $\tau$ , we considered the number of patients at risk, the number of events during at-risk windows, and the clinical relevance of this time. We chose  $\tau = 60$  for the first analysis and  $\tau' = l_k + 10$  to investigate time-varying coefficients.

#### 2.2.4. Covariates

We are interested in the effect of both multiple childhood cancer treatments and of those treatments' late effects (SMN, CD, diabetes, and chronic renal failure). LYL being positive and bound by tau, it's natural to consider a linear model to study the effect of the exposure (health-related late effects) on LYL, while adjusting for confounders (childhood cancer treatments):

$$\text{LYL}(\tau|X, Z(\cdot), l_k) = \tau - \mathbb{E}[\min(T, \tau)|X] = \beta_0 + \beta_1 X + \beta_2 Z_{l_k} \quad (2)$$

The covariates (childhood cancer treatment and late effects) were chosen using literature [9,20–22]. We retained the following childhood cancer treatment covariates for the analysis: use of radiotherapy, chemotherapy, or both radiotherapy and chemotherapy, sex, and age at childhood cancer diagnosis.

We dichotomized exposure to radiotherapy and chemotherapy because it is unclear how to adjust using doses for a population with a wide range of childhood cancer diagnoses and, therefore, treatment regimens and body areas treated.

We estimated univariable and multivariable models adjusted for all childhood cancer treatment covariates and zero or one late effect. We also estimated multivariable models adjusted for all childhood cancer treatment covariates, two late effects, and the two-way interaction of those late effects.

#### 2.2.5. Life years lost by cardiac death

We defined cardiac death as death after a CD of grade 5 (corresponding to death). We used cause-specific life years lost to perform a site analysis on cardiac death [14]. All other methodological details of

the cardiac-death analysis are the same as the main analysis performed on all-cause death.

#### 2.3. Reproducibility

The code used for the analyses is given in Appendix B. All analysis were performed using R 4.2.0. This study adhered to STROBE guidelines for observational studies [23]

### 3. Results

#### 3.1. Description of the FCCSS cohort

The demographic and treatment characteristics of the 7670 5-year childhood cancer survivors of the FCCSS cohort are presented in Table 1. The cohort includes 3469 women and 4200 men. A total of 1088 patients were treated with radiotherapy and without chemotherapy, 2574 with chemotherapy and without radiotherapy, and 3105 with both radiotherapy and chemotherapy. With a median follow-up of 30 years [Inter Quartile Range: 22.8 years, 38.7 years], we observed 1496 deaths, 828 SMN, and 380 CD of grade  $\geq 3$ . We have respectively 5371, 5110, and 1488 patients included for analysis at landmark times 16, 32, and 48 years of attained age. Cumulative incidences of death, SMN, and CD are shown in Appendix C.1.

#### 3.2. Life years lost

LYL up to 60 years by childhood cancer treatment and health-related late effects are summarized in Fig. 2 and Fig. 3. Complete results are shown in Appendix C.2.1.

Results are to be interpreted as “conditional on survival at time points 16, 32, and 48 years, exposition to covariate is associated with a reduction of restricted life expectancy of  $X$  years.” Restricted life expectancy is the life expectancy up to time  $\tau$  (here  $\tau = 60$  years old).

When adjusting for childhood cancer treatment only, we found that treating patients with radiotherapy and no chemotherapy was the treatment regimen contributing the most to LYL, with 6.0 [95 %CI: 4.7–7.4], 3.5 [95 %CI: 2.5–4.5], 1.2 [95 %CI: 0.6–1.7] LYL at landmark times 16 years old, 32 years old, and 48 years old, respectively. Use of both radiotherapy and chemotherapy was also associated with LYL, with 5.1 [95 %CI: 4.2–6.0], 2.6 [95 %CI: 1.9–3.3], 1.3 [95 %CI: 0.8–1.7] LYL at landmark times 16 years old, 32 years old, and 48 years old, respectively. Use of chemotherapy and no radiotherapy was associated with LYL, but contributed a lot less to LYL than other treatment regimen, with 0.7 [95 %CI: -0.0–1.4], 0.7 [95 %CI: 0.1–1.3], 0.6 [95 %CI: 0.1–1.0] LYL at landmark times 16 years old, 32 years old, and 48 years old, respectively.

We found no effect of sex, and only a small effect of age at diagnosis,

**Table 1**  
Description of the FCCSS cohort.

	n in FCCSS (deaths) % (% deaths)	n at 16 (deaths) % at 16 (% deaths)	n at 32 (deaths) % at 32 (% deaths)	n at 48 (deaths) % at 48 (% deaths)
<b>Overall</b>	7669 (1496) 100 % (19.5 %)	5371 (824) 100 % (15.3 %)	5110 (651) 100 % (12.7 %)	1488 (206) 100 % (13.8 %)
<b>Sex</b>				
Men	4200 (838) 54.8 % (20.0 %)	2982 (472) 55.5 % (15.9 %)	2828 (362) 55.3 % (12.8 %)	785 (105) 52.8 % (13.4 %)
Women	3469 (658) 45.2 % (19.0 %)	2389 (352) 44.5 % (14.7 %)	2282 (289) 44.7 % (12.7 %)	703 (101) 47.2 % (14.4 %)
<b>Decade of CC diagnosis</b>				
1945–1969	756 (356) 9.9 % (47.1 %)	605 (264) 11.3 % (43.6 %)	649 (252) 12.7 % (38.8 %)	531 (134) 35.7 % (25.2 %)
1970s	1676 (487) 21.9 % (29.1 %)	1275 (287) 23.7 % (22.5 %)	1433 (259) 28.0 % (18.1 %)	655 (68) 44.0 % (10.4 %)
1980s	2497 (409) 32.6 % (16.4 %)	1721 (198) 32.0 % (11.5 %)	2024 (119) 39.6 % (5.9 %)	302 (4) 20.3 % (1.3 %)
1990s	2740 (409) 35.7 % (8.9 %)	1770 (75) 33.0 % (4.2 %)	1004 (21) 19.6 % (2.1 %)	0 (0) NaN% (NaN%)
<b>Age at CC diagnosis</b>				
< 1 years old	1244 (135) 16.2 % (10.9 %)	1186 (97) 22.1 % (8.2 %)	598 (42) 11.7 % (7.0 %)	139 (13) 9.3 % (9.4 %)
1–4 years old	2482 (474) 32.4 % (19.1 %)	2322 (355) 43.2 % (15.3 %)	1410 (190) 27.6 % (13.5 %)	317 (48) 21.3 % (15.1 %)
5 – 9 years old	1677 (396) 21.9 % (23.7 %)	1584 (314) 29.5 % (19.9 %)	1159 (167) 22.7 % (14.4 %)	332 (43) 22.3 % (13.0 %)
10 – 14 years old	1623 (374) 21.2 % (23.0 %)	279 (58) 5.2 % (20.8 %)	1374 (199) 26.9 % (14.5 %)	516 (84) 34.7 % (16.3 %)
15+ years old	643 (117) 8.4 % (18.2 %)	0 (0) NaN% (NaN%)	569 (53) 11.1 % (9.3 %)	184 (18) 12.4 % (9.8 %)
<b>Treatment combination</b>				
Nor radiotherapy nor chemotherapy	902 (54) 11.8 % (6.0 %)	631 (27) 11.7 % (4.3 %)	535 (20) 10.5 % (3.7 %)	126 (5) 8.5 % (4.0 %)
Radiotherapy alone	1088 (369) 14.2 % (33.9 %)	739 (240) 13.8 % (32.5 %)	827 (222) 16.2 % (26.8 %)	371 (91) 24.9 % (24.5 %)
Chemotherapy alone	2574 (197) 33.6 % (7.7 %)	1874 (91) 34.9 % (4.9 %)	1630 (62) 31.9 % (3.8 %)	253 (8) 17.0 % (3.2 %)
Radiotherapy and chemotherapy	3105 (876) 40.5 % (28.2 %)	2127 (466) 39.6 % (21.9 %)	2118 (347) 41.4 % (16.4 %)	738 (102) 49.6 % (13.8 %)
<b>Anthracycline dose</b>				
< 100 mg/m2	5268 (1149) 68.7 % (21.8 %)	3790 (663) 70.6 % (17.5 %)	3444 (541) 67.4 % (15.7 %)	1133 (184) 76.1 % (16.2 %)
[100–250] mg/m2	1257 (129) 16.4 % (10.3 %)	863 (57) 16.1 % (6.6 %)	773 (48) 15.1 % (6.2 %)	101 (8) 6.8 % (7.9 %)
> 250 mg/m2	1144 (218) 14.9 % (19.1 %)	718 (104) 13.4 % (14.5 %)	893 (62) 17.5 % (6.9 %)	254 (14) 17.1 % (5.5 %)
<b>Radiotherapy dose at the heart</b>				
< 5 Gy	5971 (968)	4301 (546)	3955 (403)	1078 (135)

**Table 1 (continued)**

	n in FCCSS (deaths) % (% deaths)	n at 16 (deaths) % at 16 (% deaths)	n at 32 (deaths) % at 32 (% deaths)	n at 48 (deaths) % at 48 (% deaths)
	77.8 % (16.2 %)	80.1 % (12.7 %)	77.4 % (10.2 %)	72.4 % (12.5 %)
[5,20] Gy	912 (249) 11.9 % (27.4 %)	617 (155) 11.5 % (25.2 %)	676 (121) 13.2 % (17.9 %)	255 (39) 17.1 % (15.3 %)
> 20 Gy	395 (186) 5.1 % (47.1 %)	196 (83) 3.6 % (42.3 %)	300 (111) 5.9 % (37.0 %)	119 (28) 8.0 % (23.5 %)
Missing data	391 (93) 5.1 % (23.8 %)	257 (40) 4.8 % (15.6 %)	179 (16) 3.5 % (8.9 %)	36 (4) 2.4 % (11.1 %)
<b>Type of childhood cancer</b>				
Unknown	10 (0) 0.1 % (0.0 %)	7 (0) 0.1 % (0.0 %)	9 (0) 0.2 % (0.0 %)	3 (0) 0.2 % (0.0 %)
02 -Lymphomas	1278 (219) 16.7 % (17.1 %)	724 (108) 13.5 % (14.9 %)	1054 (143) 20.6 % (13.6 %)	325 (45) 21.8 % (13.8 %)
03 -CNS tumor	1140 (469) 14.9 % (41.1 %)	790 (264) 14.7 % (33.4 %)	641 (171) 12.5 % (26.7 %)	153 (39) 10.3 % (25.5 %)
04 -Peripheral nervous tumors	1034 (131) 13.5 % (12.7 %)	950 (82) 17.7 % (8.6 %)	576 (44) 11.3 % (7.6 %)	129 (17) 8.7 % (13.2 %)
05 - Retinoblastomas	619 (68) 8.1 % (11.0 %)	583 (52) 10.9 % (8.9 %)	177 (16) 3.5 % (9.0 %)	23 (6) 1.5 % (26.1 %)
06 -Renal tumors	1140 (196) 14.9 % (17.2 %)	1052 (157) 19.6 % (14.9 %)	798 (121) 15.6 % (15.2 %)	267 (42) 17.9 % (15.7 %)
07 -Hepatic tumors	79 (9) 1.0 % (11.4 %)	71 (5) 1.3 % (7.0 %)	45 (3) 0.9 % (6.7 %)	5 (0) 0.3 % (0.0 %)
08 -Bone sarcomas	686 (144) 8.9 % (21.0 %)	236 (39) 4.4 % (16.5 %)	528 (42) 10.3 % (8.0 %)	178 (18) 12.0 % (10.1 %)
09 -Soft-tissue sarcomas	859 (147) 11.2 % (17.1 %)	586 (72) 10.9 % (12.3 %)	613 (52) 12.0 % (8.5 %)	204 (20) 13.7 % (9.8 %)
10 -Germ cells and gonadal tumors	469 (50) 6.1 % (10.7 %)	244 (22) 4.5 % (9.0 %)	388 (26) 7.6 % (6.7 %)	115 (9) 7.7 % (7.8 %)
11 -Other carcinomas	344 (60) 4.5 % (17.4 %)	122 (22) 2.3 % (18.0 %)	275 (33) 5.4 % (12.0 %)	84 (10) 5.6 % (11.9 %)
12 -Other or unspecified tumors	11 (3) 0.1 % (27.3 %)	6 (1) 0.1 % (16.7 %)	6 (0) 0.1 % (0.0 %)	2 (0) 0.1 % (0.0 %)
<b>Cardiac Disease (grade &gt;= 3)</b>				
No	7289 (1332) 95.0 % (18.3 %)	5327 (810) 99.2 % (15.2 %)	4988 (616) 97.6 % (12.3 %)	1400 (186) 94.1 % (13.3 %)
Yes	380 (164) 5.0 % (43.2 %)	44 (14) 0.8 % (31.8 %)	122 (35) 2.4 % (28.7 %)	88 (20) 5.9 % (22.7 %)
<b>Second Malignant Neoplasm</b>				
No	6841 (1119) 89.2 % (16.4 %)	5285 (785) 98.4 % (14.9 %)	4866 (584) 95.2 % (12.0 %)	1273 (149) 85.6 % (11.7 %)
Yes	828 (377) 10.8 % (45.5 %)	86 (39) 1.6 % (45.3 %)	244 (67) 4.8 % (27.5 %)	215 (57) 14.4 % (26.5 %)
<b>Diabetes</b>				

(continued on next page)

Table 1 (continued)

	n in FCCSS (deaths)	n at 16 (deaths)	n at 32 (deaths)	n at 48 (deaths)
	% (% deaths)	% at 16 (% deaths)	% at 32 (% deaths)	% at 48 (% deaths)
No	7530 (1476)	5361 (824)	5064 (645)	1426 (196)
	98.2 % (19.6 %)	99.8 % (15.4 %)	99.1 % (12.7 %)	95.8 % (13.7 %)
Yes	139 (20)	10 (0)	46 (6)	62 (10)
	1.8 % (14.4 %)	0.2 % (0.0 %)	0.9 % (13.0 %)	4.2 % (16.1 %)
<b>Chronic Renal Failure</b>				
No	7557 (1460)	5352 (818)	5070 (643)	1462 (203)
	98.5 % (19.3 %)	99.6 % (15.3 %)	99.2 % (12.7 %)	98.3 % (13.9 %)
Yes	112 (36)	19 (6)	40 (8)	26 (3)
	1.5 % (32.1 %)	0.4 % (31.6 %)	0.8 % (20.0 %)	1.7 % (11.5 %)

Note: Column “n in FCCSS” includes details of all patients included in the FCCSS. Columns “n at 16”, “n at 32”, and “n at 48” includes patients included for analysis at each respective landmark time. Format is “number of patient (number of those patients who died)”, and the line below corresponds to “% of patients of this column who have this level (% of patients with this level who died)”.

with patients diagnosed at 1–4 years old faring the worst, with 1.2 [95 % CI: 0.4–2.0], 1.5 [95 % CI: 0.7–2.3], 0.6 [95 % CI: –0.0–1.3] LYL at landmark times 16 years old, 32 years old, and 48 years old respectively, using patients diagnosed at < 1 year old as reference.

SMN and CD were both associated with LYL. When adjusting for childhood cancer treatment, SMN point estimates were higher than CD point estimates, which is consistent with the usual hierarchy of SMN contributing more to mortality than CD. SMN diagnosed less than 5 years before landmark time point was associated with 9.0 [95 % CI: 4.3–13.7], 3.1 [95 % CI: 1.2–5.1], 1.6 [95 % CI: 0.7–2.4] LYL at

landmark times 16 years old, 32 years old, and 48 years old respectively. CD diagnosed less than 5 years before landmark time point was associated with 8.0 [95 % CI: 1.2–14.9], 5.3 [95 % CI: 1.2–9.3], 2.3 [95 % CI: 0.3–4.3] LYL at landmark times 16 years old, 32 years old, and 48 years old respectively.

Using univariable models, we found an association between LYL and SMN (11.0 [95 % CI: 5.9–16.2] at 16 years, 3.8 [95 % CI: 0.9–6.7] at 32 years, and 2.2 [95 % CI: 0.6–3.9] at 48 years) and CD (7.7.0 [95 % CI: 0.5–14.8] at 16 years, 5.9 [95 % CI: 1.6–10.0] at 32 years, and 2.5 [95 % CI: 0.5–4.6] at 48 years), but not between LYL and diabetes and chronic renal failure. We found no two-way interaction for either combination of SMN, CD, diabetes, and chronic renal failure.

### 3.3. Time varying effects

Results across all three at-risk time periods after landmark time points ( $\tau$ ) showed similar trends of small time-varying effects. Results of LYL for  $\tau = 10$  are displayed in Fig. 4 and Fig. 5 while complete results are shown in Appendix C.2.2. A recent SMN diagnosis has a strong association with 10-LYL at all ages (1.1 [95 % CI: 0.4–1.9] at 16 years, 0.4 [95 % CI: 0.1–0.8] at 32 years, 1.3 [95 % CI: 0.6–1.9] at 48 years). A recent CD diagnosis has a stronger association with 10-LYL at older ages, and no clear association at early ages (0.4 [95 % CI: – 0.5 – 1.3] at 16 years, 1.2 [95 % CI: 0.2–2.3] at 32 years, 1.8 [95 % CI: 0.2–3.3] at 48 years). 10-LYL by radiotherapy appears to increase as patients age (0.2 [95 % CI: 0.1–0.3] at 16 years, 0.4 [95 % CI: 0.3–0.6] at 32 years, 0.8 [95 % CI: 0.4–1.2] at 48 years). We observe the same trend for use of radiotherapy with chemotherapy, and for use of chemotherapy alone.

### 3.4. Cardiac death

LYL by cardiac death up to 60 years are shown in Appendix D. Our analysis of the cohort observations yielded an average of 0.3 LYL by cardiac death, out of the average 8 LYL by all-cause death.

When adjusting for average radiotherapy dose at the heart and brain, cumulative dose of anthracyclines, sex, and age at diagnosis, we found

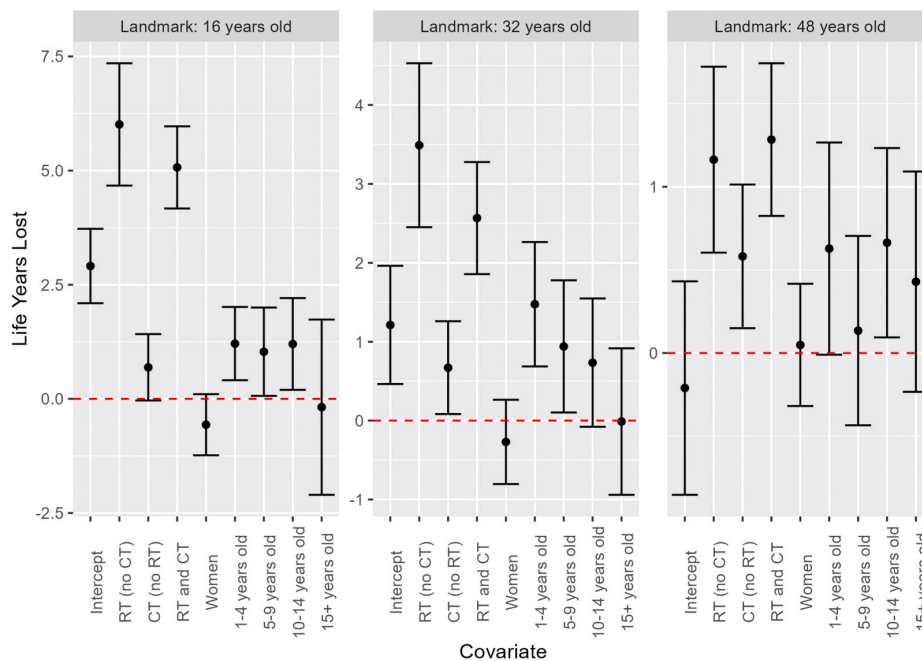
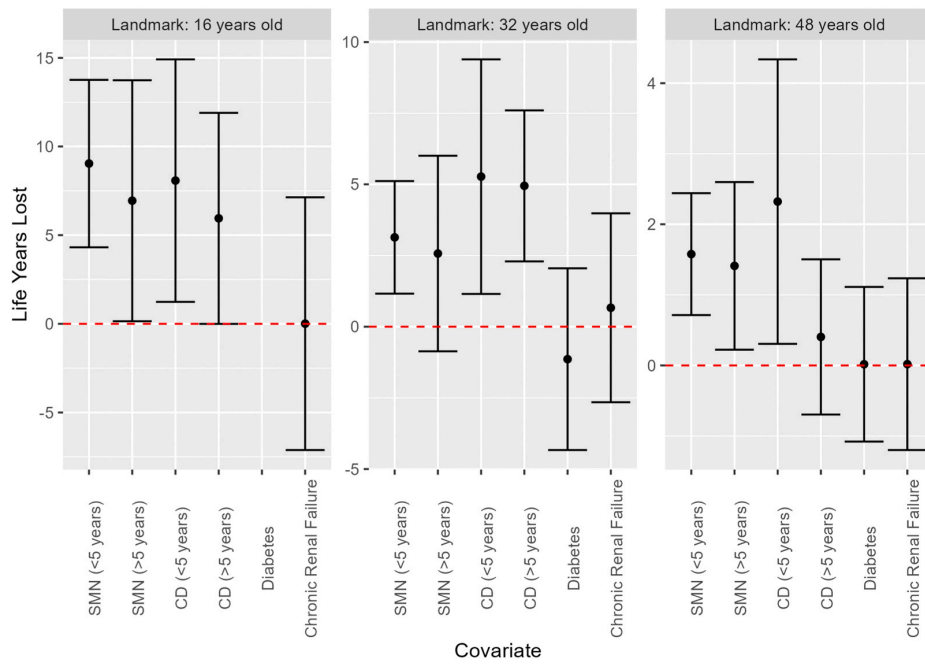
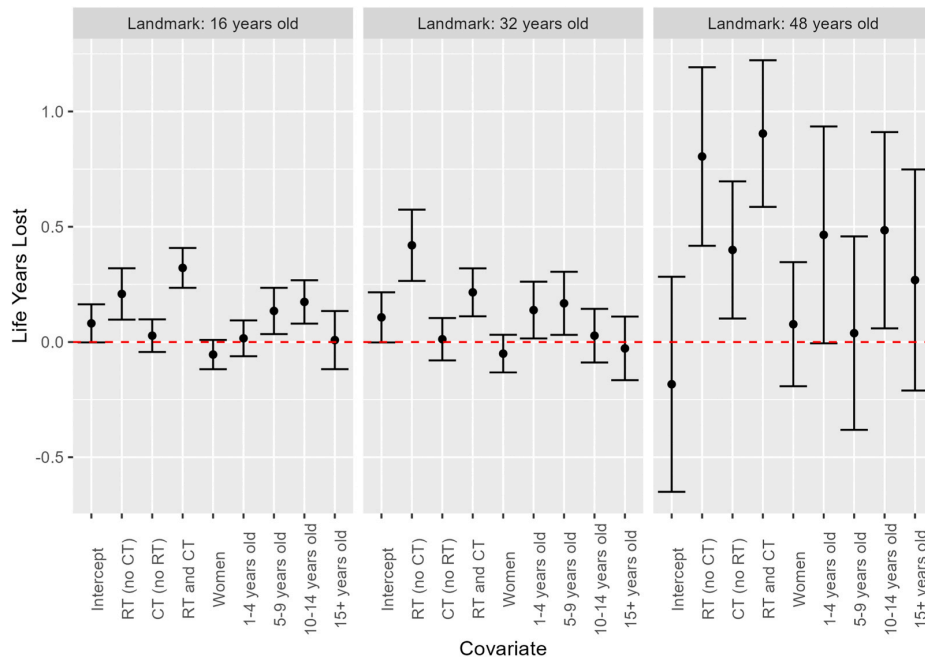


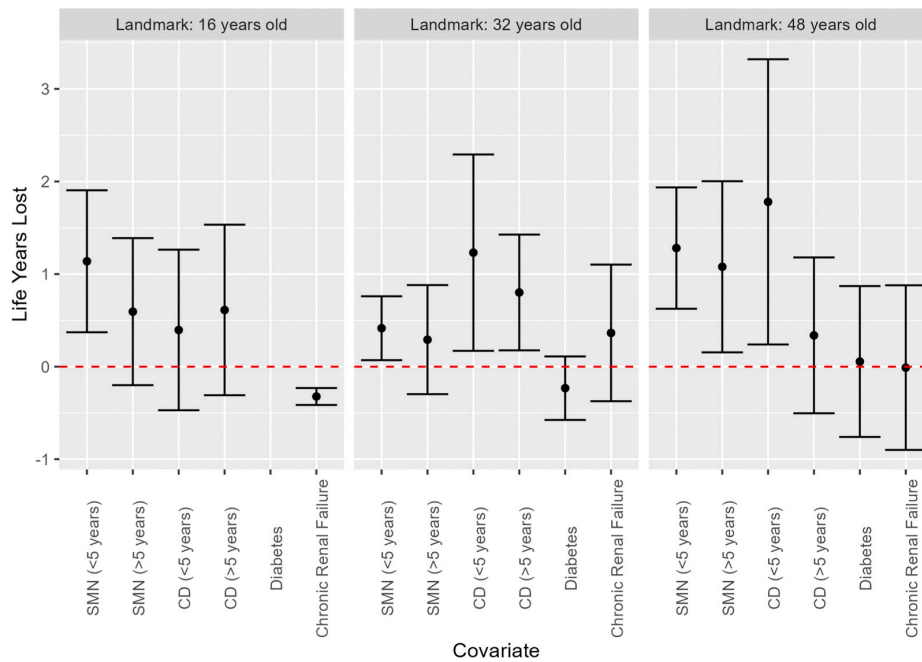
Fig. 2. Life Years Lost by all cause death up to 60 years old, conditional on survival at each landmark time. Errorbars represent 95 % confidence interval. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, Sex, and age at diagnosis. CT = Chemotherapy, RT = Radiotherapy.



**Fig. 3.** Life Years Lost by all cause death up to 60 years old, conditional on survival at each landmark time. Errorbars represent 95 % confidence interval. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, Sex, and age at diagnosis, and one of either Subsequent Malignant Neoplasm, Cardiac Disease, Diabetes, or Chronic Renal Failure. Subsequent Malignant Neoplasm and Cardiac Disease are separated conditional on occurring within the 5 years prior to landmark time. CD = Cardiac Disease, SMN = Subsequent Malignant Neoplasm.



**Fig. 4.** Life Years Lost by all cause death within 10 years of landmark time, conditional on survival at each landmark time. Errorbars represent 95 % confidence interval. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, Sex, and age at diagnosis. CT = Chemotherapy, RT = Radiotherapy.



**Fig. 5.** Life Years Lost by all cause death within 10 years of landmark time, conditional on survival at each landmark time. Errorbars represent 95 % confidence interval. All estimates correspond to a model adjusted on intercept, radiotherapy and chemotherapy, Sex, and age at diagnosis, and one of either Subsequent Malignant Neoplasm, Cardiac Disease, Diabetes, or Chronic Renal Failure. Subsequent Malignant Neoplasm and Cardiac Disease are separated conditional on occurring within the 5 years prior to landmark time. CD = Cardiac Disease, SMN = Subsequent Malignant Neoplasm.

LYL up to 60 years old by cardiac death to be associated with average radiotherapy dose at the heart higher than 20 Gy to an important extent (2.6 [95 %CI: 1.5–3.6] at 16 years, 1.8 [95 %CI: 1.0–2.6] at 32 years, 0.0 [95 %CI: – 0.1 – 0.3] at 48 years), but found no association with other variables, including cumulative dose of anthracyclines higher than 250 mg/m<sup>2</sup>.

**4. Discussion**

In this retrospective with a prospective follow-up cohort study of childhood cancer survivors, we examined: 1) the associations between childhood cancer treatments and Life Years Lost, 2) the associations between health-related late effects of childhood cancer treatments and LYL, 3) trends across survivors’ life span of those associations. We found that radiotherapy was the childhood cancer treatment most strongly associated with LYL and that other childhood cancer treatments’ associations with LYL were very small in comparison. We also showed that both SMN and CD were important contributors to the LYL by childhood cancer survivors. We found no two-way interaction between two health related effects and LYL. A supplementary analysis showed some small difference of life years lost by radiotherapy and chemotherapy when stratifying on decade of childhood cancer diagnosis, without a clear trend (Appendix C). Those findings suggest that radiotherapy is the principal root cause of early mortality among CCS, supporting the need of long-term follow-up of all patients treated with RT in childhood. Personalized long-term follow-up care plan should be established for all survivors, especially those exposed to RT, with prevention and screening procedures according to published guidelines. Moreover patients experiencing a SMN or a CD should be monitored closely after the event, as comorbidity is common and causes premature deaths.

Previous works on this population focused on epidemiological measures such as Standardized Mortality Ratios, Absolute Excessive Risk, and Proportional Hazard Ratios (PHR). These are useful for clinicians and public health decision-makers, but require extra work and care

to be interpreted and communicated to patients who may lack the knowledge to interpret relative metrics. Moskalewicz et al. summarized SMR and AER results of previous studies on Childhood Cancer Survivors, with SMR ranging from 5.6 to 17.2 and AER ranging from 34.1 to 70.1 excess deaths per 10,000 person years [4].

Although SMR, AER, and PHR are the most common metrics to study childhood cancer survivors, previous studies did analyze the LYL by health conditions among children, teenagers, and young adult cancer survivors [9]. Similarly to us, they assessed health-related late effects associations with Life Years Lost depending on the age at the health condition onset. They did not adjust for childhood cancer treatment, or use a  $\tau$ , but found results coherent with ours regarding Life Years Lost by SMN and CDs at 32 and 48 years old. They found CD to be associated with 10.13 (95 %CI: [7.13–14.30]) LYL at 32.5 years old, and –0.29 (95 %CI: [– 0.29 – 0.29]) LYL at 45 years old. They also found Subsequent Malignant Neoplasm to be associated with 11.67 (95 %CI: [9.29–15.27]) LYL at 32.5 years old, and 1.06 (95 %CI: [0.00–2.98]) LYL at 45 years old. The differences with our results can be attributed to differences in population included (solid tumors and lymphomas survivors vs all childhood cancer survivors), in at-risk periods (1950–2024 vs 1998–2020), in control populations (healthy childhood cancer survivors vs community control), methodology used regarding LYL computation, time since health related late effect onset, and inclusion of additional covariates.

LYL are often combined with causes of death and studied as Cause Specific LYL. This allows for a simple visualization ranking causes of death by the magnitude of LYL. This also enables the study of risk factors associated with each cause of death, such as investigating if anthracyclines contribute to non-cardiac death. Causes of death were not usable in our case; therefore, we studied all-cause death. Nonetheless, we still provided valuable results on all-cause mortality and did a side analysis on cardiac death.

A limitation of the study is the lack of access to alcohol consumption, smoking status, and socio-economic data. Those factors are known to be

associated with mortality, and previous works have shown an association between those and childhood cancer type (and therefore treatment regimen and health-related late effects) [24–26]. This may bias the results towards a higher association of health-related late effects and death. Nonetheless, we believe the marginal associations of secondary cancer and cardiac disease with LYL to be of interest. The FCCSS population was treated prior to 2000, and treatment regimens have evolved with time. Therefore, our results may not be translatable to patients being treated for childhood cancer nowadays. We also suggest that using a landmark strategy and looking at the effect of SMN/CD already diagnosed at that time provide estimates less susceptible to those biases than Cause Specific LYL. The design of the study (analysis at various landmark time points) shifts away from survival bias to results conditional on survival. The interpretation of results still must be made by keeping in mind that patients surviving up to a landmark time point are survivors of survivors, and therefore a special case of initial survivors. This is especially true regarding health-related late effects, which often occurred months before each landmark time point.

Further work is needed to investigate with more significant details how radiotherapy doses are linked to disparities of mortality by childhood cancer type, especially to determine which combination of radiotherapy doses and organ exposure contributes the most to late mortality.

## 5. Conclusions

Using Life Years Lost, we provided a new perspective on the mortality of Childhood Cancer Survivors, detailing the significant impact of radiotherapy use during childhood cancer, Subsequent Malignant Neoplasms, and Cardiac Diseases on the life expectancy of patients at various stages of their life.

### List of abbreviations

AER: Absolute Excessive Risk  
 CD: Cardiac Disease  
 CI: Confidence Interval  
 LYL: Life Years Lost  
 PHR: Proportional Hazard Ratio  
 SMN: Subsequent Malignant Neoplasm  
 SMR: Standardized Mortality Rate

### Ethics approval

The FCCSS received approval from a National committee on ethics and the French national agency regulating data protection (Commission Nationale Informatique et Liberté, agreements no. 902287 and no. 12038829).

### Consent to participate

Written informed consent was obtained from all patients, parents, or guardians, following national research ethics requirements

### Consent to publish

Not applicable.

### Funding

This work was supported and funded by the Gustave Roussy Foundation (Pediatric Program ‘Guérir le Cancer de l’Enfant’), the ITMO (Instituts thématiques multiorganismes) Cancer d’Aviesan Program (RadioPrediTool project no. 20CM112–00), the INCa/ARC (Institut national du cancer) foundation (CHART project), the Foundation ARC for Cancer Research (grant no. Pop-HaRC 201401208), the ‘START’ PAIR Research Program (grant no. INCa-Fondation ARC-LNCC 11902), and

the ‘Ligue Nationale Contre le Cancer’ association. These funding agencies had no role in the design and conduct of the study, in the collection, management, analysis, and interpretation of the data, or in the manuscript’s preparation, review, and approval.

### CRedit authorship contribution statement

**Médée Locquet:** Writing – review & editing, Conceptualization. **Ibrahima Diallo:** Writing – review & editing, Investigation, Data curation. **Claire Ducos:** Writing – review & editing, Conceptualization. **Charlotte Demoor-Goldschmidt:** Writing – review & editing, Investigation, Data curation. **Vincent Souhard:** Writing – review & editing, Resources, Data curation. **Florent de Vathaire:** Supervision, Investigation, Funding acquisition, Data curation. **François Doz:** Writing – review & editing, Investigation. **Aurélien Latouche:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Giao Vu-Bezin:** Writing – review & editing, Investigation, Data curation. **Rodrigue S. Alodji:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Sarah Lemler:** Writing – review & editing, Methodology. **Thibaud Charrier:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Formal analysis, Conceptualization. **Anne Laprie:** Writing – review & editing, Investigation. **Carole Rubino:** Writing – review & editing, Investigation, Data curation. **Brice Fresneau:** Writing – review & editing, Investigation, Data curation, Conceptualization. **Véronique Letort:** Writing – review & editing, Methodology. **Nadia Haddy:** Writing – review & editing, Funding acquisition, Data curation. **Naïla Myriam Aba:** Writing – review & editing, Conceptualization. **Neige Journy:** Writing – review & editing, Funding acquisition, Data curation. **Kaniav Kamary:** Writing – review & editing, Methodology. **Boris Schwartz:** Writing – review & editing, Conceptualization.

### Declaration of Competing Interest

None

### Data Availability

The datasets generated and/or analysed during the current study are not publicly available due to patients confidentiality, but are available from the corresponding author on reasonable request.

### Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.canep.2024.102692](https://doi.org/10.1016/j.canep.2024.102692).

### References

- [1] G. Gatta, L. Botta, S. Rossi, T. Aareleid, M. Bielska-Lasota, J. Clavel, et al., Childhood cancer survival in Europe 1999–2007: results of EUROcare-5—a population-based study, *Lancet Oncol.* 15 (2014) 35–47.
- [2] R.L. Harrington, D.M. Qato, J.W. Antoon, R.N. Caskey, G.T. Schumock, T.A. Lee, Impact of multimorbidity subgroups on the health care use of early pediatric cancer survivors, *Cancer* 126 (2020) 649–658.
- [3] L.L. Robison, M.M. Hudson, Survivors of childhood and adolescent cancer: Life-long risks and responsibilities, *Nat. Rev. Cancer* 14 (2014) 61–70.
- [4] Moskalewicz A., Martinez B., Uleryk E.M., Pechlivanoglou P., Gupta S., Nathan P.C. Late mortality among 5-year survivors of childhood cancer: A systematic review and meta-analysis. *Cancer*.2024; n/a.
- [5] G.T. Armstrong, Y. Chen, Y. Yasui, W. Leisenring, T.M. Gibson, A.C. Mertens, et al., Reduction in late mortality among 5-year survivors of childhood cancer, *N. Engl. J. Med* 374 (2016) 833–842.
- [6] S.B. Dixon, Q. Liu, E.J. Chow, K.C. Oeffinger, P.C. Nathan, R.M. Howell, et al., Specific causes of excess late mortality and association with modifiable risk factors among survivors of childhood cancer: a report from the childhood cancer survivor study cohort, *Lancet* (2023) 0.

- [7] J.M. Yeh, L. Nekhlyudov, S.J. Goldie, A.C. Mertens, L. Diller, A model-based estimate of cumulative excess mortality in survivors of childhood cancer, *Ann. Intern Med* 152 (2010) 409–417.
- [8] P.K. Andersen, Life years lost among patients with a given disease, *Stat. Med.* 36 (2017) 3573–3582.
- [9] W.H. Chang, M. Katsoulis, Y.Y. Tan, S.H. Mueller, K. Green, A.G. Lai, Late effects of cancer in children, teenagers and young adults: Population-based study on the burden of 183 conditions, in-patient and critical care admissions and years of life lost, *Lancet Reg. Health - Eur.* 12 (2022) 100248.
- [10] K.C. Oeffinger, T. Kawashima, D.L. Friedman, N.S. Kadan-Lottick, L.L. Robison, Chronic health conditions in adult survivors of childhood cancer, *N. Engl. J. Med* (2006) 11.
- [11] I. Mansouri, R.S. Allodji, C. Hill, C. El-Fayech, F. Pein, S. Diallo, et al., The role of irradiated heart and left ventricular volumes in heart failure occurrence after childhood cancer, *Eur. J. Heart Fail.* 21 (2019) 509–518.
- [12] Center of epidemiology on medical causes of death. (<https://www.cepdc.inserm.fr/>). Accessed May 2022.2022.
- [13] SNDS. (<https://www.snds.gouv.fr/SNDS/Composantes-du-SNDS>). Accessed May 2022.2022.
- [14] P.K. Andersen, Decomposition of number of life years lost according to causes of death, *Stat. Med.* 32 (2013) 5278–5285.
- [15] J.P. Klein, P.K. Andersen, Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function, *Biometrics* 61 (2005) 223–229.
- [16] M.K. Grand, H. Putter, Regression models for expected length of stay, *Stat. Med.* 35 (2016) 1178–1192.
- [17] G. Cortese, P.K. Andersen, Competing risks and time-dependent covariates, *Biom. J.* 52 (2010) 138–158.
- [18] C.J. Morgan, Landmark analysis: a primer, *J. Nucl. Cardiol.* 26 (2019) 391–393.
- [19] O.O. Aalen, A linear regression model for the analysis of life times, *Stat. Med.* 8 (1989) 907–925.
- [20] M. Tukenova, C. Guibout, O. Oberlin, F. Doyon, A. Mousannif, N. Haddy, et al., Role of cancer treatment in long-term overall and cardiovascular mortality after childhood cancer, *JCO* 28 (2010) 1308–1315.
- [21] K.F. Wong, R.C. Reulen, D.L. Winter, J. Guha, M.M. Fidler, J. Kelly, et al., Risk of adverse health and social outcomes up to 50 years after wilms tumor: The british childhood cancer survivor study, *J. Clin. Oncol.* 34 (2016) 1772–1779.
- [22] M.J. Ehrhardt, K.R. Krull, N. Bhakta, Q. Liu, Y. Yasui, L.L. Robison, et al., Improving quality and quantity of life for childhood cancer survivors globally in the twenty-first century, *Nat Rev Clin. Oncol.* (2023) 1–19.
- [23] E. vonElm, D.G. Altman, M. Egger, S.J. Pocock, P.C. Gotsche, J.P. Vandenbroucke, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies, *J. Clin. Epidemiol.* 61 (4) (2008) 344–349. PMID: 18313558.
- [24] A. Dumas, I. Cailbault, C. Perrey, O. Oberlin, F. De Vathaire, P. Amiel, Educational trajectories after childhood cancer: when illness experience matters, *Soc. Sci. Med.* 135 (2015) 67–74.
- [25] M. Ernst, A. Hinz, E. Brähler, H. Merzenich, J. Faber, P.S. Wild, et al., Quality of life after pediatric cancer: comparison of long-term childhood cancer survivors' quality of life with a representative general population sample and associations with physical health and risk indicators, *Health Qual. Life Outcomes* 21 (2023) 65.
- [26] N. Langeveld, H. Stam, M. Grootenhuys, B. Last, Quality of life in young adult survivors of childhood cancer, *Support Care Cancer* 10 (2002) 579–600.