



HAL
open science

Étude de la survenue de maladies cardiaques à long terme après le traitement du cancer dans l'enfance : Identification de marqueurs biologiques associés

Naïla Aba

► **To cite this version:**

Naïla Aba. Étude de la survenue de maladies cardiaques à long terme après le traitement du cancer dans l'enfance : Identification de marqueurs biologiques associés. Cancer. Université Paris-Saclay, 2024. Français. NNT : 2024UPASR012 . tel-04753458

HAL Id: tel-04753458

<https://theses.hal.science/tel-04753458v1>

Submitted on 25 Oct 2024

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.

Étude de la survenue de maladies cardiaques à long terme après le traitement du cancer dans l'enfance : Identification de marqueurs biologiques associés

*Occurrence of long-term cardiac diseases after childhood cancer treatment:
Identification of associated biomarkers*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°570, Santé Publique (EDSP)

Spécialité de doctorat : Epidémiologie

Graduate school : Santé publique, Référent : Faculté de Médecine

Thèse préparée dans l'unité de recherche CESP (**Université Paris-Saclay, UVSQ, Inserm**), sous la direction de **Gwénaél Le Teuff** (PhD, HDR) et la co-direction de **Nadia Haddy** (CR)

Thèse soutenue à Paris-Saclay, le 23 septembre 2024 par

Naila ABA

Composition du Jury

Membres du jury avec voix délibérative

Pierre BLANCHARD PU-PH, INSERM U1018 - Université Paris-Saclay	Président
Céline BOURGIER PU-PH, INSERM U1194 – Université de Montpellier	Rapporteur & Examinatrice
Karen LEFFONDRE Professeure, INSERM U1219 - Université de Bordeaux	Rapporteur & Examinatrice
Jacqueline CLAVEL DR, INSERM U1153 - Université Paris Cité	Examinatrice
Marie-Catherine VOZENIN Professeure, Hôpitaux universitaires de Genève	Examinatrice

Titre : Étude de la survenue de maladies cardiaques à long terme après le traitement du cancer dans l'enfance: Identification de marqueurs biologiques associés.

Mots clés : Cardiotoxicité; Biomarqueurs; Survivants d'un cancer de l'enfant; Transcriptome; Radiothérapie; Anthracyclines

Résumé : Les progrès en oncologie pédiatrique ont conduit à améliorer le pronostic des patients traités pour un cancer de l'enfant ou de l'adolescent avec une survie à 5 ans supérieure à 80%. Cependant, ces patients présentent un risque accru de complications tardives secondaires au traitement du cancer. Les pathologies cardiaques sont une préoccupation majeure au sein de cette population de survivants ayant été exposés à des traitements cardiotoxiques comme les anthracyclines et/ou les radiations au cœur. La capacité des modèles à prédire les pathologies cardiaques à partir de facteurs cliniques et des traitements reste encore insuffisante. Chez des survivants traités de façon similaire, une variabilité interindividuelle de la susceptibilité à la cardiotoxicité a été observée, suggérant l'existence de facteurs génétiques modulant cette susceptibilité. Dans une première partie, nous avons étudié les risques associés aux traitements cardiotoxiques sur la survenue des pathologies cardiaques tardives chez les enfants traités pour un cancer au sein de la cohorte FCCSS. Nous avons mis en évidence le rôle de l'irradiation à de faibles volumes cardiaques et la présence d'interaction

entre l'irradiation cardiaque et l'administration d'anthracyclines. Dans une deuxième partie, nous avons rapporté à travers une revue systématique de la littérature des facteurs génétiques associés aux pathologies cardiaques chez les survivants d'un cancer dans l'enfance. Dans une troisième partie, une étude cas-témoins appariée nichée dans la cohorte FCCSS a été conduite pour identifier des biomarqueurs du transcriptome associés aux pathologies cardiaques et à l'insuffisance cardiaque en particulier. Le gène NFE2L2 a été identifié avec une expression élevée associée à une diminution du risque de ces événements. En conclusion, les travaux de cette thèse contribuent à mieux comprendre les facteurs influençant les risques de complications cardiaques tardives chez les enfants traités pour cancer. Nos résultats pourraient être intégrés après validation dans la pratique clinique courante pour l'adaptation de la prise en charge thérapeutique des enfants atteints de cancer et pour les recommandations de leur suivi à long terme.

Title : Occurrence of long-term cardiac diseases after *treatment for* childhood cancer: Identification of associated biomarkers

Keywords : Cardiotoxicity; Biomarkers; Childhood cancer survivors; Transcriptome; Radiotherapy; Anthracyclines

Abstract : Advances in pediatric oncology have improved the prognosis of patients treated for childhood or adolescent cancer, with a 5-year survival rate of over 80%. However, these patients are at increased risk of late following cancer treatment. Cardiac diseases are a major concern in this population of survivors who have been exposed to cardiotoxic treatments such as anthracyclines and/or radiation to the heart. The ability of models to predict cardiac diseases from clinical and treatment factors is still insufficient. In similarly treated survivors, interindividual variability in cardiotoxicity susceptibility has been observed, suggesting the existence of genetic factors modulating this susceptibility. In a first part, we studied the risks associated with cardiotoxic treatments on the occurrence of late cardiac diseases in children treated for cancer within the French Childhood Cancer Study cohort. We highlighted the role of low-volume cardiac irradiation

and the interaction between cardiac irradiation and anthracycline administration. In the second part, we report on a systematic of the literature on genetic factors associated with cardiac diseases in survivors of childhood cancer. In the third part, a matched case-control study nested in the FCCSS cohort was conducted to identify transcriptome biomarkers associated with cardiac diseases and heart failure in particular. The NFE2L2 gene was identified, with high expression associated with a reduced risk of these events. In conclusion, the work of this thesis contributes to a better understanding of the factors impacting the risk of late cardiac complications in children treated for cancer. Once validated, our findings could be incorporated into current clinical practice to adapt the therapeutic management of children with cancer, as well as recommendations for their long-term follow-up

PRODUCTIONS SCIENTIFIQUES

Articles en cours de révision

- [Naïla Aba](#), Claire Ducos, Eric Morel, Chiraz El Fayech, Brice Fresneau, Florent de Vathaire, Gwénaël Le Teuff, Nadia Haddy, Influence of Genetic Biomarkers on Cardiac Diseases in Childhood Cancer Survivors: a Systematic Review. *Soumis en décembre 2023 à: The Pharmacogenomics Journal.*
- [Naïla Aba](#), Gwénaël Le Teuff, Brice Fresneau, Serge Koscielny, Shaima Belhechmi, Boris Schwartz, Chiraz El-Fayech, Carole Rubino, Rodrigue S. Allodji, Eric Morel, Delphine Dayde, Pierre De la Grange, Ariane Jolly, Noémie Pata-Merci, Camille Cordero, Isabelle Aerts, François Pein, Giau Vu-Bezin, Florent De Vathaire, Nadia Haddy, Genetic biomarkers related to cardiac disease occurrence in childhood cancer survivors: a matched Etude cas-témoins nested in the French Childhood Cancer Survivors Study. *Soumis en juin 2024 à: Journal of Cinical Oncology.*

Articles en cours de préparation

- [Naïla Aba](#), Gwénaël Le Teuff, Brice Fresneau, Boris Schwartz, Charried Thibaud, Chiraz El-Fayech, Carole Rubino, Rodrigue S. Allodji, Ibrahima Diallo, Isabelle Aerts, Giau Vu-Bezin, Florent De Vathaire, Nadia Haddy, Risk of cardiac diseases and heart failure associated with cardiotoxic ancer treatments in the French Childhood Cancer Study.
- Boris Schwartz*, Imene Mansouri*, [Naïla Aba](#), Brice Fresneau, Giau Vu-Bezin, Daniel Bejarano-Quisoboni, Christelle Dufour, Stéphanie Bolle, Aurore Surun, Daniel Orbach, Rodrigue S Allodji, Ibrahima Diallo, Charlotte Demoor-Goldschmidt, Chiraz El-Fayech, Florent de Vathaire, Gwénaël Le Teuff, Nadia Haddy, Insights into Cardiovascular Hospitalization Risk Among Survivors of Childhood Cancer: Findings from the French Childhood Cancer Survivor Study.
*Les auteurs ont participé à parts égales.

Autres publications

- Ducos C, Aba N, Rosselli F, Fresneau B, Al Ahmad Nachar B, Zidane M, de Vathaire F, Benhamou S, Haddy N. Genetic risk of second malignant neoplasm after childhood cancer treatment: A systematic review. *Cancer Epidemiol Biomarkers Prev.* 2024 May 27.
- Bentriou M, Letort V, Chounta S, Fresneau B, Do D, Haddy N, Diallo L, Journy N, Zidane M, Charrier T, Aba N, Ducos C, Zossou V, de Vathaire F, Allodji R, Lemler S. Combining Dosiomics and Machine Learning Methods for predicting severe Cardiac Diseases in Childhood Cancer Survivors: the French Childhood Cancer Survivor Study. *Frontiers in Oncology.*

-

Communications orales à des conférences nationales et internationales:

- Aba N, Belhechmi S, Fresneau B, El-Fayech C, Rubino C, Allodji RS, Morel E, De la Grange P, Jolly A, Pata-Merci N, Koscielny S, Vu-Bezin G, de Vathaire G, Le Teuff G, Haddy N. Identification de biomarqueurs génétiques potentiels liés à l'apparition de pathologies cardiaques chez les enfants ayant survécu d'un cancer : étude cas-témoins. Conférence francophone d'Épidémiologie CLINique (EPICLIN), Mai 2023, Nancy, France. *Prix de la meilleure présentation orale.*

Communications affichées à des conférences nationales et internationales

- Aba N, Belhechmi S, Fresneau B, El-Fayech C, Rubino C, Allodji RS, Morel E, De la Grange P, Jolly A, Pata-Merci N, Koscielny S, Vu-Bezin G, de Vathaire G, Le Teuff G, Haddy N. A case-control study to identify potential genetic biomarkers related to cardiac diseases occurrence in childhood cancer survivors. ISORED, Mai 2023, Sitges, Espagne. *Prix de la meilleure présentation affichée.*
- Aba N, Belhechmi S, Fresneau B, Allodji RS, Koscielny S, Vu-Bezin G, de Vathaire G, Le Teuff G, Haddy N. High-dimensional selection in a matched case-control

study : cardiotoxicity in childhood cancer survivors. Conference of the International Society for Clinical Biostatistics (ISCB). Aout 2023, Milan, Italie.

Communications orales hors des conférences nationales et internationales

- Aba N, Le Teuff G, Haddy N. Identification de marqueurs biologiques prédictifs de la survenue de pathologies cardiaques après traitement d'un cancer dans l'enfance. Réunion du Comité de suivi à long terme, Décembre 2021, visioconférence.
- Aba N, Chounta S, Les pathologies cardiaques dans la FCCSS : études de dosimétrie et de transcriptome. Mardis Scientifiques du CESP, Septembre 2022, Villejuif, France.
- Aba N, Belhechmi S, Fresneau B, El-Fayech C, Rubino C, Allodji RS, Morel E, De la Grange P, Jolly A, Pata-Merci N, Koscielny S, Vu-Bezin G, de Vathaire G, Le Teuff G, Haddy N., Pata-Merci N, Koscielny S, Vu-Bezin G, de Vathaire G, Le Teuff G, Haddy N. Identification of potential genetic biomarkers related to cardiac diseases occurrence in childhood cancer survivors: a case control-study. Les Journées de la Recherche, Gustave Roussy, Novembre 2023, Orsay, France.

TABLE DES MATIERES

Productions scientifiques	i
Table des matières	v
Liste des figures.....	vii
Liste des tables.....	ix
Liste des abréviations.....	xii
Chapitre 1. Introduction générale	1
1.1 <i>Cancers de l'enfant</i>	1
1.2 <i>Les pathologies cardiaques à long terme après traitement d'un cancer dans l'enfance</i>	7
1.3 <i>Prédiction et stratification du risque de pathologie cardiaques</i>	11
1.4 <i>Objectifs de la thèse</i>	13
Chapitre 2. La cohorte French Childhood Cancer Survivors Study (FCCSS)	15
2.1 <i>Définition de la cohorte</i>	15
2.2 <i>Banque biologique</i>	25
Chapitre 3. Méthodologie statistique	27
3.1 <i>Analyse de survie dans une étude de cohorte</i>	27
3.2 <i>Les taux standardisés d'hospitalisation</i>	29
3.3 <i>Modèle de régression pour une étude cas-témoins appariée et méthodes de sélection de variables en grande dimension</i>	32
Chapitre 4. Risque de pathologies et d'hospitalisation cardiaque chez les survivants d'un cancer de l'enfant 48	
4.1 <i>Introduction</i>	48
4.2 <i>Risques de pathologies cardiaques et d'insuffisance cardiaque chez les survivants d'un cancer de l'enfant</i>	49
4.3 <i>Risques d'hospitalisations pour cause cardiaque chez les survivants d'un cancer de l'enfant</i>	73
4.4 <i>Discussion</i>	84

Chapitre 5. Influence de la génétique sur l'apparition de pathologies cardiaques chez les survivants d'un cancer dans l'enfance : revue systématique de la littérature	91
5.1 Contexte et objectifs de la revue systématique de la littérature	91
5.2 Matériel et méthode.....	93
5.3 Résultats	97
5.4 Discussion	124
Chapitre 6. Identification de biomarqueurs génétiques associées à l'apparition de pathologies cardiaques chez des patients traités pour un cancer de l'enfant : étude cas-témoin appariée	127
6.1 Introduction	127
6.2 Matériels et Méthodes.....	128
6.3 Résultats	134
6.4 Discussion	152
Chapitre 7. Discussion générale.....	157
7.1 Résumé des principaux résultats.....	157
7.2 Points forts et limites de la thèse.....	161
7.3 Perspectives futures.....	163
7.4 Conclusion générale.....	166
Références bibliographiques.....	168
Annexes.....	203
Annexe A. Annexes en lien avec les risques de pathologies cardiaques et d'hospitalisation cardiaque	203
Annexe B. Annexes en lien avec la revue systématique.....	211
Annexe C. Annexes en lien avec l'étude des biomarqueurs associés aux pathologies cardiaques	222
Publications.....	228

LISTE DES FIGURES

Figure 1. Répartition des proportions de types de cancer par groupe d'âge, 2001-2010, toutes régions du monde. Source (Steliarova-Foucher et al., 2017).....	2
Figure 2. Taux bruts de mortalité chez les enfants avec un diagnostic de cancer de l'enfant (0-14 ans) au cours des années par sexe en France en cas par 100 000 (Source : Frelay et al., 2024).....	4
Figure 3. Résumé de la cardiotoxicité induite par les anthracyclines (Source : Vuong et al., 2022).....	9
Figure 4. Pathophysiologie de la cardiotoxicité induite par l'exposition aux radiations cardiaques (Source : Podlesnikar et al., 2022).....	10
Figure 5 Représentation de la reconstruction voxélisée de la dose de radiations au cœur d'un survivant d'un cancer de l'enfant (vue de face, de dos, de gauche et de droite). Les nuances de couleur représentent le niveau de la dose. (Figure réalisée par Duyen Do, équipe Epidémiologie des radiations).....	18
Figure 6. Diagramme de flux des patients de la cohorte FCCSS ayant fourni des échantillons biologiques pour la constitution de la banque biologique.....	26
Figure 7. Représentation de l'algorithme Percentile Lasso adapté à la régression logistique conditionnelle pénalisée Lasso.....	40
Figure 8. Représentation de l'algorithme Bolasso adapté à la régression logistique conditionnelle pénalisée Lasso.....	41
Figure 9. Représentation de l'algorithme Sublasso adapté à la régression logistique conditionnelle pénalisée Lasso.....	42
Figure 10. Diagramme de flux des patients inclus dans les analyses pour l'évènement première pathologie cardiaque tous grades et grades ≥ 3	52
Figure 11. Diagramme de flux de l'inclusion des patients dans les analyses pour l'évènement première insuffisance cardiaque.....	52
Figure 12. Modèle à risque concurrent utilisé pour l'étude de l'estimation du risque de premier évènement cardiaque dans la FCCSS.....	54

Figure 13. Incidence cumulée de pathologie cardiaque et de décès sans avoir développé de pathologie cardiaque parmi les survivants à 5 ans de la cohorte FCCSS	60
Figure 14. Diagramme de flux de l'inclusion des patients dans l'étude de cohorte sur les hospitalisations pour cause cardiaque.....	73
Figure 15. Nombre cumulé moyen et intervalle de confiance à 95 % correspondants (ombres) des hospitalisations cardiaques parmi les survivants de la cohorte FCCSS de 2011 à 2018, en fonction de l'âge atteint.....	77
Figure 16. Diagramme de flux PRISMA décrivant la stratégie de recherche et les études incluses et exclues	97
Figure 17. Rôles des gènes présentés associés à la cardiotoxicité liée à la doxorubicine, au sein d'une cellule cardiomyocytaire.	110
Figure 18. Organigramme de l'inclusion des patients dans l'étude cas-témoin nichée dans la cohorte FCCSS.....	129
Figure 19. Diagramme indiquant l'inclusion des gènes dans l'étude.....	131
Figure 20. A. Diagramme de Venn représentant les gènes, identifiés par les approches Percentile Lasso ($\theta=95\%$), Bolasso et Sublasso (seuil 50 %) ; B. Box plots de l'expression des gènes associés aux pathologies cardiaques chez les patients présentant une pathologie cardiaque (cas) et les patients n'en présentant pas (témoins), gènes qui résultent de l'intersection des trois approches de stabilisation	139
Figure 21. Histogramme de l'expression de NFE2L2 chez les patients présentant une pathologie cardiaque (cas) et les patients n'en présentant pas (témoins).....	139
Figure 22. Boîtes à moustaches de l'AUC calculée sur les 500 répliquions du jeu de données d'entraînement et du jeu de données de validation.....	145
Figure 23. A. Diagramme de Venn représentant le gène identifié par les approches Percentile Lasso ($\theta=95\%$), Bolasso et Sublasso (seuil 50 %) ; B. Box plots de l'expression du gène NFE2L2 associés à l'insuffisance cardiaque chez les patients présentant une insuffisance cardiaque (cas) et les patients n'en présentant pas (témoins).....	148

LISTE DES TABLES

Tableau 1. Caractéristiques des survivants inclus dans la cohorte FCCSS.....	16
Tableau 2. Sources des données recueillies pour le suivi à long terme des patients inclus dans la cohorte FCCSS	21
Tableau 3. Grades de pathologies cardiaques selon la version 4.0 des CTCAE (Source : US Department of Health and Human Services, 2010)	24
Tableau 4. Caractéristiques démographiques et cliniques des survivants inclus dans les analyses et des survivants ayant développé une première pathologie cardiaque tous grades ou de grade 3 à 5, ou une insuffisance cardiaque dans la FCCSS	57
Tableau 5. Effet des doses cumulées d'anthracyclines et des doses moyennes de radiations cardiaques sur le risque de pathologies cardiaques tous grades et grades \geq 3 dans la cohorte FCCSS	62
Tableau 6. Effet des pourcentages de volume du cœur recevant des doses spécifiques de radiothérapie sur le risque de pathologie cardiaque tous grades et grades \geq 3 dans la cohorte FCCSS.....	63
Tableau 7. Effet et incidence du traitement par anthracyclines et des doses moyennes de radiations cardiaque sur le risque de première pathologie cardiaque tous grades et grades \geq 3 dans la cohorte FCCSS	66
Tableau 8. Effet des doses cumulées d'anthracyclines et des doses moyennes de radiations cardiaques sur le risque d'insuffisance cardiaque tous grades dans la cohorte FCCSS.....	68
Tableau 9. Effet des pourcentages de volume du cœur recevant des doses spécifiques de radiothérapie sur le risque d'insuffisances cardiaques dans la cohorte FCCSS.....	69
Tableau 10. Effet et incidence du traitement par anthracyclines et des doses moyennes de radiations cardiaque sur le risque de première insuffisance cardiaque tous grades dans la cohorte FCCSS.....	71
Tableau 11. Effet des pathologies cardiaques préexistantes sur les insuffisances cardiaques tous grades	72

Tableau 12. Caractéristiques démographiques et cliniques des patients inclus dans l'étude, et des hospitalisations liées à des causes cardiaques de 2011 à 2018	76
Tableau 13. Nombre d'hospitalisations cardiaques observées et attendues, ratios d'hospitalisation standardisés et excès de risque absolu, chez les survivants de la cohorte FCCSS de 2011 à 2018.	78
Tableau 14. Effet des traitements et des comorbidités sur les hospitalisations cardiaques dans la cohorte FCCSS entre 2011 et 2018	82
Tableau 15. Stratégie de recherche des articles sur NCBI/PubMed et Scopus	94
Tableau 16. Formulaire d'extraction des données	96
Tableau 17. Caractéristiques des 19 études incluses ordonnées selon le design de l'étude (étude cas-témoins puis étude de cohorte).....	99
Tableau 18. Evaluation de la qualité des études incluses à partir de l'outil Q-GENIE	102
Tableau 19. Principaux résultats des 19 études incluses concernant les associations génétiques. Les études sont triées selon la méthode de génotypage (gènes candidates, GWAS, WES puis WGS)	105
Tableau 20. Effet des gènes et des variants sur la cardiomyopathie en fonction des doses cumulées d'anthracycline.....	107
Tableau 21. Gènes et familles de gènes présentés, et les voies biologiques auxquels ils appartiennent	108
Tableau 22. Caractéristiques démographiques et cliniques des patients inclus dans l'étude cas-témoins appariée (n=330).....	135
Tableau 23. Nombre de gènes sélectionnés par les approches Percentile lasso, Bolasso et Sublasso avec ou sans contrainte.....	137
Tableau 24. Fréquence de sélection des gènes sélectionnés par les approches Bolasso et Sublasso avec un seuil de 50 % à partir de 1 000 réplifications. Ces gènes ont également été sélectionnés pour l'approche Percentile Lasso.....	138
Tableau 25. Effet de l'expression du gène NFE2L2 sur les pathologies cardiaques estimé à partir d'une régression logistique conditionnelle ajustée sur l'année de	

diagnostic du cancer et l'exposition aux anthracyclines et à la radiothérapie incluant le cœur	141
Tableau 26. Effet de l'expression gène NFE2L2 en fonction des niveaux de doses cumulées d'anthracyclines et des doses moyennes d'irradiation du cœur (termes d'interaction) à l'aide d'une régression logistique conditionnelle ajustée sur l'année de diagnostic du cancer et les variables liées au anthracycline et à l'irradiation.....	142
Tableau 27. Estimation de la corrélation entre l'expression du gène NFE2L2 et les doses cumulées d'anthracycline ou les doses moyennes de radiations au cœur.	143
Tableau 28. Comparaison de différents modèles de régression logistique conditionnelle sur la base de la statistique du rapport de vraisemblance et de l'aire sous la courbe (AUC) pour le critère pathologie cardiaque.....	144
Tableau 29. Effet de l'expression du gène NFE2L2 sur les pathologies cardiaques de grade ≥ 2 , estimé à partir d'une régression logistique conditionnelle ajustée sur l'année de diagnostic du cancer et l'exposition aux anthracyclines et à la radiothérapie au cœur	146
Tableau 30. Nombre de gènes sélectionnés par les approches Percentile lasso, Bolasso et Sublasso avec ou sans contrainte.....	147
Tableau 31. Effet de l'expression du gène NFE2L2 sur l'insuffisance cardiaque estimé à partir d'une régression logistique conditionnelle ajustée sur l'année de diagnostic du cancer et l'exposition aux anthracyclines et à la radiothérapie incluant le cœur	149
Tableau 32. Estimation de la corrélation entre l'expression du gène NFE2L2 et les doses cumulées d'anthracycline ou les doses moyennes de radiations du cœur. Les corrélations sont testées une par une.....	149
Tableau 33. Comparaison de différents modèles de régression logistique conditionnelle sur la base de la statistique du rapport de vraisemblance et de l'aire sous la courbe (AUC) pour le critère insuffisance cardiaque	150
Tableau 34. Comparaison des niveaux d'expression du gène NFE2L2 avant et après exposition aux anthracyclines dans des cardiomyocytes de rats adultes.....	151

LISTE DES ABREVIATIONS

CCSS	Childhood Cancer Survivor Study
CepiDC	Centre d'épidémiologie sur les causes médicales de décès
CIM-10	Classification internationale des maladies, 10e révision
DCIR	Données de Consommation Inter-Régimes
EI	Etendue Interquartile
ERA	Excès de Risque Absolu
FCCSS	French Childhood Cancer Survivor Study
FEVG	Fraction d'ejection du ventricule gauche
FRVG	Fraction de raccourcissement du ventricule gauche
Gy	Gray
HR	Hazard Ratio
IC	Intervalle de Confiance
ICCC	International Classification of Childhood Cancer
INSEE	Institut National de la Statistique des Etudes Economiques
INSERM	Institut national de la santé et de la recherche médicale
OR	Odds Ratio
PMSI	Programme de Médicalisation des Systèmes d'Information
RNCE	Registre National des Cancers de l'Enfant
RR	Risque Relatif
SEER	Surveillance, Epidemiology and End Results
SHR	Ratio standard d'hospitalisation
SNC	Système nerveux central
SNDS	Système National des Données de Santé
V01_5	Volume du cœur ayant reçu 0,1 Gy lorsque la V5 =0%
V15	Volume du cœur ayant reçu 15 Gy
V20	Volume du cœur ayant reçu 20 Gy
V30	Volume du cœur ayant reçu 30 Gy
V5	Volume du cœur ayant reçu 5 Gy
V5_20	Volume du cœur ayant reçu 5 Gy lorsque la V20 =0%

Chapitre 1. INTRODUCTION GENERALE

Dans ce chapitre d'introduction générale, nous décrivons dans un premier temps l'épidémiologie et la survie des cancers de l'enfant, nous faisons ensuite le point sur les maladies cardiaques en tant que pathologies tardives secondaires aux traitements du cancer de l'enfant, nous abordons également les méthodes de prédiction du risque de ces maladies cardiaques, et finalement nous détaillons les différents objectifs de recherche de la thèse.

1.1 CANCERS DE L'ENFANT

1.1.1 Incidence des cancers de l'enfant

Chaque année en France métropolitaine, plus de 2250 enfants et adolescents âgés de 0 à 17 ans sont diagnostiqués d'un cancer, soit un taux d'incidence standardisé sur l'âge de 166,6 cas par million d'enfants par an. Cela fait du cancer la principale cause de décès non accidentel chez l'enfant (Lacour et al., 2014). De 2014 à 2020, les leucémies (n=579), les tumeurs du système nerveux central, intracrâniennes et spinales (n=560), et les lymphomes et néoplasmes réticulo-endothéliaux (n=329) ont été les types de cancer de l'enfant les plus fréquents. Pour tous types de cancers combinés, le nombre de cas étaient légèrement supérieur chez les garçons que chez les filles avec un sex-ratio global de 1,1. Ce ratio diffère cependant selon les types de cancer (*source RNCE, Registre National des Cancers de l'Enfant*).

Les cancers de l'enfant sont des cancers rares, et sont une cause importante de morbidité et de mortalité. Leurs types, leurs localisations et leurs classifications diffèrent de celles des cancers de l'adulte (Johnston et al., 2021). La troisième édition de la classification internationale des cancers de l'enfant (*International Classification of Childhood Cancer, ICC-3*) classifie les cancer de l'enfant en 12 groupes principaux, divisés en 47 sous-groupes, et se base principalement sur leur morphologie, leur topologie (siège), et leur développement (Steliarova-Foucher et al., 2005). Les types de

tumeur varie fortement en fonction de la tranche d'âge. Chez les enfants dans la tranche d'âge de 0 à 4 ans, les leucémies représentaient 36,1 % des diagnostics de cancers ; cependant, chez les jeunes âgés de 15 à 19 ans la proportion de cas de leucémies passe à 15,4 %. Les lymphomes représentaient 5,3 % des cas de cancers chez les enfants de 0-4 ans, et 22,5 % chez les 15-19 ans. Les tumeurs du système nerveux central étaient le deuxième type de tumeur le plus fréquent après la leucémie chez les enfants âgés de 0 à 14 ans (Steliarova-Foucher et al., 2017). La **Figure 1** montre la proportion des types de tumeur en fonction de la tranche d'âge entre 2000 et 2010, toutes régions du monde confondues.

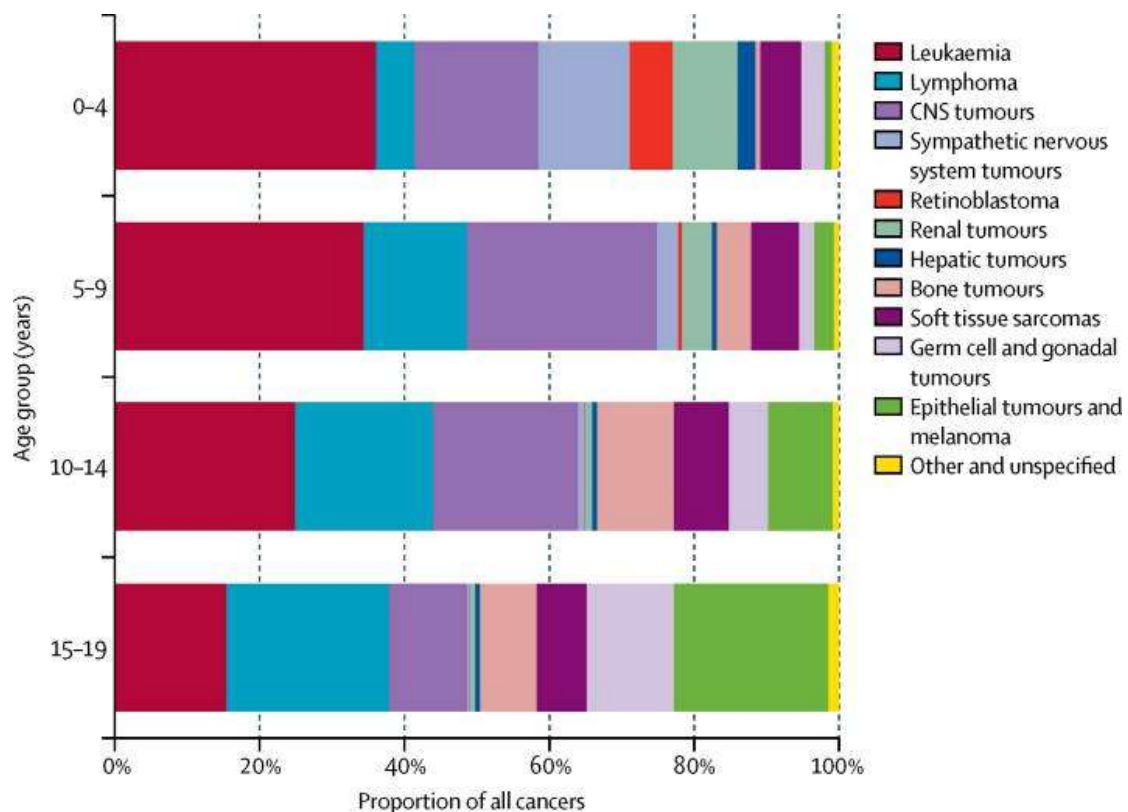


Figure 1. Répartition des proportions de types de cancer par groupe d'âge, 2001-2010, toutes régions du monde. Source (Steliarova-Foucher et al., 2017).

En 2015, le programme SEER (Surveillance, Epidemiology and End Results) a estimé un taux d'incidence standardisé sur l'âge de 178,0 cas par million d'enfant par an, sur l'ensemble de la population infantile, soit un total mondial de plus de 344 500 enfants diagnostiqués avec un cancer dans le monde (taux estimé en Amérique du Nord et en Europe) (Johnston et al., 2021). Cependant, il existe des disparités dans l'accès aux soins

selon les régions du monde, et le taux d'incidence présenté peut alors varier en fonction de la localisation géographique, car tous les pays ne disposent pas de registres du cancer de l'enfant, et que ceux qui en disposent ne couvrent qu'une partie des enfants dans le monde (Lam et al., 2019; Steliarova-Foucher et al., 2017).

1.1.2 Taux de survie après cancer de l'enfant

La survie des cancers de l'enfant s'est fortement améliorée au cours des dernières décennies. Cependant, celle-ci peut être influencée par plusieurs facteurs, incluant le sexe, l'âge au moment du diagnostic du cancer, le sous-type, le site, l'histologie, le grade, et le stade du cancer. De plus, certaines études suggèrent que les facteurs indiquant un faible statut socio-économique pourraient être associés à une survie plus faible des patients diagnostiqués d'un cancer de l'enfant. En effet, la survie n'a pas augmenté de la même manière dans toutes les régions du monde, et les estimations suggèrent que la survie dans les pays à revenu faible est nettement inférieure à celle des pays à revenu élevé (Erdmann et al., 2021).

En Europe, pour tous les cancers combinés diagnostiqués chez les enfants entre 2000-2014, la survie à cinq ans était de 81 % (95 % IC= 81 – 82) (Botta et al., 2022). En France, entre 2010 et 2016, la survie à 5 ans, tous cancers confondus, était de 84,6 % (95 % IC= 83,9 - 85,2) (*source RNCE, Registre National des Cancers de l'Enfant*). La **Figure 2** présente la diminution des taux brut de mortalité d'un cancer de l'enfant en France entre 1950 et 2015. Ces progrès importants ont sans doute été rendus possibles par les avancées récentes en matière de traitements des cancers pédiatriques (Robison & Hudson, 2014).

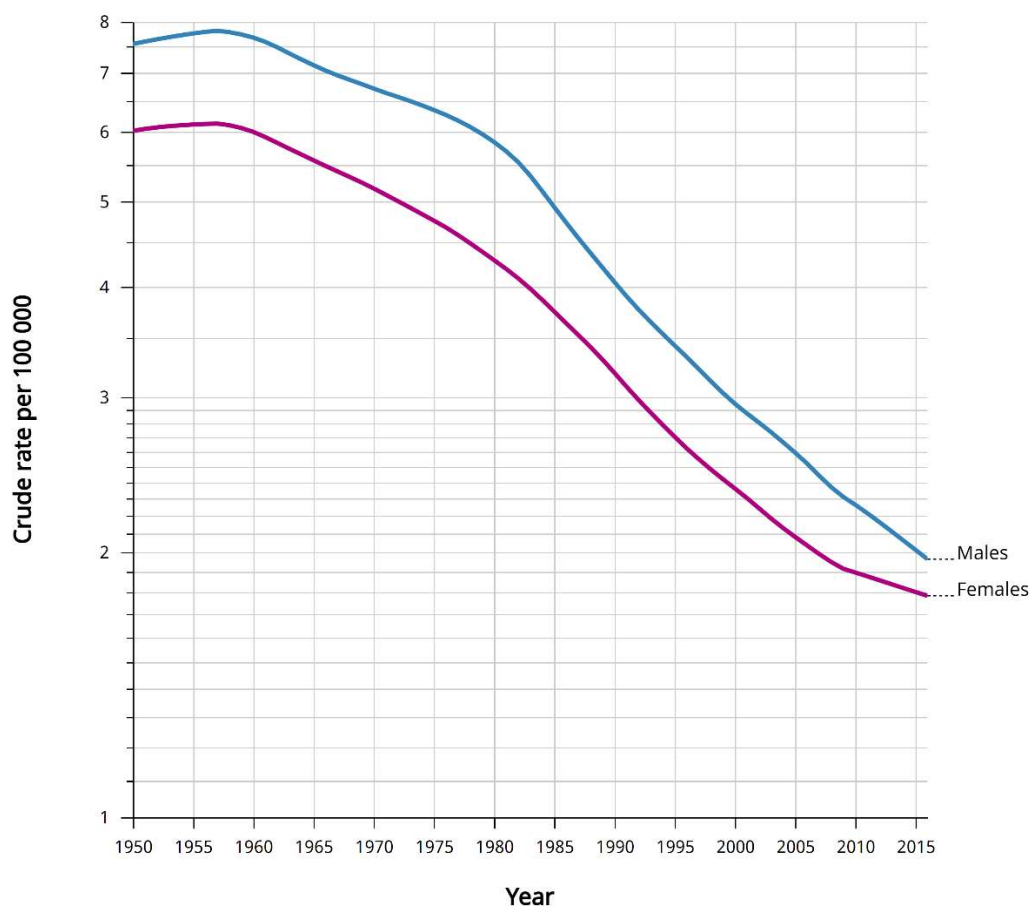


Figure 2. Taux bruts de mortalité chez les enfants avec un diagnostic de cancer de l'enfant (0-14 ans) au cours des années par sexe en France en cas par 100 000 (Source : Frelay et al., 2024)

1.1.3 Traitements des cancers chez l'enfant

Les progrès médicaux dans le domaine de la biologie du cancer, du développement thérapeutique, de la technologie des radiations, de l'imagerie diagnostique et des soins de support ont permis de faire évoluer les protocoles de traitement des cancers de l'enfant (Pritchard-Jones et al., 2013; Robison & Hudson, 2014). Conventionnellement, les stratégies de traitement du cancer les plus utilisées sont la résection chirurgicale des tumeurs, la radiothérapie et la chimiothérapie. Cependant ces traitements ont des limites importantes. La chirurgie est efficace à un stade précoce de la progression de la maladie ; la radiothérapie peut endommager les cellules, les organes et les tissus sains autour de la tumeur ; et la chimiothérapie, quel que soit les agents utilisés,

endommagement souvent les cellules saines (Arruebo et al., 2011; Debela et al., 2021).

1.1.3.1 *Radiothérapie*

La radiothérapie est un élément important du traitement de nombreuses tumeurs chez l'enfant. Actuellement, le traitement standard est la radiothérapie conformationnelle tridimensionnelle et la radiothérapie à modulation d'intensité (IMRT). Ces technologies modernes permettent d'identifier avec plus de précision les volumes tumoraux et leur relation spatiale avec les organes dans le corps. Ainsi ces technologies permettent de moduler l'intensité du faisceau de radiations, et d'administrer des radiations de façon précise, et/ou protéger les tissus normaux, et donc d'améliorer les résultats cliniques de ce traitement (Allen et al., 2017; Purdy, 2008; Steinmeier et al., 2019). Ces progrès sont très importants, en particulier chez les enfants, qui sont très sensibles à la radiothérapie en raison de l'immaturation de leurs tissus (Steinmeier et al., 2019).

Ces dernières années, la proton thérapie a été de plus en plus utilisée, particulièrement pour les tumeurs nécessitant des doses de traitement élevées, et les tumeurs malignes pour lesquelles il est essentiel d'épargner les organes à risques afin d'éviter les effets tardifs (Steinmeier et al., 2019).

1.1.3.2 *Chimiothérapie*

Depuis les années 1960, des groupes d'études multidisciplinaires ont été organisés par les oncologues, permettant à la plupart des enfants d'Europe et d'Amérique du Nord atteints de cancer de participer à des essais cliniques dans les protocoles disponibles. Ceci a finalement contribué à une amélioration de la survie des patients traités. Ainsi, l'introduction de la chimiothérapie dans les protocoles de traitement du cancer comprenant déjà la chirurgie et la radiothérapie, a permis de réduire de façon importante de la mortalité des patients traités (Debela et al., 2021; Erdmann et al., 2021).

Cependant, ces traitements ont aussi des effets secondaires cytotoxiques non négligeables. Il est donc nécessaire de trouver un équilibre entre l'efficacité des traitements administrés et leur toxicité tardive. Des études ont été menées afin

d'améliorer les soins de ces patients en intensifiant les traitements de manière plus sûre, par le biais de protocoles fournissant des conseils détaillés sur l'administration et la modification appropriée de la dose de traitement, en réponse à une toxicité excessive. Ainsi, les traitements actuels des cancers de l'enfant sont adaptés de sorte à diminuer les risques de pathologies secondaires. L'évaluation de l'état de santé tardif des survivants traités avec ces protocoles de traitements modifiés est importante pour caractériser leur état de santé ultérieur (Hudson et al., 2012).

Bien que des thérapies biologiques et moléculaires innovantes fassent l'objet de recherches continues, la chimiothérapie cytotoxique et l'irradiation continuent d'être utilisées pour le traitement de la majorité des tumeurs pédiatriques (Debela et al., 2021; Hudson et al., 2012).

1.1.4 Effets tardifs des traitements du cancer de l'enfant

La population des patients ayant survécu d'un cancer dans l'enfance, est généralement désignée sous le nom de Childhood Cancer Survivors en anglais, et nous la désignerons tout au long de ce manuscrit par le terme « survivants ».

Les taux de survie des patients traités pour un cancer dans l'enfance ayant largement augmenté au cours des dernières années grâce aux progrès thérapeutiques, il existe aujourd'hui une population croissante de survivants qui se retrouve confrontée à un large éventail d'effets indésirables sur leur santé. Ces effets secondaires tardifs peuvent apparaître plusieurs années après le traitement du cancer, et dégrader la qualité de vie des patients. Ainsi, environ trois survivants sur quatre sont susceptibles de développer des effets tardifs, qui peuvent être graves ou mortels dans 25 % des cas (Geenen et al., 2007; Kremer et al., 2013; Landier et al., 2015). Les causes de ces effets indésirables peuvent provenir de multiples facteurs, et impliquer des combinaisons de facteurs en plus de l'exposition aux traitements, comme le type de première tumeur maligne, l'âge au diagnostic de la tumeur, les conditions prémorbides, les prédispositions génétique, ou encore comportements liés à la santé (Robison & Hudson, 2014). Parmi ces effets indésirables, nous retrouvons les tumeurs secondaires, les troubles cognitifs, les

pathologies cardiaques, les dysfonctions rénales, l'infertilité, ainsi que de nombreuses autres dysfonctions (Landier et al., 2015). Cependant, ce sont généralement des facteurs spécifiques au traitement qui déterminent principalement le risque d'effets indésirables tardifs (Robison & Hudson, 2014)

1.2 LES PATHOLOGIES CARDIAQUES A LONG TERME APRES TRAITEMENT D'UN CANCER DANS L'ENFANCE

Les complications cardiaques telles que les maladies coronariennes, les pathologies valvulaires, l'insuffisance cardiaque ou encore les troubles du rythme et de la conduction cardiaque, faisant suite aux traitements par anthracyclines ou radiothérapie, sont devenues l'une des principales causes de morbidité et de mortalité chez les survivants d'un cancer de l'enfant (Armenian et al., 2018; Lipshultz, Adams, et al., 2013; Van Kalsbeek et al., 2021). Dans cette partie, nous présentons d'abord l'incidence des pathologies cardiaques chez les survivants d'un cancer de l'enfant, puis nous décrivons la physiopathologie des différentes pathologies cardiaques rencontrées chez ces patients.

1.2.1 Incidence et causes des pathologies cardiaques chez les survivants d'un cancer dans l'enfance

Les pathologies cardiaques émergent comme la principale cause de décès non cancéreux chez les survivants d'un cancer de l'enfant, les plaçant à un risque sept fois plus élevé de mourir d'une maladie cardiovasculaire par rapport à la population générale (Armenian et al., 2018). Les traitements des cancers de l'enfant par radiothérapie impliquant le cœur, et par chimiothérapie à base d'anthracyclines, contribuent au développement de dysfonctions et pathologies cardiaques (Bates et al., 2019; Haddy et al., 2016; Lipshultz et al., 2005; Mansouri et al., 2019; Mulrooney et al., 2009). Les anthracyclines (doxorubicine, daunorubicine, mitoxantrone, épirubicine, and idarubicine) sont un traitement anticancéreux découvert il y a plus de 50 ans et toujours très utilisés dans le traitement de divers types de tumeurs solides et de cancer

hématologiques. Comme les autres chimiothérapies, elles endommagent les cellules saines, en particulier les cellules qui se divisent et se développent rapidement comme les cellules cardiaques (cardiomyocytes), et sont responsables d'effets tardifs à long terme dont la cardiotoxicité (Debela et al., 2021 ; Mattioli et al., 2023). Ainsi, près de 20 % des survivants traités par anthracyclines subissent des effets cardiotoxiques, avec un risque d'évènements cardiovasculaires 3,5 fois plus élevé que les personnes non traitées (Geenen et al., 2007; Mulrooney et al., 2009). Ainsi, près d'un survivant sur huit développe une pathologie cardiaque grave (Haddy et al., 2016).

Cependant, ce ne sont pas les seuls facteurs de risque associés aux évènements cardiotoxiques. Le jeune âge au moment de l'exposition aux traitements, l'augmentation du temps depuis l'exposition au traitement, le sexe féminin, et un dysfonctionnement cardiaque préexistant sont aussi des facteurs de risques importants du développement de pathologies cardiaques chez les survivants d'un cancer de l'enfant (Bansal et al., 2017).

1.2.2 Pathophysiologie des pathologies cardiaques

La cardiotoxicité des traitements anticancéreux peut se manifester physiquement de différentes manières comme le dysfonctionnement de la contractilité cardiaque, la cardiomyopathie, le dysfonctionnement ventriculaire, le syndrome péricardite-myocardite, l'arythmie ou l'insuffisance cardiaque (Mattioli et al., 2023). Elle peut également se manifester à des moments différents : durant le traitement, dans les premières années qui suivent ou encore, et c'est ce qui nous intéresse particulièrement dans le contexte de la thèse, plusieurs années, voir décennies, après le traitement. Ce dernier type de cardiotoxicité est appelé cardiotoxicité chronique tardive, et est irréversible (Qiu et al., 2023).

A l'origine, la cardiotoxicité associée aux anthracyclines se manifeste par des lésions des cardiomyocytes, et une fibrose cardiaque, conduisant à une détérioration de la fonction systolique du ventricule gauche et à un déclin asymptomatique de la fraction d'éjection ventriculaire gauche, pouvant éventuellement évoluer vers une insuffisance

cardiaque symptomatique (Armstrong et al., 2012, 2015; Qiu et al., 2023; Samuel et al., 2023). Par ailleurs, la cardiotoxicité chronique tardive se présente par des lésions cardiaques superposées, et se manifeste par une cardiomyopathie, qui se caractérise par une vacuolisation cytoplasmique des cardiomyocytes due à un gonflement du cytoplasme et des mitochondries, une destruction des organites, la mort des cardiomyocytes, ainsi que des troubles des fibres musculaires (**Figure 3**) (Qiu et al., 2023).

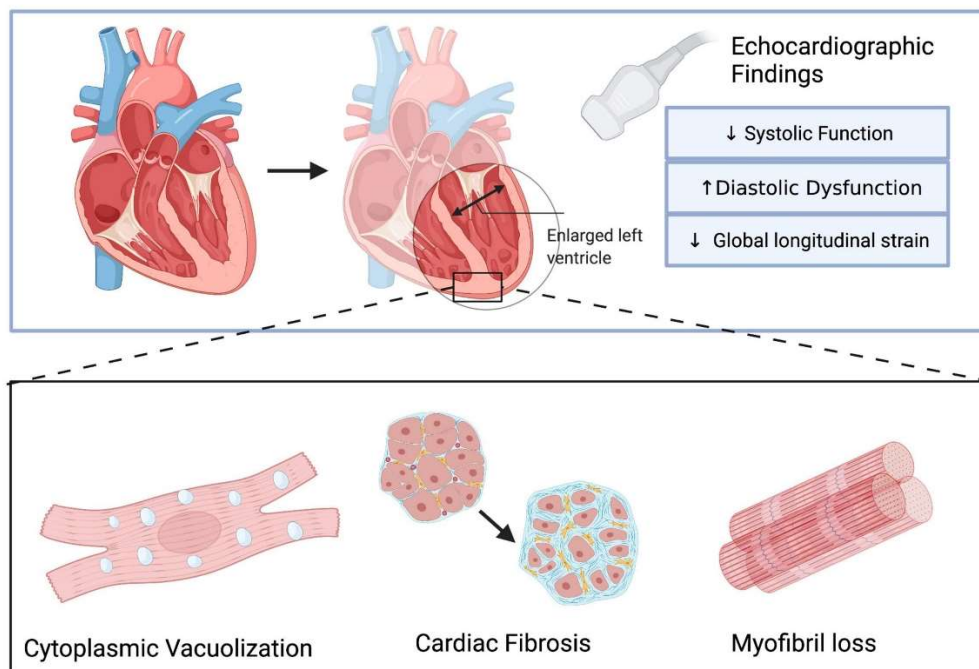


Figure 3. Résumé de la cardiotoxicité induite par les anthracyclines (Source : Vuong et al., 2022)

D'un autre côté, la radiothérapie incluant le cœur peut induire des lésions macrovasculaires, microvasculaires, endothéliales et valvulaires, entraînant ainsi des ischémies myocardiques, des péricardites, des valvulopathies, des troubles du rythme cardiaques mais aussi des insuffisances cardiaques (**Figure 4**) (Niska et al., 2018). A savoir que les pathologies cardiaques radio-induites peuvent mettre plusieurs décennies avant de se manifester cliniquement (Siaravas et al., 2023).

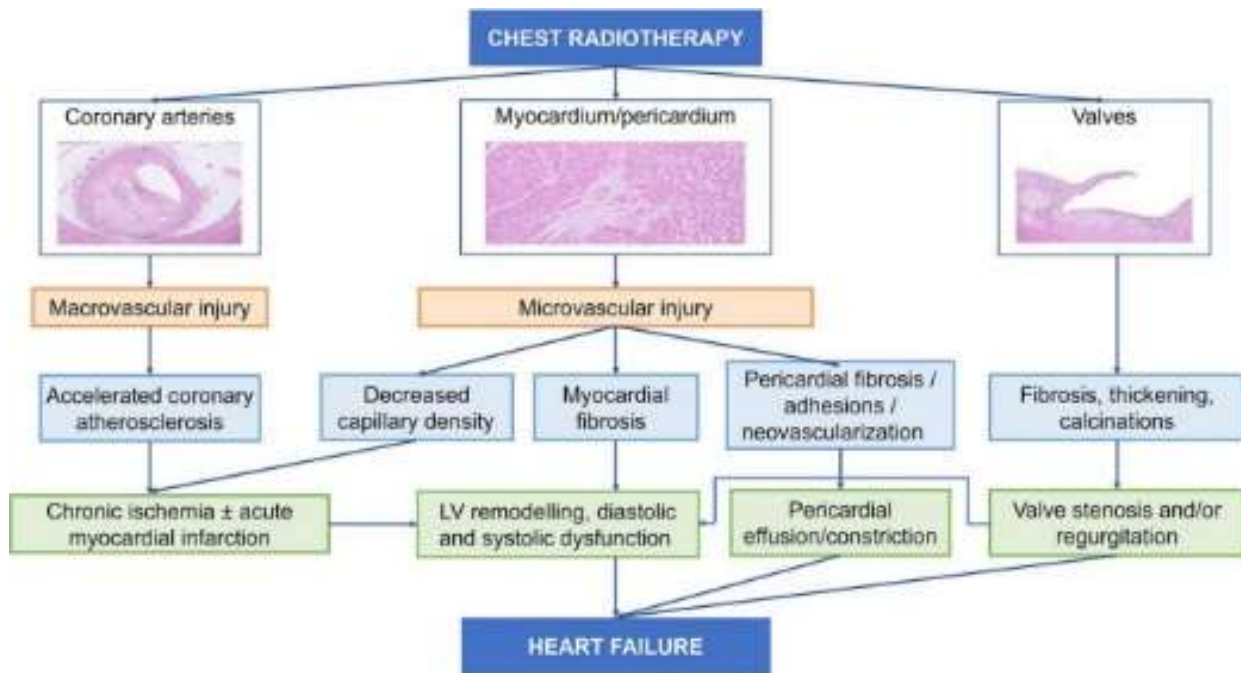


Figure 4. Pathophysiologie de la cardiotoxicité induite par l'exposition aux radiations cardiaques (Source : Podlesnikar et al., 2022)

Bien que plusieurs pathologies cardiaques puissent survenir suite à l'exposition aux radiations cardiaques et aux anthracyclines, l'insuffisance cardiaque reste l'une des pathologies cardiaques secondaires au traitement du cancer les plus fréquentes chez les survivants d'un cancer de l'enfant (Leerink et al., 2020).

La Société européenne de cardiologie définit l'insuffisance cardiaque comme un syndrome clinique caractérisé par des symptômes typiques (par exemple, essoufflement, gonflement des chevilles et fatigue) qui peuvent être accompagnés de signes cliniques (par exemple, pression veineuse jugulaire élevée, crépitements pulmonaires et œdème périphérique) causés par une anomalie cardiaque structurelle et/ou fonctionnelle (Ponikowski et al., 2016). Son diagnostic clinique est basé sur un examen minutieux des antécédents médicaux, et un examen physique, et ne peut pas être caractérisé par un seul test diagnostique. Les principales manifestations cliniques de l'insuffisance cardiaque sont la dyspnée, la fatigue et la rétention d'eau (Berg et al., 2022)

1.3 PREDICTION ET STRATIFICATION DU RISQUE DE PATHOLOGIE CARDIAQUES

En raison du jeune âge des survivants d'un cancer dans l'enfance, et donc de leur espérance de vie potentielle, il est nécessaire que ces patients aient accès à des soins complets et à un suivi à long terme personnalisé pour surveiller et traiter les effets tardifs associés aux thérapies anticancéreuses. Il est crucial que les praticiens médicaux et les survivants soient conscients des effets néfastes des traitements du cancer de l'enfant afin d'optimiser leur qualité de vie. Le risque cardiotoxiques doit alors être soigneusement évalué par rapport aux bénéfices potentiels du traitement (Samuel et al., 2023). La minimisation du risque de cardiotoxicité chez les survivants d'un cancer de l'enfant est possible en effectuant une stratification des risques de manière précoce, en optimisant la prise en charge des pathologies cardiaques pré-existantes et en identifiant rapidement les complications cardiaques dès leur apparition (Samuel et al., 2023). L'évaluation du risque de cardiotoxicité peut être réalisé selon des critères de la Heart Failure Association-International Cardio-Oncology Society (HFA-ICOS), qui prend en compte des facteurs tels que les antécédents cardiovasculaires, les expositions thérapeutiques cardiotoxiques antérieures, le mode de vie, l'âge, les facteurs de risque cardiaque et les biomarqueurs circulant cardiaques tels que le peptide natriurétique de type B (BNP), sa partie N-terminale (NT-pro-BNP), et les troponines. Elle permet de stratifier les survivants en quatre groupes : risque faible, risque modéré, risque élevé et risque très élevé (Lyon et al., 2020). Ainsi, le suivi à long terme et les stratégies de dépistage des pathologies tardives chez les survivants doivent alors être individualisés en fonction du groupe de risque d'appartenance des patients (Leerink et al., 2024).

1.3.1 Les biomarqueurs sanguins

Les biomarqueurs sanguins circulants sont très importants lors de l'évaluation des risques et de l'identification précoce de la cardiotoxicité induite par les anthracyclines. La troponine cardiaque, le BNP et le NT-pro-BNP sont des outils permettant d'identifier de façon précoce l'altération des cardiomyocytes, afin de prévenir les dysfonctionnements du ventricule gauche après administration d'anthracyclines.

L'augmentation des taux de troponine accompagne souvent les traitements par anthracyclines, et engendre un risque accru de dysfonctionnement du ventricule gauche et de la réduction de la fraction d'éjection du ventricule gauche qui en résulte. Le rôle des biomarqueurs est bien connu dans la détection des pathologies cardiaques durant le traitement du cancer (Qiu et al., 2023). Cependant, le rôle de ces protéines dans la détection des dysfonctionnements cardiaques tardifs suivant les traitements du cancer de l'enfant n'est pas clair (Leerink et al., 2024). En effet, ces protéines ont été retrouvées dans la cardiomyopathie aiguë mais sans effets sur les événements cardiotoxiques à long terme (Brouwer et al., 2011). Chez les survivants du cancer de l'enfant, certaines études ont cherché à évaluer le rôle du BNP et du NT-pro-BNP dans l'apparition de pathologies cardiaques tardives, mais leurs résultats restent limités (Brouwer et al., 2011; Mladosevicova et al., 2012).

Une étude récente a évalué la capacité de ces biomarqueurs à diagnostiquer une dysfonction cardiaque tardive chez les survivants d'un cancer de l'enfant en comparant trois modèles : (1) modèle comprenant les variables cliniques seules, (2) modèle comprenant les biomarqueurs sanguins seuls et (3) modèle comprenant à la fois les variables cliniques et les biomarqueurs sanguins. Cela a permis de mettre en évidence que ces biomarqueurs sanguins seuls ne permettent pas de diagnostiquer les patients souffrant d'une dysfonction cardiaque tardive, mais a mis en évidence la supériorité d'un modèle diagnostique combinant les biomarqueurs sanguins et les caractéristiques cliniques des survivants (Leerink et al., 2024).

1.3.2 Les marqueurs génétiques associés aux pathologies cardiaques

Certaines études se sont également intéressées aux prédispositions génétiques pour expliquer pourquoi certains survivants d'un cancer de l'enfant développent une cardiotoxicité et que d'autres non. Plusieurs variants génétiques impliqués dans les lésions de l'ADN, le stress oxydatif, le métabolisme du fer, le dysfonctionnement du sarcomère, le métabolisme et le transport des anthracyclines ont été décrits (Visscher et al., 2012, 2013, 2015).

Dans une revue de la littérature non spécifique aux cancers pédiatriques (Berkman et al., 2021), les prédispositions génétiques associées à la cardiotoxicité induite par les anthracyclines ont été étudiées. Aujourd'hui, les recommandations de surveillance de la cardiomyopathie spécifiques aux survivants du cancer de l'enfant et de l'adolescent préconisent la surveillance des patients traités par des doses d'anthracyclines ≥ 100 mg/m², et/ou de radiations ≥ 15 Gy par mesure de la fraction d'éjection du ventricule gauche mesurée par échocardiographie 2D ou 3D. Cependant, ces recommandations ne prennent pas en compte les prédispositions génétiques dans la stratification du risque de pathologie cardiaque à long terme (Ehrhardt et al., 2023). Cette stratification est pourtant essentielle pour garantir des stratégies de suivi plus efficaces basées sur la personnalisation du suivi des patients. Il est donc nécessaire de lister les gènes potentiellement impliqués dans la survenue d'effets tardifs des traitements anticancéreux chez les survivants d'un cancer de l'enfant.

1.4 OBJECTIFS DE LA THESE

L'objectif global de la thèse est de faire progresser les connaissances sur les facteurs de risque liés au traitement du cancer de l'enfant et l'adolescent, et aux facteurs de prédisposition conduisant au développement de pathologies cardiaques tardives. Pour atteindre cet objectif, nous avons mené différentes études en utilisant la cohorte French Childhood Survivor Study (FCCSS).

Les trois objectifs principaux de la thèse sont :

- Évaluer le risque de pathologies cardiaques et d'hospitalisation pour cause cardiaque chez les survivants d'un cancer de l'enfance, lié à l'exposition aux anthracyclines et aux radiations cardiaques,
- Passer en revue l'ensemble résultats publiés sur les facteurs génétiques associés aux pathologies cardiaques chez les survivants d'un cancer dans l'enfance,

- Identifier des biomarqueurs du transcriptome associés aux pathologies cardiaques dans un échantillon de patient niché dans la cohorte FCCSS.

Cette thèse comprend au total sept chapitres. Le **Chapitre 1** introduit la problématique de recherche. Le **Chapitre 2** décrit les données de la cohorte FCCSS utilisées pour répondre à cette problématique au travers de différentes analyses. Le **Chapitre 3** détaille l'ensemble des méthodes statistiques utilisées dans les différentes études de la thèse. Le **Chapitre 4** présente les résultats d'une étude sur le risque de pathologies cardiaques et d'hospitalisation cardiaque chez les survivants d'un cancer de l'enfant dans la cohorte FCCSS. Ce chapitre est divisé en deux parties : la première cherche à étudier les risques cardiaques associés aux traitements par anthracyclines et radiations incluant le cœur chez les survivants d'un cancer pédiatrique en explorant le rôle de l'interaction entre l'irradiation cardiaque et l'anthracycline. La seconde partie cherche à explorer l'effet des traitements cardiotoxiques sur l'incidence des hospitalisations cardiaques dans cette même population. Le **Chapitre 5** est une revue systématique de la littérature des marqueurs génétiques identifiés comme étant associés à l'apparition de pathologies cardiaques tardives chez des patients traités pour un cancer de l'enfant. Le **Chapitre 6** décrit une étude sur l'identification de biomarqueurs de l'expression génétique issue du transcriptome associés aux pathologies cardiaques après traitement d'un cancer dans l'enfance, à travers une étude cas-témoins appariée, nichée dans la cohorte FCCSS. Le **Chapitre 7** discute les principaux résultats obtenus à travers les différentes études menées, et leurs implications. La discussion reconnaît à la fois les forces et les limites de la recherche. Quelques recommandations pour des recherches futures sont incluses, ainsi qu'une conclusion générale.

Chapitre 2. LA COHORTE FRENCH CHILDHOOD CANCER SURVIVORS STUDY (FCCSS)

Dans ce chapitre, nous décrivons la cohorte *French Childhood Cancer Survivors Study* (FCCSS) utilisée pour répondre aux différents objectifs de la thèse, puis la banque de données biologiques associée à cette cohorte. Les parties spécifiques aux différents objectifs de la thèse sont détaillées dans les chapitres correspondants.

2.1 DEFINITION DE LA COHORTE

La French Childhood Cancer Survivors Study (FCCSS) est une cohorte de patients ayant survécu au moins 5 ans après le diagnostic d'un cancer dans l'enfance ou l'adolescence, traités avant l'âge de 21 ans pour un cancer solide ou un lymphome entre 1942 et 2000. Initialement, cette cohorte faisait partie de la cohorte rétrospective EURO2K (cohorte franco-anglaise), qui incluait des patients traités entre 1942 et 1986 dans cinq centres de lutte contre le cancer en France : Institut Gustave Roussy, Institut Curie, Institut Jean Godinot, Centre Claudius Régaud, et Centre Antoine Lacassagne. Son extension incluant des enfants diagnostiqués entre 1986 et 1999 dans les deux centres français suivants : Institut Gustave Roussy, Institut Curie. Aujourd'hui, la FCCSS inclut 7 670 patients diagnostiqués avant 2000 pour un cancer solide ou lymphome, avant l'âge de 21 ans, et suivi dans l'un des 5 centres français cités au-dessus. Le **Tableau 1** présente les principales caractéristiques cliniques et démographiques des patients inclus dans cette cohorte.

Les données de cette cohorte sont utilisées afin (i) d'étudier les effets à long terme des traitements anti-cancéreux chez les patients ayant survécu au moins 5 ans à des tumeurs malignes de l'enfant ou de l'adolescent, (ii) de construire des modèles de prédiction de risque des effets iatrogènes des traitements, (iii) d'acquérir les connaissances nécessaires pour aider à personnaliser les traitements par chimiothérapie et/ou radiothérapie, en étudiant les susceptibilités individuelles

généétique ou génomiques pouvant influencer la survenue de pathologies secondaires aux traitements du cancer.

Tableau 1. Caractéristiques des survivants inclus dans la cohorte FCCSS.

Caractéristiques	N = 7670	%
Genre		
Masculin	4201	54,8
Féminin	3469	45,2
Centres		
Institut Gustave Roussy	5504	71,8
Institut Curie	1917	25,0
Centre Claudius Régaud	159	2,07
Institut Jean Godinot	75	0,98
Centre Antoine Lacassagne	15	0,2
Année de diagnostic de la première tumeur		
Médiane (min – max)	1986 (1946 - 2000)	
Avant 1970	658	8,6
1970–1979	1549	20,2
1980–1989	2485	32,4
1990 - 2000	2978	38,8
Âge au diagnostic du premier cancer		
Médiane (EI)	6.6 (2 -11)	
0–1	1727	22,5
2–4	1783	23,2
5–9	1736	22,6
10–14	1609	21,0
≥15	815	10,6
Diagnostic de la tumeur primaire		
Lymphome	1279	16,7
Tumeur du Système Nerveux Central	1138	14,8
Tumeur rénale	1128	14,7
Neuroblastome	1035	13,5
Sarcome des tissus mous	819	10,7
Sarcome osseux	701	9,1
Rétinoblastome	619	8,1

Tumeur des gonades	439	5,7
Autres cancers solides	512	6,7
Traitement		
Pas de Radiothérapie ni de Chimiothérapie	905	11,8
Radiothérapie seule	1089	14,2
Chimiothérapie seule	2569	33,5
Radiothérapie et Chimiothérapie	3107	40,5

Abréviations : FCCSS : French Childhood Cancer Survivors Study ; EI : Etendue Interquartiles ; min : minimum ; max : maximum

Le protocole de la cohorte FCCSS a été approuvé par le comité d'éthique de l'Institut National de la Santé et de la Recherche Médicale (INSERM), et par la Commission Nationale de l'Informatique et des Libertés (CNIL).

2.1.1 Recueil des données et suivi

Les données démographiques, les caractéristiques cliniques et histopathologiques des cancers des patients de la cohorte FCCSS, ainsi que les informations détaillées sur les traitements administrés ont été recueillies à partir des dossiers médicaux des patients. La classification des tumeurs s'est faite sur la base de la troisième édition des classification internationale des cancers de l'enfant (ICCC-3) (Steliarova-Foucher et al., 2005).

Les données de chimiothérapie extraites des dossiers médicaux comprennent : le nom des traitements administrés, les dates et voies d'administration, ainsi que la dose totale administrée par unité de surface corporelle en milligrammes par mètre carré (mg/m²) pour chaque type de traitement. Les traitements antinéoplasiques ont été classés selon leur groupe pharmacologique et en fonction de leurs mécanismes d'action cellulaire : anthracyclines, agents alkylants, vinca alcaloïdes, antimétabolites, autres antibiotiques et autres traitements cytotoxiques.

La particularité de la cohorte FCCSS est qu'elle est la seule à disposer d'une reconstruction dosimétrique voxélisée du corps entier pour la quasi-totalité des patients traités par radiothérapie. Les doses de radiations ont été estimées à partir d'un

fantôme anthropomorphe voxélisé afin de générer pour chaque sujet une tomographie de substitution assistée par ordinateur, pour la planification de la radiothérapie du corps entier, et par conséquent, une anatomie de fantôme spécifique générée comme substitut pour chaque patient en fonction du sexe, de l'âge et de la position du traitement par radiothérapie (Veres et al., 2014; Vű Bezin et al., 2017). Les métriques dose-volume peuvent être calculées pour chaque organe ou région anatomique sélectionné dans le fantôme voxélisé tridimensionnel.

Dans le cadre la thèse, nous nous sommes spécifiquement intéressés au cœur (**Figure 5**). Plusieurs métriques de doses ont été recueillies pour chaque patient : la dose moyenne, la dose médiane, la dose minimale et maximale de radiations reçues au cœur, ainsi que les paramètres dose-volume indiquant le volume du cœur ayant reçu au moins d Gy (s'écrit Vd).

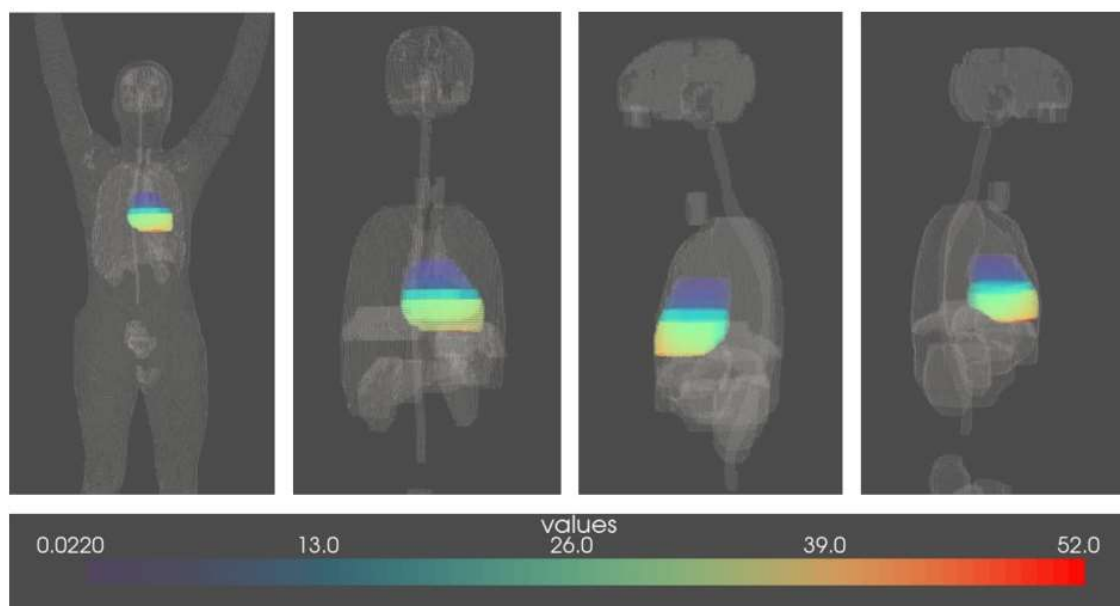


Figure 5 Représentation de la reconstruction voxélisée de la dose de radiations au cœur d'un survivant d'un cancer de l'enfant (vue de face, de dos, de gauche et de droite). Les nuances de couleur représentent le niveau de la dose. (Figure réalisée par Duyen Do, équipe Epidémiologie des radiations).

Encore aujourd'hui, les patients de la cohorte FCCSS bénéficient d'un suivi prospectif à partir de l'accès aux dossiers médicaux, des auto-questionnaires, des consultations de suivi à long terme, et des données du Système National des Données de Santé (SNDS). Le **Tableau 2** regroupe l'ensemble des sources à partir desquelles les données concernant le suivi à long terme des patients inclus dans la cohorte FCCSS ont été recueillies.

Un questionnaire d'auto-évaluation portant sur divers évènements indésirables d'ordre sanitaire, social et psychologique a été envoyé en deux vagues (une première en 2005 et une seconde en 2010), à tous les patients encore vivants dans la cohorte FCCSS, dont les adresses étaient disponibles dans le registre de la Caisse Nationale d'Assurance Maladie (CNAM). La lettre de contact comprenait un accord de participation à signer et une autorisation pour contacter le médecin et les établissements médicaux qui suivaient les patients. Parmi les 5023 patients encore en vie avant l'envoi des questionnaires, 3293 (65,5 %) ont répondu au questionnaire, dont 2961 (59,0 %) ont répondu à tous les éléments du questionnaire.

Dans un premier temps, les évènements indésirables ont été déclarés par les patients eux-mêmes, puis les diagnostics ont été confirmés par les médecins concernés. Depuis 2012, un programme de suivi à long terme a été mis en place à l'Institut Gustave Roussy et à l'Institut Curie. Le chaînage des données de la cohorte FCCSS avec celles du SNDS a été approuvé par un arrêté ministériel. Le SNDS regroupe les données de soins de ville et des établissements de santé privés (appelées Données de Consommation Inter-Régimes, DCIR) et des données des séjours hospitaliers issues du Programme de Médicalisation des Systèmes d'Information (PMSI). Les données du SNDS pour les patients de la cohorte FCCSS sont disponibles depuis 2006 et sont régulièrement mises à jour.

Les informations médicales extraites à partir du SNDS sont nombreuses. Elles incluent les pathologies chroniques donnant lieu à une affection de longue durée (ALD), dont le motif d'exonération est codé selon la Classification Internationale de Maladies

version 10 (CIM-10), ainsi que les motifs d'hospitalisation, également codés selon la CIM-10. Les causes du décès pour les patients décédés sont obtenues par croisement avec le registre national des causes de décès, coordonné par le Centre d'épidémiologie sur les causes médicales de Décès (CepiDC). Récemment les données des causes de décès sont également disponibles dans le SNDS.

2.1.2 Identification des pathologies cardiaques et des hospitalisations pour cause cardiaque

Les pathologies cardiaques (insuffisance cardiaque, valvulopathie, péricardites, les troubles du rythme et de la conduction) ont été identifiées à partir de multiples sources: (i) rapportées par les patients à l'aide d'auto-questionnaires, (ii) depuis les dossiers médicaux, (iii) lors des consultations suivi à long terme des survivants du cancer, (iv) à partir des bases de données du SNDS. Le médecin généraliste ou le cardiologue de tous les patients en vie ayant signalé au moins une des maladies cardiaques étudiées a été invité à remplir un questionnaire confirmant le diagnostic, précisant la date d'apparition, si nos critères de validation étaient remplis et quel traitement médicamenteux et interventionnel avait été administré. Une copie des résultats médicaux a été demandée le cas échéant. Chaque diagnostic était validé sur la base des critères de la Société Européenne de Cardiologie et des études Framingham et PRIME (Ducimetière et al., 2001; Ho et al., 1993; Ponikowski et al., 2016; Truong et al., 2012).

Les données sur le devenir des patients et la cause du décès pour les patients décédés, sont obtenues par couplage avec le registre national des causes de décès coordonné par le CepiDC, et par couplage au SNDS, croisé avec la base de données du PMSI. Les informations relatives aux hospitalisations pour cause cardiaque, dont les dates et la durée du séjour, les diagnostics (affections principales et associées) et les actes médicaux réalisés, ont été extraites des données du SNDS, qui fournit des données de suivi entre 2006 et 2018, incluant des données ambulatoires dont l'attribution et type d'affections de longue durée codées selon la CIM-10.

Tableau 2. Sources des données recueillies pour le suivi à long terme des patients inclus dans la cohorte FCCSS

Sources	Données extraites
Questionnaires d'auto-évaluation	<p>Questionnaires comprenant des questions sur les éléments suivants :</p> <ul style="list-style-type: none"> • Situation générale et vie socio-professionnelle : situation conjugale, scolarité et niveau scolaire, déroulement de la vie professionnelle, ressources financières, situation administrative, mode de vie, entourage, activités et loisirs. • Qualité de vie : perception de la santé, état de santé au quotidien, accomplissement des activités de la vie courante. • Etat de santé physique et psychologique à la suite du traitement du cancer primaire : diagnostic de tumeurs secondaires malignes ou bénignes, diagnostic de pathologies liées au traitement du cancer primaire, suivi psychologique ou psychiatrique, suivi médical (généraliste et spécialiste), prise régulière de traitements. • Déroulement de la vie reproductive : début de la puberté, suivi gynécologique et prévention des cancers du sein (chez la femme), désir d'enfant, conception des enfants, déroulement des grossesses et ménopause (chez la femme), déroulement des grossesses des conjointes (chez l'homme). • Famille : date de naissance et de décès des parents, enfants, grands-parents, fratrie, développement cancers dans la famille élargie • Prise en charge médicale : contact des médecins qui les ont suivis durant l'enfance, apparition de problèmes de santé importants après la fin du traitement du cancer primaire. • Environnement de la maladie : besoin d'informations sur le cancer primaire ou ses traitements, suivi médical particulier relatif au cancer primaire depuis la fin du traitement, contact avec des associations de malades et d'anciens malades.
Consultations de suivi à long terme	<p>Depuis 2012, un programme de suivi à long terme a été mis en place à l'Institut Gustave Roussy et à l'Institut Curie, qui permet d'extraire des données sur :</p> <ul style="list-style-type: none"> • L'état de santé générale des patients : poids, taille, degrés d'autonomie et de dépendance, niveau d'études, profession, situation professionnelle. • L'état de santé le jour de la consultation : diagnostic de : pathologie thyroïdienne, cardiaque, cérébrovasculaire, vasculaires, pulmonaires, rénales, endocrinienne, osseuse, et sphinctérienne ; diagnostic de troubles sensoriels, neurocognitifs, digestifs, et de la sexualité ; diagnostic de problème de peau ou de phanères, d'anomalies dentaires, diagnostic de secondes tumeurs ; et identification de syndrome de prédisposition génétique aux cancers. Si diagnostic d'une pathologie ou d'un trouble secondaire au traitement du cancer primaire, remplissage d'une fiche clinique dédiée : diagnostic exact, date du diagnostic, actes de biologie et d'imagerie liés au

	<p>diagnostic, actes médicaux liés au diagnostic, grade de la pathologie, antécédents et facteurs de risques, traitement de la pathologie, statut vital, date du décès, cause du décès</p> <ul style="list-style-type: none"> • Le déroulement de la puberté et sur la fertilité • Le dernier bilan sanguin • La dernière échographie abdominale • Les examens complémentaires : radiologique, ophtalmologique, stomatologique, endocrinologique, et ORL.
SNDS	<p>Fusion de plusieurs bases de données pseudonymisées collectées par des organismes publics. Elle contient l'ensemble des données de remboursement de l'assurance maladie en soin ambulatoire :</p> <ul style="list-style-type: none"> • Les données extraites du DCIR : consultations médicales en ambulatoire, actes biologiques réalisés hors secteur hospitalier, médicaments délivrés (identifiés à partir du code club inter-pharmaceutique (CIP) et de la classification anatomique, thérapeutique et chimique (ATC)), pathologies chroniques donnant lieu à une ALD, dont le motif d'exonération est codé selon la CIM-10. • Les données extraites du PMSI : motif d'hospitalisation renseigné par le diagnostic principal ou le diagnostic relié au séjour (codés à partir de la CIM-10), actes médicaux réalisés au cours du séjour, comorbidités ayant conduit à des soins au cours du séjour (diagnostics associés significatifs) ou autres diagnostic principal issus des résumés d'unités médicales, médicaments et dispositifs implantables coûteux, destination de la sortie (autre service du même hôpital, autre hôpital, décès, ou sortie vers le domicile). • Les données extraites du CepiDC : date de décès, âge au décès, enchainement des évènements ayant conduit au décès, causes du décès (codés selon la CIM-10).

Les pathologies cardiaques ont été classées à l'aide des critères terminologiques communs pour les événements indésirables du CTCAE (Common Terminology Criteria for Adverse Events) version 4.03 (US Department of Health and Human Services, 2010). La pathologie est : de grade 1 si le patient est asymptomatique ou présente des symptômes légers, et qu'aucune intervention n'est indiquée ; de grade 2 s'il présente une pathologie modérée, et qu'aucune intervention invasive n'est indiquée ; de grade 3 s'il présente une pathologie sévère ou médicalement significative mais qui ne met pas immédiatement sa vie en danger, mais peut nécessiter une hospitalisation, et limiter son autonomie ; de grade 4 si la pathologie a des conséquences mettant en jeu le pronostic vital du patient, ou si une intervention urgente est indiquée ; de grade 5 si le patient est décédé d'une pathologie cardiaque. Le grade pour chaque pathologie est présenté dans le **Tableau 3**.

Tableau 3. Grades de pathologies cardiaques selon la version 4.0 des CTCAE (Source : US Department of Health and Human Services, 2010) .

Pathologie	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Insuffisance cardiaque	Asymptomatique avec des anomalies de biologie (ex : BNP) ou d'imagerie cardiaque	Symptômes lors d'une activité ou d'un effort léger à modéré	Sévère avec des symptômes au repos ou lors d'une activité ou d'un effort minimal ; intervention indiquée	Conséquences mettant en jeu le pronostic vital; intervention urgente indiquée (ex: traitement intraveineux continu ou soutien hémodynamique mécanique).	Décès
Valvulopathies	Épaississement valvulaire asymptomatique avec ou sans régurgitation valvulaire légère ou sténose par imagerie	Asymptomatique; régurgitation ou sténose modérée par imagerie	Symptomatique ; régurgitation ou sténose sévère à l'imagerie; symptômes contrôlés par une intervention médicale	Conséquences mettant en jeu le pronostic vital ; intervention urgente indiquée (ex: remplacement de la valve, valvuloplastie)	Décès
Péricardites	Asymptomatique, ECG ou signes physiques (p. ex. frottement) compatibles avec une péricardite	Péricardite symptomatique (ex: douleur thoracique)	Péricardite avec conséquences physiologiques (ex: constriction péricardique)	Conséquences mortelles ; intervention urgente indiquée	Décès
Arythmie	Asymptomatique, intervention non indiquée	Intervention médicale non urgente indiquée	Intervention médicale indiquée	Conséquences mettant en jeu le pronostic vital ; compromis hémodynamique ; intervention urgente indiquée	Décès
Trouble de la conduction	Symptômes légers ; intervention non indiquée	Symptômes modérés	Symptôme sévères : intervention indiquée	Conséquences mettant en jeu le pronostic vital ; intervention urgente indiquée	Décès
Autres troubles cardiaques	Asymptomatique ou symptômes légers ; observations cliniques ou diagnostiques uniquement ; intervention non indiquée	Modéré ; intervention locale ou non invasive minimale indiquée ; limitation des AVQ instrumentales adaptées à l'âge.	Sévère ou médicalement significative mais ne mettant pas immédiatement la vie en danger ; hospitalisation ou prolongation d'une hospitalisation existante indiquée ; invalidante ; limitant l'autonomie dans les AVQ	Conséquences mortelles ; intervention urgente indiquée	Décès

Abréviations : ECG : Electrocardiogramme, AVQ : activités de la vie quotidienne

2.2 BANQUE BIOLOGIQUE

Une banque biologique à visée de recherche pour la cohorte FCCSS a été établie. Cette banque biologique a été constituée par une demande d'échantillons salivaires et sanguins chez les survivants adultes de la cohorte FCCSS, prélevés lors de visites médicales de routine ou à domicile. Le recrutement a débuté en mars 2011. En 2021, 3032 survivants de la FCCSS ont accepté de fournir un échantillon biologique : 2421 survivants ont fourni au moins un échantillon de salive, et 1571 survivants ont fourni au moins un échantillon de sang (**Figure 6**).

Des kits de prélèvement d'échantillons sanguins ont été envoyés par courrier aux survivants de la FCCSS qui ont accepté de participer à l'étude. Ces kits comprennent tous les tubes nécessaires à la prise de sang dans un laboratoire d'analyses médicales. Tous les tubes sont pré-étiquetés avec un code-barres unique afin d'assurer une traçabilité complète à partir de l'inclusion du patient. Les échantillons de sang sont conservés à température ambiante et expédiés sous 48h au laboratoire Généthon. Généthon assure le traitement de l'échantillon sanguin reçu (isolement des lymphocytes, établissement des lignées lymphoblastiques, isolement du plasma), stocke les échantillons sanguins et les produits dérivés, enregistre les données associées à chaque échantillon, duplique les collections afin d'assurer une sécurité maximale en les stockant dans un autre lieu géographique. A partir des culots cellulaires expédiés par Généthon, la plateforme génomique Gustave Roussy réalise l'extraction de l'ARN à partir des lignées cellulaires et la mesure de l'expression des gènes en utilisant la puce Affymetrix Human Clariom-D.

Les échantillons de salive prélevés lors des visites de suivi des patients ayant accepté de participer à l'étude ont été transportés par les Assistant de Recherche Clinique (ARC) de l'équipe Epidémiologie des Radiations à l'Institut Gustave Roussy, tandis que les échantillons de salive prélevés à domicile ont été scellés dans une enveloppe adressée, que le patient a postée depuis son domicile. Tous les participants reçoivent un mode d'emploi et le kit de prélèvement d'ADN salivaire. La traçabilité est assurée par

l'apposition d'un code-barres sur le tube de salive. La salive est conservée et transportée à température ambiante. A l'arrivée à l'Institut Gustave Roussy, les échantillons ont été stockés de manière sécurisée avant d'être centralisés au centre de ressources biologiques de Gustave Roussy. Le prélèvement d'échantillons d'ADN à partir de la salive est effectué avec Oragene DNA OG-600 de DNAgentok TM un système tout-en-un pour le prélèvement, la stabilisation et le transport.

Tous les échantillons sont étiquetés dans le respect de la pseudonymisation à l'aide d'un code-barres.

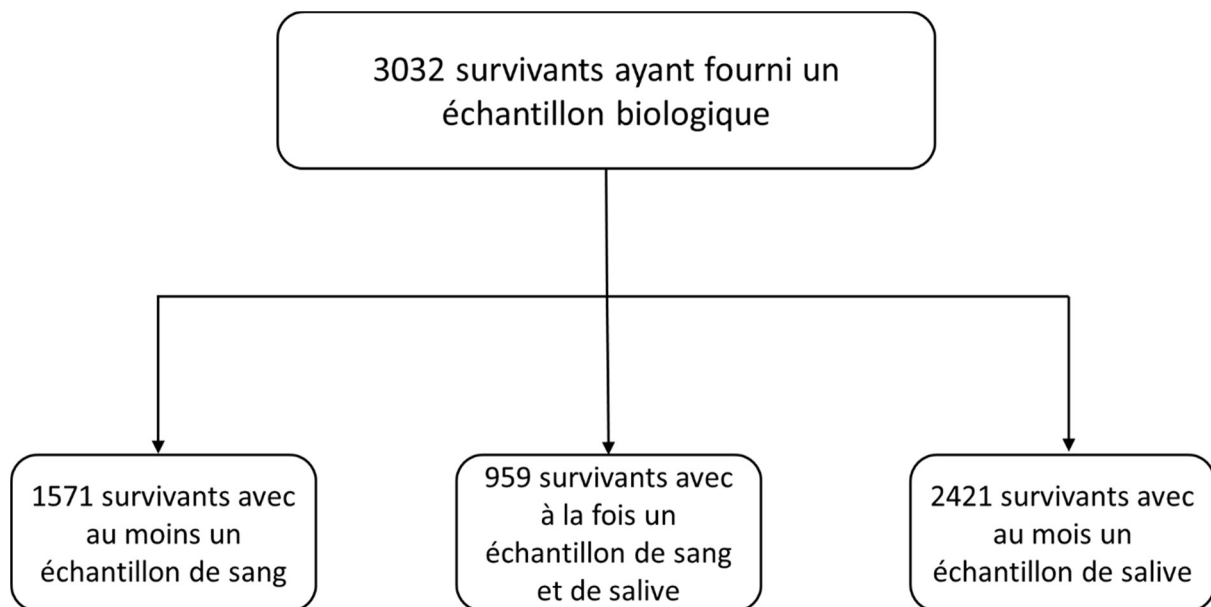


Figure 6. Diagramme de flux des patients de la cohorte FCCSS ayant fourni des échantillons biologiques pour la constitution de la banque biologique.

Chapitre 3. METHODOLOGIE STATISTIQUE

Ce chapitre introduit les éléments généraux expliquant les choix méthodologiques qui ont été faits pour réaliser les différents travaux de la thèse. Les modèles de Cox cause-spécifique et de Poisson ont été utilisés dans le Chapitre 4 pour analyser les risques associés aux traitements du cancer sur les pathologies et hospitalisations cardiaques. La régression logistique conditionnelle adaptée à la procédure de sélection Lasso pour des données en grande dimension dans une étude cas-témoins appariée, nichée dans la cohorte FCCSS, a été utilisée dans le Chapitre 6, pour identifier des biomarqueurs associés aux pathologies cardiaques après traitement d'un cancer dans l'enfance.

3.1 ANALYSE DE SURVIE DANS UNE ETUDE DE COHORTE

Cette section du chapitre présente les deux méthodes statistiques d'analyse de survie menées dans le Chapitre 4. Nous découpons donc cette partie en deux sous-parties correspondant chacune à une analyse dans l'ordre dans lequel elles sont présentées dans le chapitre correspondant.

3.1.1 Prise en compte des risques concurrents

Un risque concurrent ou compétitif est défini comme un évènement qui (i) exclu la survenue d'un autre évènement étudié, ou (ii) modifie fondamentalement la probabilité de survenue de cet autre évènement (Gooley et al., 1999).

En présence d'évènements concurrents, l'estimation de l'incidence cumulée de l'évènement d'intérêt doit tenir compte des risques concurrents : l'estimateur de Kaplan-Meier est alors biaisé (Satagopan et al., 2004). L'estimateur non paramétrique de Aalen-Johansen est une solution possible pour l'estimation de la probabilité de se trouver dans chaque état à chaque moment du suivi.

Dans le cadre de l'analyse d'un évènement en présence d'un évènement concurrent, deux types de régressions peuvent être utilisées, et le choix de la méthode adéquate

dépend du critère de jugement choisi.

Le modèle de Fine-Gray (Fine & Gray, 1999) est souvent utilisé, mais les sub-hazard-ratios dérivés de ce modèle n'ont pas d'interprétation simple, même si elles reflètent l'ordre des courbes d'incidence cumulée (Z. Zhang, 2017). Dans le cas où il est nécessaire d'obtenir une mesure interprétable, comme le rapport de risques instantanés (Hazard-Ratio en anglais), l'utilisation d'un modèle de Cox cause-spécifique est plus adéquate (Z. Zhang, 2017).

Le modèle de régression de Cox cause-spécifique peut être effectué à travers un modèle de régression de Cox en traitant les évènements d'intérêt comme des évènements et les évènements concurrents comme des observations censurées. Le risque associé aux traitements est ensuite estimé pour l'évènement concurrent, ainsi, des hazard-ratio cause-spécifiques sont obtenus pour chacun des deux évènements (Latouche et al., 2013; Z. Zhang, 2017). C'est donc ce dernier modèle qui a été utilisé dans la première partie du Chapitre 4 de la thèse.

Le modèle de Cox à risques proportionnels est un modèle de régression fréquemment employé en recherche médicale pour étudier l'association entre la survenue d'un évènement au cours du temps, et une ou plusieurs variables indépendantes (Cox, 1972). Le modèle de Cox s'écrit par la fonction de risque instantané $h(t)$, qui correspond au risque de la survenue de l'évènement en question, en fonction d'un ensemble de p covariables $X = (X_1, X_2, \dots, X_p)$:

$$h(t) = h_0(t) \times \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p) \quad (1)$$

Où $h_0(t)$ représente le risque de base lorsque toutes les variables X sont égales à 0, et l'ensemble des coefficients $\beta_1, \beta_2, \dots, \beta_p$ sont les coefficients correspondants aux covariables X_1, X_2, \dots, X_p . Il est possible d'estimer le risque associé à une variable indépendante X en calculant un Hazard-Ratio (HR), c'est-à-dire l'exponentielle du coefficient β correspondant à la variable. Un HR significativement supérieur à 1 indique un risque élevé d'apparition de l'évènement au cours du temps, tandis qu'un HR inférieur à 1 démontre un risque inverse.

L'estimation des paramètres du modèle de Cox est obtenue en utilisant l'estimation du maximum de la vraisemblance partielle, et les intervalles de confiance correspondant.

Dans le modèle présenté (équation (1)), toutes les variables sont indépendantes du temps, c'est-à-dire qu'elles ne varient pas au cours du suivi des patients. Cependant, lorsque les données sont collectées de manière longitudinale, celle-ci peuvent évoluer au cours du temps, les variables sont donc considérées comme dépendantes du temps, et le modèle s'écrit comme suit :

$$h(t) = h_0(t) \times \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i + \beta_{i+1} X_{i+1}(t) + \beta_p X_p(t)) \quad (2)$$

Où les covariables allant de 1 à i sont fixées dans le temps et la covariable p est dépendante du temps.

3.2 LES TAUX STANDARDISÉS D'HOSPITALISATION

3.2.1 Calcul des taux standardisés

Les taux d'hospitalisation pour une pathologie spécifique dans la population générale sont calculés à partir de l'ensemble des données extraites des registres du PMSI, en divisant le nombre de personnes hospitalisées chaque année par les effectifs de la population générale en milieu d'année, par strate d'année, de sexe et d'âge. Les effectifs de population générale sont issus des données de l'Institut National de la Statistique et des Etudes Economiques (INSEE). Les nombres attendus d'hospitalisations dans la population étudiée sont obtenus en multipliant les taux dans la population générale par le nombre de personnes-années dans chacune des strates.

Les taux d'hospitalisation standardisés (SHR), sont calculés en divisant la somme des nombres observés d'hospitalisations dans notre population d'intérêt par la somme des nombres attendus. L'excès absolu de risque (AER) est calculé en soustrayant le nombre attendu du nombre observé et en divisant par le nombre de personnes-années, puis en le multipliant par 10 000. Les SHR correspondent à un facteur multiplicatif de risque, et les AER à un facteur additif, comparé à la population générale.

3.2.2 Modèle de Poisson

Pour étudier l'effet simultané des facteurs de risque cardiaque, un modèle de régression de Poisson multivarié a été utilisé pour estimer les risques relatifs, en y ajoutant le logarithme du nombre attendu d'évènement étudié comme « offset », (Breslow & Day, 1991). Les risques relatifs (RR) peuvent ainsi être interprétés comme le ratio des SHR ajustés pour les facteurs de confusion potentiels.

Le modèle de régression de Poisson est un type de modèle linéaire généralisé, souvent utilisé en épidémiologie pour modéliser des données de comptage. Il est possible d'utiliser les modèles de régression de Poisson pour modéliser le nombre d'évènements observés par intervalle de temps et par dose d'exposition en supposant que le rapport de taux associé à une exposition d'intérêt X est constant sur une série d'intervalles de temps. Les données individuelles sont alors regroupées dans un tableau de contingence qui comprend k catégories de l'exposition de X, et j strates de temps. A chaque cellule du tableau est attribué le nombre d'évènements observés n_{jk} , le nombre de personnes-années PA_{jk} , et le taux d'occurrence des évènements λ_{jk} .

Étant donné que le nombre d'évènements observés n_{jk} suit une distribution de Poisson de paramètre λ_{jk} , le nombre attendu d'évènements N_{jk} dans chaque cellule peut être estimé comme le produit du nombre de personnes-années par strate et par niveau d'exposition, et du taux d'occurrence des évènement par strate et par niveau d'exposition :

$$N_{jk} = PA_{jk} \lambda_{jk} \quad (3)$$

Ainsi, le RR associé à cette exposition est donc le rapport entre le taux d'occurrence des évènements par strate j et par niveau d'exposition k. La comparaison du nombre d'évènements observés au nombre d'évènements attendus dans une population standard se fait par le calcul du ratio d'incidence standardisé (dans ce contexte il s'agit de l'incidence des hospitalisations, soit le SHR):

$$SHR = RR_{jk} = \frac{\lambda_{jk}}{\lambda_j} \quad (4)$$

Pour étudier le risque associé à un effet-dose de l'exposition, le RR_{jk} prend souvent la forme $RR_{jk} = \exp(\beta d_{jk})$, avec d_{jk} la dose moyenne d'exposition. En résumé, le risque relatif associé à X dans la population est modélisé comme suit :

$$RR = w_{jk} \lambda_j \exp(\beta d_{jk}) \quad (5)$$

où w_{jk} représente l'offset, soit le logarithme du nombre attendu d'évènement étudié :
 $w_{jk} = \log(N_{jk}) = \log(PA_{jk} \lambda_{jk})$

3.3 MODELE DE REGRESSION POUR UNE ETUDE CAS-TEMOINS APPARIEE ET METHODES DE SELECTION DE VARIABLES EN GRANDE DIMENSION

3.3.1 Le design cas-témoins niché et apparié

En épidémiologie, les études cas-témoins sont utilisées pour identifier les facteurs qui peuvent contribuer à la survenue d'un évènement de santé en comparant un groupe de cas, c'est-à-dire de personnes présentant l'évènement étudié, à un groupe de témoins qui ne présentent pas l'évènement mais qu'on considère comme similaires aux cas sur d'autres facteurs (Avalos et al., 2015). Les cas et les témoins sont parfois appariés de sorte que chaque cas est apparié à un nombre prédéfini de témoins qui partagent des similarités sur les facteurs d'appariement qui sont eux-mêmes liés à l'évènement de santé étudié, mais qui ne présentent pas d'intérêt particulier. L'appariement des cas et des témoins sur un ou plusieurs facteurs de confusion peut se traduire par des gains d'efficacité significatifs (Avalos et al., 2015; Rothman et al., 2008; Xu et al., 2019). Les covariables courantes d'appariement comprennent les caractéristiques du patient, les facteurs de risque de la maladie, ainsi que d'autres marqueurs (Janes & Pepe, 2008). Les études cas-témoins sont souvent nichées dans une importante cohorte prospective et sont utiles lorsque l'évènement de santé à étudier est rare. Dans de nombreux contextes, la mesure de ou des biomarqueurs sur tous les individus d'une cohorte prospective a un coût excessif et rend difficilement réalisable les études de cohorte. L'étude cas-témoin permet d'équilibrer le nombre de cas et de témoins au sein de strates d'appariement définies par des facteurs de confusion préalablement connus, et de ne mesurer les biomarqueurs que sur le sous-ensemble de sujets sélectionné (Xu et al., 2019). Ces designs ont récemment été introduits dans des études d'imagerie et de génomique, qui sont caractérisées par des covariables à grande dimension (Qian et al., 2014).

3.3.2 La régression logistique conditionnelle

Les réponses binaires sont couramment étudiées dans de nombreux domaines, notamment la présence ou l'absence d'une maladie particulière (Harrell , 2015). La régression logistique est l'une des plus importantes méthodes statistiques en épidémiologie pour analyser les données issues d'une étude cas-témoins. Elle permet de prendre en compte les facteurs de confusion potentiels (facteurs indépendamment associés à la fois à l'évènement de santé et aux facteurs de risque d'intérêt principal) et, si le modèle logistique est correct, d'éliminer leur effet (Avalos et al., 2015). Cependant, utiliser la régression logistique pour des données appariées peut être à l'origine d'un problème de dispersion des données, et donc d'un biais de l'estimation par rapport à la valeur réelle (Greenland et al., 2000). C'est pour remédier à ce problème que la régression logistique conditionnelle a été conçue (Kuo et al., 2018). Ainsi, dans le cadre des études cas-témoins appariées, le modèle de régression logistique conditionnelle est l'outil standard utilisé (Avalos et al., 2015). Ce type d'analyse fournit des estimations de paramètres qui peuvent être interprétés comme des rapports de cotes, aussi appelés Odds Ratio (OR). Au moyen d'une modélisation multivarié, les estimations des OR peuvent être ajustées en fonction des facteurs de risque qui n'ont pas été pris en compte dans la procédure d'appariement. Généralement, les estimations des paramètres sont obtenues en maximisant une log-vraisemblance conditionnelle (Breslow & Day, N E, 1980; Heinze & Puhr, 2010).

Soit Y une variable binaire codée 0 - 1, et pour laquelle une association est recherchée avec les variables du vecteur $X = (X_1, \dots, X_p)$. Les patients sont groupées par un appariement sur les variables du vecteur $X_m = (X_{m1}, \dots, X_{mj})$ et forment des strates notées $s = 1, \dots, S$, comprenant un cas (Y_{is}) et M témoins ($Y_{ls} = 0, l \neq i$). Dans une régression logistique conditionnelle, le modèle s'écrit :

$$\text{logit}(\pi_{is}) = \beta_{0s} + \sum \beta^T x \quad (6)$$

avec π_{is} la probabilité d'avoir l'évènement pour le i^{eme} sujet de la s^{ieme} strate, β_{0s} la contribution des termes d'appariement dans la s^{ieme} strate, et β les coefficients de

régression associés aux variables du vecteur X . Lorsque les différences entre les strates ne sont pas pertinentes, dans le sens où les facteurs d'appariement sont des facteurs de confusion potentiels, mais ne sont pas les facteurs de risque potentiels qui nous intéressent, il nous suffit d'estimer seulement β . Par conséquent, le paramètre β_{0s} est éliminé de la vraisemblance en conditionnant le fait que chaque ensemble cas-témoins apparié comporte exactement un sujet qui est un cas (Avalos et al., 2015). La vraisemblance du modèle pour le i^{eme} sujet s'écrit alors :

$$L_i(\beta) = \mathbb{P}(Y_{is} = 1 | X_{m_{is}}) = \frac{\exp(\sum \beta^T X_{is})}{1 + \exp(\sum \beta^T X_{is})} \quad (7)$$

Sont indexés par $i = 0$ et tous les témoins par $i \in (1, \dots, M)$. Lorsque l'appariement des cas se fait avec un seul témoin $M = 1$ (c'est à dire un appariement 1 :1), la formule de vraisemblance évaluée est simplifiée de sorte que vraisemblance conditionnelle est identique à la vraisemblance non-conditionnelle du modèle logistique avec X comme vecteur de covariables, pas d'ordonnée à l'origine et une réponse constante égale à 1 (Avalos et al., 2015). En d'autres termes, les données des témoins sont soustraites à celle des cas auxquels ils sont appariés, et la formule de vraisemblance s'écrit ainsi :

$$L(\beta) = \mathbb{P}(Y_{is} = 1 | X_{m_{is}}) = \frac{\exp(\sum \beta^T X_{0s})}{\exp(\sum \beta^T X_{0s} + \sum \beta^T X_{1s})} \quad (8)$$

$$L(\beta) = \frac{1}{1 + \exp(\sum -\beta (X_{0s} - X_{1s}))}$$

Par ailleurs, il se trouve que la log-vraisemblance d'un modèle de régression logistique conditionnelle équivaut à la log-vraisemblance d'un modèle de Cox stratifié (une strate correspond à un groupe cas-témoins apparié) dont le temps est fixé à une constante (Avalos et al., 2015; Breslow & Day, N E, 1980).

3.3.3 Analyses de données en grande dimension

Les statistiques en grande dimension se réfèrent à des méthodes d'inférence statistique lorsque le nombre de paramètres inconnus p est beaucoup plus important que la taille de l'échantillon n . Ces méthodes englobent les modèles de régression et de

classification supervisée et non-supervisée avec plus de variables que d'observations (Bühlmann & Van De Geer, 2011). Il est de plus en plus fréquent de se retrouver confronté à l'analyse de données en grande dimension dans différents domaines scientifiques dont l'épidémiologie grâce à la révolution de la génomique, et donc de l'augmentation de la quantité d'information disponible (Fan & Lv, 2010).

Comme expliqué dans la section précédente, les modèles de régression classiques ont pour objectif d'estimer des coefficients de régression (β) permettant d'évaluer l'association entre des prédicteurs d'intérêt et un évènement (dans ce contexte un évènement de santé). Cependant, dans les études classiques, le nombre d'observations n est largement supérieur au nombre de prédicteurs d'intérêt p , et il est alors possible de trouver une solution unique β qui maximise la fonction de vraisemblance du modèle. Dans le cadre de l'analyse de données en grande dimension ($p \gg n$), le modèle devient non identifiable, et il est impossible de trouver une solution unique β .

Une autre problématique se pose lorsqu'il s'agit de la sélection des variables. Un des objectifs principaux des analyses de données en grande dimension est d'identifier parmi un grand nombre de variables, des "variables informatives". D'un point de vue clinique, cela consiste à sélectionner les biomarqueurs qui seraient associées à l'évènement de santé étudié, de sorte à identifier les biomarqueurs les plus discriminants (Rahnenführer et al., 2023).

De nombreuses approches ont été décrites dans la littérature pour prendre en compte le nombre très élevé de variables de sorte à sélectionner les plus pertinentes. Nous en présenterons quelques-unes, dont la régression pénalisée Lasso qui est la plus détaillée car centrale dans ce travail de thèse.

3.3.3.1 *Approche univariée*

Il est possible d'étudier indépendamment chaque biomarqueur en évaluant son association avec l'évènement de santé étudié. Cela consiste donc à effectuer un nombre d'analyses statistiques égal au nombre de biomarqueurs à tester (p). Il s'agit là d'une situation de tests multiples qui rend cette approche peu puissante. L'approche

classique consiste à contrôler le nombre de faux positifs en contrôlant la probabilité globale de tout résultat faussement positif, c'est à dire en diminuant très fortement le seuil de signification (α) du test. Cette approche, est souvent utilisée dans le cadre des données en grande dimension, mais limite la puissance statistique des tests effectués. L'approche la plus simple et la plus utilisée en génomique pour le contrôle des faux positifs est la correction de Bonferroni, où le seuil de signification est ajusté en le divisant par le nombre de tests effectués. Il permet de contrôler la probabilité (test conservateur) que l'un des très nombreux variants génomiques inclus dans le modèle soit faussement identifié comme étant associé à l'évènement de santé étudié (Rahmenführer et al., 2023).

3.3.3.2 Approches pénalisées

La régularisation, ou régression pénalisée, est un modèle de régression qui introduit un terme de pénalisation noté λ dans l'écriture de la vraisemblance du modèle. Cela permet de minimiser l'effet de certaines variables sans les supprimer complètement de l'équation. Ce terme ne peut prendre que des valeurs positives ou nulles. La vraisemblance pénalisée s'écrit alors :

$$L_p(\lambda, \beta, X) = L(\beta, X) - p(\lambda, \beta) \quad (9)$$

Ces méthodes constituent une stratégie appropriée permettant de choisir des indicateurs essentiels à partir d'un ensemble de données beaucoup plus important. Elles permettent d'équilibrer le compromis biais-variance tout en évitant de sur-ajuster le modèle aux données. Ainsi, le modèle réduit l'ampleur des coefficients estimés sur les données d'apprentissage (c'est-à-dire en ajoutant un biais et en réduisant la précision des données d'apprentissage), ce qui entraîne moins de variations dans les prédictions lors de l'estimation du même modèle dans les données de test (réduction de la variation et augmentation de la généralisation). La régression pénalisée cherche donc à contraindre les coefficients de régression β à tendre vers zéro, de sorte que plus la pénalisation est forte, plus la contrainte est forte (Jaafar & Ismail, 2022). Les trois méthodes de pénalisation présentées ci-dessous diffèrent selon le terme de

pénalisation.

La régression pénalisée Lasso, proposée par Tibshirani (1996) est un acronyme pour Least Absolute Shrinkage and Selection Operator. Il s'agit d'une méthode souvent utilisée dans le contexte des données en grande dimension pour sa précision statistique pour la prédiction et la sélection des variables, associée à sa faisabilité informatique. Dans ce travail de thèse, nous nous sommes concentrés sur l'utilisation de la pénalisation Lasso pour effectuer la sélection de biomarqueurs. Ce type de régression est aussi appelé régularisation L_1 et est notée comme suit :

$$p(\lambda, \beta) = \lambda \sum_{j=1}^p |\beta_j| \quad (10)$$

Cette méthode permet de contraindre les coefficients de régression des biomarqueurs les moins pertinents en les faisant tendre vers zéro jusqu'à atteindre une valeur exacte de zéro lorsque des valeurs élevées sont prises par le paramètre de pénalisation λ . Ainsi, lorsque λ tend vers l'infini, toutes les variables sont exclues du modèle, et si $\lambda = 0$ l'estimateur Lasso correspond à l'estimateur du maximum de vraisemblance du modèle de régression utilisé (Tibshirani, 1996). La pénalisation Lasso permet donc de faire de la sélection de variables de sorte à ne sélectionner que les biomarqueurs dont les coefficients de régression n'ont pas été contraints à zéro. Le paramètre de pénalisation optimal (λ_{opt}) est habituellement choisi par validation croisée (k-fold, avec $k = 10$) afin de minimiser le risque d'erreur quadratique (Belhechmi et al., 2020; Bühlmann & Van De Geer, 2011). La validation-croisée consiste à diviser l'ensemble des données en plusieurs parties (appelées "folds"). Ensuite, sur une séquence de valeurs de paramètres λ , des modèles pénalisés sont appliqués à toutes les sous-parties sauf une. La performance prédictive de chaque modèle est évaluée sur la sous-partie mise de côté. Ce processus est répété jusqu'à ce que chaque sous-partie ait été exclue au moins une fois. La valeur du paramètre d'ajustement est alors, en général, choisie pour être la valeur de la séquence qui présente la plus petite erreur de validation croisée. Par conséquent, le choix du λ_{opt} dépend donc de la séparation au hasard des sous-

parties, ce qui implique souvent des sélections de faux positifs (S. Roberts & Nowak, 2014). Il est donc nécessaire de stabiliser la sélection des biomarqueurs de sorte à avoir plus de chances de sélectionner les biomarqueurs réellement pertinents (Bühlmann & Van De Geer, 2011). Différentes méthodes de stabilisation sont ainsi présentées dans la section 3.3.5. Par ailleurs, en présence de corrélations importantes entre les variables, comme c'est souvent le cas dans les études génomiques, une limite importante du Lasso pour la sélection de variables est le fait que l'algorithme a tendance à ne sélectionner arbitrairement qu'un seul biomarqueur parmi l'ensemble des biomarqueurs corrélés.

La régression pénalisée Ridge introduite par Hoerl & Kennard (1970), aussi appelée régularisation L_2 , pénalise le carré de la magnitude des coefficients et non la valeur absolue des coefficients, et s'écrit donc :

$$p(\lambda, \beta) = \lambda \sum_{j=1}^p \beta_j^2 \quad (11)$$

Cette pénalisation ne permet pas d'éliminer les biomarqueurs non pertinents mais minimise leur impact sur le modèle contraignant les coefficients de régression en les faisant tendre vers zéro (mais n'atteignent jamais strictement 0) au fur et à mesure que la valeur de λ augmente. Ainsi, la pénalisation Ridge ne permet pas de faire de la sélection de variables.

La régression pénalisée Elastic-net, est une méthode qui utilise une combinaison pondérée des méthodes de régularisation Lasso (L_1) et Ridge (L_2). Elle a été introduite par Zou & Hastie (2005), et le terme de pénalisation est notée :

$$p(\lambda_1, \lambda_2, \beta) = \lambda_1 \sum_{j=1}^p |\beta_j| + \lambda_2 \sum_{j=1}^p \beta_j^2 \quad (12)$$

La pénalisation Elastic-net combine les avantages offerts par les deux méthodes Lasso et Ridge, et permet donc de faire de la sélection de variables. En comparaison à la régression Lasso, l'ajout de la pénalisation Ridge permet de sélectionner plus de

variables tout en étant plus performant en termes de prédiction, particulièrement lorsque les variables sont très corrélées entre elles.

3.3.4 La régression logistique conditionnelle pénalisée Lasso

Comme écrit dans la section 3.3.1, le design cas-témoins appariée est souvent utilisée en épidémiologie, surtout dans le contexte où le nombre de biomarqueurs mesuré est très élevé, et qu'on se trouve dans un contexte de données en grande dimension (Qian et al., 2014; Xu et al., 2019). Ainsi, il est nécessaire de prendre en compte le design cas-témoins appariés de l'étude tout en utilisant des méthodes adaptées aux données en grande dimension.

Avalos et al. (2015), proposent dans un premier temps d'utiliser la similarité entre la vraisemblance du modèle logistique conditionnel et la vraisemblance partielle du modèle de Cox stratifié et à temps discret, et d'effectuer un modèle de Cox stratifié pénalisé Lasso. Ainsi, les algorithmes proposés pour résoudre le Lasso pour le modèle de Cox permettant la stratification peuvent être utilisés comme le package *glmnet* ou *penalized* sur le logiciel R. Dans un second temps, les auteurs proposent aussi d'appliquer une régression logistique non-conditionnelle pénalisée dans le cas particulier où chaque strate d'appariement est composé d'un seul cas et d'un seul témoin (1 :1), ainsi le package *penalized* permet d'effectuer cette régression logistique avec la variable événement d'intérêt fixée à 1, mais ce n'est pas le cas du package *glmnet* qui ne converge pas lors que la réponse de la variable événement est constante.

Selon Reid & Tibshirani, (2014), l'utilisation du package *penalized* pour effectuer un modèle de Cox stratifié pénalisé Lasso a des limites importantes, ne permettant pas d'obtenir des estimations à partir de la vraisemblance exacte pour un appariement prenant en compte plusieurs témoins (1:M, $M > 1$). Ils proposent donc l'utilisation du package *clogitL1* qui semble fonctionner plus efficacement que le package *penalized* même pour un appariement (1 :1), présentant un temps de calcul inférieur.

Nous avons testé les deux packages cités ci-dessus pour implémenter la régression

logistique conditionnelle pénalisée Lasso, et nous avons choisi de travailler avec le package *clogitL1* (Reid & Tibshirani, 2014).

3.3.5 Les approches de stabilisation de la sélection de biomarqueurs

Comme décrit dans la section 3.3.3.2, le choix du λ_{opt} par validation croisée rend la sélection des biomarqueurs instable, et donc la construction du modèle. Ce qui nécessite l'utilisation d'approches permettant d'obtenir une sélection stable de variables. Trois approches sont présentées ci-dessous. Elles sont basées sur le sous-échantillonnage ou le bootstrapping intégrés aux algorithmes de grande dimension pour améliorer la stabilité de ceux-ci. L'expression « sélection stable de variables » laisse entendre que les variables sélectionnées sont incluses dans une grande partie des modèles sélectionnés par la régression Lasso appliqué aux données sous-échantillonnées. Ces approches ont été adaptées à la régression logistique conditionnelle pénalisée Lasso.

3.3.5.1 L'approche Percentile Lasso

L'approche Percentile Lasso consiste à lancer m fois une régression logistique conditionnelle pénalisée Lasso sur la base initiale afin d'estimer pour chaque régression un paramètre de pénalisation λ (**Figure 7**).

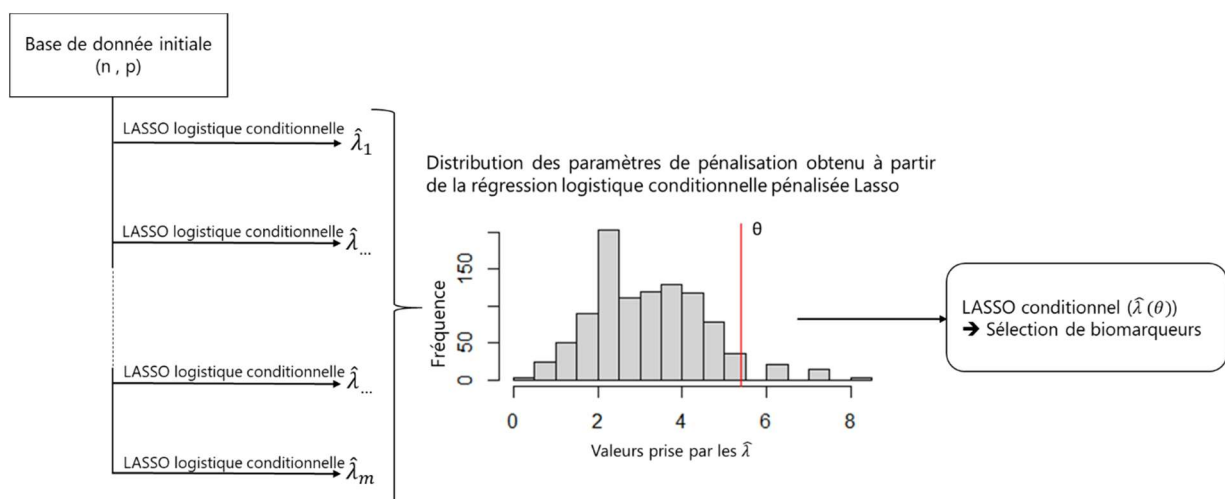


Figure 7. Représentation de l'algorithme Percentile Lasso adapté à la régression logistique conditionnelle pénalisée Lasso

Ainsi, un ensemble de m paramètres de régularisation optimaux noté $\Lambda(m) = (\lambda_1, \dots, \lambda_m)$ est récupéré. A partir de cette séquence de valeurs, un paramètre noté $\hat{\lambda}(\theta)$ et correspondant au $\theta^{\text{ième}}$ percentile de la distribution de l'ensemble $\Lambda(m)$ est choisi comme paramètre de pénalisation optimal, et une nouvelle régression pénalisée est effectuée sur le jeu de données initial. Ainsi pour une valeur fixe de θ , les variables sélectionnées par le modèle correspondent aux biomarqueurs dont les coefficients n'ont pas été rapportés à zéro pour le $\hat{\lambda}(\theta)$ choisi (S. Roberts & Nowak, 2014).

3.3.5.2 L'approche Bolasso

L'approche Bolasso qui est l'acronyme bootstrap-enhanced least absolute shrinkage operator, consiste à perturber le jeu de données original en créant m échantillons bootstrap (rééchantillonnage avec remise), sur lesquels est réalisée une régression pénalisée Lasso. Nous avons effectué le bootstrap sur les strates de sorte à toujours conserver le design apparié de l'étude (**Figure 8**). Un ensemble de biomarqueurs sélectionnés au moins une fois est obtenu, et chaque gène a sa fréquence de sélection pouvant aller de 1 à m (Bach, 2008).

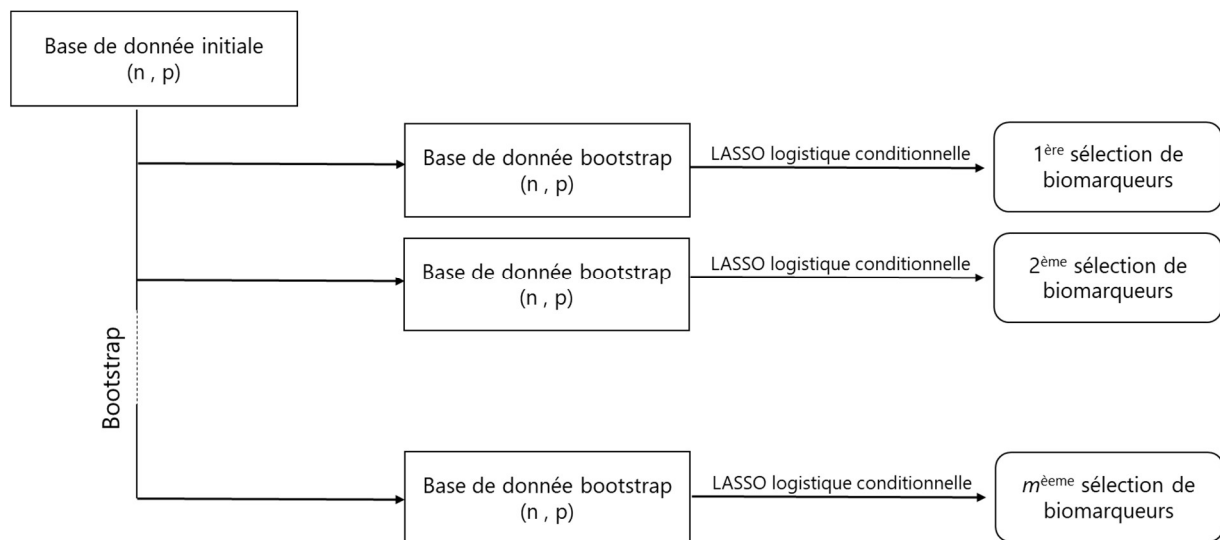


Figure 8. Représentation de l'algorithme Bolasso adapté à la régression logistique conditionnelle pénalisée Lasso

3.3.5.3 L'approche Sublasso

L'approche Sublasso, acronyme de *subsample-enhanced least absolute shrinkage operator*, consiste à créer m sous-échantillons du jeu de données d'origine sur lesquels la régression logistique conditionnelle pénalisée Lasso est appliquée (**Figure 9**). Comme pour l'approche Bolasso un ensemble de biomarqueurs sélectionnés au moins une fois est obtenu, et chaque gène a sa fréquence de sélection pouvant aller de 1 à m (Meinshausen & Bühlmann, 2009).

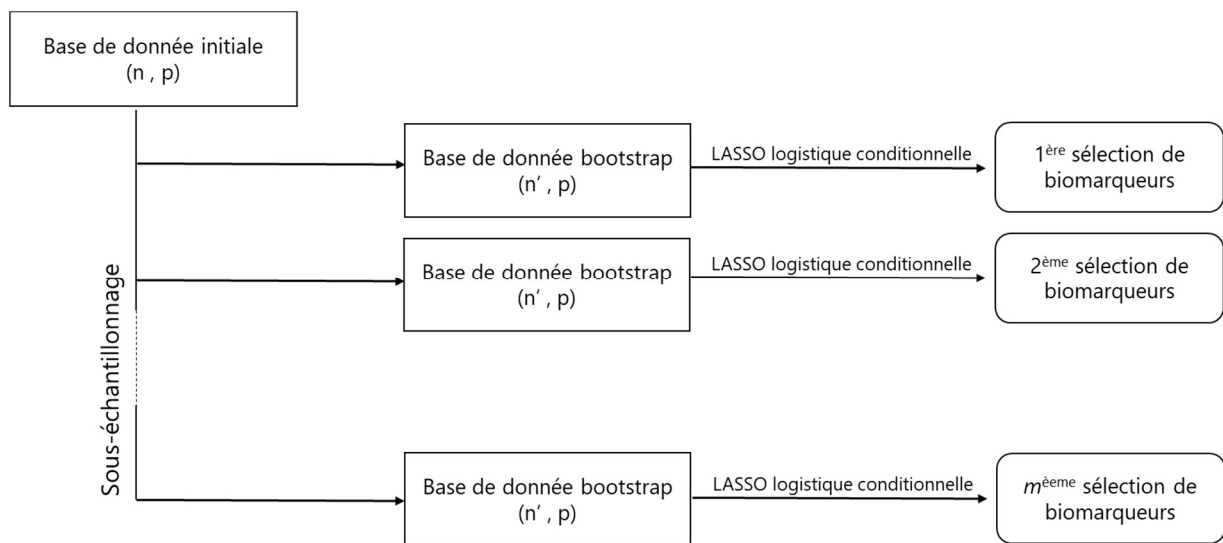


Figure 9. Représentation de l'algorithme Sublasso adapté à la régression logistique conditionnelle pénalisée Lasso

3.3.6 Evaluation de la performance discriminante d'un modèle estimé dans une étude cas-témoins appariée

Il est courant d'évaluer la performance discriminante d'un modèle de régression logistique non conditionnelle par l'aire sous la courbe ROC, c'est-à-dire l'AUC.

Dans une étude cas-témoins appariée, il est nécessaire que l'analyse statistique tienne compte de l'appariement comme l'utilisation régression logistique conditionnelle pour l'estimation de l'association entre des prédicteurs et l'évènement de santé étudié. Dans le cadre d'un modèle de régression logistique conditionnelle, différentes approches sont possibles pour estimer cette métrique. Elles diffèrent par leur façon d'estimer le prédicteur linéaire.

Reprenons les notations décrites dans la section 3.3.2, avec Y dénotant l'évènement de santé (0 -1), $X = (X_1, \dots, X_p)$ les variables pour lesquelles nous cherchons l'association avec Y , et $X_m = (X_{m1}, \dots, X_{mj})$ les variables utilisées pour l'appariement des cas aux témoins.

L'approche 1 est basée sur une régression logistique non conditionnelle incluant les variables du vecteur X , ajustées sur les variables d'appariement (X_m). Les coefficients de régression du modèle sont estimés tel que (β_0) correspond à l'ordonnée à l'origine, β correspond aux covariables X et γ aux variables d'appariement X_m . L'AUC est ensuite calculée à l'aide du prédicteur linéaire estimé : $\widehat{\beta}_0 + \widehat{\beta}X + \widehat{\gamma}X_m$.

L'approche 2 repose sur une régression logistique conditionnelle incluant les variables du vecteur X . Seuls les coefficients β estimés sur X sont obtenus ($\widehat{\beta}_0 = \widehat{\gamma} = 0$). L'AUC est ensuite estimée à l'aide du prédicteur linéaire $\widehat{\beta}X$.

Deux autres approches (approches 3 et 4) ont été développées par Qian et al. (2014) pour l'estimation de l'AUC dans des études cas-témoins appariées tout en prenant en compte l'aspect niché du design. La cohorte dans laquelle est nichée notre échantillon apparié sera dénommé cohorte parent. Les auteurs proposent tout d'abord d'estimer un offset noté $f(X_m)$ à partir des données de la cohorte parent. Cet offset est calculé à partir d'une régression logistique prenant Y comme évènement d'intérêt et $X_m = (X_{m1}, \dots, X_{mj})$ comme prédicteurs comme suit :

$$f(X_m) = \log \frac{\mathbb{P}(Y = 0|X_m)}{\mathbb{P}(Y = 1|X_m)} + \log(m) \quad (13)$$

L'approche 3 consiste à estimer les coefficients β_0 , β et γ se fait en deux étapes (d'où son nom *Two-stage procedure*). Le coefficient β est estimé en premier à partir d'une régression logistique conditionnelle appliquée sur l'échantillon cas-témoins apparié, qui prend Y comme évènement d'intérêt et X comme prédicteurs. β_0 et γ sont quant à eux estimés dans une seconde étape par le biais d'une régression logistique non-conditionnelle appliquée sur l'échantillon cas-témoins apparié. Ce modèle prend comme évènement d'intérêt Y , et X_m comme prédicteurs, tout en incluant l'offset estimé, et les prédicteurs linéaires $\widehat{\beta}X$ estimés dans l'étape précédente comme offset

du modèle : $f(X_m) + \hat{\beta} X$. L'AUC est finalement estimé à partir des prédicteurs linéaires obtenus dans ce modèle.

L'approche 4 revient à estimer les coefficients β_0 , β et γ en une seule étape de façon simultanée (*Simutaneous procedure*), à partir d'une régression logistique appliquée à l'échantillon cas-témoins apparié, incluant l'offset auparavant estimé, selon le modèle suivant :

$$\text{logit}(\mathbb{P}(Y = 1|X, X_m)) = \beta_0 + \beta^T X + \gamma^T X_m \quad (14)$$

Comme pour les approches précédentes, l'estimation des AUC se fait à partir des prédicteurs linéaires obtenus dans le modèle.

Pour associer une précision à l'AUC estimé par l'une des approches décrite précédemment, un intervalle de confiance est estimé. Il est estimé en générant de nombreux échantillons bootstrap (sur les strates de sorte à toujours conserver le design apparié de l'étude) à partir du jeu de données d'origine. À partir de la distribution des AUC estimés à partir des échantillons bootstrap, les percentiles 0,025 et 0,975 sont utilisés comme limites inférieures et supérieures de l'intervalle de confiance à 95 % (95 % IC).

3.3.6.1 Calcul d'AUC pondérée

Les méthodes citées précédemment peuvent présenter d'important biais dans l'estimation directe de l'AUC dans le cas d'études cas-témoins appariées nichées dans une cohorte parent (comportant plus de sujets). Cela est dû au fait que l'ensemble des témoins sélectionnés peut ne pas être représentatif de la population parent de témoins. Ce biais est d'autant plus important si les biomarqueurs sont fortement corrélés aux variables d'appariement. Xu et al., (2019) proposent donc une approche permettant d'estimer un AUC pondéré sur l'inverse du poids estimé dans chaque strate d'appariement. Cette approche par pondération permet de reproduire l'appariement sur l'ensemble de la cohorte parents puis d'y récupérer la proportion relative de cas et de témoins spécifique à chaque strate. Les poids calculés permettent de rééquilibrer la

proportion pondérée de cas et de témoins dans chaque strate d'une étude appariée, pour qu'elle soit équivalente à celle de la cohorte d'origine. Soit n_{s0} le nombre de témoins n_{s1} le nombre de cas dans la $s^{\text{ième}}$ strate d'appariement de la cohorte parent. Dans l'étude cas-témoins appariée (1 : m), on note j_s et m_j_s le nombre de cas et de témoins sélectionnés dans la $s^{\text{ième}}$ strate. Les poids proposés pour chaque cas et chaque témoin dans la $s^{\text{ième}}$ strate sont calculés ainsi :

$$w_{s,case} = \frac{n_{s1}}{j_s} ; w_{s,témoins} = \frac{n_{s0}}{mj_s} \quad (15)$$

Cette pondération est ensuite appliquée aux quatre approches présentées dans la section précédente.

3.3.6.2 Calcul de l'AUC avec correction des biais

Un modèle dont la capacité discriminante n'est valable que dans le jeu de données dans lequel il a été estimé ne garantit pas forcément sa capacité à bien discriminer les patients présentant ou non un évènement sur un jeu de données externes. Il est donc essentiel de disposer d'estimations non biaisées de l'AUC pour prendre en compte un possible sur ajustement (ou overfitting en anglais). Une estimation de l'AUC dit corrigé (*biais-corrected AUC* en anglais) peut être obtenue pour l'ensemble des approches décrites précédemment, et ce à travers trois méthodes. Ces méthodes permettent d'estimer le biais dû au sur-ajustement appelé aussi "optimisme", dans l'ajustement final du modèle. Une fois l'optimisme estimé, il peut être soustrait de l'AUC dérivé de l'échantillon original pour obtenir une estimation de l'AUC avec correction du biais ou de l'overfitting (Harrell , 2015).

Le bootstrap permet d'obtenir des estimations quasi non-biaisées des performances discriminantes (AUC) d'un modèle. À partir du jeu de données original de taille n , un échantillon bootstrap de taille n est construit. Le modèle est ensuite développé sur l'échantillon bootstrap sur lequel l' $AUC_{\text{Bootstrap}}$ est estimé, puis ce même modèle est appliqué sans modification à l'échantillon original pour obtenir l' AUC_{Apparent} . L'optimisme est ensuite estimé à partir de la différence entre $AUC_{\text{Bootstrap}}$ et AUC_{Apparent} . Ce processus

est répété pour un nombre B de répliques bootstrap afin d'obtenir un optimisme moyen qui est soustrait de l'AUC obtenue à partir du modèle estimé sur l'échantillon d'origine afin d'obtenir l'estimation de l'AUC corrigé. L'intervalle de confiance à 95 % de l'AUC corrigé est obtenu en appliquant les percentiles 0.025 et 0.975 à la distribution des AUC corrigés obtenus pour chaque échantillon bootstrap.

La méthode par échantillonnage en deux parties repose sur un principe de base identique à la méthode par bootstrap, mais au lieu de travailler sur des échantillons bootstraps, cette méthode consiste à diviser de manière aléatoire un jeu de données en deux groupes, un échantillon d'entraînement et un échantillon de validation. Les divisions typiques sont 50 % - 50 % ou 2/3 - 1/3. Le modèle est construit sur le jeu de données d'entraînement puis l'AUC est estimé à la fois sur le jeu de données d'entraînement et sur le jeu de données de validation. L'optimisme résulte alors de la différence entre les AUC obtenus à partir de ces deux échantillons. La procédure est répétée B fois, et l'AUC corrigé correspond à la différence entre l'AUC original et l'optimisme estimé (Steyerberg, 2009).

La validation croisée est une extension de la validation par échantillonnage en deux parties, où l'ensemble des données est divisé en un nombre k de groupes (appelés folds) contenant le même nombre de sujets. Un modèle est estimé sur $k - 1$ fold, puis testé sur une partie aléatoire qui a été exclue de l'échantillon. Ce processus est répété k fois pour des fractions consécutives de patients. De cette manière, tous les patients ont servi une fois à tester le modèle. L'AUC est estimé comme étant la moyenne de toutes les évaluations obtenue. L'optimisme correspond alors à la différence de l'AUC obtenu sur l'échantillon d'entraînement ($k - 1$ sous-parties) et l'AUC obtenu sur l'échantillon de validation (dernier fold restant). La procédure est répétée B fois, et la différence entre l'AUC original et l'optimisme estimé correspond à l'AUC corrigé (Steyerberg, 2009). La validation croisée la plus extrême consiste à exclure chaque patient une fois, ce qui équivaut à la méthode jack-knife.

3.3.7 Validation d'un modèle

Lorsqu'un modèle est développé au sein d'un échantillon représentatif de patients d'une population spécifique, cette dernière présente des caractéristiques spécifiques, Il est donc nécessaire de déterminer la « reproductibilité » du modèle pour cette population (validation interne), et sa « généralisabilité » à d'autres populations (validation externe) (Steyerberg, 2009).

3.3.7.1 *La validation interne*

La validation interne consiste à évaluer la qualité d'un modèle développé dans un échantillon de données, sur un échantillon plus ou moins similaire. Il existe différentes techniques de validation interne, dont la validation par échantillonnage en deux parties (utilisée dans le Chapitre 6 de la thèse).

3.3.7.2 *La validation externe*

La validation externe consiste à tester un modèle développé sur un échantillon de données sur un échantillon de données externe. Ces deux échantillons sont totalement indépendants et proviennent d'environnements différents, mais plausiblement liés. Plus le modèle est validé sur des échantillons externes et plus les populations sont diversifiées, plus nous sommes confiants dans la capacité de généralisation du modèle. Cette méthode est plus rigoureuse que la validation interne, mais suppose évidemment d'avoir à disposition des données externes ce qui n'est pas toujours envisageable.

Chapitre 4. RISQUE DE PATHOLOGIES ET D'HOSPITALISATION CARDIAQUE CHEZ LES SURVIVANTS D'UN CANCER DE L'ENFANT

L'objectif de ce chapitre est d'identifier et d'évaluer les facteurs de risques associés aux pathologies cardiaques et aux hospitalisations cardiaques dans la cohorte FCCSS. Un premier article sur le risque de pathologies cardiaques est en cours de préparation, et un second, réalisé en collaboration avec les membres de l'équipe Epidémiologie des Radiations, se focalisant uniquement sur les hospitalisations est en cours de soumission.

4.1 INTRODUCTION

Les pathologies cardiaques sont la cause non cancéreuse la plus fréquente de décès chez les survivants d'un cancer dans l'enfance, et la mortalité due à des causes cardiaques reste plus élevée chez ces patients en comparaison à la population générale (Bansal et al., 2017; Bottinor et al., 2024; Mulrooney et al., 2009; Tukenova et al., 2010). L'exposition aux anthracyclines et l'irradiation thoracique augmentent les risques de ces pathologies, mais avec des effets différents sur la fréquence et le type de pathologie cardiaque (Armenian et al., 2018; Bansal et al., 2017; Bates et al., 2019). Les anthracyclines sont associées à une toxicité cardiaque se traduisant par un dysfonctionnement du ventricule gauche (VG). Par ailleurs, l'effet à long terme de l'irradiation est associé à diverses pathologies cardiaques, qui se manifestent souvent par la valvulopathie, l'ischémie, ou l'insuffisance cardiaque. De plus, l'exposition du cœur à des doses moyennes de radiations supérieures à 15 Gy augmente le risque de développer des pathologies cardiaques, même lorsque de faibles volumes sont exposés. Cependant, les conséquences de l'irradiation d'un volume important du cœur à des doses plus faibles doivent être davantage étudiées (Bates et al., 2019; Chounta et al., 2023; Mansouri et al., 2019). Il est alors nécessaire d'analyser la relation entre les doses de rayonnement reçues à différents volumes du cœur et le risque de pathologies

cardiaques chez les survivants d'un cancer de l'enfant.

Le risque de pathologies cardiaques chez les survivants d'un cancer de l'enfance a été étudié à différents niveaux de sévérité, conformément aux grades de la CTCAE. Certaines recherches ont évalué ce risque en considérant tous les grades de sévérité, tandis que d'autres se sont concentrées uniquement sur les pathologies de grade 2 ou 3 et plus. Par ailleurs, certaines études ont spécifiquement évalué le risque d'hospitalisation pour des causes iatrogènes, équivalentes au grade 4 de la CTCAE (Reulen et al., 2010). L'accès aux données médico-administratives, notamment les demandes de remboursement, a facilité l'étude de ces risques d'hospitalisation à long terme liés aux traitements anticancéreux administrés aux enfants survivants d'un cancer. Ces résultats sont donc primordiaux pour la mise en œuvre de stratégies de suivi à long terme ciblées pour les patients présentant un risque élevé d'évènements cardiaques (Armenian et al., 2015; Robison & Hudson, 2014).

Ce chapitre est structuré autour de deux objectifs principaux. Le premier objectif est d'étudier les conséquences de l'exposition aux anthracyclines et aux radiations cardiaques, en mettant l'accent sur les effets des doses moyennes de radiation au cœur inférieures à 15 Gy, pour la survenue de la première pathologie cardiaque et plus spécifiquement l'insuffisance cardiaque. Pour cette dernière pathologie, on distinguera la prise en compte ou non des autres évènements cardiaques survenant avant celle-ci. Le second objectif vise à évaluer les risques d'hospitalisation liés aux pathologies cardiaques chez ces patients.

4.2 RISQUES DE PATHOLOGIES CARDIAQUES ET D'INSUFFISANCE CARDIAQUE CHEZ LES SURVIVANTS D'UN CANCER DE L'ENFANT

4.2.1 Population d'étude

Cette étude a été menée dans la cohorte FCCSS, qui comprend 7670 survivants à 5 ans de cancers de l'enfant pour tout type de tumeur maligne, à l'exception de la leucémie (cohorte décrite dans le Chapitre 2). L'identification des pathologies cardiaques est

également détaillée dans le Chapitre 2

Parmi les survivants de la cohorte FCCSS, 43 ont développé une première pathologie cardiaque dans les 5 ans suivant le diagnostic du cancer et 10 ont été perdus de vue durant cette période. Ces patients ont donc été exclus des analyses, réduisant ainsi le nombre total de sujets à 7617. Parmi les sujets inclus dans cette étude, 449 ont développé une première pathologie cardiaque tous grades, et 313 ont développé une première pathologie cardiaque de grade 3 ou plus. De plus, 1166 sont décédés d'autres causes (**Figure 10**). Concernant l'insuffisance cardiaque, 29 survivants ayant développé cette pathologie avant la période des 5 ans suivant le diagnostic du cancer ainsi que les 10 patients perdus de vue durant cette période ont été exclus des analyses. Les analyses ont donc été menées sur 7631 survivants, parmi lesquels 236 ont développé une première insuffisance cardiaque tous grades, et 1224 sont décédés sans avoir développé d'insuffisance cardiaque (**Figure 11**).

Concernant l'exposition aux traitements anticancéreux pendant l'enfance, les doses de radiations reçues au cœur ainsi que les doses de chimiothérapies, notamment les anthracyclines, ont été étudiées. Pour chaque sujet inclus, les doses d'irradiation reçues au cœur ont été estimées rétrospectivement et les mesures dose-volume ont été reconstituées. Parmi les métriques dose-volume, le pourcentage du volume cardiaque recevant au moins une certaine dose a été calculé. Dans cette étude nous avons considéré les métriques suivantes :

- V5 : le pourcentage du volume cardiaque recevant au moins 5 Gy,
- V15 : le pourcentage du volume cardiaque recevant au moins 15 Gy,
- V20 : le pourcentage du volume cardiaque recevant au moins 20 Gy,
- V30 : le pourcentage du volume cardiaque recevant au moins 30 Gy.

Pour se concentrer spécifiquement sur les survivants n'ayant reçu qu'une dose faible ou modérée de radiations au cœur, deux mesures supplémentaires ont été utilisées :

V01_5 : cette mesure représente le pourcentage du volume cardiaque recevant au moins 0,1 Gy (V01) chez les survivants dont la dose maximale au cœur était inférieure

à 5 Gy. En d'autres termes, c'est le pourcentage du volume cardiaque recevant une dose comprise entre 0,1 et 5 Gy exclu. Ceci permet d'évaluer les effets des très faibles doses de radiation.

V5_20 : Cette mesure représente le pourcentage du volume cardiaque recevant au moins 5 Gy (V5) chez les survivants dont la dose maximale au cœur était inférieure à 20 Gy. En d'autres termes, c'est le pourcentage du volume cardiaque recevant une dose comprise entre 5 et 20 Gy exclu. Ceci permet d'évaluer les effets des doses modérées de radiation.

La majorité des patients traités par anthracyclines dans la FCCSS ont reçu de la doxorubicine, et la dose cumulée d'anthracycline a été estimées en utilisant la dose cumulée équivalente de doxorubicine (Feijen et al., 2019) pour l'ensemble des survivants.

Figure 10. Diagramme de flux des patients inclus dans les analyses pour l'évènement première pathologie cardiaque tous grades et grades ≥ 3

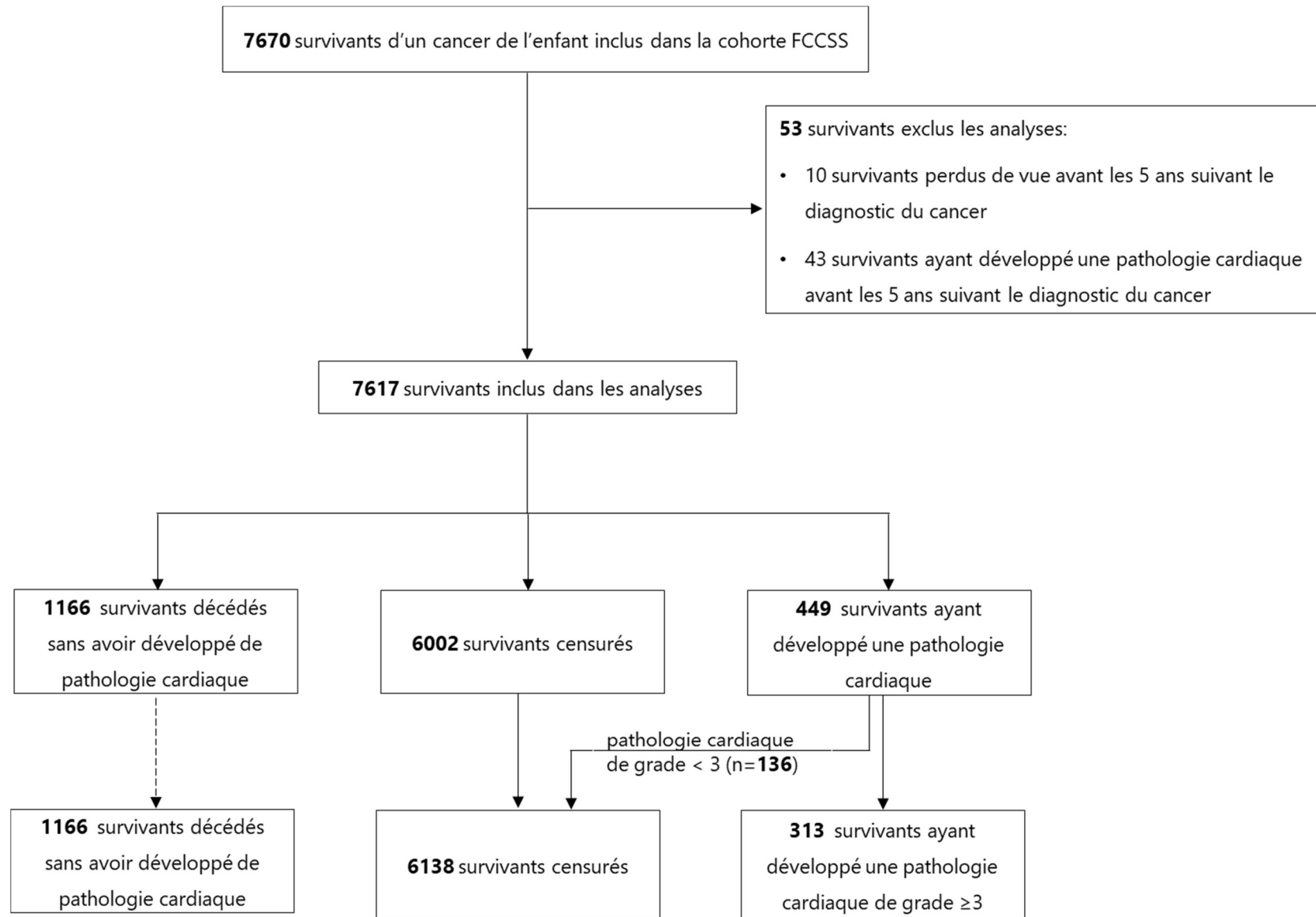
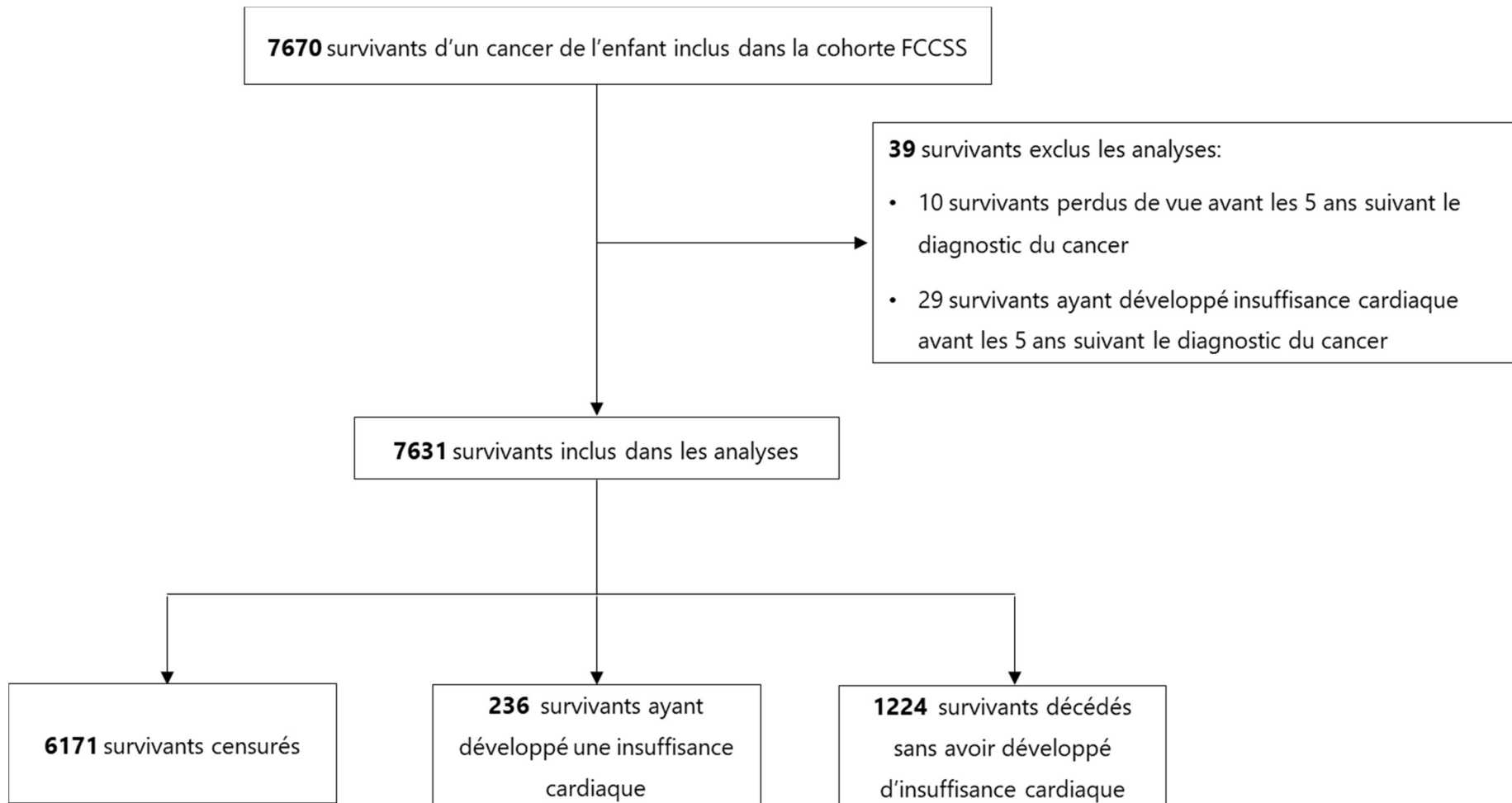


Figure 11. Diagramme de flux de l'inclusion des patients dans les analyses pour l'évènement première insuffisance cardiaque



4.2.2 Méthode statistique

L'incidence cumulée des pathologies cardiaques a été estimée pour l'évènement première occurrence de pathologie cardiaque en tant que critère composite (toute pathologie cardiaque confondue). Les décès non cardiaques ont été traités comme des risques concurrents (**Figure 12**) selon la méthode Aalen-Johansen (Aalen, 1978).

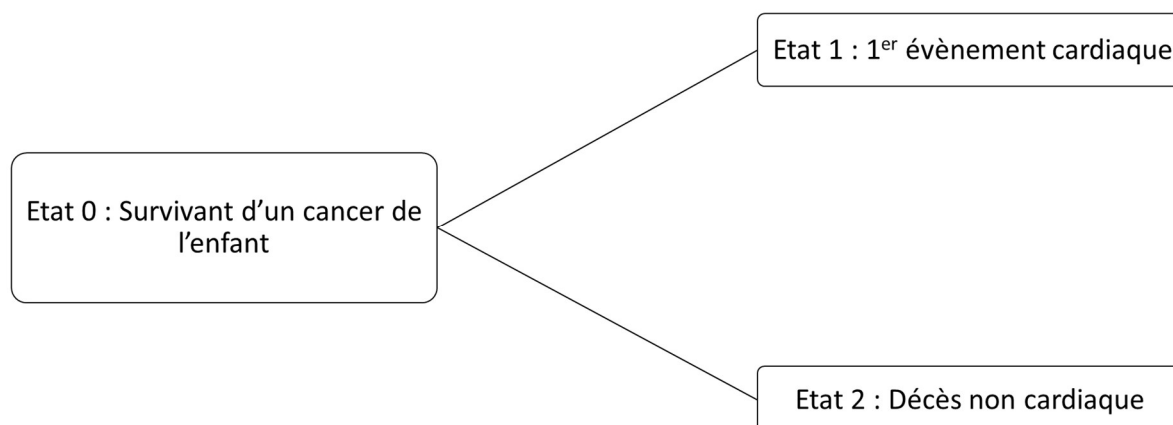


Figure 12. *Modèle à risque concurrent utilisé pour l'étude de l'estimation du risque de premier évènement cardiaque dans la FCCSS.*

Le suivi des survivants a été utilisé comme échelle de temps, et calculé à partir de la date de diagnostic du premier cancer jusqu'à (1) la date de diagnostic du premier évènement cardiaque, (2) la date de décès non cardiaque, et (3) la date des dernières nouvelles médicales, en fonction de ce qui se produit en premier. L'évaluation des effets des traitements cardiotoxiques (doses cumulées d'anthracyclines, doses moyennes de radiations au cœur, et mesures dose-volume) sur le taux d'incidence d'une première pathologie cardiaque s'est faite en utilisant un modèle de Cox cause spécifique (voir Chapitre 3 pour plus de détails). Les modèles ont été ajustés sur l'année calendaire et l'âge au moment du diagnostic de cancer, ainsi que sur le sexe. Les interactions entre les doses moyennes de radiations reçues au cœur et l'administration d'anthracyclines ont aussi été recherchées. Les mêmes analyses ont été reproduites en considérant comme évènement d'intérêt le premier évènement cardiaque de grade 3 à 5.

Pour l'évènement correspondant à la première occurrence d'insuffisance cardiaque, des analyses similaires ont été effectuées. De plus, le risque associé au développement antérieur d'une pathologie cardiaque différente (valvulopathie, ischémie, péricardite, et trouble du rythme et de la conduction) a été analysé en considérant ces variables comme étant dépendantes du temps.

Les doses de traitements reçues ont été catégorisées en se basant sur les données de la littérature. Les doses cumulées d'anthracyclines en mg/m^2 ont été catégorisées comme suit : pas d'anthracyclines,]0-250],]250-360], > 360 (Haddy et al., 2016; Mansouri et al., 2019). Concernant les variables de doses de radiothérapie et de mesures dose-volume, 174 sujets avaient des données manquantes. Pour éviter toute perte d'information sur les effets des autres variables, nous avons choisi de créer une classe « données manquantes » supplémentaire dans nos analyses. Pour les doses moyennes de radiations au cœur en Gy, même si nous nous sommes basé sur la littérature (Chounta et al., 2023; de Baat et al., 2023; Haddy et al., 2016), des analyses de sensibilités ont été effectuées pour le choix des catégories de doses qui ont été construites comme suit : pas de radiations au cœur,]0-1],]1-5],]5-15],]15-25],]25-35], > 35 Gy, données manquantes. Pour les volumes du cœur irradiés en pourcentages, nous avons utilisé les catégories suivantes : pas de radiothérapie, 0,]0-10],]10-50], > 50 %, données manquantes. Cependant, pour l'analyse d'interaction, il était nécessaire de supprimer ces patients des analyses car cette classe introduisait de l'instabilité dans le modèle.

4.2.3 Résultats

4.2.3.1 Description des pathologies cardiaques et insuffisance cardiaque

Les caractéristiques des patients avec une première pathologie cardiaque sont rapportées dans le **Tableau 4**. Pour la majorité des patients (32,5 %), le diagnostic du cancer de l'enfant s'est fait entre 1 et 5 ans. La durée médiane du suivi des patients était de 26,9 ans (EI, 19,6 – 35,2), et l'âge atteint jusqu'à la fin du suivi était en médiane de 37 ans (EI, 28 – 45). Nous avons observé 449 premières pathologies

cardiaques tous grades, dont 313 étaient de grade 3 à 5. Les types de pathologies cardiaques les plus fréquents étaient l'insuffisance cardiaque (49 %), la valvulopathie (24 %), les troubles du rythme et de la conduction (14 %), la péricardite (6 %), et l'ischémie (5 %). Parmi les patients ayant développé une première pathologie cardiaque, le type de tumeur maligne primaire le plus fréquent était le lymphome (35 %), suivi des néphroblastomes (17 %) et des neuroblastomes (13 %).

Bien que dans la cohorte entière seuls 34 % des patients ont reçu des anthracyclines, lorsqu'on s'intéresse aux patients ayant développé une pathologie cardiaque, la tendance s'inverse, avec 51 % traités par anthracyclines, dont 24 % ont reçu une dose cumulée d'anthracyclines de plus de 360 mg/m². Pour ce qui est de la radiothérapie incluant le cœur sur l'ensemble des patients inclus, la même proportion de patients (49 %) a reçu ou non des radiations cardiaques, alors que 77 % des patients ayant développé une pathologie cardiaque ont reçu des radiations au cœur.

Tableau 4. Caractéristiques démographiques et cliniques des survivants inclus dans les analyses et des survivants ayant développé une première pathologie cardiaque tous grades ou de grade 3 à 5, ou une insuffisance cardiaque dans la FCCSS

Caractéristiques	Pathologies cardiaques						Insuffisance cardiaque			
	Tous les survivants inclus (n=7617)		Pathologies cardiaques tous grades (n=452)		Pathologies cardiaques grade ≥3 (n=316)		Tous les survivants inclus (n=7631)		Insuffisance cardiaque tous grades (n=236)	
	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)
Genre										
Masculin	4182	(54,9)	240	(53,5)	178	(56,9)	4187	(54,9)	130	(55,1)
Féminin	3435	(45,1)	209	(46,6)	135	(43,1)	3444	(45,1)	106	(44,9)
Type de tumeur primaire										
Néphroblastome	1120	(14,7)	76	(16,9)	52	(16,6)	1121	(14,7)	43	(18,2)
Neuroblastome	1023	(13,4)	60	(13,4)	42	(13,4)	1025	(13,4)	38	(16,1)
Lymphome	1269	(16,7)	156	(34,7)	104	(33,2)	1275	(16,7)	70	(29,7)
Sarcome des tissus mous	817	(10,7)	49	(10,9)	33	(10,5)	817	(10,7)	29	(12,3)
Sarcome osseux	695	(9,1)	47	(10,5)	33	(10,5)	697	(9,1)	32	(13,6)
Tumeur du SNC	1132	(14,9)	14	(3,1)	12	(3,9)	1133	(14,8)	4	(1,7)
Tumeur des gonades	438	(5,7)	22	(4,9)	18	(5,7)	438	(5,7)	10	(4,2)
Tumeur de la thyroïde	70	(0,9)	4	(0,9)	1	(0,3)	70	(0,9)	,	
Rétinoblastome	617	(8,1)	4	(0,9)	3	(1,0)	617	(8,1)	1	(0,4)
Autres	436	(5,7)	17	(3,8)	15	(4,8)	438	(5,7)	9	(3,8)
Âge au diagnostic du premier cancer (années)										
Médiane (EI)	5,2	(1,7-11,2)	6,3	(2,8 -11,7)	6,3	(2,8-11,7)	5,3	(1,7-11,2)	5,2	(2,5-9,9)
<1	1240	(16,3)	36	(8,0)	24	(7,7)	1241	(16,3)	24	(10,2)
[1-5[2457	(32,3)	147	(32,7)	103	(32,9)	2459	(32,2)	91	(38,6)
[5-10[1669	(21,9)	117	(26,1)	85	(27,2)	1674	(21,9)	64	(27,1)
[10-15[1614	(21,2)	108	(24,1)	70	(22,2)	1618	(21,2)	45	(19,1)
>15	637	(8,4)	41	(9,1)	31	(9,9)	639	(8,4)	12	(5,1)
Année de diagnostic du premier cancer										
Médiane (min - max)	1986	(1946-2000)	1978	(1949 - 2000)	1978	(1950 - 2000)	1986	(1946-2000)	1978	(1951-1999)
Avant 1970	658	(8,6)	77	(17,1)	56	(17,9)	658	(8,6)	34	(14,4)
[1970-1980[1548	(20,3)	176	(39,2)	122	(39,0)	1548	(20,3)	101	(42,8)
[1980-1990[2460	(32,3)	151	(33,6)	102	(32,6)	2466	(32,3)	86	(36,4)

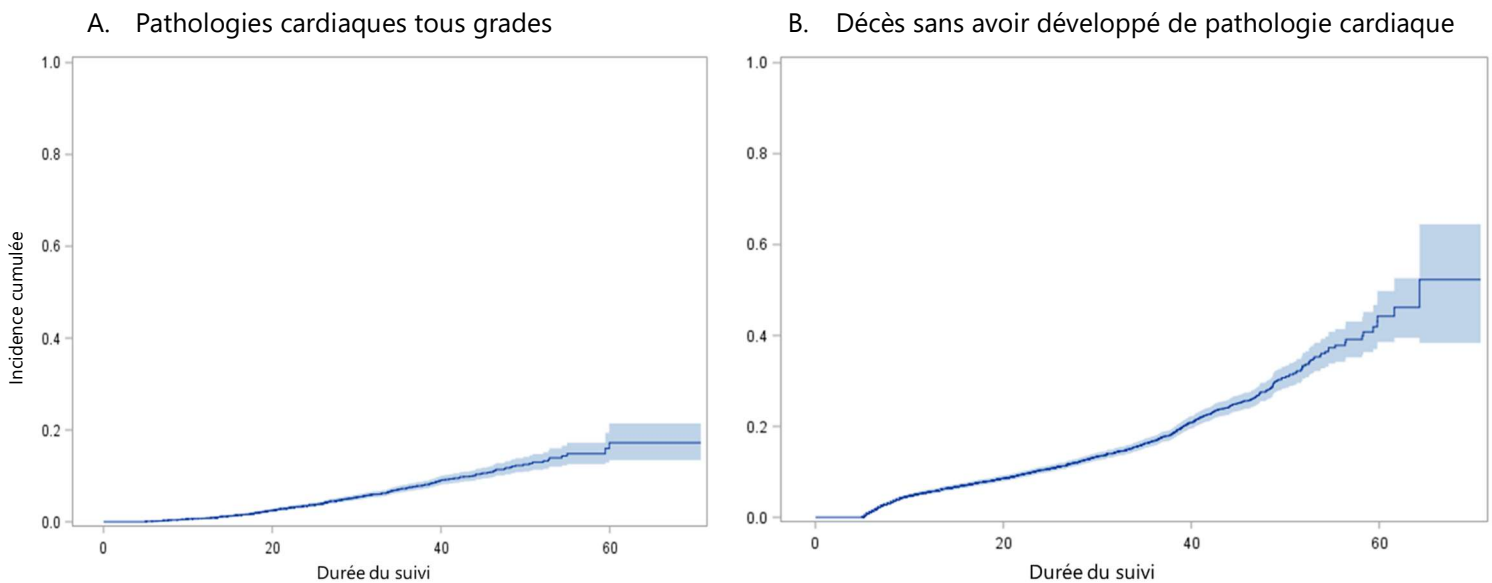
[1990 -2000]	2951	(38,7)	45	(10,0)	33	(10,5)	2959	(38,8)	15	(6,4)
Suivi après diagnostic du cancer (années)										
Médiane (EI)	26,9	(19,6 – 35,2)	22,9	(16,4 – 31,5)	23,6	(16,9 – 32,0)	27,1	(19,8 – 35,6)	21,4	(15,8-28,5)
[0-10[474	(6,2)	45	(10,0)	32	(10,2)	457	(6,0)	24	(10,2)
[10 - 20[1538	(20,2)	133	(29,6)	84	(26,8)	1503	(19,7)	80	(33,9)
[20 - 30[2573	(32,8)	141	(31,4)	101	(32,3)	2553	(33,5)	78	(33,1)
[30 - 40[1971	(25,9)	92	(20,5)	68	(21,7)	2014	(26,4)	40	(16,9)
≥ 40	1061	(13,9)	38	(8,5)	28	(8,9)	1104	(14,5)	14	(5,9)
Âge atteint à la fin du suivi (années)										
Médiane (EI)	37,0	(28,0-45,0)	31,6	(22,4 - 39,8)	32,1	(22,8 - 40,4)	37,4	(28,3-45,5)	28	(20,7-36,8)
[0 - 10[102	(1,3)	8	(1,8)	7	(2,2)	66	(5,4)	6	(2,5)
[10 - 20[489	(6,4)	70	(15,6)	47	(15,0)	321	(26,2)	45	(19,1)
[20 - 30[1708	(22,4)	131	(29,2)	84	(26,8)	328	(26,8)	82	(34,7)
[30- 40[2286	(30,0)	132	(29,4)	94	(30,0)	232	(19,0)	66	(28,0)
≥ 40	3032	(39,8)	108	(24,1)	81	(25,9)	277	(22,6)	37	(15,7)
Pathologie cardiaque										
Insuffisance cardiaque	218	(2,9)	218	(48,6)	178	(56,9)	236	(3,1)		
Valvulopathie	109	(1,4)	109	(24,3)	50	(16,0)	164	(2,1)	55	(23,4)
Ischémie	23	(0,3)	23	(5,1)	23	(7,3)	52	(0,7)	14	(5,9)
Troubles du rythme et de la conduction	63	(0,8)	63	(13,9)	39	(12,5)	117	(1,5)	38	(15,9)
Péricardite	25	(0,3)	25	(5,6)	18	(5,7)	43	(0,6)	7	(2,9)
Autres	11	(0,1)	11	(2,5)	5	(1,6)	19	(0,3)	4	(1,6)
Chimiothérapie										
Non	1984	(25,5)	56	(12,5)	40	(12,8)	1984	(26,0)	15	(6,4)
Oui	5633	(74,5)	393	(87,5)	273	(87,2)	5647	(74,0)	221	(93,6)
Anthracyclines										
Non	5043	(66,2)	220	(49,0)	150	(47,9)	5045	(66,1)	86	(35,6)
Oui	2574	(33,8)	229	(51,0)	163	(52,1)	2586	(33,9)	152	(64,4)
Dose cumulée d'anthracyclines (mg/m²)										
Médiane (EI)	227,82	(150,0 - 355,9)	352,76	(245,9 - 453,8)	352,8	(245,9 - 454,0)	227,8	(150-356,2)	363,4	(285,74 – 460,2)
Aucune	5043	(66,2)	220	(49,0)	150	(47,9)	5045	(66,1)	84	(35,6)
]0-250]	1451	(19,0)	62	(13,8)	43	(13,7)	1459	(19,1)	29	(12,3)
]250-360]	524	(6,9)	60	(13,4)	43	(13,7)	524	(6,9)	45	(19,1)
>360	599	(7,9)	107	(23,8)	77	(24,6)	603	(7,9)	78	(33,1)

Radiothérapie										
Non	3448	(45,3)	97	(21,6)	74	(23,6)	3457	(45,3)	51	(21,6)
Oui	4169	(54,7)	352	(78,4)	239	(76,4)	4174	(54,7)	185	(78,4)
Radiothérapie au cœur										
Non	3712	(48,7)	103	(23,0)	77	(24,6)	3722	(48,8)	56	(23,7)
Oui	3731	(49,0)	345	(76,8)	235	(75,1)	3735	(48,9)	180	(76,3)
Manquant	174	(2,3)	1	(0,2)	1	(0,3)	174	(2,3)	0	(0,0)
Doses moyennes de radiations au cœur (Gy)										
Médiane (EI)	1,6	(0,3-12,0)	12,4	(1,1- 27,0)	15,5	(1,4 - 28,3)	1,6	(0,3-12)	10,8	(0,9 – 22,8)
Aucune	3712	(48,7)	103	(22,9)	77	(24,6)	3722	(48,8)	56	(23,7)
]0-1]	1641	(21,5)	83	(18,5)	50	(16,0)	1642	(21,5)	46	(19,5)
]1-5]	741	(9,7)	56	(12,5)	33	(10,5)	741	(9,7)	32	(13,6)
]5-15]	580	(7,6)	40	(8,9)	32	(10,2)	582	(7,6)	17	(7,2)
]15-25]	501	(6,6)	76	(16,9)	55	(17,6)	501	(6,6)	46	(19,5)
]25-35]	172	(2,3)	54	(12,0)	41	(13,1)	172	(2,3)	23	(9,7)
>35	96	(1,3)	36	(8,0)	24	(7,7)	97	(1,3)	16	(6,8)
Manquant	174	(2,3)	1	(0,2)	1	(0,3)	174	(2,3)	0	(0,0)

Abréviations : N (%) : Effectif (pourcentage), EI : étendue interquartile ; Gy : Gray ; SNC : Système nerveux central

L'incidence cumulée du premier diagnostic de pathologie cardiaque au bout de 40 ans de suivi était de 9,0 % (95 % IC, 8,1 – 10,0) lorsque tous les grades étaient pris en compte (**Figure 13**), et de 6,5 % (95 % IC= 5,7 – 7,3) lorsque seules les pathologies cardiaques de grade ≥ 3 étaient prises en compte.

Figure 13. Incidences cumulées de pathologies cardiaques et de décès sans avoir développé de pathologie cardiaque parmi les survivants à 5 ans de la cohorte FCCSS



Chez les patients atteints d'insuffisance cardiaque, 30 % avaient été traités pour un lymphome, et les diagnostics du cancer de l'enfant ont été émis entre 1951 et 1999. La durée médiane du suivi des patients était de 21,4 ans (EI, 15,8 - 28,5), et l'âge atteint médian à la fin du suivi était de 28 ans (EI, 21 – 37). Par ailleurs, parmi les 236 patients, 87 avaient déjà développé auparavant une ou plusieurs autres pathologies cardiaques, dont la plus répandue était la valvulopathie (23 %). Certains patients avaient développé plusieurs pathologies cardiaques auparavant : 1 patient avait été diagnostiqué avec 4 pathologies, 5 patients avec 3 pathologies, 18 patients avec 2 pathologies et 63 patients avec une seule pathologie. Les patients ayant développé une insuffisance cardiaque et traités par anthracyclines (64 %) ont pour la plupart reçu des doses supérieures à 360 mg/m² (33 % des patients). Ceux traités par radiothérapie incluant le cœur représentent 76 % des patients malades, et ont reçu en médiane, une dose moyenne de radiations au cœur de 10,8 Gy (EI, 0,9 – 22,8).

4.2.3.2 *Rôles des doses de traitement sur les pathologies cardiaques et l'insuffisance cardiaque*

Pour l'évènement pathologie cardiaque tous grades et grades ≥ 3 , l'effet des traitements anticancéreux cardiotoxiques a été évalué par un modèle de Cox cause-spécifique, incluant les doses cumulées d'anthracyclines et les doses moyennes de radiations au cœur, ajusté sur le sexe, l'âge et l'année calendaire de diagnostic du cancer (**Tableau 5**). En comparaison aux patients non exposés aux anthracyclines, le risque de développer une pathologie cardiaque tous grades a été multiplié par 4,4 (95 % IC= 3,2 - 6,0) chez les patients ayant reçu entre 250 et 360 mg/m² d'anthracycline, et par 5,9 (95 % IC= 4,6 - 7,6) chez les patients traités à des doses supérieures à 360 mg/m². Des résultats similaires sont observées pour l'analyse des patients ayant développé une pathologie cardiaque de grade ≥ 3 , avec un effet encore plus élevé des doses d'anthracyclines > 360 mg/m² (csHR = 6,7 ; 95 % IC= 4,9 -9,0).

Tableau 5. Effet des doses cumulées d'anthracyclines et des doses moyennes de radiations cardiaques sur le risque de pathologies cardiaques tous grades et grades ≥ 3 dans la cohorte FCCSS

Caractéristiques	N	Pathologies cardiaques					
		Tous grades (n= 449)			Grade ≥ 3 (n=313)		
		n	csHR (95 % IC)	p-value	n	csHR (95 % IC)	p-value
Âge au diagnostic du premier cancer (années)			1,0 (1,0 -1,0)	0,2145		1,0 (1,0-1,0)	0,1911
Genre				0,3320			0,7128
Masculin	4182	240	Ref		178	Ref	
Féminin	3435	209	1,1 (0,9-1,3)		135	0,9 (0,8-1,2)	
Année de diagnostic du premier cancer				0,0325			0,1138
Avant1970]	658	77	Ref		56	Ref	
[1970-1980[1548	176	0,9 (0,7-1,3)		122	0,8 (0,5-1,2)	
[1980-1990[2460	151	0,9 (0,6-1,3)		102	0,7 (0,5-1,1)	
[1990 – 2000[2951	45	0,6 (0,4-0,9)		33	0,5 (0,3-0,9)	
Doses cumulées d'anthracyclines (mg/m²)				<0,0001			<0,0001
Aucune	5043	220	Ref		150	Ref	
]0-250]	1451	62	1,7 (1,2 - 2,3)		43	1,8 (1,2-2,6)	
]250-360]	524	60	4,4 (3,2 - 6,0)		43	4,8 (3,3-7,0)	
>360	599	107	5,9 (4,6 - 7,6)		77	6,7 (4,9-9,0)	
Doses moyennes de radiations au cœur (Gy)				<0,0001			<0,0001
Aucune	3712	103	Ref		77	Ref	
]0-1]	1641	83	1,3 (0,9-1,7)		50	1,0 (0,7-1,4)	
]1-5]	741	56	1,7 (1,2- 2,4)		33	1,3 (0,9-2,0)	
]5-15]	580	40	2,1 (1,4- 3,0)		32	2,2 (1,4-3,5)	
]15-25]	501	76	4,9 (3,6- 6,7)		55	4,6 (3,2-6,6)	
]25-35]	172	54	11,1 (7,9- 15,7)		41	11,1 (7,5-16,5)	
>35	96	36	16,6 (11,1- 24,9)		24	15,4 (9,4-25,2)	

Abréviations : N : nombre total de survivants ; n : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence.* Les csHR ont été estimés à partir d'un modèle de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), e le genre, et l'année calendaire du premier cancer

Une relation dose-réponse entre la dose moyenne de radiations cardiaques et les pathologies cardiaques a été observée à des doses moyennes au cœur supérieures ou égales à 1 Gy. En effet, par rapport à des patients n'ayant pas reçu de radiations cardiaques, avoir reçu en moyenne une dose de radiations comprise entre 1 et 5 Gy au cœur augmente le risque de pathologie cardiaque (csHR=1,7 ; 95 % IC= 1,2 – 2,4), et ce risque est multiplié par 16,9 (11,2 - 25,3) pour des doses supérieures à 35 Gy. Cependant, chez patients ayant développé une pathologie cardiaque de grade ≥ 3 , l'effet de la dose moyenne de radiations au cœur n'est observable que pour des doses supérieures à 5 Gy, avec une augmentation du risque de 2,2 fois (1,4–3,5) pour les patients traités à des doses comprises entre 5 et 10 Gy comparé aux patients n'ayant pas reçu de radiations au cœur (**Tableau 5**). Des analyses similaires à celles présentées ci-dessus ont été réalisées en remplaçant la dose moyenne au cœur par différentes mesures dose-volume. Les résultats sont présentés dans le **Tableau 6**.

Tableau 6. Effet des pourcentages de volume du cœur recevant des doses spécifiques de radiations sur le risque de pathologie cardiaque tous grades et grades ≥ 3 dans la cohorte FCCSS

Caractéristiques	N	Pathologies cardiaques					
		Tous grades			Grade ≥ 3		
		n	csHR (95 % IC)	p-value	n	csHR (95 % IC)	p-value
Volume (%) du cœur ayant reçu au moins 5 Gy				<0,0001			<0,0001
Pas de RT	3448	97	Ref		74	Ref	
0	2154	105	1,2 (0,9-1,6)		61	0,9 (0,6-1,3)	
]0-10]	247	19	1,7 (1,0-2,8)		13	1,4 (0,8-2,6)	
]10-50]	449	32	1,9 (1,3-2,9)		21	1,6 (1,0-2,7)	
>50	1145	195	5,5 (4,2-7,1)		143	5,2 (3,9-7,0)	
Volume (%) du cœur ayant reçu au moins 15 Gy				<0,0001			<0,0001
Pas de RT	3448	97	Ref		74	Ref	
0	2507	132	1,3 (1,0-1,7)		82	1,0 (0,7-1,4)	
]0-10]	280	25	2,1 (1,4-3,3)		16	1,8 (1,0-3,0)	
]10-50]	426	34	2,4 (1,6-3,5)		25	2,2 (1,4-3,6)	
>50	782	160	7,0 (5,2-9,1)		115	6,4 (4,7-8,8)	

Volume (%) du cœur ayant reçu au moins 20 Gy				<0,0001		<0,0001	
Pas de RT	3448	97	Ref	74	Ref		
0	2707	155	1,4 (1,1-1,8)	95	1,1 (0,8-1,5)		
]0-10]	351	31	2,4 (1,6-3,6)	22	2,2 (1,3-3,5)		
]10-50]	424	38	2,7 (1,9-4)	29	2,6 (1,7-4,1)		
>50	513	127	8,4 (6,4-11,2)	92	8,0 (5,8-11,0)		
Volume (%) du cœur ayant reçu au moins 30 Gy				<0,0001			
Pas de RT	3448	97	Ref	74	Ref		
0	3241	208	1,7 (1,3-2,2)	135	1,4 (1,0-1,9)		
]0-10]	320	33	2,7 (1,8-4,0)	24	2,5 (1,6-4,1)		
]10-50]	205	26	3,6 (2,3-5,6)	20	3,5 (2,1-5,9)		
>50	229	84	14,7 (10,7-20,1)	59	13,4 (9,3-19,3)		
Volume (%) du cœur ayant reçu 0,1 Gy quand V5 = 0%				0,5061		0,8659	
Pas de RT	3448	97	Ref	74	Ref		
0	444	19	1,2 (0,7-1,9)	11	0,9 (0,5-1,8)		
]0-10]	45	2	1,0 (0,2-4,0)	2	1,3 (0,3-5,4)		
]10-50]	68	6	1,8 (0,8-4,1)	4	1,6 (0,6-4,4)		
>50	1597	78	1,3 (0,9-1,8)	44	1,0 (0,7-1,5)		
Volume (%) du cœur ayant reçu 5 Gy quand V20 = 0%				<0,0001		<0,0001	
Pas de RT	3448	97		74	Ref		
0	2154	105	1,3 (0,9-1,7)	61	1,0 (0,7-1,4)		
]0-10]	231	18	1,9 (1,1-3,1)	12	1,7 (0,9-3,1)		
]10-50]	162	12	2,3 (1,2-4,2)	8	2,1 (1,0-4,4)		
>50	160	20	4,0 (2,4-6,7)	14	3,8 (2,1-6,9)		

Abréviations ; N : nombre total de survivants ; n : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence ; V5 : Volume du cœur ayant reçu au moins 5Gy ; V20 : Volume du cœur ayant reçu au moins 20Gy ; RT : Radiothérapie.

Toutes les sections de ce tableau proviennent de modèle distincts, exclusifs l'un de l'autre. Les csHR ont été estimés à partir de modèles de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), le genre, l'année calendaire du premier cancer, et la dose cumulée d'anthracyclines.

A des doses de radiations ≥ 5 Gy sur des volumes cardiaques moyens ($[10-50]$ %) et en comparaison à des survivants non exposés à la radiothérapie, une augmentation du risque de pathologie cardiaque tous grades (csHR= 1,9 ; 95 % IC= 1,3 - 2,8) a été observée. A des doses plus élevées (≥ 15 Gy), une association significative a été observée (csHR= 2,1 ; 95 % IC= 1,4 - 3,3) pour des volumes cardiaques compris entre 0 et 10 % comparé à des patients non exposés. Concernant le risque de pathologies cardiaques de grade ≥ 3 , le risques associé à des doses de radiations ≥ 5 Gy n'est observé qu'à partir d'un volume irradié supérieur à 50 %, et pour les doses plus élevées (≥ 15 Gy), l'augmentation du risque est observée à partir d'un volume de 10% du cœur irradié ($[10 - 50]$ %, csHR= 2,2 ; 95 % IC= 1,4 - 3,6).

Par ailleurs, pour des doses comprises entre 0,1 et 5 Gy, aucune association n'a été trouvée avec les pathologies cardiaques tous grades et de grade ≥ 3 , même sur des grands volumes du cœur (> 50 %). Cependant, pour des doses allant de 5 à 20 Gy au cœur, une association avec la pathologie cardiaque tous grades a été mise en évidence même pour de petits volumes du cœur irradiés ($[0 - 10\%]$), avec un risque multiplié par 1,9 (95 % IC= 1,1 - 3,1) comparé aux patients n'ayant pas été exposés à la radiothérapie. Cette association n'a pas été retrouvée pour les grades ≥ 3 , pour lesquels l'association n'apparaît qu'à partir d'un grande volume du cœur irradié (> 50 % ; csHR= 3,8 ; 95 % IC= 2,1 - 6,9).

Des interactions importantes ont été observées entre l'administration d'anthracyclines et les doses moyennes de radiations reçues au cœur. Ces résultats sont présentés dans le **Tableau 7**. Que ce soit avec ou sans anthracyclines, les patients qui avaient reçu une dose moyenne de radiations au cœur de plus de 25 Gy avaient un risque de pathologie cardiaque très élevé, multiplié respectivement par 23,3 (95 % IC= 14,8 - 36,6) et 27,1 (95 % IC= 14,9 - 49,5), par rapport aux patients qui n'étaient ni exposés aux radiations cardiaques ni aux anthracyclines. Lorsque seules les pathologies cardiaques de grades ≥ 3 sont prises en compte, une augmentation de risque similaire est observée pour ces deux catégories de traitements.

Tableau 7. Effet et incidence du traitement par anthracyclines et des doses moyennes de radiations cardiaques sur le risque de première pathologie cardiaque tous grades et grades ≥ 3 dans la cohorte FCCSS

Administration d'Anthracyclines	Dose moyenne de radiations au cœur, Gy	Pathologies cardiaques						
		Tous grades (n= 449)				Grade ≥ 3 (n=313)		
		N	n	Incidence à 40 ans de suivi (95 % IC)	csHR* (95 % IC)	n	Incidence à 40 ans de suivi (95 % IC)	csHR* (95 % IC)
Non								
	Aucune	2349	28	2,5 (1,5 – 4,0)	Ref	20	1,7 (0,9 – 2,3)	Ref
]0-1]	1121	33	3,5 (2,3 – 5,2)	1,6 (1,0-2,7)	16	1,6 (0,9 – 2,7)	1,0 (0,5-2,0)
]1-5]	522	28	4,4 (2,7 – 6,8)	2,1 (1,3-3,7)	18	2,8 (1,4 – 4,8)	1,8 (1,0-3,5)
]5-15]	433	25	6,3 (3,8 – 9,5)	2,9 (1,7-4,9)	20	4,7 (2,6 – 7,6)	3,1 (1,7-5,8)
]15-25]	285	33	14,6 (10,1 – 19,9)	6,9 (4,1-11,5)	21	10,2 (6,4 – 15,1)	6,0 (3,2-11,1)
	> 25†	188	72	38,8 (31,2 – 46,3)	23,3 (14,8-36,6)	54	30,9 (23,5 – 38,5)	23,7 (13,9-40,3)
Oui								
	Aucune	1363	75	10,2 (7,0 – 14,1)	5,9 (3,8-9,1)	57	8,1 (5,2 – 22,9)	6,5 (3,8-10,9)
]0-1]	520	50	13,6 (9,9 – 18,0)	7,6 (4,7-12,2)	34	10,1 (6,7 – 14,3)	7,4 (4,2-13,0)
]1-5]	219	28	17,3 (11,3 – 24,5)	9,9 (5,8-16,9)	15	11,0 (5,8 – 18,1)	7,7 (3,9-15,1)
]5-15]	147	15	14,8 (8,2 – 23,3)	9,3 (4,9-17,4)	12	11,7 (6,0 – 19,4)	10,7 (5,2-22,0)
]15-25]	216	43	32,2 (21,5 – 43,3)	20,1 (12,4-32,5)	34	28,0 (17,4 – 39,6)	22,5 (12,8-39,4)
	> 25†	80	18	44,9 (18,5 – 68,4)	27,1 (14,9-49,5)	11	37,1 (9,8 – 65,5)	23,4 (11,1-49,3)

Abréviations ; N : nombre total de survivants ; n : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence.

* Les csHR ont été estimés à partir d'un modèle de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), le genre, et l'année calendaire du premier cancer. **Les catégories]25 – 35] et >35 Gy ont été regroupée à cause du faible nombre de patients dans la catégorie des patients traités par anthracyclines et des doses de radiations >35 Gy..

En s'intéressant au risque d'insuffisance cardiaque chez les survivants d'un cancer dans l'enfance, il en ressort que les patients traités par anthracyclines, même à des doses < 250 mg/m² ont un risque plus élevé de développer une première insuffisance cardiaque. Ainsi, ce risque est multiplié par 2,4 (95 % IC= 1,5-3,9) pour des doses comprises entre 0 et 250 mg/m², et par 12,1 (95 % IC= 8,43-17,4) pour les doses > 250 mg/m² par rapport aux patients non traités par anthracyclines (**Tableau 8**). Une augmentation du risque d'insuffisance cardiaque a été observée chez les patients ayant reçu des radiations au cœur entre 1 et 5 Gy comparativement aux patients sans radiothérapie (csHR=1,8, 95 % IC= 1,1-2,8) pour des doses supérieures à 35 Gy, ce risque est multiplié par 16,8 (95 % IC= 9,3-30,3).

De même, sur des volumes cardiaques moyens ([10-50] %) et à des doses de radiations ≥5 Gy, et en comparaison avec des survivants non exposés à la radiothérapie, on a constaté une augmentation du risque d'insuffisance cardiaque de 2,1 fois (95 % IC= 1,2 – 3,7). Les patients ayant été irradiés à au moins 20 Gy pour des volumes supérieurs à 10%, voient leur risque de développer une insuffisance cardiaque augmenté de 2,7 fois (95 % IC= 1,6– 4,6) en comparaison aux patients non irradiés. En outre, les associations entre les indicateurs de doses bornés et l'insuffisance cardiaque n'étaient pas significatifs (**Tableau 9**).

Tableau 8. Effet des doses cumulées d'anthracyclines et des doses moyennes de radiations cardiaques sur le risque d'insuffisance cardiaque tous grades dans la cohorte FCCSS

Caractéristiques	Insuffisance cardiaque tous grades (n= 236)			
	N	n	csHR (95 % IC)	p-value
Âge au diagnostic du premier cancer (années)			1,0 (0,9-1,0)	0,0094
Genre				0,5449
	Masculin	4187	130	Ref
	Féminin	3444	106	1,1 (0,8-1,4)
Année de diagnostic du premier cancer				<,0001
	Avant1970]	1986	34	Ref
]1970-1980]	658	101	0,7 (0,4-1,1)
]1980-1990]	1548	86	0,6 (0,4-1,1)
]1990 - 2000	2466	15	0,2 (0,1-0,4)
Doses cumulées d'anthracyclines (mg/m²)				<,0001
	Aucune	5045	84	Ref
]0-250]	1459	29	2,4 (1,5-3,9)
]250-360]	524	45	9,6 (6,3-14,5)
	>360	603	78	12,1 (8,4-17,4)
Doses moyennes de radiations au cœur (Gy)				<,0001
	Aucune	3666	56	Ref
]0-1]	1596	46	1,3 (0,9-2)
]1-5]	709	32	1,8 (1,1-2,8)
]5-15]	565	17	1,6 (0,9-2,8)
]15-25]	455	46	5,4 (3,6-8,2)
]25-35]	149	23	9,6 (5,8-16,0)
	>35	81	16	16,8 (9,3-30,3)

Abréviations ; N : nombre total de survivants ; n : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence. * Les csHR ont été estimés à partir d'un modèle de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), e le genre, et l'année calendaire du premier cancer

Tableau 9. Effet des pourcentages de volume du cœur recevant des doses spécifiques de radiations sur le risque d'insuffisance cardiaque dans la cohorte FCCSS

Caractéristiques	Insuffisance cardiaque tous grades (n=236)			
	N	n	csHR (95 % IC)	p-value
Volume (%) du cœur ayant reçu au moins 5 Gy				<0,0001
Pas de RT	3406	51	Ref	
0%	2097	59	1,3 (0,9-2,0)	
]0-10] %	236	11	1,9 (1,0-3,7)	
]10-50] %	431	19	2,1 (1,2-3,7)	
>50 %	1051	96	5,4 (3,8-7,8)	
Volume (%) du cœur ayant reçu au moins 15 Gy				<0,0001
Pas de RT	3406	51	Ref	
0%	2437	73	1,4 (0,9-2,0)	
]0-10] %	264	16	2,6 (1,4-4,6)	
]10-50] %	412	15	1,9 (1,1-3,5)	
>50 %	702	81	7,3 (5,0-10,6)	
Volume (%) du cœur ayant reçu au moins 20 Gy				<0,0001
Pas de RT	3406	51	Ref	
0	2628	83	1,5 (1,0-2,1)	
]0-10] %	333	18	2,6 (1,5-4,4)	
]10-50] %	404	20	2,7 (1,6-4,6)	
>50 %	450	64	9,2 (6,2-13,6)	
Volume (%) du cœur ayant reçu au moins 30 Gy				<0,0001
Pas de RT	3406	51	Ref	
0	3132	113	1,8 (1,3-2,5)	
]0-10] %	301	19	3,2 (1,8-5,4)	
]10-50] %	186	19	4,9 (2,9-8,5)	
>50 %	196	34	12,9 (8,1-20,5)	
Volume (%) du cœur ayant reçu 0,1 Gy quand V5 = 0%				0,0866
Pas de RT	3457	51	Ref	
0%	445	14	1,8 (1,0-3,3)	
]0-10] %	45	2	2,0 (0,5-8,3)	
]10-50] %	68	5	3,1 (1,2-8,1)	
>50 %	1598	38	1,3 (0,8-2,0)	

Volume (%) du cœur ayant reçu 5 Gy quand V20 = 0%

0,0049

Pas de RT	3457	51	Ref
0%	2156	59	1,5 (1,0-2,2)
]0-10] %	231	10	2,2 (1,1-4,2)
]10-50] %	163	5	2,0 (0,8-5,2)
>50 %	161	9	3,7 (1,8-7,6)

Abréviations ; N : nombre total de survivants ; n : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence ; V5 : Volume du cœur ayant reçu au moins 5Gy ; V20 : Volume du cœur ayant reçu au moins 20Gy ; RT : Radiothérapie.

Toutes les sections de ce tableau proviennent de modèle distincts, exclusifs l'un de l'autre. Les csHR ont été estimés à partir de modèles de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), le genre, l'année calendaire du premier cancer, et la dose cumulée d'anthracyclines.

Le **Tableau 10** présente les résultats des analyses interactions entre l'administration d'anthracyclines et les doses de radiations reçues au cœur. Les patients n'ayant pas reçu d'anthracyclines et ayant reçu une dose moyenne de radiations au cœur de plus de 15 Gy ont 21,0 fois (95 % IC= 9,8-44,9) plus de risque de développer une insuffisance cardiaque comparés aux patients non exposés aux anthracyclines ni aux radiations. Pour les patients exposés aux anthracyclines avec la même dose moyenne de radiations, ce risque est multiplié par 49,4 (95 % IC= 22,7-107,5) par rapport à ceux non exposés aux anthracyclines et aux radiations cardiaques.

Tableau 10. Effet et incidence du traitement par anthracyclines et des doses moyennes de radiations cardiaques sur le risque de première insuffisance cardiaque tous grades dans la cohorte FCCSS

Administration d'Anthracyclines	Dose moyenne de radiations au cœur, Gy	Insuffisance cardiaque tous grades (n= 239)			
		N	n	Incidence à 40 ans de suivi (95 % IC)	csHR* (95 % IC)
Non					
	Aucune	2350	8	0,8 (0,3 – 2,0)	Ref
]0 – 5]	1644	20	1,4 (0,8 – 2,1)	1,9 (0,8-4,4)
]5 – 15]	433	7	1,6 (0,6 – 3,5)	2,4 (0,9-6,7)
	> 15	473	49	10,9 (7,9 – 14,4)	21,0 (9,8-44,9)
Oui					
	Aucune	1372	48	8,5 (3,5 – 5,9)	15,2 (7,1-32,5)
]0 – 5]	739	58	10,6 (7,9 – 13,7)	24,1 (11,4-51,1)
]5 – 15]	149	10	10,1 (4,9– 17,6)	23,4 (9,1-59,6)
	> 15	297	36	10,9 (21,9 – 30,3)	49,4 (22,7-107,5)

Abréviations ; N : nombre total de survivants ; n : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence.

Certaines catégories de doses ont été regroupées à cause du faible nombre d'évènement dans les catégories précédentes.

Les csHR ont été estimés à partir d'un modèle de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), le genre, et l'année calendaire du premier cancer

4.2.3.3 Effet des pathologies cardiaques intermédiaires sur le risque d'insuffisance cardiaque

Le risque d'insuffisance cardiaque était associé à différents évènements cardiaques préexistants, et ces risques sont résumés dans le **Tableau 11**. Le risque d'insuffisance cardiaque suite à une valvulopathie était augmenté de 8,9 fois (95 % IC= 5,0 – 15,9) en comparaison aux patients n'ayant pas développé de pathologie cardiaque auparavant. De même, les pathologies ischémiques préexistantes semblaient multiplier le risque d'insuffisance cardiaque par 8,8 (95 % IC= 4,3 – 17,9), et pour les troubles du rythme et de la conduction ce risque était multiplié par 3,6 (95 % IC= 1,7 – 7,6) comparé aux patients n'ayant pas été diagnostiqués d'une pathologie cardiaque préexistante. Seule

la péricardite ne semble pas être associée à l'insuffisance cardiaque.

Les tendances observées pour les effets des doses cumulées d'anthracyclines ne semblent pas être modifiées par l'ajout de ces variables au modèle. Cette tendance reste similaire aussi pour les doses moyennes de radiations au cœur, mais se stabilise pour des doses moyennes de radiations au cœur > 25 Gy. Cette dernière catégorie semble avoir un effet plus faible sur l'apparition de l'insuffisance cardiaque.

Tableau 11. Effet des pathologies cardiaques préexistantes sur les insuffisances cardiaques tous grades

Caractéristiques	Insuffisance cardiaque tous grades (n= 239)			
	N	n	csHR (95 % IC)	p-value
<i>Valvulopathie</i>	164	55	8.9 (5,0-15.9)	<0,0001
<i>Trouble du rythme et de la conduction</i>	117	38	3.6 (1.7-7.6)	0,0007
<i>Ischémie</i>	52	14	8.8 (4.3-17.9)	<0,0001
<i>Péricardite</i>	43	7	0,4 (0,2 - 1,2)	0,1240

Abréviations ; N : nombre total de survivants ; n : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence.

Les csHR ont été estimés à partir d'un modèle de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), le genre, l'année calendaire du premier cancer, les doses cumulées d'anthracyclines et les doses moyennes de radiations au cœur

Les tableaux exposant les risques de décès sans avoir développé de pathologie cardiaque (évènement concurrent) sont présentés en **Annexe A. 1** à **Annexe A. 5**.

4.3 RISQUES D'HOSPITALISATIONS POUR CAUSE CARDIAQUE CHEZ LES SURVIVANTS D'UN CANCER DE L'ENFANT

4.3.1 Population d'étude

Pour étudier le risque d'hospitalisations pour cause cardiaque après cancer de l'enfant dans la cohorte FCCSS, nous avons sélectionné les patients en vie en janvier 2011 (date de disponibilité des données d'hospitalisations dans la population générale) et résidant en France, pour lequel le chaînage au SNDS a été effectué. Les données d'hospitalisation de ces patients couvrant la période du 1er janvier 2011 au 31 décembre 2018 ont été analysées. Le suivi de notre population d'étude a commencé le 1er janvier 2011 et s'est terminé à la date de fin de l'étude, le 31 décembre 2018. Les patients ont été censurés aux dates suivantes : décès, perte de suivi, hospitalisation pour cause cardiaque ou date de fin de l'étude, selon ce qui est survenu en premier. Les dates et les causes de décès ont été identifiées à partir de la base nationale des causes médicales de décès. La cohorte finale comprend 5498 patients, soit 83 % des survivants en vie en 2011 (**Figure 14**).

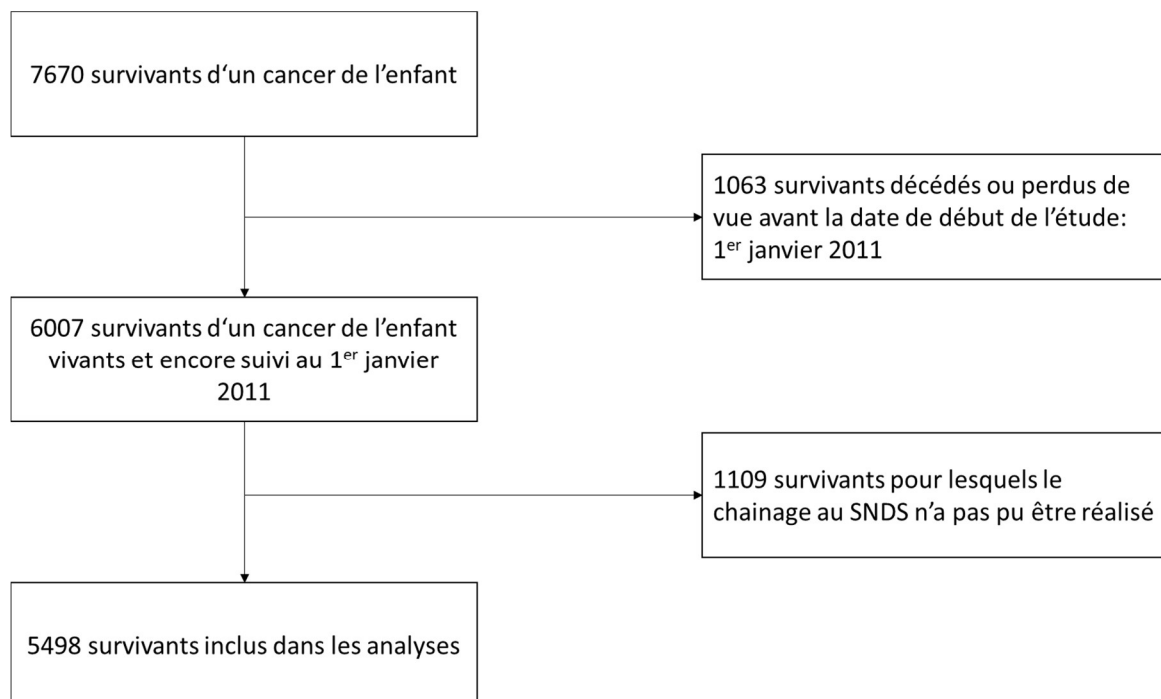


Figure 14. Diagramme de flux de l'inclusion des patients dans l'étude de cohorte sur les hospitalisations pour cause cardiaque

Les données relatives aux traitements anticancéreux ont été recueillies comme décrit dans le Chapitre 2. Les doses des différentes chimiothérapies et les doses moyennes de radiations au cœur ainsi que les estimateurs dose-volume ont été estimés et calculés. Les pourcentages du volume cardiaque ayant reçu au moins 5 (V5), 10 (V10), 15 (V15), 20 (V20) ou 30 (V30) Gy ont été utilisés dans les analyses.

Des variables spécifiques telles que les comorbidités incluant les maladies rénales, et le diabète sucré diagnostiqués avant les hospitalisations pour pathologies cardiaques ont aussi été recueillies. Les tumeurs malignes secondaires survenant 5 ans ou plus après le premier diagnostic de cancer ont été considérées dans les analyses. Le statut de la splénectomie (Oui / Non) a été également pris en compte.

4.3.2 Méthode

Les taux d'hospitalisation ont été calculés en combinaison avec les données démographiques. Les nombres attendus ont été estimés en combinant les personnes-années à risque au sein de strates spécifiques selon le sexe, l'âge, et l'année calendaire et multipliés par les taux correspondants dans la population générale. Le nombre observé d'hospitalisations cardiaques et le nombre attendu dans la population générale ont été comparés par le calcul des ratios standardisés d'hospitalisation (SHR), ce qui revient à calculer le rapport entre le nombre observé et le nombre attendu d'hospitalisation cardiaque. Les excès de risque absolus (ERA) ont été calculés comme le nombre observé moins le nombre attendu d'hospitalisations cardiaques, divisé par le nombre de personnes-années (PA) à risque et multiplié par 10 000. La procédure de calcul est présentée dans le Chapitre 3 section 3.1.2. Les SHR peuvent être interprétés comme des facteurs multiplicatifs d'hospitalisations pour causes cardiaques par rapport à la population générale, tandis que les ERA peuvent être interprétés comme des facteurs additifs. Les tests d'hétérogénéité étaient réalisés à l'aide de tests du rapport de vraisemblance.

L'estimation des risques relatifs (RR) d'être hospitalisé pour une cause cardiaque s'est faite à l'aide d'une régression de Poisson univariée et multivariée prenant en compte le nombre attendu d'hospitalisations pour cause cardiaque comme décrit dans le

Chapitre 3 section 3.1.2. De plus, le choix des variables inclus dans le modèle multivarié s'est fait selon une procédure stepwise. Les modèles multivariés ont été ajustés sur les facteurs de risque démographiques, cliniques et liés au traitement. Ces variables d'ajustement ont été conservées dans le modèle final même si elles n'étaient pas significatives. Nous avons inclus les différents indicateurs de dose de radiation dans des modèles séparés afin d'éviter des sur-ajustements. Les comorbidités ont été incluses en tant que variables dépendantes du temps. Les modèles multivariés ont inclus des variables sur la base des connaissances cliniques et de la littérature, ainsi que celles significativement associées dans l'analyse bivariée.

De plus, pour tenir compte de la récurrence des hospitalisations le nombre moyen cumulé d'hospitalisations ont été calculés et peuvent être interprétés comme le nombre moyen d'hospitalisation cardiaque par survivant, à l'instant t.

4.3.3 Résultats

4.3.3.1 Description

Les analyses ont été menées sur 5498 survivants, dont les caractéristiques démographiques et cliniques sont décrites dans le **Tableau 12**. L'âge médian au diagnostic de cancer de l'enfant est de 4 ans (EI, 0-20) et le délai médian entre le cancer de l'enfant et 2011 est de 22,9 ans (EI, 16,0-30,5). Entre 2011 et 2018, 647 hospitalisations cardiaques ont été recensées sur 247 patients hospitalisés pour une cause cardiaque (**Annexe A. 6**). L'âge médian lors de la première hospitalisation est de 41 ans (EI,34-48). Les causes cardiaques les plus fréquentes sont d'abord l'insuffisance cardiaque (n=152 ; 23,5 %), puis la cardiomyopathie (n=106 ; 16,4 %) et enfin la cardiopathie ischémique chronique (n=80 ; 12,4 %) (**Annexe A. 6**).

Tableau 12. Caractéristiques démographiques et cliniques des patients inclus dans l'étude, et des hospitalisations liées à des causes cardiaques de 2011 à 2018

	Survivants inclus (n= 5489) N (%)	Hospitalisations (n= 647) N (%)	Patients hospitalisés (n= 247) N (%)
Genre			
Masculin	2999 (54.5)	377 (69)	149 (60.3)
Féminin	2499 (45.5)	169 (31)	98 (39.7)
Âge au diagnostic du premier cancer (années)			
Médiane (min - max)	4 (0-20)	6 (0-19)	6 (0-19)
[0 – 5[2442 (44.4)	170 (31.1)	88 (35.6)
[5 – 10[1283 (23.3)	181 (33.2)	68 (27.5)
[10 – 15[638 (11.6)	82 (15)	39 (15.8)
[15 – 22[681 (12.4)	58 (10.6)	31 (12.6)
Médiane (El)	454 (8.3)	55 (10.1)	21 (8.5)
Type de tumeur primaire			
Néphroblastome	834 (15.2)	84 (15.4)	38 (15.4)
Neuroblastome	767 (14)	40 (7.3)	29 (11.7)
Lymphome	943 (17.2)	221 (40.5)	78 (31.6)
Sarcome des tissus mous	598 (10.9)	84 (15.4)	33 (13.4)
Sarcome osseux	485 (8.8)	31 (5.7)	22 (8.9)
Tumeur du SNC	687 (12.5)	34 (6.2)	20 (8.1)
Tumeur des gonades	340 (6.2)	10 (1.8)	8 (3.2)
Tumeur de la thyroïde	48 (0.9)	2 (0.4)	2 (0.8)
Rétinoblastome	477 (8.7)	10 (1.8)	5 (2)
Autres	319 (5.8)	30 (5.5)	12 (4.9)
Année de diagnostic du premier cancer			
Médiane (min - max)	1988 (1946-2000)	1980 (1948-2000)	1981 (1948-2000)
Avant 1970	339 (6.2)	61 (11.2)	39 (15.8)
[1970, 1979[967 (17.6)	194 (35.5)	73 (29.6)
[1980, 1989]	1828 (33.2)	220 (40.3)	96 (38.9)
[1990, 2000]	2364 (43)	71 (13.0)	39 (15.8)
Suivi après diagnostic du cancer (années)			
Médiane (min - max)	30.1 (11.1-72.3)	32.5 (16-64.6)	33.8 (14-64.6)
<20	604 (11)	38 (7.0)	19 (7.7)
[20 – 30[2136 (38.9)	156 (28.6)	70 (28.3)
[30 – 40[1723 (31.3)	252 (46.2)	91 (36.8)
[40 – 50[778 (14.2)	80 (14.7)	50 (20.2)
≥ 50	257 (4.7)	20 (3.7)	17 (6.9)
Âge atteint au 01/01/2011 (années)			
Médiane (min - max)	31 (11-71)	38 (13-68)	39 (13-68)
<20	854 (15.5)	11 (2)	6 (2.4)
[20;30[1669 (30.4)	92 (16.8)	47 (19)
[30;40[1787 (32.5)	207 (37.9)	76 (30.8)
[40;50[911 (16.6)	185 (33.9)	87 (35.2)
≥ 50	277 (5)	51 (9.3)	31 (12.6)

Abréviations : N (%) : Effectifs (pourcentages), El : Etendue interquartile ; SNC : Système nerveux central

En comparaison au nombre d'hospitalisation cardiaque attendu en population générale, ce nombre est 4,6 fois (95 % IC= 4,0 - 5,2) plus élevé chez les survivants d'un cancer de l'enfant. Cela se traduit par 45,8 hospitalisations supplémentaires (95 % IC= 38,7 - 53,6) pour 10 000 PA. Le nombre cumulé moyen à 60 ans pour les hospitalisations cardiaques était de 0,86 (95 % IC= 0,69 - 1,06 ; **Figure 15**).

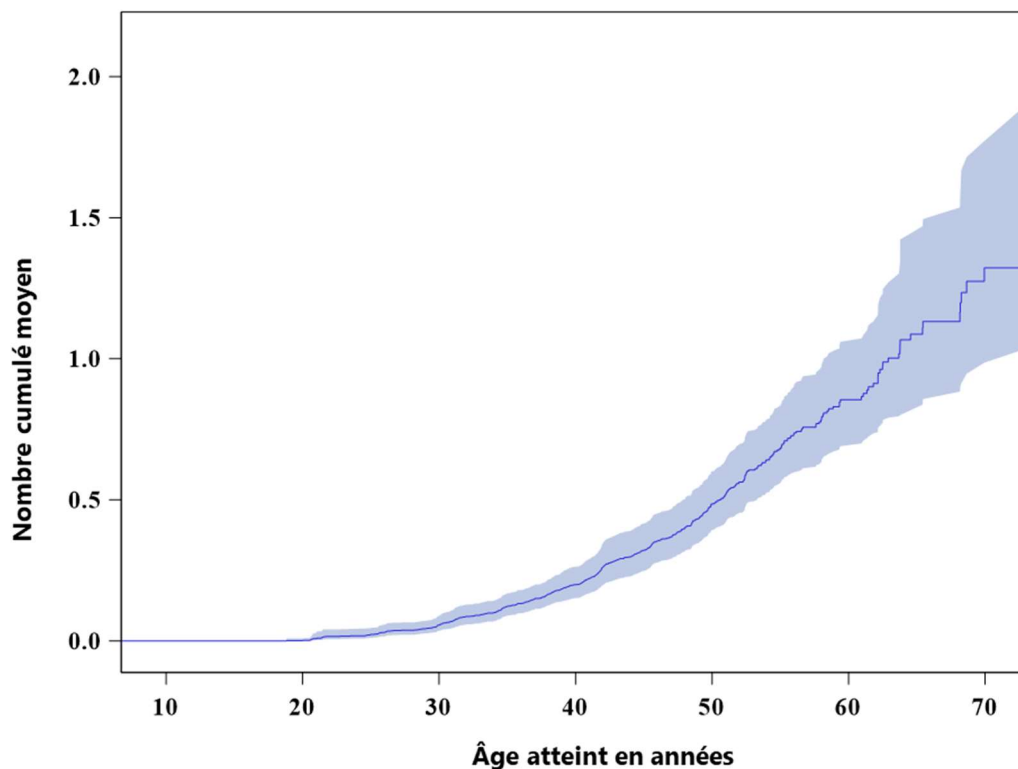


Figure 15. Nombre cumulé moyen d'hospitalisations pour cause cardiaque dans la cohorte FCCSS de 2011 à 2018, en fonction de l'âge atteint.

4.3.3.2 Données démographiques et cliniques

La description des SHR et des ERA selon les différentes caractéristiques cliniques sont présentées dans le **Tableau 13**. L'insuffisance cardiaque est associée aux SHR les plus élevés (SHR=9,8 ; 95 % IC= 7,5-12,5) et est suivie des maladies valvulaires (SHR=8,4 ; 95 % IC= 5,6-12,1). La cardiomyopathie et l'arythmie, l'insuffisance cardiaque et les maladies ischémiques présentent respectivement des ERA de 17,4 (95 % IC= 12,8-22,7), 13,4 (95 % IC= 10,0-17,5) et 10,9 (95 % IC= 7,0-15,5) pour 10 000 PA.

Tableau 13. Nombre d'hospitalisations pour cause cardiaque observées et attendues, ratios d'hospitalisation standardisés et excès de risque absolu, chez les survivants de la cohorte FCCSS de 2011 à 2018.

Caractéristiques	N	Hospitalisations cardiaques			
		PA	Obs/Att	SHR (95 % IC)	ERA (95 % IC)
Genre					
Masculin	2999	22973,1	149/41,1	3,6 (3,1-4,3)	47 (37-58,3)
Féminin	2499	19229,4	98/12,7	7,7 (6,3-9,4)	44,4 (34,8-55,5)
<i>p</i> -value d'hétérogénéité				< 0,001	0,696
Type de tumeur primaire					
Tumeur du SNC	687	5173,5	20/6,2	3,2 (2-5)	26,6 (11,6-47,7)
Tumeur des gonades	340	2669,3	8/3,6	2,2 (1-4,4)	16,5 (-0,6-45,6)
Lymphome	943	7119	78/11,4	6,8 (5,4-8,5)	93,5 (70,6-120,7)
Néphroblastome	834	6375,3	38/8,6	4,4 (3,1-6,1)	46,1 (28,7-68,3)
Neuroblastome	767	5965,2	29/4,8	6,0 (4,0-8,6)	40,5 (24,4-61,7)
Sarcome osseux	485	3746,7	22/6,6	3,3 (2,1-5)	41,1 (19,1-71,3)
Rétinoblastome	477	3750,5	5/1,9	2,7 (0,9-6,3)	8,4 (-0,6-26,1)
Tumeur de la Thyroïde	48	377,3	2/0,8	2,4 (0,3-8,7)	30,9 (-15,7-169,4)
Sarcome des tissus mous	598	4590,7	33/7,1	4,7 (3,2-6,6)	56,5 (34,1-85,6)
Autres	319	2435,1	12/2,7	4,4 (2,3-7,8)	38,2 (14,4-75)
<i>p</i> -value d'hétérogénéité				0,003	< 0,001
Âge au diagnostic du premier cancer (années)					
]0-3]	2442	18915,5	88/17,5	5 (4-6,2)	37,3 (28,1-48,1)
[4-8]	1283	9770,4	68/12,5	5,4 (4,2-6,9)	56,8 (41,2-75,4)
[9-11]	638	4820	39/7,9	5 (3,5-6,8)	64,6 (41,2-94,3)
[12-14]	681	5201,6	31/9,8	3,2 (2,1-4,5)	40,7 (21,6-65,7)
≥ 15	454	3495	21/6,1	3,5 (2,1-5,3)	42,7 (19,8-74,5)
<i>p</i> -value d'hétérogénéité				0,068	0,05
Année de diagnostic du premier cancer					
Avant 1970	339	2424,5	39/14,1	2,8 (2-3,8)	102,5 (56,1-161,6)
[1970-1980[967	7159,1	73/16,9	4,3 (3,4-5,4)	78,4 (56,4-104,7)
[1980-1990[1828	13976,6	96/14,9	6,4 (5,2-7,9)	58 (45-73,2)
[1990 -2000]	2364	18642,4	39/7,9	5 (3,5-6,8)	16,7 (10,7-24,4)
<i>p</i> -value d'hétérogénéité				< 0,001	< 0,001
Années depuis le diagnostic du cancer (années)					
<20	604	4679,5	19/1,4	13,5 (8,1-21)	37,6 (21,4-60,4)
[20-29]	2136	16543,2	70/8,3	8,5 (6,6-10,7)	37,3 (28-48,5)
[30-39]	1723	13157,8	91/15,9	5,7 (4,6-7)	57,1 (43,6-72,8)
[40-49]	778	5887,3	50/16,1	3,1 (2,3-4,1)	57,6 (35,7-84,7)
≥ 50	257	1934,8	17/12,1	1,4 (0,8-2,2)	25,1 (-11,6-77,9)
<i>p</i> -value d'hétérogénéité				< 0,001	0,036
Âge atteint au 01/01/2011 (années)					
<20	854	6808,1	6/1,7	3,5 (1,3-7,7)	6,3 (0,7-16,7)
[20 -30[1669	13008,9	47/5,6	8,4 (6,2-11,2)	31,8 (22,3-43,8)

[30 -40[1787	13745,7	76/14	5,4 (4,3-6,8)	45,1 (33,4-59)
[40 -50[911	6661,8	87/18,8	4,6 (3,7-5,7)	102,4 (76,4-132,9)
≥50	277	1978	31/13,7	2,3 (1,5-3,2)	87,3 (37,1-153,1)
<i>p-value d'hétérogénéité</i>				< 0,001	< 0,001
Traitement du premier cancer					
Chirurgie seule	715	5661	9/5,3	1,7 (0,8-3,2)	6,5 (-2,1-20,8)
Radiothérapie sans Chimiothérapie	658	4964,2	31/12,5	2,5 (1,7-3,5)	37,3 (17,3-63,5)
Chimiothérapie sans Radiothérapie	2023	15849,6	64/13,9	4,6 (3,6-5,9)	31,6 (22,4-42,8)
Chimiothérapie et Radiothérapie	2102	15727,7	143/22,2	6,5 (5,4-7,6)	76,8 (62,5-93)
<i>p-value d'hétérogénéité</i>				< 0,001	< 0,001
Chimiothérapie					
Non	1373	10625,1	40/17,8	2,3 (1,6-3,1)	20,9 (10,2-34,5)
Oui	4125	31577,4	207/36	5,7 (5-6,6)	54,2 (45,5-63,7)
<i>p-value d'hétérogénéité</i>				< 0,001	< 0,001
Anthracyclines					
Non	3587	27590,5	138/38	3,6 (3-4,3)	36,2 (28,2-45,3)
Oui	1911	14612	109/15,7	6,9 (5,7-8,4)	63,8 (50,5-79,2)
<i>p-value d'hétérogénéité</i>				< 0,001	< 0,001
Doses cumulées d'anthracyclines (mg/m²)					
Aucune	3587	27590,5	138/38	3,6 (3-4,3)	36,2 (28,2-45,3)
]0-250[979	7569,6	36/6,4	5,6 (3,9-7,8)	39,1 (24,9-57,4)
[250-360[501	3816,9	29/4,5	6,4 (4,3-9,2)	64,1 (39-97,3)
≥ 360	431	3225,4	44/4,8	9,1 (6,6-12,2)	121,5 (84,2-168,2)
<i>p-value d'hétérogénéité</i>				< 0,001	< 0,001
Dose moyenne de radiations au cœur (Gy)					
0	3192	25042,4	87/22,3	3,9 (3,1-4,8)	25,8 (18,9-33,9)
]0 -5[1479	11142,4	70/22	3,2 (2,5-4)	43,1 (29,2-59,6)
]5 -15	405	3069,8	21/4,7	4,4 (2,7-6,8)	53 (26,9-89,1)
]15 -30[354	2531,3	44/3,9	11,2 (8,1-15)	158,3 (110,8-217,8)
> 30	68	416,5	25/0,8	30,8 (19,9-45,4)	580,7 (368,9-866,5)
<i>p-value d'hétérogénéité</i>				< 0,001	< 0,001
Volume (%) du cœur ayant reçu au moins 20 Gy					
0	4731	36665,3	160/45,2	3,5 (3-4,1)	31,3 (24,8-38,6)
]0- 10[193	1443,8	12/2,1	5,6 (2,9-9,8)	68,3 (28,2-130,4)
]10 -50[309	2281,8	22/3,6	6,2 (3,9-9,4)	80,8 (44,8-130,4)
> 50	265	1811,6	53/2,9	18,5 (13,9-24,2)	276,8 (203,3-366,9)
<i>p-value d'hétérogénéité</i>				< 0,001	< 0,001
Volume (%) du cœur ayant reçu au moins 30 Gy					
0	5070	39201,4	182/48,6	3,7 (3,2-4,3)	34 (27,5-41,3)
]0- 10[172	1286,6	10/2,2	4,5 (2,1-8,2)	60,4 (19,9-125,6)
]10 -50[153	1092,1	18/1,8	9,8 (5,8-15,5)	148 (80,9-243,7)
> 50	103	622,5	37/1,1	32,3 (22,7-44,5)	576 (400,1-800,9)
<i>p-value d'hétérogénéité</i>				< 0,001	< 0,001
Tumeur maligne secondaire *					
Non	5110	39938,1	220/50	4,4 (3,8-5)	42,6 (35,5-50,4)
Oui	388	2264,4	27/3,8	7,1 (4,7-10,3)	102,4 (61,8-156,7)

				<i>p</i> -value d'hétérogénéité	0,02	< 0,001
Pathologie cardiaque*						
Non	5216	40599,5	151/50,7		3 (2,5-3,5)	24,7 (19-31,1)
Oui	282	1602,9	96/3,1		30,6 (24,8-37,4)	579,4 (465,6-711,8)
				<i>p</i> -value d'hétérogénéité	< 0,001	< 0,001
Diabète sucré *						
Non	5191	40001,7	208/48,4		4,3 (3,7-4,9)	39,9 (33,1-47,5)
Oui	307	2200,8	39/5,4		7,3 (5,2-10)	152,9 (101,7-217,9)
				<i>p</i> -value d'hétérogénéité	0,003	< 0,001
Splénectomie*						
Non	5452	41911,9	235/52,8		4,4 (3,9-5,1)	43,5 (36,5-51,1)
Oui	46	290,6	12/0,9		12,7 (6,6-22,2)	380,5 (180,9-688,9)
				<i>p</i> -value d'hétérogénéité	< 0,001	< 0,001
Pathologie rénale chronique*						
Non	5410	41614	230/52,7		4,4 (3,8-5)	42,6 (35,7-50,2)
Oui	88	588,5	17/1,1		16,1 (9,4-25,8)	270,9 (150,3-444,6)
				<i>p</i> -value d'hétérogénéité	< 0,001	< 0,001

Abréviations : N : nombre de survivants, Gy : Gray, SNC : Système nerveux central, PA : Personne-Année, Obs/Att : Observés/Attendus, SHR : Ratio standardisé d'hospitalisation, ERA : Excès de risque absolu pour 10 000 PA. *comorbidités avant l'hospitalisation cardiaque,

Le taux d'hospitalisation selon le type de tumeur primaire diffère peu d'une tumeur à l'autre, avec des SHR situés entre 2 et 5 pour la majorité, sauf pour les lymphomes (SHR= 6,8 ; 95 % IC= 5,4–8,5) et les neuroblastomes (SHR= 6,0 ; 95 % IC= 4,0–8,6). Les ERA pour 10 000 PA sont aussi particulièrement élevés pour les lymphomes (ERA= 93,5 ; 95 % IC= 70,6–120,7), et les sarcomes des tissus mous (56,5 (95 % IC= 34,1–85,6). Chez les femmes, les taux d'hospitalisations sont plus élevés que chez les hommes (SHR=7,7 ; 95 % IC=6,3 – 9,4 vs. SHR= 3,6 ; 95 % IC= 3,1 – 4,3, respectivement). De plus, dans un modèle multivarié, le risque d'hospitalisation pour cause cardiaque était 2,3 fois (95 % IC= 1,8 – 3,0) plus élevé chez les femmes que chez les hommes (**Tableau 14**).

Un risque plus élevé d'hospitalisations pour cause cardiaque a été constaté chez les patients diagnostiqués d'un premier cancer entre l'âge de 4 et 8 ans, par rapport à la population française (SHR=5,4 ; 95 % IC= 4,2 - 6,9). En comparaison aux patients âgés de plus de 14 ans au diagnostic du premier cancer, et après ajustement, le risque d'hospitalisation pour cause cardiaque des patients diagnostiqués avant l'âge de 3 ans est 2,0 fois plus élevé (95 % IC= 1.0-3.7). Ce risque diminue légèrement pour les

patients pour les patients diagnostiqués entre 4 et 8 ans (RR=1,7 ; 95 % IC= 1,0–3,1) (**Tableau 14**).

4.3.3.3 Rôles des doses de traitement

Le **Tableau 14** décrit l'ensemble des risques associés aux hospitalisations pour cause cardiaques en fonction des doses de traitement et des comorbidités chez les survivants de la FCCSS.

Les patients traités par chimiothérapie ont un taux d'hospitalisation (SHR= 5,7 ; 95 % IC= 5,0 – 6,6) supérieur à celui des patients non exposés à la chimiothérapie. Ainsi, ces patients ont un risque ajusté 1,7 fois (95 % IC= 1.1 - 2.5) plus élevé d'être hospitalisé pour cause cardiaque que ceux qui n'ont pas été traités par chimiothérapie. Pour ce qui est des dose cumulées d'anthracyclines, le risque ajusté est 3,4 fois (95 % IC= 2,2 – 5,1) plus élevé chez les patients ayant reçu au moins 360 mg/m² comparé à ceux qui n'en ont pas reçu. Le SHR pour les patients appartenant à cette catégorie de dose très élevée est de 9,1 (94% IC= 6,6 – 12,2), et leur ERA de 121,2 (95 % IC= 84,2 – 168,2), soit largement supérieur aux valeurs retrouvées chez les patients n'ayant pas été traités par anthracyclines.

En observant les dose moyenne de radiations cardiaques, il s'avère que pour les patients dont le cœur n'est pas exposé aux radiations, le SHR est de 3,9 (95 % IC= 4,1–4,8), et que celui-ci augmente pour des doses plus élevées, jusqu'à atteindre un SHR à 30,8 (95 % IC= 19,9–45,4) chez les patients exposés à plus 30 Gy. En outre, une augmentation du risque d'hospitalisation est observé à partir des doses supérieures à 15 Gy : les patients ayant reçu entre 15 et 30 Gy et ceux exposés à plus de 30 Gy, voient leur risque d'hospitalisations être multiplié respectivement par, 3,5 (94% IC= 2,4–5,3) et 13,2 (95 % IC= 8,1–21,7) comparé aux patients non exposés aux radiations cardiaques.

Tableau 14. Effet des traitements et des comorbidités sur les hospitalisations cardiaques dans la cohorte FCCSS entre 2011 et 2018

Caractéristiques	Risque Relatif (95 % IC)	p-value
Genre		<i><0,0001</i>
Masculin	Ref	
Féminin	2,3 (1,8-3,0)	
Année de diagnostic du premier cancer		0,1
Avant 1970	Ref	
[1970-1980[1,3 (0,5-3,4)	
[1980-1990[1,2 (0,7-2,1)	
[1990 -2000]	1,7 (0,9-3,6)	
Âge atteint au 01/01/2011 (années)		0,07
<20	Ref	
[20 -30[2,1 (0,8-5,1)	
[30 -40[1,2 (0,4-3,2)	
[40 -50[1 (0,3-3,3)	
≥50	0,9 (0,2-3,6)	
Âge au diagnostic du premier cancer (années)		0,02
]0-3]	2,0 (1- 3,7)	
[4-8]	1,7 (1-3,1)	
[9-11]	1,6 (0,9-2,8)	
[12-14]	0,9 (0,5-1,5)	
≥ 15	Ref	
Chimiothérapie		0,02
Non	Ref	
Oui	1,7 (1,1-2,5)	
Doses cumulées d'antracyclines (mg/m²)		<i><0,0001</i>
Aucune	Ref	
]0-250[1,4 (0,9-2,2)	
[250-360[2,2 (1,4-3,5)	
≥ 360	3,4 (2,2-5,1)	
Doses cumulées de vinca alcaloïdes (g/m²)		0,01
Aucune	Ref	
]0-11,5[0,5 (0,3-0,8)	
[11,5-24[0,6 (0,4-0,9)	
≥ 24	0,7 (0,5-1,1)	

Diabète sucré *			<i><0,0001</i>
	Non	Ref	
	Oui	2,2 (1,5-3,1)	
Pathologie rénale chronique*			<i><0,0001</i>
	Non	Ref	
	Oui	3,4 (2-5,6)	
Dose moyenne de radiations au cœur (Gy)			<i><0,0001</i>
	0	Ref	
]0 -5]	1,0 (0,7-1,4)	
]5 -15]	1,2 (0,7-2,0)	
]15 -30]	3,5 (2,4-5,3)	
	> 30	13,2 (8,1-21,7)	
Volume (%) du cœur ayant reçu au moins 30 Gy			<i><0,01</i>
	0	Ref	
]0 -10]	1,3 (0,7-2,4)	
]10 -50]	3,0 (1,9-5,0)	
	> 50	13,3 (8,9-19,9)	

Abréviations : Gy : Gray ; IC : Intervalle de confiance; Ref : Référence ; *: comorbidités avant l'hospitalisation cardiaque.

Le Risque Relatif est calculé à partir d'un modèle de Poisson multivarié incluant toutes les variables ajusté sur le genre, l'âge et l'année de diagnostic du cancer primaire et l'âge atteint.

Chez les patients dont moins de 10% du cœur est irradié par des doses d'au moins 30 Gy, le SHR est de 4,5 (95 % IC= 2,1-8,2), et augmente à 9,8 (95 % IC= 5,8-15,5) et 32,3 (95 % IC= 22,7-44,5) pour des volumes cardiaques situés respectivement entre 10 et 50 %, et de plus de 50 %. Dans un modèle multivarié, les patients dont la V30 est comprise entre 10 et 50 %, et les patients dont la V30 est supérieure à 50 % ont respectivement un RR de 3,0 (95 % IC= 1,9-5,0) et de 13,3 (95 % IC= 8,9-19,9) par rapport aux patients dont la V30 est nulle.

4.3.3.4 Rôle des comorbidités

Il ressort des analyses que pour les patients ayant des antécédents de diabète sucré antérieurs à l'hospitalisation, le SHR est de 7,3 (95 % IC= 5,2 – 10,0) et l'ERA est de 152,9 pour 10000 PA (95 % IC=101,7-217,9). Les patients atteints de cette comorbidité

ont un RR de 2,2 (95 % IC= 1,5 - 3,1) par rapport aux autres. Quant aux patients atteints de maladies rénales chroniques, leur SHR esy de 16,1 (95 % IC=9,4-25,8), tandis que l'ERA pour 10000 PA est de 270,9 (95 % IC=150,3-444,6). Les hospitalisations pour cause cardiaque présentent un RR ajusté de 3,4 (95 % IC=2,0-5,7) par rapport aux patients qui ne présentent aucune maladie rénale chronique. Aucune association significative n'a été trouvée entre l'hospitalisation pour cause cardiaque et les antécédents de tumeurs malignes secondaires, ni avec la splénectomie.

4.4 DISCUSSION

Ce chapitre rapporte l'effet des doses de traitements anticancéreux reçus par les survivants lors du traitement d'un premier cancer de l'enfant sur le développement de pathologies cardiaques (tous grades, et grade ≥ 3), d'insuffisance cardiaque, et sur les hospitalisations pour cause cardiaque dans la cohorte FCCSS. Ce travail confirme que les doses de radiations reçues au cœur et des doses cumulées d'anthracyclines sont fortement associées à ces trois événements, en augmentant leurs risques respectifs. Pour l'évènement hospitalisations cardiaques, le rôle des comorbidités a aussi été évalué, mettant en évidence une augmentation du risque d'hospitalisation pour les patients ayant un diabète sucré, et une pathologie rénale chronique. De plus, les patients présentant des antécédents de pathologies cardiaques dont la valvulopathie, l'ischémie, et les trouble du rythme et de la conduction, présentent un risque d'insuffisance cardiaque largement supérieur comparé aux patients ne présentant aucun antécédent cardiaque.

Le traitement d'un cancer de l'enfant par anthracyclines et/ou radiothérapie incluant le cœur est associé à un risque très élevé de pathologies cardiaques tardives. Les patients exposés aux anthracyclines, même à faibles doses présentent un risque élevé de pathologies cardiaques de tous grades, de grade ≥ 3 , et d'insuffisance cardiaque en comparaison aux patients non exposés aux anthracyclines. Ce risque augmente à mesure que les doses d'anthracyclines auxquels les survivants ont été exposés augmentent, avec un risque de pathologie cardiaque tous grades multiplié par 4,4 (95 % IC= 3,2 - 6,0) chez les patients ayant reçu entre 250 et 360 mg/m²

d'anthracycline, et par 5,9 (95 % IC= 4,6 - 7,6) chez les patients traités à des doses supérieures à 360 mg/m². Des résultats similaires ont été observés pour les risques associés aux pathologies cardiaques de grade ≥ 3 et aux insuffisances cardiaques. Pour des doses moyennes d'irradiation du cœur supérieures à 35 Gy, le risque de pathologie cardiaque est multiplié par 16,6 (95 % IC= 11,1- 24,9), et le risque d'insuffisance cardiaque par 16,8 (95 % IC= 9,3–30,3) comparé aux patients non exposés aux radiations cardiaques. De plus, lorsque n'importe quel volume du cœur était exposé à des doses de radiations d'au moins 20 Gy, le risque pathologie cardiaque tous grades, grade ≥ 3 et d'insuffisance cardiaque était augmenté de manière modérée (csHR < 3) pour des volumes de moins de 50 %, et très élevée (csHR autour de 8 pour les pathologies cardiaques, et 9 pour les insuffisances cardiaques) pour les volumes au-dessus de 50 % en comparaison aux patients non exposés à la radiothérapie. Une étude Pan-Européenne a rapporté que le risque d'insuffisance cardiaque pour des doses moyennes de radiothérapie au cœur allant de 5 à 15 Gy était cinq fois plus élevé que les survivants qui n'ont pas reçu de radiothérapie cardiaque, tout particulièrement pour un volume de plus de 50 % du cœur irradié (de Baat et al., 2023). Cependant, dans notre étude, un effet significatif des radiations n'est observé qu'à partir des doses supérieures à 15 Gy, de sorte qu'une irradiation de plus de 50 % du cœur multiplie par 7,3 le risque d'insuffisance cardiaque. Ces résultats confirment la survenue de toxicité pour des doses de radiation au cœur > 15 Gy , même pour de petits volumes cardiaques (Bates et al., 2019; Haddy et al., 2016; Mansouri et al., 2019), et pour des doses de radiation au cœur comprises entre 5 et 15 Gy pour des plus grands volumes irradiés (de Baat et al., 2023).

Nous avons aussi mis en évidence un risque 2,2 fois (95 % IC= 1,4–3,5) plus élevé de pathologie cardiaque grade ≥ 3 , pour les patients traités à des doses comprises entre 5 et 15Gy comparé aux patients n'ayant pas reçu de radiations au cœur. Ce résultat confirme celui obtenu dans la cohorte américaine des survivants d'un cancer de l'enfant CCSS (Childhood Cancer Survivor Study) lorsque le cœur a été exposé à des doses de radiations comprises entre 0,1 et 9,9 Gy, où le risque de pathologie cardiaque grade ≥ 3 était multiplié par 1,6 par rapport à ceux qui n'ont pas reçu de radiothérapie sur le

cœur (Bates et al., 2019). Un effet des doses faibles à modérées de radiations cardiaques sur le risque de pathologie cardiaque tous grades a été observé dans notre étude. Ce risque est multiplié par 1,7 (95 % IC= 1,2 – 2,4) chez les patients ayant reçu entre 1 et 5 Gy au cœur comparé aux patients non irradiés. A notre connaissance, cette étude est la première à mettre en évidence un effet des doses de radiations cardiaques comprises entre 1 et 5 Gy. Un autre résultat important de ce travail est que pour des patients ayant reçu au moins 5 Gy pour un volume du cœur irradié situé entre 10 et 50 %, le risque de pathologie cardiaque tous grades était augmenté de 1,9 fois (95 % IC= 1,3-2,9) comparé aux patients n'ayant pas été traités par anthracyclines, alors que cette augmentation n'a pas été observée pour les pathologies cardiaques de grade ≥ 3 . De plus, une association significative a été mise en évidence entre l'augmentation du risque de pathologies cardiaques tous grades, et les doses allant de 5 à 20 Gy de radiations au cœur quel que soit le volume irradié, comparé aux patients n'ayant pas été exposés à la radiothérapie. Contrairement à l'étude de Bates et al. (2019) qui rapporte une association similaire pour les pathologies cardiaques de grades ≥ 3 , cette dernière n'a été retrouvée significative dans notre étude que pour les grands volumes du cœur irradiés (>50 %). Cette différence peut être expliquée par le nombre d'évènements cardiaques de grades ≥ 3 , moins important dans notre étude (658 vs. 316). L'ensemble de ces résultats montrent l'importance de la prise en compte des volumes du cœur irradiés, même à petits volumes, et confirment que la dose moyenne de radiations reçue par le cœur donne une image incomplète du risque d'insuffisance cardiaque. Ces résultats ont déjà été mis en évidence auparavant chez les survivants à 5 ans d'un lymphome de Hodgkin de l'adolescent ou de l'adulte (van Nimwegen et al., 2017), et chez les survivants d'un cancer de l'enfant (Bates et al., 2019; de Baat et al., 2023; Mansouri et al., 2019).

Un autre point important de cette étude est que le risque de pathologie cardiaque dans cette population n'est pas seulement lié au traitement par anthracyclines. En effet, le risque de pathologie cardiaque associé à des doses de radiothérapie supérieures à 25 Gy est similaire chez les patients traités ou non par anthracyclines. Même si les résultats obtenus dans notre étude diffèrent en termes de risques par rapport à l'étude

Euro2K (Haddy et al., 2016), les tendances restent similaires. Ces différences peuvent être attribuées au faible nombre d'évènements dans les strates utilisées par l'étude précédente. Cependant, dans le cas précis des insuffisances cardiaques, les patients non exposés aux anthracyclines et recevant en moyenne moins de 15 Gy au cœur ne présentent pas de risque significatif d'insuffisance cardiaque, et l'administration d'anthracyclines semble jouer un rôle plus important que les radiations cardiaques, avec un risque d'insuffisance cardiaque 2,4 fois plus élevé chez les patients traités par anthracyclines par rapport aux patients n'y étant pas été exposés pour des doses moyennes reçues au cœur supérieures à 15 Gy. Nous avons pu observer que les patients traités par anthracyclines présentent un risque de base d'insuffisance cardiaque très élevé (cHR= 15,2 (95 % IC= 7,1–32,5), qui a augmenté proportionnellement aux doses moyennes d'irradiation du cœur.

Par ailleurs, nous avons observé une augmentation du risque d'insuffisance cardiaque chez les patients ayant développé d'autres types de pathologies cardiaques au cours du temps (pathologies intermédiaires). La prise en compte de ces événements ne modifient pas l'effet des doses cumulées d'anthracyclines ni l'effet des doses moyennes de radiations au cœur <25 Gy. En revanche, le risque associé aux doses moyennes de radiations au cœur > 25 est atténué. Cela pourrait suggérer que l'effet des doses moyennes de radiations au cœur > 25 Gy est médié par l'effet de ces pathologies intermédiaires. Pour confirmer cette hypothèse, une analyse de médiation pourrait être envisagée pour la suite de ce travail.

Il est important de souligner que le risque de pathologie cardiaque induit par les radiations augmente très fortement avec le temps, avec une incidence cumulée à 40 ans de suivi très élevée chez les patients recevant des doses de radiations >25Gy (37% chez les patients non traités par anthracyclines et 33% chez ceux traités par anthracyclines) comparé aux patients non traités par radiothérapie ni anthracyclines (Incidence cumulée à 40 ans de suivi = 1,9% (95 % IC= 1,2 – 3,1). Cela pose un réel problème quant à la détérioration qualité de vie de ces patients, dont l'âge médian au diagnostic de la pathologie cardiaque était de 32 ans, ce qui est très jeune en

comparaison à la population générale française (fréquence élevée dès 50 ans, (Ducimetière, 2004)).

Dans notre étude, nous avons pu observer une diminution significative du risque de pathologie cardiaque chez les patients traités après 1970. Ce résultat est rassurant car suggère que les effets cardiotoxiques des traitements du cancer de l'enfant tendent à diminuer dans les protocoles de traitements plus récents. En effet, au cours des dernières décennies, les efforts se sont concentrés sur l'amélioration continue, ou du moins le maintien, de la guérison et de la survie des patients atteints de tumeurs malignes pédiatriques, tout en ajustant l'intensité des protocoles thérapeutiques en fonction du pronostic et de la gravité du cancer primaire, et en ne perdant pas de vue l'objectif de minimiser les effets à long terme de ces traitements. Cela s'est traduit par une réduction des doses de radiations et l'administration de traitements cardioprotecteurs pour les patients atteints de lymphome de Hodgkin (Lo et al., 2024). Cependant, les risques restent toujours élevés, même si les doses moyennes de radiations administrées actuellement sont moins fortes et mieux ciblées. Dans notre étude, nous avons montré que même à des doses de radiations comprises entre 1 et 5 Gy, il existe un risque faible mais significatif de pathologies cardiaques (csHR=1,7 ; 95 % IC= 1,2- 2,4), et que même pour des volumes irradiés moyens (10 à 50 %) et à des doses comprises entre 5 et 20 Gy un faible risque subsiste (csHR=2,3 ; 95 % IC= 1,2- 4,2). Nous avons également observé le risque d'insuffisance cardiaque lié aux anthracyclines pour des doses ≤ 250 mg/m² est 2,4 fois plus élevé que chez les patients non exposés aux anthracyclines.

Pour ce qui est des hospitalisations cardiaques, nous avons observé que chez les survivants d'un cancer de l'enfant ce nombre est 4,6 fois (95 % IC= 4,0 - 5,2) plus élevé que le nombre attendu en population générale. Cela revient à 45,8 hospitalisations supplémentaires (95 % IC= 38,7 - 53,6) pour 10000 personnes-années. Les femmes ayant survécu d'un cancer dans l'enfance ont un risque 7,7 fois plus élevé d'être hospitalisé pour cause cardiaque que la population générale, alors que pour les hommes ce risque est multiplié par 3,6 par rapport à la population générale. Le risque

d'hospitalisation pour cause cardiaque est 2,3 fois élevé chez les femmes que chez les hommes. Par ailleurs, les survivants diagnostiqués avant l'âge de 9 ans présentent un risque plus élevé d'hospitalisation cardiaque par rapport à la population générale française, et les survivants ayant été diagnostiqué d'un lymphome ou d'un neuroblastome comme tumeur primaire ont un taux d'hospitalisation cardiaque plus élevé que les survivants diagnostiqués d'autres tumeurs. Cela peut s'expliquer par l'administration de doses d'anthracyclines et de radiations incluant le cœur élevées lors du traitement de ce type de cancers, particulièrement avant la prise de conscience des effets à long termes associés aux traitements anticancéreux. Le taux d'hospitalisation cardiaque après traitement par chimiothérapie était largement supérieur à celui des patients n'ayant pas été exposés à la chimiothérapie, et le risque associé à la chimiothérapie était 1,7 fois plus élevé chez les patients traités par chimiothérapie à ceux qui n'ont pas été traités par chimiothérapie. De plus, les patients traités par des doses modérées d'anthracyclines (entre 250 et 360 mg/m²) avaient un risque 2,2 fois (95 % IC= 1,4- 3,5) plus élevé d'être hospitalisé que les patients non exposés aux anthracyclines. En observant les dose moyenne de radiations cardiaques, une augmentation du risque d'hospitalisation est observé à partir des doses supérieures à 15 Gy, avec un risque multiplié par 3,5 (94% IC= 2,4–5,3 chez les patients ayant reçu entre 15 et 30 Gy et par 13,2 (95 % IC= 8,1–21,7) chez ceux exposés à plus de 30 Gy. Nous avons également mis en lumière que les survivants ayant reçu une irradiation cardiaque moyenne d'au moins 30 Gy présentaient un risque très élevé d'hospitalisation, même pour des petits volumes (< 10 %) du cœur irradié, et que ce risque augmente à mesure que le volume du cœur irradié est plus grand.

Même si pour des doses d'anthracyclines comprises entre 0 et 250 mg/m², et de doses moyennes de radiations au cœur entre 5 et 15 Gy le risque d'hospitalisation cardiaque n'est pas significatif (RR= 1,4 (95 % IC= 0,9 – 2,2) et RR = 1,2 (95 % IC= 0,7 - 2,0), respectivement), nous observons un risque similaire à celui associé aux pathologies cardiaques tous grades et grades ≥ 3 (csHR =1,8 (95 % IC= 1,2 - 2,6) et csHR= 2,2 (95 % IC= 1,4 - 3,5), respectivement). Cette différence peut s'expliquer par des différences dans la population étudiée. En effet, même si nous travaillons pour les deux études sur

la cohorte FCCSS, dans l'étude des hospitalisations cardiaques, seul 83 % des patients sont chaînés au SNDS donc inclus dans les analyses.

Outre les points forts de notre étude, qui nous a permis de fournir des estimations précises du risque de pathologie cardiaque et l'insuffisance cardiaque pour les doses moyennes de radiations inférieures à 15 Gy, et les volumes du cœur irradiés inférieurs à 10 %, notre étude présente certaines limites. Les risques de pathologies cardiaques et d'insuffisance cardiaque peuvent être sous-estimés du fait que certains patients (et même leurs médecins) peuvent ne pas être conscients de problèmes latents s'ils ne font pas l'objet d'examens précoces, et que les symptômes cardiaques peuvent être attribués à tort à des problèmes non cardiaques. De plus, au vu du manque d'information sur certains facteurs de risques cardiaques traditionnels comme le tabagisme, l'origine ethnique, l'indice de masse corporelle et l'activité physique, dans notre cohorte, ceux-ci n'ont pas été pris en compte dans nos analyses. Cependant, une étude cas-témoins précédemment réalisée sur la cohorte FCCSS a testé l'association de certains de ces facteurs avec l'insuffisance cardiaque et n'a pas trouvé de résultats significatifs (Mansouri et al., 2019). Par ailleurs, certaines hospitalisations antérieures à 2011 ont été omises, ce qui peut induire une sous-estimation des risques associés aux hospitalisations pour cause cardiaque.

En conclusion, cette étude a complété les résultats de l'étude précédente de Haddy et al. (2016) et a mis en lumière l'importance de la prise en compte des indicateurs dose-volume dans l'analyse des effets des traitements des cancers de l'enfant sur l'occurrence de pathologies cardiaques. Cette étude met en exergue la nécessité de la prise en compte des indicateurs doses-volumes dans le suivi des survivants d'un cancer dans l'enfance. Par ailleurs, un suivi plus étroit peut être recommandé chez les survivants à haut risque, c'est-à-dire traités par des doses fortes de radiations et d'anthracyclines et ayant des comorbidités à risque comme le diabète sucré, ou pathologies rénales chroniques.

Chapitre 5. INFLUENCE DE LA GENETIQUE SUR L'APPARITION DE PATHOLOGIES CARDIAQUES CHEZ LES SURVIVANTS D'UN CANCER DANS L'ENFANCE : REVUE SYSTEMATIQUE DE LA LITTERATURE

Ce chapitre présente une revue systématique de littérature décrivant l'ensemble des articles publiés avant avril 2023 traitant des marqueurs génétiques associés aux pathologies cardiaques chez les survivants d'un cancer de l'enfant. Ce travail a été soumis en Décembre 2023 à la revue *The Pharmacogenomics Journal*, et est en attente de réponse.

5.1 CONTEXTE ET OBJECTIFS DE LA REVUE SYSTEMATIQUE DE LA LITTERATURE

Bien que l'augmentation du risque de pathologies cardiaques chez les patients survivants traités par anthracyclines soit connue, la radiothérapie peut également contribuer au développement de troubles cardiaques. Ces troubles peuvent se manifester par l'apparitions de pathologies telles que les cardiomyopathies, les maladies coronariennes, les valvulopathies, les troubles du rythme ou de la conduction, ou encore les péricardites (Haddy et al., 2016; Mansouri et al., 2019; Mulrooney et al., 2009). Plusieurs autres facteurs de risque associés aux évènements cardiotoxiques ont été identifiés parmi lesquels le jeune âge au moment de l'exposition aux traitements, le sexe féminin, l'origine ethnique africaine, le statut tabagique, le manque d'exercice, et un dysfonctionnement cardiaque préexistant (Bansal et al., 2017). Cependant, la capacité des modèles incluant ces facteurs à prédire les patients à risque de cardiotoxicité reste limitée (Chow et al., 2015; Leerink et al., 2020). En effet, chez des survivants d'un cancer dans l'enfance traités de façon similaire, une variabilité interindividuelle de la susceptibilité à la cardiotoxicité a été observée, suggérant l'existence de facteurs génétiques modulant cette susceptibilité (Aminkeng et al., 2016; Clemens et al., 2018).

Une des premières études recherchant des prédispositions génétiques associées aux

pathologies tardives chez les patients ayant survécu d'un cancer dans l'enfance a été publié en 2000, et portait sur les effets ototoxiques des traitements par cisplatine (Peters et al., 2000). En 2008, une des premières études s'intéressant au potentiel lien entre des prédispositions génétiques et l'apparition de cardiomyopathies après traitement d'un cancer de l'enfant par anthracyclines a été publiée (Blanco et al., 2008). Par la suite, plusieurs études ont examiné le rôle de la génétique dans la cardiotoxicité induite par les anthracyclines chez les patients traités pour un cancer pédiatrique. Ces études sont basées sur différentes approches de génotypage et visent à identifier les séquences d'ADN (polymorphisme ou variants mononucléotidiques (SNP ou SNV)) associées à la cardiotoxicité induite par les traitements anticancéreux.

En 2021, une revue publiée par Berkman et al., a étudié les prédispositions génétiques à la cardiotoxicité induite par les anthracyclines, mais cette revue n'était pas spécifique aux cancers pédiatriques, qui sont histopathologiquement et moléculairement différents des cancers chez l'adulte. De plus, les recommandations de suivi des survivants d'un cancer pédiatrique actuelles ne préconisent pas l'utilisation des tests pharmacogénomiques dans la prédiction du risque de cardiomyopathie chez les patients traités pour un cancer (Ehrhardt et al., 2023).

Il est donc nécessaire de dresser une liste complète des gènes impliqués dans les effets tardifs des traitements anticancéreux, et de concentrer sur les survivants d'un cancer de l'enfant, afin afin d'améliorer la performance des modèles de prédiction de la susceptibilité aux toxicités des traitements, d'identifier les individus à haut risque de pathologies cardiaques tardives, et de prévenir la cardiotoxicité en adaptant les thérapies (Berkman et al., 2021; Clemens et al., 2018). Nous avons donc effectué une revue systématique de la littérature dont l'objectif est de résumer les résultats publiés sur les facteurs génétiques associés aux pathologies cardiaques chez les survivants d'un cancer dans l'enfance et de mettre en évidence les interactions potentielles des gènes et/ou variants avec les traitements anticancéreux.

5.2 MATERIEL ET METHODE

Le protocole de cette revue systématique a été enregistrée sur PROSPERO (CRD N°2023388689). Les données ont été rapportées conformément aux recommandations PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) (Page et al., 2021).

5.2.1 Stratégie de recherche

Une recherche de la littérature a été menée à partir des librairies en ligne PubMed/MEDLINE et Scopus afin d'identifier les études publiées jusqu'en Avril 2023. Dans un second temps, nous avons consulté une à une les références bibliographiques des articles éligibles afin d'identifier d'autres études pertinentes. Les mots-clés ont été élaborés à partir de termes de Medical Subject Headings (MeSH) et d'autres termes en texte libre (stratégie complète en **Annexe B**)

Nous nous sommes appuyés sur le cadre PICOS (Population, Intervention, Comparaison, Résultat, Conception de l'étude) (Amir-Behghadami & Janati, 2020) pour construire la stratégie de recherche. La population d'étude comprend les patients ayant survécu d'un cancer dans l'enfance ; l'intervention concerne les traitements du cancer administrés : anthracyclines et/ou radiations reçues au cœur ; la pathologie étudiée inclut l'ensemble des pathologies cardiaques existantes ; et le design de l'étude comprend les études observationnelles de type cas-témoins ou de cohorte (**Tableau 15**).

Tableau 15. Stratégie de recherche des articles sur NCBI/PubMed et Scopus

PICOS		Libellé de la stratégie de recherche
Population	Childhood Cancer Survivors	#1: Cancer OR Oncology OR neoplasms OR malignant OR tumor OR carcinoma OR leukemia OR retinoblastoma OR sarcoma OR osteosarcoma OR Wilms tumor OR nephroblastoma OR neuroblastoma OR brain neoplasms OR CNS OR central nervous system OR lymphoma OR Hodgkin disease OR Hodgkin Diseases #2: Cancer survivors OR "long-term survival" OR "followed up" OR Survivor* OR childhood cancer survivors #1 AND #2
Intervention	Radiotherapy Chemotherapy	#1: Radiotherapy OR Radiations OR "Dose-volume" #2: Chemotherapy OR Antimetabolites OR alkylating OR Topoisomerase OR "Vinca Alkaloid" OR "Tyrosine kinase" OR Anthracyclines OR Daunorubicin OR Adriamycin OR Doxorubicin OR Mitoxantrone OR Antineoplastic agent #1 OR #2
Comparator	None	
Outcome	Genetic factors associated with Cardiotoxicity	#1: Heart diseases OR Heart Disorders OR Cardiac Diseases OR Cardiac Disorders OR cardiovascular OR Cardiotoxicit* OR Cardiomyopathy* OR Heart failure OR arrhythmia OR Conduction disorders OR Valvular diseases OR heart valve disease OR ischemia OR myocardial infarction OR Coronary artery disease OR Ischemic heart diseases OR Angina OR Heart attack OR Pericarditis OR cardiac toxicity #2: genetics OR Genotype OR gene expression OR Transcript* OR "SNPs" OR "single nucleotid polymorphism" OR "copy number variations" OR "Sequencing" OR "rna seq" OR "GWAS" OR "genetic association studies" OR "Biomarkers" OR "genome-wide association study" #1 AND #2
Study Design	Cohorts, case-control	Cohort Stud* OR Observational Stud* OR Case control studi* OR Prospective Stud* OR Retrospective Stud*

5.2.2 Procédure de sélection

Tous des articles identifiés ont été examinés indépendamment par deux investigateurs lors de la recherche initiale sur la base du titre et du résumé. Un seul investigateur a analysé le texte intégral des articles potentiellement éligibles.

Les principaux critères d'inclusion des études étaient les suivants : (1) les patients sont traités pour un cancer pédiatrique avant l'âge de 21 ans, (2) les patients ont été traités par chimiothérapie et/ou radiothérapie, (3) les patients ont développé une pathologie cardiaque après le traitement du cancer, (4) la pathologie cardiaque est toujours apparente après la fin du traitement du cancer, (5) l'étude porte sur des facteurs génétiques associés à l'apparition de la pathologie cardiaque suite au traitement anti-cancéreux, (6) l'étude est observationnelle (étude de cohorte, étude observationnelle ou étude cas-témoins). Ont été exclues (1) les études portant sur des patients traités pour un cancer chez l'adulte, (2) les études ne s'intéressant pas aux pathologies cardiaques, (3) les études incluant des patients ayant une pathologie cardiaque préexistante, (4) les études ne rapportant pas de résultats génétiques et (5) les études incluant moins 100 patients.

5.2.3 Extraction des données et contrôle qualité des études

Les informations relatives aux études éligibles ont été extraites du texte principal et des documents complémentaires à l'aide d'un formulaire d'extraction de données préalablement construit présenté dans le **Tableau 16**.

La qualité des études a été évaluée à l'aide de l'outil Q-Genie (Sohani et al., 2015) conçu pour apprécier la qualité des études d'association génétique, qui comprend un total de 11 caractéristiques notées de "faible" (*note* = 1) à "excellent" (*note* = 7). Le score de qualité est la somme de ces notes par étude, tel qu'un score ≤ 35 indique une étude de mauvaise qualité, un score > 35 et ≤ 45 est associé à une étude de qualité modérée, et un score > 45 désigne une étude de bonne qualité.

Tableau 16. Formulaire d'extraction des données

Reference citation of the article	
Study design	<input type="checkbox"/> Case-control study <input type="checkbox"/> Cohort study
Sequencing method	<input type="checkbox"/> candidate gene, <input type="checkbox"/> genome-wide association studies (GWAS) <input type="checkbox"/> whole-exome sequencing (WES) <input type="checkbox"/> whole-genome sequencing (WGS)
Sample size (case/controls)	
Country of origin and/ or ethnicity	
Cancer type	
Method and follow-up duration	
Cancer treatment information	
Cancer treatment or cancer diagnosis area	
Age at cancer diagnosis	
Cardiac outcome definition	
Definition of cases and controls	
Gene selection method	
Statistical analysis method	
Selected genes and variants, and risk-associated with CD	

5.3 RESULTATS

La stratégie de recherche a identifié un total de 978 études (196 dans PubMed and 781 dans Scopus, et 1 dans les recherches manuelles) parmi lesquelles 56 étaient des doublons. Sur les 921 études restantes, 884 n'ont pas satisfait les critères d'inclusion après vérification des titres et résumés, aboutissant à 38 études éligibles à une lecture entière de l'article.

Finalement, 19 études ont été sélectionnées et 19 études exclues principalement car les études étaient menées sur des patients ayant eu un cancer à l'âge adulte (n=8), ou encore à cause du nombre trop faible de patients inclus dans l'étude (n=4). Le détail de la sélection des études est présenté dans le diagramme de flux **Figure 16**.

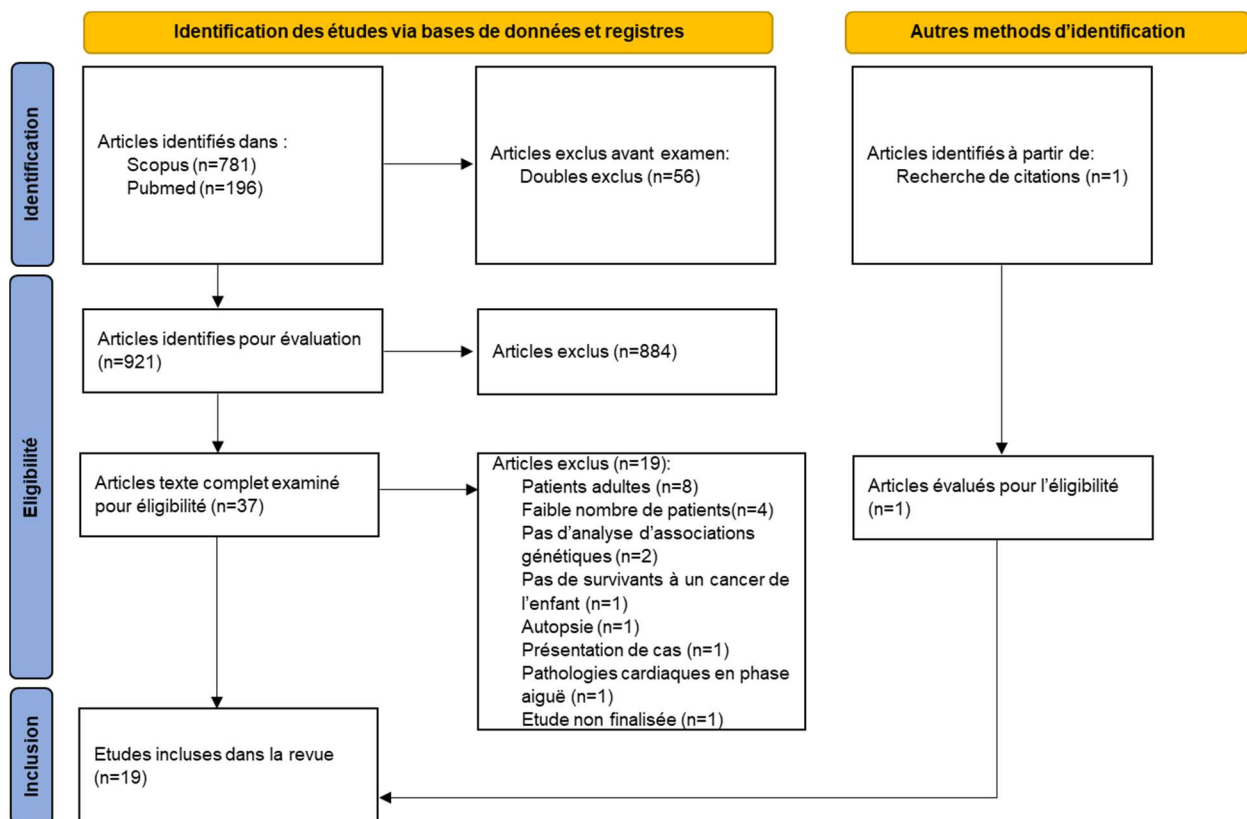


Figure 16. Diagramme de flux PRISMA décrivant la stratégie de recherche et les études incluses et exclues

5.3.1 Description des études

Les 19 études incluses ont été publiées entre 2008 et 2022, et concernent des patients ayant été diagnostiqués d'un cancer avant l'âge de 21 ans, pour divers types de tumeurs malignes : leucémie aigüe lymphoblastique (LAL, n=4) (Krajinovic et al., 2016; Lipshultz, Lipsitz, et al., 2013; Petrykey et al., 2021; Semsei et al., 2012) et leucémie aigüe lymphoblastique et ostéosarcome (n=1) (Sági et al., 2018), tout type de cancer (n=14) (Aminkeng et al., 2015; Blanco et al., 2008, 2012; Chaix et al., 2020; Hildebrandt et al., 2017; Sapkota et al., 2021, 2022; Singh et al., 2020; Visscher et al., 2012, 2013, 2015; X. Wang et al., 2014, 2016, 2022). Ces études sont décrites dans le **Tableau 17**.

Toutes mentionnent l'administration d'anthracyclines dans le traitement du cancer, dont cinq détaillent l'utilisation de doxorubicine, avec les doses cumulées spécifiées pour chaque patient. Quatorze études décrivent l'utilisation de radiothérapie thoracique (incluant le cœur) soit en catégorisant l'exposition de façon binaire (exposé / non-exposé) (n=9), soit en spécifiant les doses reçues en utilisant des paramètres tels que la dose cumulée et la dose moyenne. Aussi, cinq études indiquent l'utilisation de Dexrazoxane, un médicament utilisé pour la prévention de la cardiotoxicité induite par les anthracyclines. Seules huit études ont évalué l'effet des variants génétiques en fonction des doses cumulées d'anthracyclines, et aucune n'a étudié cet effet en fonction des doses de radiations au cœur.

Tableau 17. Caractéristiques des 19 études incluses ordonnées selon le design de l'étude (étude cas-témoins puis étude de cohorte)

Etude	Pays/ Ethnie	Type de tumeur	Age au diagnostic ou au traitement du cancer (années)	Suivi (années)	Traitements (anthracyclines/ radiothérapie)	Méthode de génotypage	Design de l'étude	Taille de la cohorte (cas vs, témoin)
Sági et al., 2018	Hongrie	Leucémie lymphoblastique aiguë et ostéosarcome	Médiane (min-max) LAL: 5,2 (0 - 18) OSC: 13,2 (5 - 18)	Manquant	Anthracyclines (tous les patients)	Gène candi-dat	Etude cas-témoins	661 (20 vs, 641)
Blanco et al., 2012	Amérique du Nord	Tout type de cancer	Médiane (min - max) Cas: 7,3 (0-20,7) Témoins: 7,6 (0 - 21,1)	Médiane (min - max) 9,7 (0,1 - 40,3)	Anthracyclines et/ ou radiothérapie	Gène candi-dat	Etude cas-témoins appariée	487 (170 vs, 317)
Visscher et al., 2013	Pays-Bas et Canada	Tout type de cancer	Médiane (min - max) Cas: 9,9 (0,5-17,0) Témoins: 8,16 (0,5 - 17,7)	Médiane (min - max) 13,8 (0,4 - 31,6)	Anthracyclines +/- radiothérapie	Gène candi-dat	Etude cas-témoins	344 (78 vs, 266)
Singh et al., 2020	Toute ethnicité	Tout type de cancer	Médiane (min - max) Cas: 7,8 (3,8 - 11,5) Témoins: 9,6 (3,3 - 14,8)	Médiane (min - max) 9,3 (1,3 - 17,2)	Anthracyclines +/- ou radiothérapie	Gène candi-dat	Etude cas-témoins appariée	167 (75 vs, 92)
Blanco et al., 2008	Amérique du Nord	Tout type de cancer	Moyenne (SD) Cas: 10,3 (6,5) Témoins: 9,1 (5,8)	≥ 5 ans après le diagnostic du cancer	Anthracyclines +/- radiothérapie	Gène candi-dat	Etude cas-témoins nichée, appariée	140 (29 vs, 111)
Hildebrandt et al., 2017	Toute ethnicité	Tout type de cancer	Moyenne (SD) Cas: 9,2 (4,7) Témoins: 9,3 (5,7)	Moyenne (SD) Cas: 21,2(11,2) Témoins: 15,7(7,6)	Anthracyclines +/- radiothérapie	Gène candi-dat	Etude cas-témoins appariée	108 (46 vs, 62)
Wang et al., 2022	Amérique du Nord: Blancs non-Hispaniques	Tout type de cancer	Médiane (min - max) Cas: 7 (0 - 20) Témoins: 11 (0 -20)	Médiane (min - max) Cas: 22 (0 - 39) Témoins: 31 (14 - 45)	Anthracyclines +/- radiothérapie	GWAS	Etude cas-témoins	1866 (126 - 1740)
Aminkeng et al., 2015	Canada: Ascendants européens et non européens	Tout type de cancer	Médiane (EI) Cas: 8,5 (4 -14) Témoins: 5 (2 - 10)	Manquant	Anthracyclines +/- radiothérapie	GWAS	Etude cas-témoins	376 (54 vs, 322)
Visscher et al., 2015	Pays-Bas et Canada	Tout type de cancer	Médiane (min-max) Cas: 9,9 (0,5-17,0) Témoins: 4,9 (0,1-17,7)	Médiane (min-max) 13,8 (0,4 - 31,6)	Anthracyclines +/- radiothérapie	GWAS	Etude cas-témoins	335 (78 vs, 257)

Wang et al., 2016	Blancs non-Hispaniques	Tout type de cancer	Médiane (min-max) Cas: 7,5 (0 - 20) Témoins: 8,3 (5 - 8)	Médiane (min-max) 11,7 (0,1 - 41)	Anthracyclines et/ ou radiothérapie	GWAS	Etude cas-témoins appariée	331 (112 vs, 219)
Wang et al., 2014	Blancs non-Hispaniques	Tout type de cancer	Médiane (min-max) Cas: 6,9 (0 - 20,2) Témoins: 6,3 (0 - 20,6)	Médiane (min-max) 10,87 (0,1 - 41,0)	Anthracyclines et/ ou radiothérapie	GWAS	Etude cas-témoins appariée	287 (93 vs, 194)
Visscher et al., 2012	Canada: Ascendants européens et non-européens	Tout type de cancer	Médiane (min-max) Cas: 6,1 (0,04-17,6) Témoins: 3,8 (0,05 - 16,9)	Médiane (min-max) 8,2 (0,1 - 18,6)	Anthracyclines +/- radiothérapie	GWAS	Etude cas-témoins	156 (38 vs, 118)
Chaix et al., 2020,	Amérique du Nord	Tout type de cancer	Médiane (EI) Cas: 4,0 (2,0 - 7,0) Témoins: 6,0 (2,0 - 10,0)	Médiane (EI) Cas: 9,0 (6,0-12,3) Témoins: 8,5 (5,0-12,3)	Anthracyclines +/- radiothérapie	WES	Etude cas-témoins nichée	289 (183 vs, 106)
Krajinovic et al., 2016	Canada: Caucasiens	Leucémie aiguë lymphoblastique	Médiane (min - max) 5,0 (1,0 - 18,0)	Médiane (min - max) 8,00 (1,00 - 18)	Anthracyclines +/- radiothérapie	Gène candidat	Etude de cohorte	251
Semsei A., 2012	Hongrie	Leucémie aiguë lymphoblastique	Moyenne (min-max) 5,7 (1,2 - 18,0)	Médiane (min - max) 6,3 (2,4 - 13,7)	Anthracyclines (tous les patients)	Gène candidat	Etude de cohorte	235
Lipshultz S., 2013	Manquant	Leucémie aiguë lymphoblastique à haut risque	59% : <10 ans 41% : ≥10 ans	Médiane (min - max) 7,0 (1,1 - 17,1)	Anthracyclines	Gène candidat	Etude de cohorte	184
Petrykey et al., 2021	Canada: Ascendants Européens	Leucémie aiguë lymphoblastique	Médiane (min - max) 4 (0 - 18)	Médiane (min - max) 13 (3 - 24)	Anthracyclines (tous les patients)	WES	Etude de cohorte	233
Sapkota et al., 2022	EUA : Ascendants Européens; Ascendants Africains;	Tout type de cancer	Médiane (min - max) Cas: 10,5 (0,0 - 20,5) Témoins: 6,9 (0,0 - 23,6)	Manquant	Anthracyclines et/ ou radiothérapie	WGS	Etude de cohorte	1870 (227 - 1643)
Sapkota et al., 2021	EUA: Ascendants Africains	Tout type de cancer	Médiane (min - max) 8,8 (0,1-19,7)	Manquant	Anthracyclines et/ ou radiothérapie	WGS	Etude de cohorte	246 (40 vs, 206)

Abbreviations: SD, Standard Deviation= Ecart-type; EI, Étendue Interquartile; GWAS, Genome-Wide Association Studies ; WGS, Whole Genome Sequencing ; WES, Whole Exome Sequencing

Parmi les 19 études sélectionnées, 14 sont des études cas-témoins (Aminkeng et al., 2015; Blanco et al., 2008, 2012; Chaix et al., 2020; Hildebrandt et al., 2017; Sági et al., 2018; Singh et al., 2020; Visscher et al., 2012, 2013, 2015; Wang et al., 2014, 2016, 2022), dont six études appariées, avec un nombre de cas allant de 38 à 277. Les cinq autres études sont des études de cohorte (Krajinovic et al., 2016; Lipshultz, Lipsitz, et al., 2013; Petrykey et al., 2021; Sapkota et al., 2021, 2022; Semsei et al., 2012), incluant entre 184 et 1870 survivants d'un cancer de l'enfant.

Dans 63% des études (n=12), les résultats ont été reproduits sur des données indépendantes provenant soit d'une cohorte externe (n=4), soit de patients de la cohorte globale non inclus dans l'échantillon analysé. La qualité des 19 études incluses dans cette revue a été évaluée avec 15 études considérées de bonne qualité, et quatre de qualité modérée (**Tableau 18**).

5.3.2 Méthodes de génotypage

Des échantillons biologiques provenant de diverses sources ont été utilisés pour effectuer le génotypage dans les différentes études incluant le sang, la salive, les cellules buccales, et les tissus germinaux. Trois études n'ont pas précisé la source des échantillons biologiques utilisés dans le génotypage (Aminkeng et al., 2015; Krajinovic et al., 2016; Visscher et al., 2012). Quatre différentes méthodes de génotypage ont été utilisées : neuf analyses de gènes candidats, les études d'association pangénomiques (Genome-Wide Association Studies, GWAS, n=6), le séquençage de génome entier (Whole Genome Sequencing, WGS, n=2), et le séquençage d'exome entier (Whole Exome Sequencing, WES, n=2).

Tableau 18. Evaluation de la qualité des études incluses à partir de l'outil Q-GENIE

Author, year	Rationale for study	Outcome of interest	Comparison Group	Technical classification of exposure	Non-technical classification of exposure	Other sources of bias	Sample size and power	A priori planning of analyses	Statistical methods and control for confounding	Testing assumptions and inferences for genetic analyses	Appropriateness of inferences drawn from results	Score
Sàgi et al., 2018	6	6	5	4	0	5	6	5	5	1	5	48
Blanco et al., 2012	4	7	6	3	0	5	4	7	6	2	6	50
Visscher et al., 2013	5	7	5	5	0	4	5	6	6	4	5	52
Singh et al., 2020	6	5	5	2	0	5	3	4	6	4	4	44
Blanco et al., 2008	4	7	5	5	0	5	3	3	4	5	5	46
Hildebrandt et al., 2017	4	4	5	3	0	5	3	5	3	2	5	39
Wang et al 2022	6	6	5	4	0	5	5	6	6	6	5	54
Aminkeng et al., 2015	5	5	6	2	0	5	6	6	5	4	5	49
Visscher et al., 2015	6	6	5	5	0	5	4	6	6	4	6	53
Wang et al., 2016	6	5	6	5	0	5	6	6	5	5	6	55
Wang et al., 2014	5	5	6	4	0	4	4	4	5	4	6	47
Visscher et al., 2012	5	7	6	5	0	4	4	5	6	6	5	53
Chaix et al., 2020	5	6	5	5	0	5	4	6	5	4	6	51
Krajcinovic et al., 2016	6	5	2	1	0	5	4	5	5	3	5	41
Semsei et al., 2012	5	7	2	3	0	4	4	3	5	3	4	40
Lipshultz et al., 2013	6	5	3	5	0	4	4	5	5	3	6	46
Petrykey, 2022	7	6	4	4	0	5	4	5	5	4	6	50
Sapkota et al., 2022	7	5	5	5	0	6	5	5	6	7	6	57
Sapkota et al., 2021	6	6	4	5	0	4	4	6	5	4	5	49

Source: Sohani ZN, et al., Assessing the quality of published genetic association studies in meta-analyses: the quality of genetic studies (Q-Genie) tool. BMC Genet. Dec 2015;16(1):50.

5.3.3 Evaluation des pathologies cardiaques

Bien qu'aucune restriction n'ait été apportée à la stratégie de recherche concernant le type de pathologies cardiaques, seule la cardiomyopathie a été examinée dans l'ensemble des études incluses. Cette dernière a été définie et dénommée de différentes manières en fonction des études. Quatre études font directement référence à la cardiomyopathie (Blanco et al., 2012; X. Wang et al., 2014, 2016, 2022), une étude évoque la cardiomyopathie induite par les traitements anticancéreux (Sapkota et al., 2022) quatre études indiquent une cardiotoxicité induite par les anthracyclines (Aminkeng et al., 2015; Visscher et al., 2012, 2013, 2015), et deux une cardiotoxicité clinique (Chaix et al., 2020; Hildebrandt et al., 2017). Par ailleurs, dans une étude l'insuffisance cardiaque congestive a été considérée comme pathologie d'intérêt (Blanco et al., 2008), cette dernière étant une conséquence de la cardiomyopathie. D'autres auteurs ont défini la cardiomyopathie à partir de la surveillance de paramètres échocardiographiques révélant une diminution de la fraction d'éjection du ventricule gauche (FEVG) et/ou de la fraction de raccourcissement du ventricule gauche (FRVG) dans 5 études (Lipshultz, Lipsitz, et al., 2013; Petrykey et al., 2021; Sági et al., 2018; Sapkota et al., 2021; Semsei et al., 2012). Par souci de simplicité, le terme « cardiomyopathie » est utilisé dans le reste de notre étude pour couvrir les différentes dénominations rapportées ci-dessus.

5.3.4 Analyses statistiques

La mise en évidence d'associations entre des gènes et/ou variants et la cardiomyopathie a été faite de différentes manières en fonction des méthodes de génotypages utilisées et de la définition utilisée pour la cardiomyopathie (dichotomique ou continue). Dans les études gène candidat avec la cardiomyopathie représentée par une variable binaire (cardiomyopathie : oui/non), la régression logistique (non-conditionnelle ou conditionnelle) a été la méthode la plus utilisée sinon des analyses de variance ont été utilisées lorsque la cardiomyopathie est représentée par des valeurs de la fraction d'éjection ou de raccourcissement du ventricule gauche.

La régression linéaire a également été utilisée dans une étude analysant l'évolution des paramètres échocardiographiques tels que la FEVG, et la FRVG. Dans les études GWAS, seuls les évènements binaires (cardiomyopathie oui/ non) ont été analysés, la sélection de gènes/variants associés aux cardiomyopathies s'est faite majoritairement par des régressions logistiques, et une étude a utilisé un modèle de Cox. Quelle que soit la méthode d'analyse utilisée pour ces analyses GWAS, une correction pour multiplicité des tests a été réalisée. Les deux études WGS recensées ont cherché (i) à sélectionner les variants communs par des régressions linéaires ou logistiques avec correction pour tests multiples, ou (ii) à sélectionner des variants rares par une méthode dédiée telle que le SKAT-O (Sequence Kernel Association Test - Optimal unified test). Pour ce qui est des études WES, l'une a utilisé la régression linéaire pour le choix des gènes/variants associés à la cardiomyopathie, tandis que la seconde étude a utilisé des méthodes SKAT, SKAT-O et Burden, plus adaptées à ce type de données.

5.3.5 Synthèse des résultats génétiques

Plusieurs variants dans 61 gènes différents et deux loci ont été identifiés comme étant significativement associées à la cardiomyopathie. Ces gènes sont présentés dans l'**Annexe B 4**. Le **Tableau 19** synthétise les principales associations trouvées dans les différentes études, et le **Tableau 20** résume l'effet des variants génétiques en fonction des doses cumulées d'anthracyclines reçues par les patients.

Tableau 19. Principaux résultats des 19 études incluses concernant les associations génétiques. Les études sont triées selon la méthode de génotypage (gènes candidates, GWAS, WES puis WGS)

Etude	Méthode de génotypage	Evènement cardiaque	Gène ou région	Variant	Effet allélique	OR (95 % IC)	p-values
Sági et al., 2018	Gène candidat (70 SNP dans 26 gènes)	Cardiotoxicité	CYP3A5	rs4646450	TT	7,25 (1,83 - 28,78)	4,81 x 10 ⁻³
			SLC28A3	rs7853758	AA	9,84 (1,73 - 56,02)	1,00 x 10 ⁻²
		Diminution de la FRVG 5 à 10 ans après le diagnostic du cancer	ABCC2	rs3740066	GG/ GA/ AA	39,1 vs, 40,6 vs, 42,4 ^c	7,11 x 10 ⁻⁴
			NQO1	rs1043470	TT/ CC + CT	38,5 vs, 40,6 ^c	5,82 x 10 ⁻³
Blanco et al., 2012	Gène candidat (2 SNP dans 2 gènes)	Cardiomyopathie	CBR3	rs1056892	GG	1,79 (1,08 - 2,96)	0,02
			CBR1	rs9024	GG	0,81 (0,45 - 1,47)	0,49
			Combine CBR3 and CBR1	Combine: rs1056892/ rs9024	GG + GG	1,53 (0,93 - 2,51)	0,09
Visscher et al., 2013	Gène candidat (23 SNP)	Cardiotoxicité induite par les anthracyclines	SLC28A3	rs7853758	A	0,36 (0,22 - 0,60)	1,6 x 10 ⁻⁵
			SLC28A3	rs885004	A	0,34 (0,20 - 0,60)	3,0 x 10 ⁻⁵
Singh et al., 2020	Gene candidate (1 gène)	Cardiomyopathie	GSTM1	Null/ Positive		2,52 (1,2 - 5,2)	0,01
Blanco et al., 2008	Gène candidat (2 SNP dans 2 gènes)	Insuffisance cardiaque congestive	CBR3	rs1056892	GG/ AA	5,63 (0,80 - 39,57)	0,083
					GA/ AA	3,66 (0,64 - 21,07)	0,15
			NQO1	rs1800566	CC/ TT	1,26 (0,14 - 11,39)	0,84
					CT/ TT	0,65 (0,06 - 6,77)	0,72
Hildebrandt et al., 2017	Gène candidat (12 SNP)	Cardiotoxicité	PLCE1	rs932764	G	0,48 (0,27-0,85)	0,012
			ATP2B1	rs17249754	G	0,33 (0,12 - 0,92)	0,034
			PLCE1	rs932764	G	0,36 (0,18 - 0,76)	0,0068
			ATP2B1	rs17249754	G	0,26 (0,07 - 0,96)	0,04
Krajinovic et al., 2016	Gène candidat (33 SNP communs dans 12 gènes)	FEVG	ABCC5	rs7627754	TT/ AA + AT	- 12% ^e	<0,0005
		FRVG	ABCC5	rs7627754	TT/ AA + AT	- 8% ^e	0,001
Lipshultz S et al., 2013	Gène candidat (2 SNP dans 1 gène)	LVSF 4 à 7 après le diagnostic du cancer	HFE	Combinaison de H63D/C282Y	Carriers	-0,93% (0,27) ^d	0,003
Semsei A et al., 2012	Gène candidat (9 SNP dans 1 gène)	FRVG à la fin du suivi	ABCC1	rs246221	TT/ CC/ TC	38,5%/ 40,7%/ 38,4% ^c	0,027
Aminkeng et al., 2015	GWAS	Cardiotoxicité induite par les anthracyclines	RARG	rs2229774		4,7 (2,7-8,3)	5,9 x 10 ⁻⁸
			RARG	rs2229774	AA + AG/ GG	5,2 (3,0 - 9,0)	5,9 x 10 ⁻¹⁰
Visscher et al., 2015	GWAS		SLC22A17	rs4982753	A/ G	0,52 (0,31 - 0,85)	0,0078

		Cardiotoxicité induite par les anthracyclines	SLC22A7	rs4149178	G/ A	0,41 (0,21 - 0,77)	0,0034
Wang et al., 2016	GWAS	Cardiomyopathie	CELF4	rs1786814	GA/ AA	3,18 (0,5 - 18,9)	0,2
					GG/ AA	6,73 (1,1 - 41,1)	0,04
					GG/AA + GA	2,26 (1,2 - 4,0)	0,006
Wang et al., 2014	GWAS	Cardiomyopathie	HAS3	rs2232228	GA/ GG	2,0 (0,8 - 4,6)	0,12
			HAS3	rs2232228	AA/ GG	1,8 (0,7 - 4,7)	0,2
Visscher et al., 2012	GWAS	Cardiotoxicité induite par les anthracyclines	SLC28A3	rs7853758	A	0,31 (0,16 - 0,60)	1,0 x 10 ⁻⁴
Chaix et al., 2020 ^a	WES	Cardiotoxicité	CELF4			0,18 (0,06 - 0,56)	<0,001
			ZNF71			0,05 (0,00 - 0,92)	<0,001
			PIK3R2			0,05 (0,00 - 0,92)	<0,001
Petrykey et al., 2021	WES ^b	FEVG	TTN	rs72648998	CC/ CT + TT	4,78 (2,4 - 7,3) ^d	0,0003
			TTN	rs3829747	CC/ CT + TT	3,33 (1,5 - 5,2) ^d	0,0003
			TTN	rs2303838	CC/ CT + TT	2,59 (0,9 - 4,2) ^d	0,0005
		FRVG	TTN	rs72648998	CC/ CT + TT	3,35 (1,6 - 5,1) ^d	0,0002
		FEVG	NOS1	rs76090928			0,0013
			ABCG2	rs199473672			0,0026
Sapkota et al., 2022	WGS	Dysfonctionnement cardiaque induit par le traitement du cancer	KCNK17	rs2815063	A	1,57 (1,19 - 2,09)	0,002
			FEVG	KCNK17	rs2815063	A	- 0,016 (0,003) ^d
				rs115537302	A	- 0,019 (0,004) ^d	8,7 x 10 ⁻⁸
				rs115057884	C	- 0,019 (0,004) ^d	9,0 x 10 ⁻⁸
				rs116181858	G	- 0,019 (0,004) ^d	9,6 x 10 ⁻⁸
			rs114566901	C	- 0,019 (0,004) ^d	1,7 x 10 ⁻⁷	
Sapkota et al., 2021	WGS	FEVG	1p13. 2	rs6689879	C/ T	-0,042 (0,007) ^d	2,8 x 10 ⁻⁸
			15q25. 3	rs9788776	G/ A	-0,059 (0,01) ^d	3,5 x 10 ⁻⁸
		Cardiomyopathie (grade 2-4)	1p13. 2	rs6689879	C/ T	3,71 (1,87 - 7,37)	1,8 x 10 ⁻⁴
			15q25. 3	rs9788776	G/ A	5,24 (2,30 - 11,92)	7,9 x 10 ⁻⁵

Abréviations : OR, Odds ratio (rapport de cotes); IC, intervalle de confiance, SE, Standard error (erreur type) ; FRVG, Fraction de raccourcissement du ventricule gauche ; FEVG, Fraction de d'éjection du ventricule gauche ^a28 gènes significativement associés au statut cas-témoins mais nous n'avons rapporté que les 3 gènes considérés comme les plus importants par les auteurs ; ^bEn se concentrant sur les variants génétiques dans les gènes sélectionnés pour les voies d'action de la doxorubicine, dans le fonctionnement du système cardiaque et dans la fonction mitochondriale ; ^cFréquence des allèles ; ^dBeta (SE) ou Beta (95 % IC) ; ^eTaille de l'effet mesurée par les paramètres écho-cardiaques.

Tableau 20. Effet des gènes et des variants sur la cardiomyopathie en fonction des doses cumulées d'anthracycline

Etude	Evènement cardiaque	Interaction ou stratification	Définition des catégories	Gènes /variants	Effet allélique	Cat 1	Cat 2	Cat 3
						OR (95 % IC)	OR (95 % IC)	OR (95 % IC)
Blanco et al., 2012	Cardiomyopathie	Interaction SNP* Dose cumulée d'anthracyclines (3 catégories)	Cat 1: non exposé Cat 2: 1 to 250 mg/m ² Cat 3: > 250 mg/m ²	CBR3 / rs1056892	GA + AA	Ref	1,7 (0,5 - 5,7)	18,9 (6,1 - 58,5)
					GG	0,8 (0,2-3,4)	5,5 (1,8 - 16,6)	25,9 (7,7 - 87,6)
				CBR1/ rs9024	GA + AA	Ref	0,7 (0,1 - 3,6)	24,6 (5,4-112,0)
					GG	0,3 (0,1-1,5)	2,3 (0,6- 8,2)	11,2 (2,9-12,6)
				CBR3 / rs1056892 and CBR1/ rs9024	GA + AA and GA + AA	Ref	1,5 (0,5- 4,8)	17,4 (5,91-51,2)
					GG and GG	0,5 (0,1-2,7)	5,4 (1,8- 16,2)	17,9 (5,6-57,1)
Singh et al., 2020	Cardiomyopathie	Stratification sur la dose cumulée d'anthracyclines	Cat 1: ≤ 250 mg/m ² Cat 2: > 250 mg/m ²	GSTM1	Null / Positive	2,3 (0,7 – 6,7)	2,6 (1,2 - 5,1)	
Wang et al., 2022	Cardiomyopathie	Interaction SNP* Dose cumulée d'anthracyclines (2 catégories)	Cat 1: ≤ 250 mg/m ² Cat 2: > 250 mg/m ²	ROBO2/ rs17736312	GG + AG	Ref	0,14 (1,1- 2,9)	
					AA	1,8 (1,1 – 2,9)	2,2 (1,2 - 4,0)	
Wang et al., 2016	Cardiomyopathie	Interaction SNP* Dose cumulée d'anthracyclines (2 catégories)	Cat 1: ≤ 300 mg/m ² Cat 2: > 300 mg/m ²	CELF4/ rs1786814	GA + AA*	Ref	1,75 (0,6- 5,0)	
					GG	1,0 (0,5 – 2,2)	10,5 (3,8 - 27,3)	
Wang et al., 2014	Cardiomyopathie	Interaction SNP* Dose cumulée d'anthracyclines (continue)	Cat 1: dose cumulée d'anthracyclines > 450 mg/m ² Cat 1: Dose > 250 mg/m ²	HAS3/ rs2232228	AA	56,5 (NA – NA)		
					GG	0,6 (NA – NA)		
				HAS3/ rs2232228	GG	Ref		
					GA	4,7 (1,4 – 16,2)		
				AA	8,9 (2,1 – 37,5)			
					Interaction SNP* Dose cumulée d'anthracyclines (2 catégories)	Cat 1: 0 -250 mg/m ² Cat 2: > 250 mg/m ²	HAS3/ rs2232228	GG
					GA	0,5 (0,2 – 1,8)	5,2 (1,6 – 17,4)	
					AA	0,2 (0,1 – 0,8)	9,9 (2,4 – 40,9)	

Abréviations : OR : Odds ratio (rapport de cotes) ; IC : intervalle de confiance ; Cat : Catégorie

Les gènes et les variants identifiés dans cette revue systématique de la littérature sont impliqués dans divers mécanismes, et les résultats sont présentés selon quatre principales voies biologiques : (i) l'oxydation des anthracyclines et la production de dérivés réactifs de l'oxygène (DRO), (ii) le transport, et/ou l'élimination des anthracyclines, (iii) le métabolisme des anthracyclines, (iv) d'autres mécanismes biologiques. Une dernière classe concerne est spécifique à l'ethnie. Le

Tableau 21 résume l'ensemble des résultats présentés, classés selon les voies biologiques auxquels ils appartiennent. De plus, la **Figure 17** présente une partie de ces gènes et résume schématiquement leurs rôles dans la cardiotoxicité liée aux anthracyclines, au sein d'une cellule cardiomyocytaire.

Tableau 21. Gènes et familles de gènes présentés, et les voies biologiques auxquels ils appartiennent

Gènes ou Famille de gènes	Nom du gène ou de la famille de gènes	Voie biologique
Famille des NOS (NOS1, NOS3)	Nitric Oxide synthase	Oxydation des anthracyclines et production de DRO
HAS3	Hyaluronane synthase 3	Oxydation des anthracyclines et production de DRO
GSTM1	Glutathione S-Transferase Mu 1	Oxydation des anthracyclines et production de DRO
NQO1	NAD(P)H Quinone Dehydrogenase 1	Oxydation des anthracyclines et production de DRO
Superfamille des SLC (SLC22A17 ; SLC22A6 ; SLC22A7 ; SLC28A3)	Superfamille de gènes des transporteurs de solutés	Transport et élimination des anthracyclines
Famille des transporteurs ABC	Famille des transporteurs à cassettes liant l'ATP	Transport et élimination des anthracyclines

(ABCA1 ;ABCC1 ; ABCC2 ; ABCC5 ; ABCC9 ; ABCG2)		
SULTB1	Sulfotransferase Family 2B Member 1	Transport et élimination des anthracyclines
Carbonyl réductase (CBR1 et CBR3)	Short-chain deshydrogenase/reductase	Métabolisme des anthracyclines
UGT1A6	UDP-glucuronosyltransférase	Métabolisme des anthracyclines
CYP450	Cytochrome P450	Métabolisme des anthracyclines
HFE	Homeostatic Iron Regulator	Autres mécanismes biologiques
RARG	Retinoic Acid Receptor Gamma	Autres mécanismes biologiques
CELF4	CUGBP Elav-Like Family Member 4	Autres mécanismes biologiques
TTN	Titine	Autres mécanismes biologiques
PLCE1	Phospholipase C epsilon 1	Autres mécanismes biologiques
Locus 15q25.3		Variants lié à une éthnie spécifique
Locus 1p13.2		Variants lié à une éthnie spécifique
KCNK17	Potassium two pore domain channel subfamily K member 17	Variants lié à une éthnie spécifique

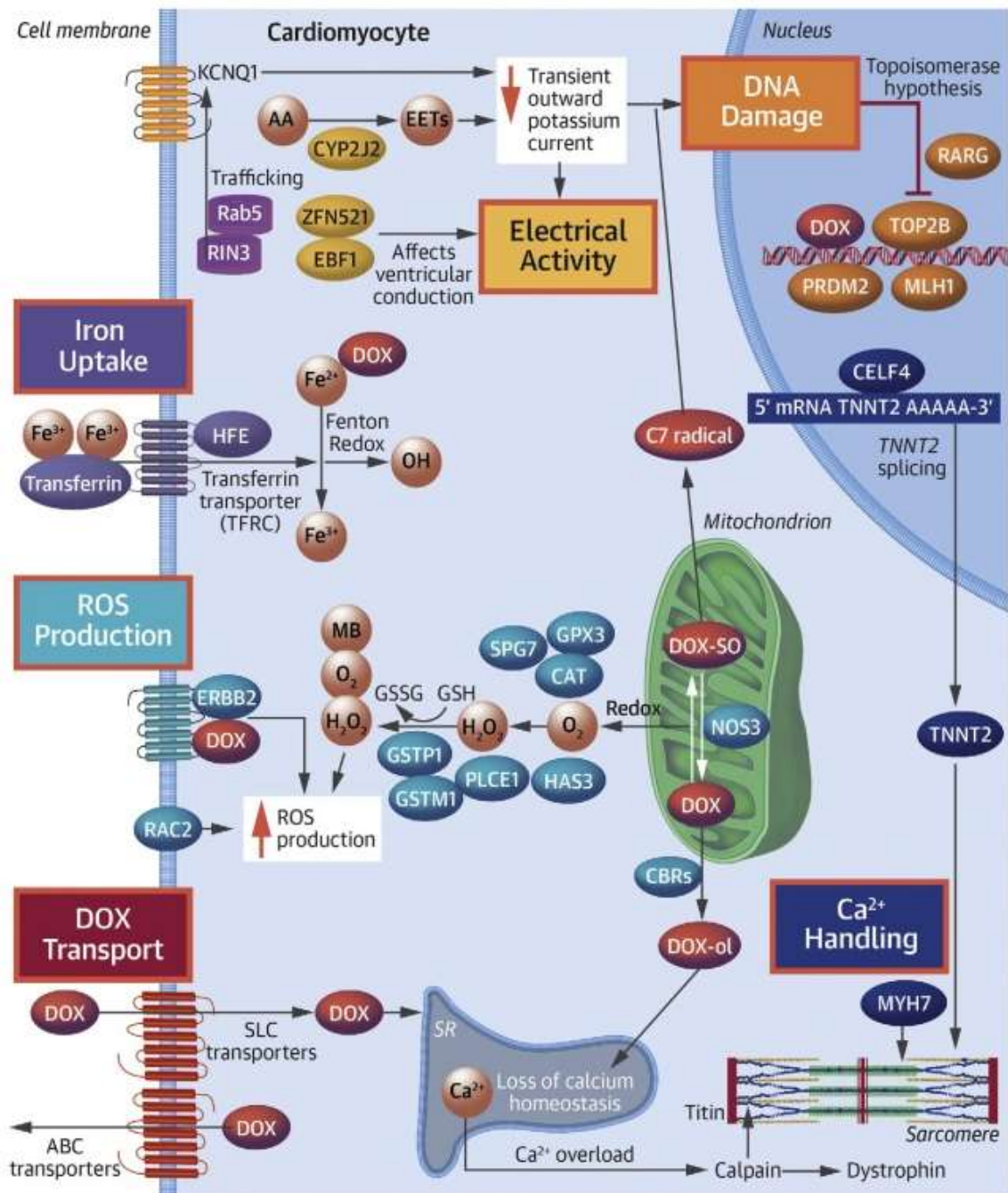


Figure 17. Rôles des gènes présentés associés à la cardiotoxicité liée à la doxorubicine, au sein d'une cellule cardiomyocytaire (source : Fonoudi et al., 2024).

Les différents mécanismes de cardiotoxicité après exposition à la doxorubicine (DOX) sont mis en évidence dans des encadrés colorés. AA = acide arachidonique ; ABC = cassette de liaison à l'adénosine triphosphosphate ; EET = acide époxyeicosatriénoïque ; ARNm = ARN messenger ; ROS = DRO = dérivés réactifs de l'oxygène ; SLC = transporteur de soluté ; SR = réticulum endoplasmique

5.3.5.1 *Oxydation des anthracyclines et production de DRO*

Un mécanisme important des anthracyclines se traduit par une réduction de l'ATP et une augmentation des dérivés réactifs de l'oxygène (DRO) dans les cellules, induisant l'apoptose des cellules tumorales (Fogli et al., 2004; Koleini & Kardami, 2017; L'Ecuyer et al., 2004; Megias-Vericat et al., 2018). Dans les cellules, à des niveaux physiologiques, ces DRO jouent un rôle fondamental dans régulation de plusieurs fonctions cellulaires. Cependant, les traitements par anthracyclines peuvent induire une augmentation forte des niveaux de DRO dans les cellules, qui peut persister et induire un déséquilibre dramatique entre les DRO et les systèmes de défense antioxydants (Rocca et al., 2020). Ainsi, un taux de DRO et une faible défense antioxydant rendent les cardiomyocytes particulièrement sensibles aux dommages causés par les radicaux libres produits par les anthracyclines (Yang et al., 2020), provoquant, directement ou indirectement, des lésions cardiaques (Rocca et al., 2020).

Nous décrivons dans cette partie le rôle de cinq gènes (NOS1, NOS3, HAS3, GSTM1 et NQO1) impliqués dans l'oxydation des anthracyclines, et associés à la cardiomyopathie chez les survivants d'un cancer dans l'enfance.

Nitric Oxide synthase (NOS)

Deux études se sont interrogées sur l'impact des gènes de la famille NOS sur le développement de la dysfonction cardiaque après traitement d'une leucémie aiguë lymphoblastique (LAL) dans l'enfance (Petrykey et al., 2022; Krajinovic et al., 2016). Dans une étude menée sur une cohorte canadienne de patients traités pour une LAL dans l'enfance, le séquençage de l'exome complet a mis en évidence un effet protecteur du variant rare *NOS1*-rs76090928 sur la FEVG (Petrykey et al., 2022). Un effet similaire a été retrouvé pour le variant *NOS3*-rs1799983, dans une étude gène candidat portant sur une population semblable (Krajinovic et al., 2016). En effet, les patients porteur du génotype TT pour ce variant commun présentaient une FEVG plus élevée que les individus portant au moins un allèle G (respectivement 61,4% vs 59,0%, $p=0.02$). Par

ailleurs, l'effet protecteur de ce génotype a également été observé chez les patients présentant une leucémie aigue lymphoblastique, et n'ayant pas reçu de Dexrazoxane avant traitement par doxorubicine (64.1% vs. 56.8%, $p=0.002$). Cependant, chez les patients ayant reçu ce traitement cardioprotecteur, l'effet protecteur du génotype *NOS3*-rs1799983-TT n'a pas été observé (Krajinovic et al., 2016).

Les gènes *NOS1* et *NOS3* codent pour la production d'une enzyme de type oxyde nitrique synthase, qui a pour rôle de catalyser la production d'oxyde nitrique (NO), une molécule clé impliquée dans la physiopathologie du cœur (Linschoten et al., 2018; Megias-Vericat et al., 2018), et dont les niveaux augmentent en cas de cardiotoxicité induite par la doxorubicine (Songbo et al., 2019). Le dérèglement de l'activité des NOS et du métabolisme du NO semble être une caractéristique commune à diverses maladies cardiaques. La contribution du NO aux lésions cardiaques causées par les anthracyclines est suggérée par les études mettant en avant l'induction de l'expression des NOS et la libération du NO dans le cœur, ainsi que la capacité des NOS à promouvoir le cycle d'oxydoréduction des anthracyclines pour produire des DRO (Folgli et al., 2004).

La hyaluronane synthase 3 (HAS3)

Une étude cas-témoins menée sur des patients issus d'un groupe d'essais cliniques menée en Amérique du nord (États-Unis et Canada) a permis de mettre en évidence une interaction entre le variant *HAS3*-rs2232228 et les doses cumulées d'anthracyclines : les patients exposés à des doses d'anthracyclines supérieures 250 mg/m² et porteurs des génotypes AA ou GA présentaient un risque de cardiomyopathie élevé (OR = 8,9 (95 % IC : 2,1 – 37,5) en comparaison aux patients homozygotes GG (OR = 4,7 (95 % IC=1,4 – 16,2) (Wang et al., 2014). Ce variant est porté par le gène *HAS3*, pour hyaluronane synthase 3, qui code pour la synthèse d'un composé de la matrice extracellulaire, l'acide hyaluronique, dont la dégradation est déterminée par les dommages oxydatifs dus aux DRO. L'acide hyaluronique peut aussi

être dégradé directement par les DRO, ces derniers étant produits en grande quantité pour des traitements par anthracyclines (Fallacara et al., 2018).

Glutathione S-Transferase (GST)

Les enzymes de la famille des GST sont connues dans le métabolisme des médicaments et sont très utiles à la caractérisation de la réponse des patients aux médicaments chimiothérapeutiques (Yang et al., 2020), de sorte que la doxorubicine a tendance à diminuer les taux de glutathione, engendrant un stress oxydatif très élevé (Songbo et al., 2019). La stimulation des gènes de la famille des GST dont le gène Glutathione S-Transferase Mu 1 (GSTM1), permet de réduire le stress oxydatif (Songbo et al., 2019). Ces derniers jouent un rôle important dans le métabolisme des DRO (Songbo et al., 2019) et leur désactivation peut augmenter le risque de cardiotoxicité (Yang et al., 2020). Ainsi, une étude gènes candidats a cherché à examiner l'impact du variant nul du gène GSTM1 dans le développement de la cardiomyopathie chez les patients traités par des anthracyclines pour un cancer pédiatrique, et a mis en évidence que la délétion complète du gène (GSTM1 variant nul) était plus fréquente chez les patients ayant développé une cardiomyopathie que chez les témoins qui n'en avaient pas développé (60,0 % vs. 38,04 %, $p = 0,005$) (Singh et al., 2020). Il a aussi été observé que les personnes présentant un génotype GSTM1 nul avaient un risque 2,7 fois plus élevé de développer une cardiomyopathie que celles présentant un génotype GSTM1 positif. L'association est restée constante après stratification sur l'exposition à la dose d'anthracyclines ($\leq 250 \text{ mg/m}^2$ ou $> 250 \text{ mg/m}^2$). En outre, la comparaison de l'expression de ce gène dans une étude cas-témoins a mis en évidence une diminution de l'expression du gène GSTM1 chez 60% des cas ($p = 0,049$) (Singh et al., 2020).

NAD(P)H Quinone Dehydrogenase (NQO1)

Le gène NQO1 participe à la réduction des anthracyclines en dérivés hydroquinones empêchant la formation de DRO (Megías-Vericat et al., 2018), évitant donc que les cardiomyocytes soient endommagés. Chez les patients porteurs de l'allèle T du variant

NQO1-rs1043470 une réduction de la fonction ventriculaire gauche à la fois en phase aiguë et 5 à 10 ans après le diagnostic a été observée. Plus précisément, pour cette dernière période, une valeur moyenne plus faible de la FRVG a été observée chez ces patients ($38,5 \% \pm 0,7$), alors que les patients ayant au moins un allèle C avaient une valeur de FRVG de $40,6 \% \pm 0,4$ ($p = 5,82 \times 10^{-3}$) (Sági et al., 2018), indiquant un risque moins élevé de cardiomyopathie chez les patients portant un allèle C de ce variant. Dans une seconde étude cherchant à mettre en évidence une association entre le polymorphisme NQO1*2 (variant rs1800566) et la cardiomyopathie, aucune association significative n'a été observée (Blanco et al., 2008).

5.3.5.2 *Transport et élimination des anthracyclines*

Les premières recherches menées sur l'évaluation des mécanismes de transport transmembranaire des anthracyclines suggèrent un rôle important de la diffusion passive des traitements non-ionisés, ces derniers étant habituellement liposoluble et se diffusant facilement à travers les membranes cellulaires. Cependant, l'existence d'une cinétique de saturation dans l'absorption indique plutôt un transport facilité par des transporteurs spécifiques (Huang et al., 2018).

Trois familles de gènes impliquées dans le transport et l'élimination des anthracyclines à travers les cellules cardiaques ont été identifiées : la superfamille de gènes des transporteurs de solutés (quatre gènes : SLC22A17 ; SLC22A6 ; SLC22A7 ; SLC28A3), la famille des transporteurs à cassettes liant l'ATP (six gènes : ABCA1 ; ABCC1 ; ABCC2 ; ABCC5 ; ABCC9 ; ABCG2), et la famille des sulfotransférases (un gène : SULTB1).

La superfamille de gènes des transporteurs de solutés (SLC)

Cette superfamille code pour des protéines membranaires capables de transporter une large variété de molécules dont des médicaments à travers les membranes cellulaires (Visscher et al., 2015) et peut contribuer à la distribution et l'élimination des anthracyclines (Huang et al., 2018). Les gènes *SLC28A3*, *SLC22A17*, *SLC22A7*, et *SLC22A6* ont été identifiés à travers différentes études comme associés aux

cardiomyopathies chez les patients traités pour un cancer pédiatrique.

Trois variants dans le gène *SLC28A3*, à savoir rs7853758, rs885004 et rs4877847, ont été identifiés comme étant associés à un risque plus faible de cardiotoxicité. Dans une cohorte de patients traités pour des cancers divers, porter un allèle A du variant *SLC28A3*-rs7853758 diminuerait le risque de développer une cardiomyopathie (OR = 0,31; 95 % IC= 0,16 - 0,60) (Visscher et al., 2012), et ce résultat a été confirmé dans une étude gène candidat (OR= 0,36; 95 % IC= 0,22 - 0,60) (Visscher et al. 2013). Cependant, dans une cohorte de patients traités pour des leucémies aiguës lymphoblastiques et des ostéosarcomes, les patients arborant deux allèles A pour ce variant ont présenté un risque 9,8 fois (95 % IC= 1,7 - 56,0) plus élevé de développer une cardiomyopathie que les patients portant d'autres génotypes (Sági et al., 2018). Dans cette même étude, les patients portant le génotype *SLC22A6*-rs6591722-AA présentaient aussi un risque plus élevé de développer une cardiomyopathie 5 à 10 ans après le diagnostic du premier cancer. De plus, les variants *SLC22A17*-rs4982753-A et *SLC22A7*-rs4149178-G ont été identifiés dans une cohorte canadienne de patients traités pour divers cancers pédiatriques comme associés à un risque plus faible de développer une cardiomyopathie, et ces résultats ont été validés dans une cohorte de réplication (Visscher et al., 2015).

Les transporteurs ABC ou transporteurs à cassettes liant l'ATP (ABC)

Il s'agit d'un groupe de gènes composé de sept sous-familles, dont les sous-familles B et C sont spécifiquement impliquées dans le transport de xénobiotiques comme les anthracyclines. Le blocage de ces transporteurs entraîne l'altération de la distribution et du métabolisme des traitements anticancéreux, augmentant les concentrations de ces traitements dans le plasma et dans les cellules (Couture et al., 2006).

Le gène *ABCC1* est fortement exprimé dans le cœur humain (Couture et al., 2016), et une étude gène candidat a identifié deux variants (rs3743527, et rs246221) comme associé aux cardiomyopathies dans une cohorte de patients traités pour leucémie aiguë

lymphoblastique (LAL) en Hongrie (Semsei et al., 2012). Le génotype TT du variant rs3743527 est associé à une diminution de la FRVG à la fin du traitement (34.0%) en comparaison aux autres génotypes (CC, 39.5%; CT 39.3%). De plus, les patients portant au moins un allèle T pour le variant rs246221 ont une FRVG plus faible que les patients ayant un génotype CC. De plus, après l'administration d'anthracyclines, la FRVG moyenne était significativement plus faible chez les patients portant à la fois les génotypes *ABCC1*-rs3743527-TT et *ABCC1*-rs246221-TC ou TT comparé aux patients portant d'autres génotypes (Semsei et al., 2012).

Dans une cohorte canadienne de patients traités pour une LAL, les individus porteurs du génotype *ABCC5*-rs7627754-TT présentaient en moyenne une réduction de 12% de la FEVG et de 8% de la FRVG. Cependant, il est important de noter que les patients porteurs du génotype TT semblaient avoir des valeurs de FEVG et de FRVG plus faibles au moment du diagnostic du cancer (avant l'exposition au traitement) (Krajinovic et al., 2016). En outre, un variant rare liée à la voie de la doxorubicine a également été identifié dans le gène *ABCC5* comme ayant un effet protecteur sur la FEVG et la FRVG (respectivement, $p = 0,0008$ et $p = 0,0014$) (Petrykey et al., 2022). Cette même étude a rapporté un effet protecteurs de variants rs199473672 et rs45605536 dans le gène *ABCG2* (Petrykey et al., 2022).

De plus, une étude gène candidat a cherché à évaluer l'association entre le variant *ABCC2*-rs3740066 et la cardiomyopathie chez des patients traités pour leucémie aiguë lymphoblastique et ostéosarcome, et a mis en évidence une augmentation du risque de cardiomyopathie 5 à 10 ans après diagnostic du traitement chez les patients présentant un génotype GG en comparaison aux autres génotypes (GA et AA).

Famille des Sulfotransférases (SULT)

Dans une étude combinant des survivants d'un cancer pédiatrique d'une cohorte canadienne et d'une cohorte néerlandaise, être porteur de l'allèle A du variant *SULT2B1*-rs10426628 était associé à un risque 1,9 fois plus élevé de développer une

cardiomyopathie (95 % IC= 1,2–2,5, $p=3,2 \times 10^{-4}$) (Visscher et al., 2015). Ce gène code pour la production d'une enzyme qui conduit à l'inactivation de composés biologiques, dont les anthracyclines font partie, contribuant à leur élimination ainsi que l'élimination de leurs métabolites du corps humain (Luo et al., 2016).

5.3.5.3 *Métabolisme des anthracyclines*

Les anthracyclines dont la doxorubicine, sont fréquemment administrées dans le cadre de diverses affections malignes, et sont connues pour leur association avec une cardiotoxicité qui se caractérise par des lésions directes irréversibles et est liée à la dose (Linschoten et al., 2018). Il est estimé qu'environ 50 % de la dose d'anthracycline administrée est éliminée dans l'organisme sans modification de sa structure (Edwardson et al., 2015), et que le métabolisme des anthracyclines comprend trois voies : (1) l'hydroxylation, (2) la réduction à un électron et (3) la déglycosylation (Megías-Vericat et al., 2018).

Quatre gènes (CBR1, CBR3, UGT1A6, CYP450) jouant un rôle dans le métabolisme des anthracyclines ont été identifiés comme associés à la cardiomyopathie dans les études incluses.

Short-chain deshydrogenase/reductase (SDR)

Les gènes Carbonyl réductase 1 et 3 (CBR1 et CBR3) codent pour des enzymes appartenant à la superfamille des SDRs. Ces deux gènes participent à l'élimination de la doxorubicine, en catalysant la réaction de réduction en métabolites : le doxorubicinole (phénomène d'hydroxylation). Ce dernier est connu pour son rôle dans l'observation de la cardiotoxicité après chimiothérapie (Choi et al., 2006; Malatkova et al., 2010). Dans une première étude, Blanco et al. (2008) a mis en évidence dans une étude gène candidat que le polymorphisme V244M dans le gène CBR3 (correspondant au variant rs1056892) n'était pas associé au développement de l'insuffisance cardiaque congestive, mais dans une seconde étude, le même auteur rapporte que le génotype GG du variant *CBR3*-rs1056892 est associé à un risque élevé de développer une

cardiomyopathie (OR=1,79, 95 % IC= 1,08 - 2,96) (Blanco et al., 2012), et une interaction a été observée entre ce génotype et l'exposition à des doses faibles à modérées d'anthracyclines (1 à 250 mg/m²), augmentant de 3,3 fois le risque de cardiomyopathie en comparaison aux génotypes GA et AA (Blanco et al., 2012). Dans cette même étude, les auteurs rapportent que les patients portant à la fois le génotype GG pour les variants *CBR3*-rs1056892 et *CBR1*-rs9024, et recevant des doses faibles à modérées d'anthracycline augmente le risque de cardiomyopathie de 3,6 fois par rapport aux génotypes AA et GA (Blanco et al., 2012).

UDP-glucuronosyltransférase (UGT)

Dans une étude cas-témoins menée sur des patients traités par anthracyclines pour un cancer de l'enfant, le variant rs6759892 dans le gène *UGT1A6* a été suggéré comme étant à risque d'augmenter les chances de développer une cardiomyopathie (OR = 2,9; 95 % IC= 1,7 - 4,9) (Visscher et al., 2012). Une seconde étude menée sur une population provenant du même échantillon de patients a mis en évidence que le variant *UGT1A6*-rs17863783 était associé à un risque 6,2 fois plus élevé (95 % IC= 2,5-15,4) de développer une cardiomyopathie sévère (Visscher et al., 2013). De plus, un risque augmenté a été observé chez les patients ayant été traités par doxorubicine et/ou daunorubicine, et qui ont reçu des radiations au cœur (OR = 8,0; 95 % IC= 1,8 - 34,4)

Ces deux variants sont situés dans le gène UDP Glucuronosyltransferase Family 1 Member A6 (*UGT1A6*), qui code pour une enzyme hépatique qui joue un rôle majeur dans le métabolisme et l'élimination des médicaments et autres xénobiotiques (Olinga et al., 2008). Il jouerait un rôle de médiateur de la cardiotoxicité et utiliserait les mêmes procédés que les gènes *CBR1* et 3 (Linschoten et al., 2018). Par ailleurs, son activité enzymatique serait altérée par la présence d'un haplotype caractérisé par le SNP rs17863783. Cependant, le rôle de cette enzyme dans l'élimination des anthracyclines et de ses métabolites n'est pas clair (Linschoten et al., 2018).

Le cytochrome P450 (CYP450)

Dans une étude cas-témoins menée sur des patients traités pour une leucémie aiguë lymphoblastique (LAL) ou un ostéosarcome en Hongrie, le génotype TT dans le variant *CYP3A5*-rs4646450 a été identifié comme associé à une fraction de raccourcissement inférieure à 28%. En effet, spécifiquement chez les patients avec un antécédent de LAL, 17% des cas (ayant une FRVG < 28%) portaient le génotype TT alors que seul 3% des témoins en étaient porteurs ($p = 5,6 \times 10^{-3}$) (Sági et al., 2018). Ce gène appartient à la superfamille d'enzymes CYP450, responsables de la catalyse de nombreuses réactions dans les xénobiotiques, et plus spécifiquement à la sous-famille CYP3A qui joue un rôle important dans le métabolisme d'environ 30% des médicaments utilisés en cliniques et dans toutes les catégories thérapeutiques, dont celle des anticancéreux (Zanger & Schwab, 2013). Par ailleurs les patients porteurs du variant *CYP3A5*-rs776746 avaient une activité enzymatique plus faible, et ce variant a été associé à la concentration plasmatique de daunorubicine et aux effets indésirables du médicament (Huang et al., 2017).

5.3.5.4 Autres mécanismes biologiques

D'autres mécanismes biologiques sont associés à la cardiotoxicité liée aux anthracyclines comme la production de protéines de régulation du fer par le gène HFE, les interactions avec l'enzyme Topoisomerase II β pour le gène RARG, la perturbation de la structure et de la fonction des sarcomères pour les gènes CELF4, ou TTN, ou encore au stress oxydatif comme le gène PLCE1

Homeostatic Iron Regulator (HFE)

Les traitements par doxorubicine ont tendance à interagir avec les protéines de transport et de fixation du fer, entraînant un dérèglement de l'homéostasie du fer dans les cellules. L'accumulation du fer dans les cellules entraîne alors une augmentation de la production de DRO, et donc des dommages sur les cardiomyocytes. Par ailleurs, l'accumulation de fer dans les cardiomyocytes active les voies de mort cellulaire de

manière indépendante de l'oxydation (Songbo et al., 2019). Ces connaissances ont poussés des chercheurs américains à émettre l'hypothèse que les conditions conduisant à des concentrations élevées de fer dans les tissus peuvent prédisposer les individus au développement de la cardiotoxicité liée à la doxorubicine (Lipshultz, Lipsitz, et al., 2013). Par conséquent, une étude a été réalisée sur une cohorte de patients traités par doxorubicine pour une leucémie aiguë lymphoblastique (LAL) à risque élevé, en séquençant les variants rs1799945 (C282Y) et rs1800562 (H63D) du gène HFE, en surveillant les paramètres du ventricule gauche du cœur tels que la fonction, la masse et l'épaisseur de la paroi. Il a été mis en évidence qu'en comparaison à une population d'enfants sains, les patients traités pour LAL et portant les variants alléliques C282Y et/ou H63D avaient des paramètres du ventricule gauche (FEVG et FRVG) significativement plus faibles deux ans après le diagnostic de la LAL, suggérant que les patients porteurs de ces génotypes pourraient présenter un risque de cardiomyopathie plus élevé (Lipshultz, Lipsitz, et al., 2013).

Retinoic Acid Receptor Gamma (RARG)

Dans une étude GWAS, le variant *RARG*-rs2229774 s'est révélé associé fortement à la cardiomyopathie après traitement d'un cancer dans l'enfance par anthracyclines dans des populations Européennes et non-Européennes (OR= 4,7; 95 % IC= 2,7– 8,3) (Aminkeng et al., 2015). Il s'agit de l'un des résultats les plus prometteurs dans notre revue, mettant en évidence une association forte avec la cardiomyopathie après traitement d'un cancer dans l'enfance par anthracyclines. Ce gène participe à la régulation de la transcription de l'enzyme Topoisomérase-2 β , qui est elle-même ciblée par la doxorubicine et est fortement exprimée dans les cellules cancéreuses. Cette protéine a une importance majeure dans la division cellulaire pour bloquer la réplication de l'ADN (Pommier et al., 2010). Par ailleurs, la topoisomérase II β interagit avec la doxorubicine dans les tissus cardiaques, entraînant chez la souris l'apoptose des cardiomyocytes et l'apparition d'une insuffisance cardiaque (Linschoten et al.,

2018). Or, chez les souris topoisomerase II β -knockout, un retard dans la tendance à la mort des cardiomyocytes induite par la doxorubicine a été observé (S. Zhang et al., 2012). Ainsi, le variant *RARG*-rs2229774 pourrait réguler à la hausse l'enzyme topoisomerase II β entraînant l'apparition de la cardiomyopathie chez les patients.

CUGBP Elav-Like Family Member 4 (CELF4)

Le gène *CELF4* a été identifié comme associé au risque de cardiomyopathie dans une étude GWAS menée sur des patients traités pour un cancer pédiatrique quel que soit la tumeur. En effet, le génotype *CELF4*-rs1786814-GG a été associé à un risque accru de cardiomyopathie par rapport aux génotypes GA et AA ($p=0,006$). Une exposition à des doses inférieures ou égales à 300 mg/m², chez les patients porteurs du génotype GG est associée à un risque 5,8 fois plus élevé de développer une cardiomyopathie par rapport aux génotypes GA et AA (Wang et al., 2016). Le gène *CELF4* code pour CUGBP Elav-Like Family Member 4, une protéine impliquée dans la régulation de l'épissage de l'ARN messager de la protéine du sarcomère cardiaque troponine T2. Cette protéine sarcomérique joue un rôle essentiel dans la régulation de la contraction musculaire. Par ailleurs, dans l'étude Wang et al. (2016), les patients porteurs du génotype *CELF4*-rs1786814-GG ont une susceptibilité accrue de coexprimer des isoformes de la troponine T2 dans le cœur. Or, la coexpression de plusieurs isoformes de la troponine T2 a été observée chez les patients ayant une faible efficacité de fonction de pompe du myocarde (Linschoten et al., 2018).

Titine (TTN)

Le gène *TTN* code pour la production de la titine, un filament sarcomérique géant qui joue un rôle majeur dans l'apparition des pathologies cardiaques. Des études ont rapportées que des mutations dans le gène *TTN* étaient associées à diverses cardiomyopathies, et certains variants de ce gène étaient associés à un risque augmenté de cardiotoxicité induite par la chimiothérapie (Kellermayer et al., 2019; Luo et al., 2016). Cependant, cet effet délétère du gène *TTN* est en contradiction avec les

résultats du travail de Petrykey et al. (2021). En effet, dans une étude de cohorte portant sur analyse de l'exome complet de patients souffrant de leucémie aiguë lymphoblastique, l'analyse des variants communs a révélé que les génotypes TTN-rs72648998-CC et TTN-rs3829747-CC étaient associés à une augmentation de la FRVG ($p=0,0002$ et $0,0004$, respectivement) et de la FEVG ($p=0,0003$ et $0,0003$, respectivement), et que le génotype TTN-rs2303838-CC était associé à une augmentation de la FEVG ($p=0,0005$) (Petrykey et al., 2021).

Phospholipase C epsilon 1 (PLCE1)

Hildebrandt et al. (2017) ont supposé que les loci de susceptibilité génétique à l'hypertension pouvaient être des prédicteurs du développement tardif de la cardiotoxicité. Ils ont alors étudié le variant rs932764 dans le gène PLCE1 qui avait été retrouvé dans une étude GWAS comme étant associé à un risque élevé d'hypertension et de pathologies cardiovasculaires (Ehret et al., 2011; Hoffmann et al., 2017; Wain et al., 2017). Par ailleurs, le gène PLCE1 encode pour l'enzyme phospholipase C epsilon (PLC ϵ), qui a été précédemment identifiée comme ayant un effet cardioprotecteur contre le stress oxydatif (Xiang et al., 2013). Ainsi, dans une cohorte de patients ayant été traités par anthracyclines pour un cancer dans l'enfance, une association entre l'allèle G du variant rs932764, situé dans le gène *PLCE1*, et une diminution du risque de cardiomyopathie a été mis en évidence (OR= 0,36 (95 % IC : 0,18–0,76) (Hildebrandt et al., 2017).

5.3.5.5 Variants lié à une ethnie spécifique

Une étude WGS (Whole Genome Sequencing) menée sur des patients d'origine africaine a mis en évidence un ensemble de 31 variants rares ou de faible fréquence et deux loci dans l'analyse des variants communs, significativement associés à des fluctuations de la fraction d'éjection du ventricule gauche (Sapkota et al., 2021). Les deux variants rs6689879 dans le locus 1p13.2 et rs9788776 dans le locus 15q25.3, sont deux variants communs identifiés comme significativement associées à une réduction

de la FEVG (OR = 0,96 ; $p = 2,8 \times 10^{-8}$ et OR = 0,94 ; $p = 3,5 \times 10^{-8}$, respectivement). Chez les patients exposés uniquement aux anthracyclines, un effet de réduction de la FEVG de 3,0% ($p = 0,044$) a été observé chez les porteurs du variant rs6689879, et de 4,4% ($p = 0,022$) chez les patients porteurs du variant rs9788776. Cependant, chez les patients exposés uniquement à la radiothérapie, le variant rs9788776 a montré une réduction de la FEVG de 8,1% ($p = 5,4 \times 10^{-4}$). Ces associations sont spécifiques aux patients d'origine africaine, car elles n'ont pas été observées dans la cohorte de patients d'origine européenne. Dans le groupe d'ascendance européenne, le variant rs6689879 n'était pas statistiquement significatif, et le variant rs9788776 n'a pas pu être analysé en raison de sa nature monomorphe (Sapkota et al., 2021).

Dans une étude menée sur des survivants d'origine européenne, le locus 6p21.2, plus précisément le gène *KCNK17*, s'est révélé être associé de manière significative à une FEVG plus faible (Sapkota et al., 2022). Le génotype rs2815063-A a été associé à une réduction de la FEVG de 1,6% ($p = 2,1 \times 10^{-8}$) dans l'ensemble de la cohorte. Après ajustement sur l'administration de traitement cardioprotecteur chez les patients présentant une cardiomyopathie de grade ≥ 3 , l'association est restée significative, avec une réduction de la FEVG de 1,1% ($p = 7,5 \times 10^{-6}$). Après classification en groupes de risque basés sur la thérapie, l'association entre rs2815063-A et la réduction de la FEVG était plus forte dans les groupes à risque élevé et modéré (1,7% ; $p = 9,6 \times 10^{-4}$ et 1,9% ; $p = 1,4 \times 10^{-4}$, respectivement), par rapport au groupe à faible risque (1,4%, $p = 0,002$). Ces résultats ont été répliqués sur 4020 patients de la cohorte Childhood Cancer Survivors Study (CCSS) d'origine européenne, et les porteurs du génotype rs2815063-A avaient un risque 1,4 fois (95 % IC= 1,02 - 1,92 ; $p = 0,004$) plus élevé de développer une insuffisance cardiaque congestive comparé aux autres génotypes. Il est intéressant de noter que cette association a également été reproduite chez les individus d'origine africaine, avec une réduction de la FEVG de 1,5 % ($p = 0,004$) (Sapkota et al., 2022). Aucune autre étude n'a pu mettre en évidence une association entre le gène *KCNK17* et l'apparition de pathologies cardiaque, ni chez les survivants d'un cancer, ni en

population générale. Cependant, ce gène a été identifié comme associé à des hémorragies cérébrales importantes et à des accidents vasculaires cérébraux dans la population chinoise (He et al., 2014).

5.4 DISCUSSION

Cette revue systématique de la littérature a identifié 19 études rapportant des associations entre des facteurs génétiques et les cardiomyopathies faisant suite au traitement d'un cancer dans l'enfance. Parmi ces études, quatorze (76 %) ont mené des études cas-témoins et cinq (26 %) des études de cohorte. Plusieurs variants génétiques dans 61 gènes et deux loci ont été identifiés comme significativement associés à la cardiomyopathie, en utilisant diverses méthodes de séquençage. Parmi ces variants, SLC28A3-rs7853758 et RARG-rs2229774 représentent les résultats les plus cohérents, montrant des associations fortes et significatives avec la cardiomyopathie suite au traitement du cancer de l'enfant comprenant des anthracyclines. Ces associations ont été retrouvées avec succès chez des patients de cohortes externes (Aminkeng et al., 2015; Visscher et al., 2012, 2013). Notre analyse a également permis d'identifier de nombreux autres gènes et variants nécessitant d'être validés dans une cohorte externe, et de mettre en évidence des interactions entre les gènes/variants et les doses cumulées d'anthracyclines, ainsi que des gènes/variants spécifiques à une ethnie ou à un sexe. Cependant, trop peu d'études ont examiné le risque associé aux doses excessives de radiothérapie au cœur sur le développement de la cardiomyopathie chez les survivants d'un cancer de l'enfance et, en outre, aucune étude n'a examiné l'existence d'interactions entre les gènes/variants identifiés et la radiothérapie.

Même si de précédentes revues se sont intéressées aux effets cardiotoxiques des traitements du cancer (Berkman et al., 2021; Linschoten et al., 2018), aucune n'a considéré spécifiquement la cardiotoxicité liée à la radiothérapie impliquant le cœur. Bien que dans notre revue nous ayons cherché à introduire l'impact des radiations au cœur dans le développement des pathologies cardiaques, aucune étude n'a recherché

les interactions entre les gènes identifiés et ce traitement anticancéreux alors que les risques de pathologies cardiaques associés à la radiothérapie sont maintenant bien connus. De plus, des données de recherche sur ce sujet sont maintenant disponibles grâce à des outils qui estiment la distribution dose-volume des radiations dans le cœur entier et les sous-structures du cœur (Bates et al., 2019; Gröbner et al., 2018; Mansouri et al., 2019). Cependant, la physiopathologie des pathologies cardiaques associées aux anthracyclines est différente de la physiopathologie des pathologies cardiaques associées aux radiations. Cela suggère donc que les voies biologiques liées à cardiotoxicité induites par ces deux types de traitements sont différentes (K. Wang & Tepper, 2021). Il serait donc intéressant de chercher à analyser l'implication des gènes appartenant à ces voies biologiques dans l'apparition des pathologies cardiaques tardives après traitement d'un cancer de l'enfant. Les orientations futures doivent se concentrer sur les interactions entre les facteurs de risque liés aux traitements du cancer de l'enfant, comme les doses de radiations au cœur, les métriques dose-volume, les doses d'anthracyclines et d'autres chimiothérapies, avec la prédisposition génétique aux pathologies cardiaques.

Au vu du nombre élevé de variants et de gènes identifiés à travers les différentes études, la formulation de conclusions définitives est très difficile. Il serait peut-être possible de mettre en évidence des cardiotoxiques des traitements en se concentrant sur des variants génétiques rares ou ultra-rare, plutôt que par l'utilisation de variants communs chez les survivants d'un cancer dans l'enfance. De même, l'ouverture de nouveaux champs de recherche, comme l'expression des gènes, pourrait être intéressante : la nature dynamique du transcriptome, sensible aux xénobiotiques, permet l'identification de biomarqueurs directement utilisables pour la détection de pathologies résultant de l'exposition à des traitements (Joseph, 2017).

Il est également important de ne pas se focaliser uniquement sur les cardiomyopathies, mais d'étudier l'ensemble des pathologies cardiaques, qui peuvent avoir des

mécanismes physiopathologiques différents impliquant différents gènes. Les études dans ces différents domaines nécessitent l'inclusion de milliers de patients et ne peuvent donc être menées que dans le cadre d'une collaboration internationale, permettant d'obtenir des résultats robustes.

En conclusion, cette revue systématique présente une vue d'ensemble des variants génétiques prédisposant à la cardiomyopathie. Ce qui nous permet de mieux comprendre les mécanismes biologiques sous-jacents, mais aussi de suggérer de nouvelles pistes de recherches telles que la recherche de biomarqueurs transcriptomiques, de variants rares ou ultra-rares et de se concentrer sur les variants associés à la radiothérapie. En effet, il est crucial de prévenir les effets tardifs des traitements anticancéreux, tels que la pathologie cardiaque, et d'améliorer la prise en charge des survivants d'un cancer de l'enfant. Malgré les avancées thérapeutiques, ces patients sont toujours plus à risque de développer des problèmes de santé tardifs. Il serait extrêmement utile d'établir un ensemble de variants génétiques et de biomarqueurs, afin d'identifier les patients les plus à risque de développer une pathologie cardiaque, et d'adapter autant que possible leur exposition aux thérapies cardiotoxiques, ainsi que de mettre en place un suivi et une prévention.

Chapitre 6. IDENTIFICATION DE BIOMARQUEURS GENETIQUES ASSOCIEES A L'APPARITION DE PATHOLOGIES CARDIAQUES CHEZ DES PATIENTS TRAITES POUR UN CANCER DE L'ENFANT : ETUDE CAS-TEMOIN APPARIEE

Dans ce chapitre, nous avons cherché à identifier des biomarqueurs du transcriptome associés aux pathologies cardiaques dans un échantillon de patients survivants d'un cancer de l'enfant niché dans la cohorte FCCSS. Ce travail a été présenté oralement aux congrès EPICLIN (prix de la meilleure communication, Nancy) et ISORED (prix de la meilleure communication, Sitges) et en communication affichée à l'ISCB (Milan) en 2023. Il a été soumis au journal *Journal of Clinical Oncology* en Juillet 2024.

6.1 INTRODUCTION

L'amélioration substantielle des traitements du cancer a considérablement augmenté la survie globale des patients, avec un taux de survie à 5 ans de plus de 80% chez les patients traités pour un cancer dans l'enfance (Hudson et al., 2013; Leerink et al., 2020). Bien que ces traitements soient fortement associés aux pathologies cardiaques, une variabilité interindividuelle de la susceptibilité à la cardiotoxicité chez des survivants d'un cancer dans l'enfance traités de façon similaire est observée. Cela laisse suggérer l'existence de facteurs génétiques modulant cette susceptibilité (Aminkeng et al., 2016; Clemens et al., 2018). Plusieurs études portant sur les biomarqueurs génétiques ont été menées, se concentrant sur (i) les caractéristiques génétiques de la cardiotoxicité induite par les anthracyclines (Hildebrandt et al., 2017) et (ii) les marqueurs pronostiques associés à cette cardiotoxicité (Aminkeng et al., 2015; Chaix et al., 2020; Petrykey et al., 2021; Visscher et al., 2012, 2013, 2015; X. Wang et al., 2014, 2016). Bien que plusieurs variants génétiques aient été identifiés comme associés aux pathologies cardiaques,

aucun consensus n'a été trouvé. Contrairement au génome, le transcriptome, c'est-à-dire l'expression des gènes, est très dynamique, et sensible aux expositions aux xénobiotiques, y compris aux traitements anticancéreux qui peuvent affecter de manière significative le profilage de l'expression des gènes (Joseph, 2017). L'étude de l'expression des gènes pourrait alors apporter de nouvelles connaissances sur les mécanismes de cardiotoxicité chez les patients traités par anthracyclines et radiothérapie impliquant le cœur. À ce jour, une seule étude a analysé les données transcriptomiques de 104 patients en identifiant les gènes différentiellement exprimés entre les survivants présentant une cardiomyopathie et les survivants n'en présentant pas (Singh et al., 2023).

L'objectif de ce chapitre est de présenter une étude cas-témoin appariée visant à identifier des biomarqueurs transcriptomiques associés à la survenue de pathologies cardiaques après traitement d'un cancer de l'enfant, en utilisant des méthodes statistiques développées pour des données de grande dimension.

6.2 MATERIELS ET METHODES

6.2.1 Etude cas-témoin

L'étude cas-témoins est nichée dans la French Childhood Cancer Survivors Study (n=7670) et inclus une partie des patients pour lesquels un prélèvement sanguin a été obtenu (n=1571), comme décrit dans le Chapitre 2. Parmi ces survivants, 191 patients ont développé une pathologie cardiaque, et le transcriptome a été obtenu pour 181 d'entre eux. Finalement, seuls 165 survivants avec données transcriptomiques et ayant développé une pathologie cardiaque (les cas) ont pu être appariés à 165 témoins sans pathologie cardiaque. L'appariement a été réalisé sur les variables suivantes : le sexe, l'âge au diagnostic du premier cancer (± 3 ans), le type de la première tumeur, et la durée du suivi (chaque témoin a un suivi supérieur ou égal à celui de son cas). La **Figure 18** présente l'organigramme de sélection des patients inclus dans l'étude cas-témoins.

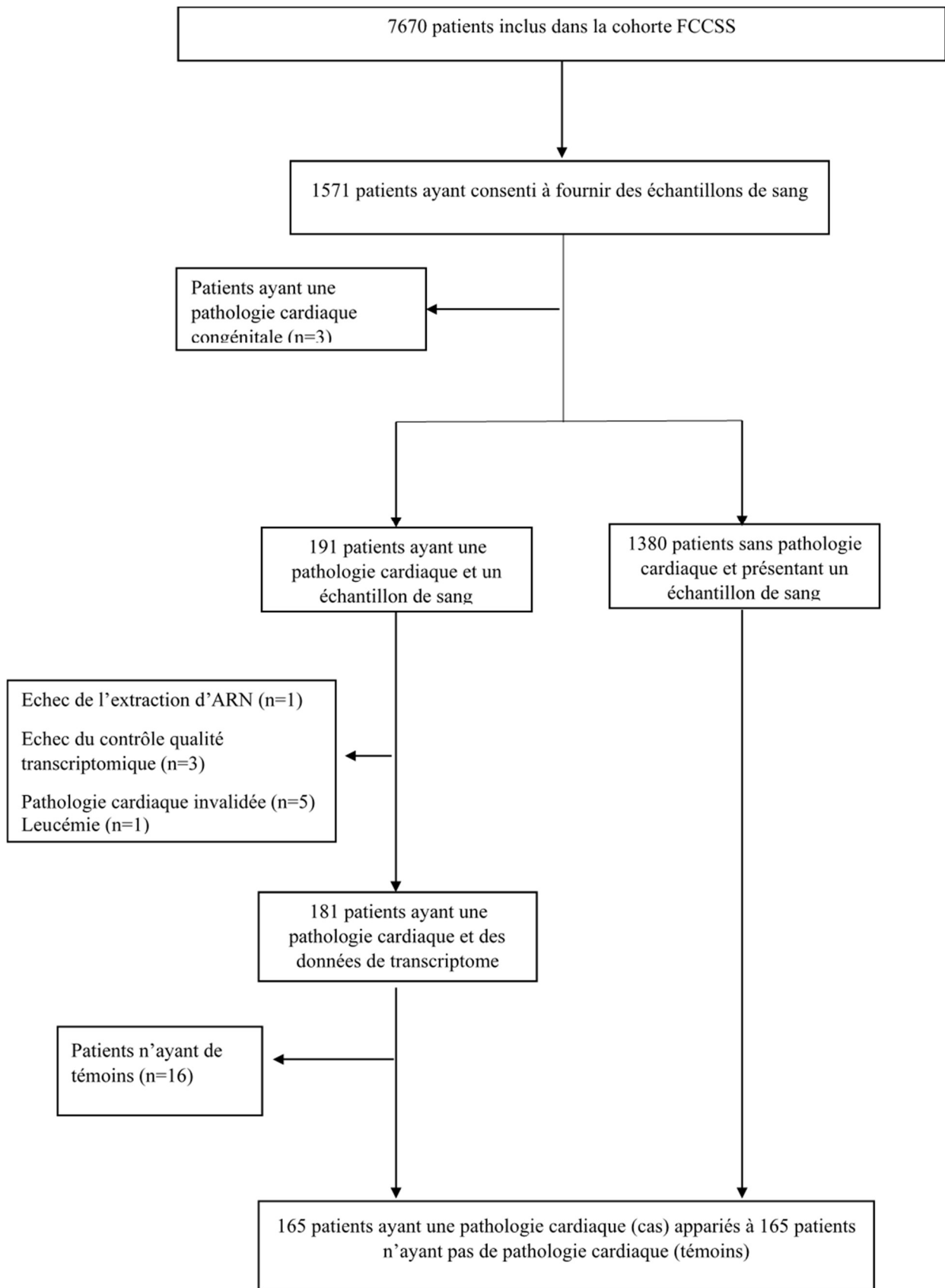


Figure 18. Organigramme de l'inclusion des patients dans l'étude cas-témoin nichée dans la cohorte FCCSS

6.2.2 Extraction de l'ARN messager (ARNm)

Les cellules mononucléaires du sang périphérique ont été isolées et des cultures de lignées cellulaires de lymphocytes ont été réalisées. L'extraction de l'ARN a été réalisée sur les cellules lymphocytaires des 165 cas et de leurs témoins en utilisant la méthode du réactif TRIzol (ThermoFischer, Invitrogen). Pour déterminer la concentration d'ARN, un spectrophotomètre Nanodrop (www.nanodrop.com) a été utilisé, et la qualité des préparations d'ARN a été évaluée à l'aide de la technologie Bioanalyser 2100 de Lab-on-a-chip (Agilent technologies). La concentration de tous les ARN a été ajustée et vérifiée par une seconde mesure sur un spectrophotomètre Nanodrop.

Pour la majorité des cas (n=160), le prélèvement sanguin a été effectué après le diagnostic de la pathologie cardiaque, et pour les cinq autres cas, le prélèvement sanguin a eu lieu moins d'un an avant le diagnostic de pathologie cardiaque.

6.2.3 Analyses bioinformatiques

Les données d'expression de gènes (transcriptome) ont été obtenues via une puce Affymetrix Human Clariom-D pour 33 267 gènes et les analyses bioinformatiques ont été effectuées par Genosplice (Paris, France). Les données des puces ont été normalisées en utilisant la normalisation par quantile. Le contrôle qualité de l'expression des gènes a été vérifié pour différentes métriques ("pos_vs_neg_auc", "all_probeset_rle_mean", métriques "%P", pics bactériens et expression des ARN de contrôle polyA) (de la Grange et al., 2007). La correction du bruit de fond a été effectuée à l'aide des sondes antigéniques. Seules les sondes ciblant les exons annotés à partir des transcriptions de la FAST DB (v2016₁) ont été sélectionnées afin de se concentrer sur les gènes bien annotés, et dont les séquences d'ARNm se trouvent dans des bases de données publiques (de la Grange et al., 2007). Parmi ces sondes sélectionnées, les sondes de mauvaise qualité (par exemple, les sondes identifiées par Affymetrix comme "cross-hybridizing") et les sondes ayant un signal d'intensité trop faible par rapport aux

sondes de fond antigéniques ayant le même contenu en Guanine-Cytosine (GC) ont été éliminées des analyses. Seules les sondes ayant une valeur significativement ($p \leq 0,05$) supérieure à la valeur de fond (Detection above background, DABG) dans au moins la moitié des puces ont été prises en compte pour les analyses statistiques. Seuls les gènes exprimés dans au moins une condition comparée ont été analysés (c'est-à-dire les échantillons de cas et de contrôle). Pour être considérée comme exprimée, la p-valeur du DABG devait être $\leq 0,05$ pour au moins la moitié des sondes génétiques.

Le contrôle qualité effectué par l'équipe bio-informatique de Genosplice indique que seulement 11 986 gènes peuvent être analysés. En outre, 3 247 gènes insuffisamment exprimés pour être confirmés biologiquement ont été exclus des analyses statistiques, ce qui donne un nombre total de 8 739 gènes (**Figure 19**).

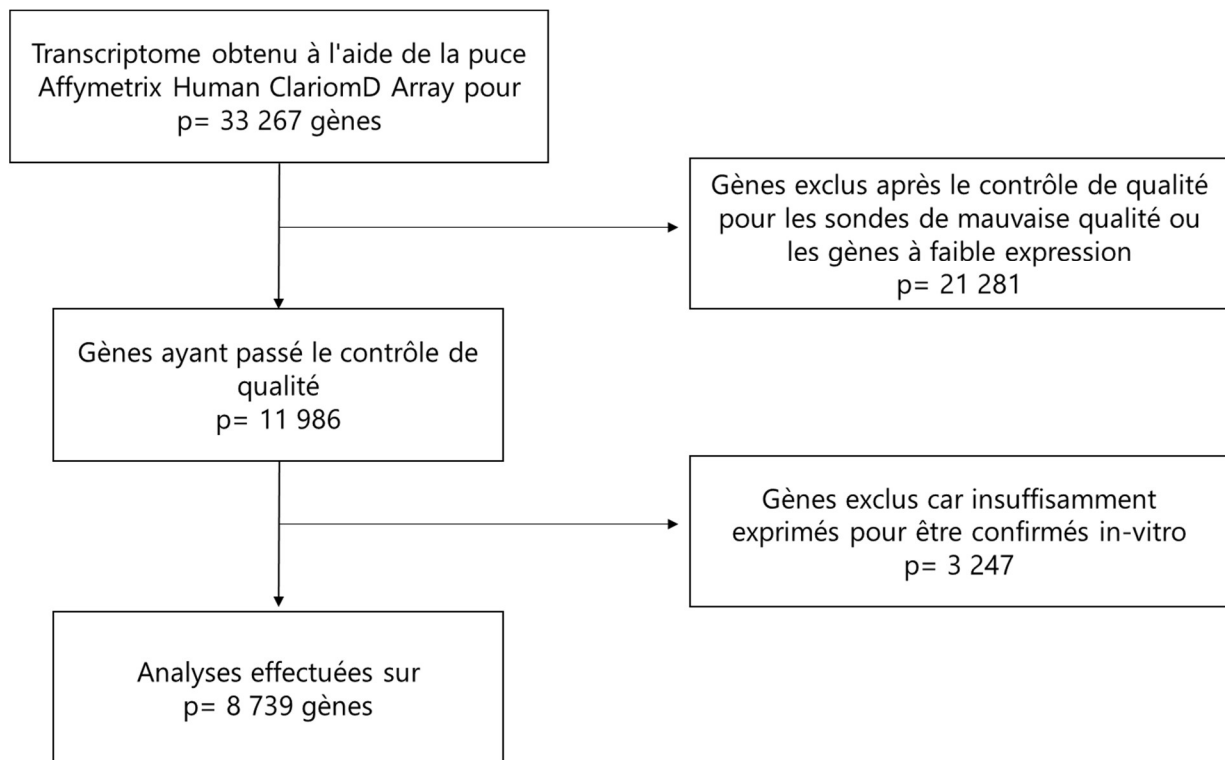


Figure 19. Diagramme indiquant l'inclusion des gènes dans l'étude

6.2.4 Etude in-vitro

Une analyse *in-vitro* a été effectuée pour mesurer l'effet des anthracyclines sur l'expression des gènes identifiés dans les cardiomyocytes dérivés de ventricules de rats adultes en utilisant la RT-qPCR. L'hypothèse sous-jacente est que l'exposition des cardiomyocytes aux anthracyclines entraînerait une modification de l'expression des gènes identifiés. Six échantillons de cardiomyocytes ont été exposés à trois concentrations d'anthracycline : 0 μ M (contrôle), 1 μ M et 10 μ M pendant 24 heures. L'expression du gène l'EPAC1 a été utilisée comme contrôle positif en raison de sa réponse dose-dépendante connue aux anthracyclines (Mazevet et al., 2023).

6.2.5 Analyses statistiques

6.2.5.1 Critère principal et secondaire

L'évènement principal est la survenue d'une première pathologie cardiaque, quel que soit son grade définissant ainsi les cas de l'étude cas-témoin. L'évènement secondaire est la survenue d'une insuffisance cardiaque.

Une analyse de sensibilité a été réalisée en se limitant aux patients ayant un premier évènement cardiaque de grade 2 et plus.

6.2.5.2 Sélection des biomarqueurs

Afin d'identifier les biomarqueurs d'expression de gènes associés aux différents critères dans le cadre de données à grande dimension et compte tenu de la nature du design de l'étude (appariement cas-témoins), la procédure de sélection Lasso (Least Absolute Shrinkage and Selection Operator) adaptée à une régression logistique conditionnelle a été utilisée. Le package R `clogitL1` a été utilisé. Les variables d'expression de gènes normalisées ont été incluses dans la régression logistique conditionnelle pénalisée Lasso, sans ajustement des covariables. Les trois approches de stabilisation décrites dans le Chapitre 3 section 3.3.5, ont été utilisées pour s'assurer de la stabilité de la sélection des biomarqueurs d'expression de gènes. Pour l'approche Percentile Lasso,

le paramètre de pénalisation θ a été fixé au 95^{ème} percentile (selon les recommandations des auteurs) de la distribution du paramètre de pénalisation obtenue à partir de 1 000 exécutions de la régression logistique conditionnelle pénalisée Lasso. Quant aux approches Bolasso et Sublasso, celles-ci sélectionnent des gènes dont la fréquence de sélection est supérieure ou égale à un seuil prédéfini, de sorte que plus la fréquence de sélection d'un gène est élevée, moins il y a de doutes sur la force de l'association entre le gène et le critère étudié. Différents seuils (50, 60 et 70 %) ont été examinés. Le nombre de rééchantillonnages est fixé à 1 000, et les sous-échantillons de l'approche Sublasso ont été construits à partir de 60 % du jeu de données d'origine.

La sélection finale des gènes est définie comme l'intersection des ensembles de gènes résultant des trois approches de stabilisation, et un seuil de 70 % a été fixé pour les approches Bolasso et Sublasso.

6.2.5.3 Estimation de l'effet des biomarqueurs sélectionnés

Les effets des gènes sélectionnés ont été estimés par une régression logistique conditionnelle ajustée sur l'année de diagnostic du premier cancer (avant 1970, [1970 - 1979], [1980 - 1989], [1990 - 2000]), les doses cumulées d'anthracycline (aucune,]0 - 250],]250 - 360], >360 mg/m²) et les doses moyennes de radiations au cœur (aucune,]0 - 5],]5 -15], >15 Gy). Dans des analyses supplémentaires, les doses moyennes de radiations reçues au cœur ont été remplacées par l'estimateur dose-volume V_{15} indiquant le volume du cœur ayant reçu au moins 15 Gy (0,]0 - 10],]10 - 50], > 50 %). Les biomarqueurs d'expression génétique ont été analysés comme des variables continues standardisées, car le test de Box-Tidwell (Box & Tidwell, 1962) n'a pas permis d'observer d'écart à l'hypothèse de log-linéarité.

La corrélation et l'interaction entre les gènes sélectionnés et les deux variables de traitement (doses cumulées d'anthracyclines et doses moyennes de radiations au cœur) ont été étudiées. Une association est considérée comme statistiquement

significative pour un seuil inférieur à 5%.

6.2.5.4 *Performance discriminante du modèle*

La valeur ajoutée des gènes sélectionnés a été quantifiée par la statistique du rapport de vraisemblance, et la performance discriminatoire a été évaluée à l'aide de l'AUC non pondéré avec correction de biais et de son intervalle de confiance en utilisant l'approche logistique conditionnelle, comme décrit dans le Chapitre 3 section 3.2.6. Ces deux métriques ont été estimées à partir de trois modèles qui diffèrent en termes de covariables : un modèle de base incluant l'année de diagnostic du premier cancer, la dose cumulée d'anthracyclines (aucune,]0 - 250],]250 - 360], >360 mg/m²), et la dose moyenne de radiations au cœur (aucune,]0 - 5],]5 - 15], > 15 Gy) (modèle 1), un second modèle incluant l'année de diagnostic du premier cancer et les gènes sélectionnés (modèle 2), et un troisième modèle incluant à la fois l'année du diagnostic du premier cancer, la dose cumulée d'anthracyclines, la dose moyenne de radiations du cœur et les gènes sélectionnés (modèle 3). Nous avons ensuite comparé les métriques obtenus pour les trois modèles afin d'évaluer la valeur ajoutée de notre sélection de gènes, le modèle 3 étant considéré comme le modèle final. Une validation interne par validation croisée a été effectuée pour évaluer la performance discriminante du modèle final.

6.2.5.5 *Analyse in vitro*

Les niveaux d'expression génétique pour les trois concentrations d'anthracycline ont été décrits et comparés à l'aide du test non paramétrique de Kruskal-Wallis.

6.3 RESULTATS

6.3.1 Description des caractéristiques de l'étude cas-témoin

Le **Tableau 22** présente les caractéristiques démographiques et cliniques des patients inclus dans l'étude cas-témoin appariée.

Tableau 22. Caractéristiques démographiques et cliniques des patients inclus dans l'étude cas-témoins appariée (n=330)

	Cas (n=165) n (%)	Témoins (n=165) n (%)	p-value
Genre*			
Féminin	78 (47,3)	78 (47,3)	
Masculin	87 (52,7)	87 (52,7)	
Âge au diagnostic du premier cancer (années) *			
Médiane (EI)	6 (2 – 11)	6 (2 – 11)	
[0 – 5[72 (43,6)	73 (44,2)	
[5 – 10[44 (26,7)	43 (26,1)	
[10 – 15[36 (21,8)	41 (24,8)	
[15 – 22[13 (7,9)	8 (4,8)	
Type de tumeur primaire *			
Lymphome	59 (35,8)	59 (35,8)	
Neuroblastome	27 (16,4)	27 (16,4)	
Sarcome osseux	23 (13,9)	23 (13,9)	
Néphroblastome	20 (12,1)	20 (12,1)	
Sarcome des tissus mous	18 (10,9)	18 (10,9)	
Tumeur gonadique	6 (3,6)	6 (3,6)	
Autres	12 (7,3)	12 (7,3)	
Pathologie cardiaque			
Insuffisance cardiaque	74 (44,8)		
Valvulopathie	41 (24,8)		
Trouble du rythme cardiaque	32 (19,4)		
Péricardite	11 (6,7)		
Pathologie ischémique	7 (4,2)		
Suivi après diagnostic du cancer (années)*			
Médiane (min - max)	20 (0 - 54)	33 (10 – 65)	
[0 – 10[*	30 (18,2)	0 (0,0)	
[10 – 20[48 (29,1)	15 (9,1)	
[20 – 30[50 (30,3)	46 (27,9)	
[30 – 40[32 (19,4)	59 (35,8)	
[40 – 50[4 (2,4)	32 (19,4)	
≥ 50	1 (0,6)	13 (7,9)	
Année de diagnostic du premier cancer			
Avant 1970	17 (10,3)	29 (17,6)	
[1970, 1979[57 (34,5)	43 (26,0)	
[1980, 1989]	67 (40,6)	65 (39,4)	
[1990, 2000]	24 (14,5)	28 (17,0)	
Anthracyclines n (%)			
Non	57 (34,5)	75 (45,5)	
Oui	108 (65,5)	90 (54,5)	
Doses cumulées d'anthracyclines (mg/m²)			0,0004
Médiane (min - max)	355,1 (100,0 – 1207,5)	248,4 (20,8 – 546,1)	
Aucune	57 (34,5)	75 (45,5)	
]0 – 250]	32 (19,4)	45 (27,3)	
]250 – 360]	25 (15,2)	23 (13,9)	

>360	51 (30,9)	22 (13,3)	
Radiations au cœur			
Non	51 (30,9)	56 (33,9)	
Oui	114 (69,1)	109 (66,1)	
Doses moyennes de radiations au cœur (Gy)			0,0410
Médiane (min - max)	9,5 (0,0 – 43,2)	0,1 (0,0 – 66,0)	
Aucune	51 (30,9)	56 (33,9)	
]0 – 5]	54 (32,7)	66 (40,0)	
]5 – 15]	10 (6,1)	14 (8,5)	
>15	50 (30,3)	29 (17,6)	
Volume du cœur (%) ayant reçu ≥ 15 (Gy)**			0.0129
Médiane (min - max)	80,1 (42,0 -96,9)	71,1 (29,7 -93,7)	
None	99 (60,0)	116 (70,3)	
]0 -10]	7 (4,2)	5 (3,0)	
]10 – 50]	11 (6,7)	16 (9,7)	
>50	48 (29,1)	28 (17,0)	

Abréviations : EI : écart interquartile ; Gy : Gray ; les p-valeurs ont été obtenues à l'aide du test exact unilatéral Tau-c de Stuart-Kendall.

* Variables utilisées pour l'appariement entre les cas et les témoins ; * Chez les cas, le suivi commence à partir du diagnostic de cancer et se termine au développement de la pathologie cardiaque, qui peut survenir avant les 5 premières années de suivi ; **V15 reflète le pourcentage du volume cardiaque qui a reçu au moins 15 Gy.

L'âge médian au moment du diagnostic du premier cancer était de 6 ans (EI, 2-11 ans) pour les cas et les témoins. Les principaux types de tumeurs malignes étaient les lymphomes (35,8 %), les neuroblastomes (16,4 %), les sarcomes osseux (13,9 %), les néphroblastomes (12,1 %) et les sarcomes des tissus mous (10,9 %). L'insuffisance cardiaque était la pathologie cardiaque la plus fréquente (44,8 %), suivie par les valvulopathies (24,8 %), des troubles du rythme cardiaque (19,4 %), des péricardites (6,7 %) et des ischémies (4,2 %). Les cas ont reçu des doses cumulées d'anthracycline significativement plus élevées, 30,9 % d'entre eux ayant reçu des doses supérieures à 360 mg/m², contre 13,3 % des témoins (p= 0,0004). En outre, les cas ont été exposés à des doses moyennes au cœur plus élevées, 30,3 % d'entre eux ayant reçu des doses supérieures à 15 Gy contre 17,6 % des témoins (p= 0,0410).

De plus, pour l'ensemble des 165 cas, le délai médian entre le diagnostic de pathologie cardiaque et le prélèvement sanguin était de 12,5 ans (étendue interquartile EI, 6,1 - 16,6). En outre, l'âge médian des cas au moment de la prise de sang (39,7, (EI, 34,4 -

45,8)) était proche de celui des témoins (41,7, (EI, 34,6 - 46,7)).

6.3.2 Sélection de gènes associés aux pathologies cardiaques (critère principal)

6.3.2.1 Sélection des gènes

La sélection des gènes par les trois approches de stabilisation, Percentile Lasso, Bolasso et Sublasso, met en évidence la réduction significative du nombre de gènes lorsque des contraintes sont appliquées (**Tableau 23**).

Tableau 23. Nombre de gènes sélectionnés par les approches Percentile lasso, Bolasso et Sublasso avec ou sans contrainte

Approche	Sans contrainte **	Contrainte appliquée					
		$\theta=95\%$ †	40 %	50 %	60 %	70 % ‡	80 % *
Percentile Lasso	61	45					
Bolasso	1 525			5	2	1	0
Sublasso	1 311			6	2	1	1

Abréviations : **Sans contrainte, correspond à $\theta=0$ pour le Percentile Lasso et aucun seuil spécifié pour les approches Bolasso et Sublasso, † $\theta=95\%$ correspond à $\hat{\lambda}(95) = 5.41$, ‡Le même gène, NFE2L2, est sélectionné pour les deux approches Bolasso et Sublasso lors de l'application du seuil de 70%, *Seule l'approche Sublasso sélectionne un gène, NFE2L2, lors de l'application d'un seuil de 80%.

L'approche Percentile Lasso sélectionne initialement 61 gènes, qui ont été réduits à 45 avec la contrainte utilisant le 95^{ème} percentile de la distribution du paramètre de pénalisation comme valeur finale de pénalisation. En revanche, les approches Bolasso et Sublasso identifient initialement 1 525 et 1 311 gènes, respectivement. Ces nombres diminuent à cinq et six gènes, respectivement, pour une fréquence de sélection supérieure ou égale à 50 % (**Tableau 24**).

Tableau 24. Fréquence de sélection des gènes sélectionnés par les approches Bolasso et Sublasso avec un seuil de 50 % à partir de 1 000 réplifications. Ces gènes ont également été sélectionnés pour l'approche Percentile Lasso

Gènes sélectionnés	Fréquence de sélection	
	Bolasso, n (%)	Sublasso, n (%)
NFE2L2	757 (75,7)	805 (80,5)
TGIF2 - RAB5IF	622 (62,2)	662 (66,2)
HK2	577 (57,7)	565 (56,5)
HDGFL3	535 (53,5)	527 (52,7)
LSP1	508 (50,8)	551 (55,1)
RGS1*		527 (52,7)

Abréviations : * : La fréquence de sélection du gène RGS1 étant inférieure à 50 % selon l'approche de Bolasso, nous n'avons pas indiqué sa fréquence.

L'intersection des ensembles de gènes résultant des trois approches de stabilisation est représentée par un diagramme de Venn (**Figure 20.A**), révélant un ensemble de cinq gènes différentiellement exprimés (NFE2L2, TGIF2-RAB5IF, HK2, HDGFL3 et LSP1) entre les cas et les témoins (**Figure 20.B**). Les gènes HDGFL3, NFE2L2 et TGIF2-RAB5IF sont sous-exprimés chez les cas par rapport aux témoins, tandis que les gènes LSP1 et HK2 sont surexprimés chez les cas par rapport aux témoins. Lorsqu'un seuil de 70 % a été appliqué aux approches Bolasso et Sublasso, NFE2L2 a été le seul gène résultant de l'intersection, et la **Figure 21** présente la distribution de son niveau d'expression entre les cas et les témoins.

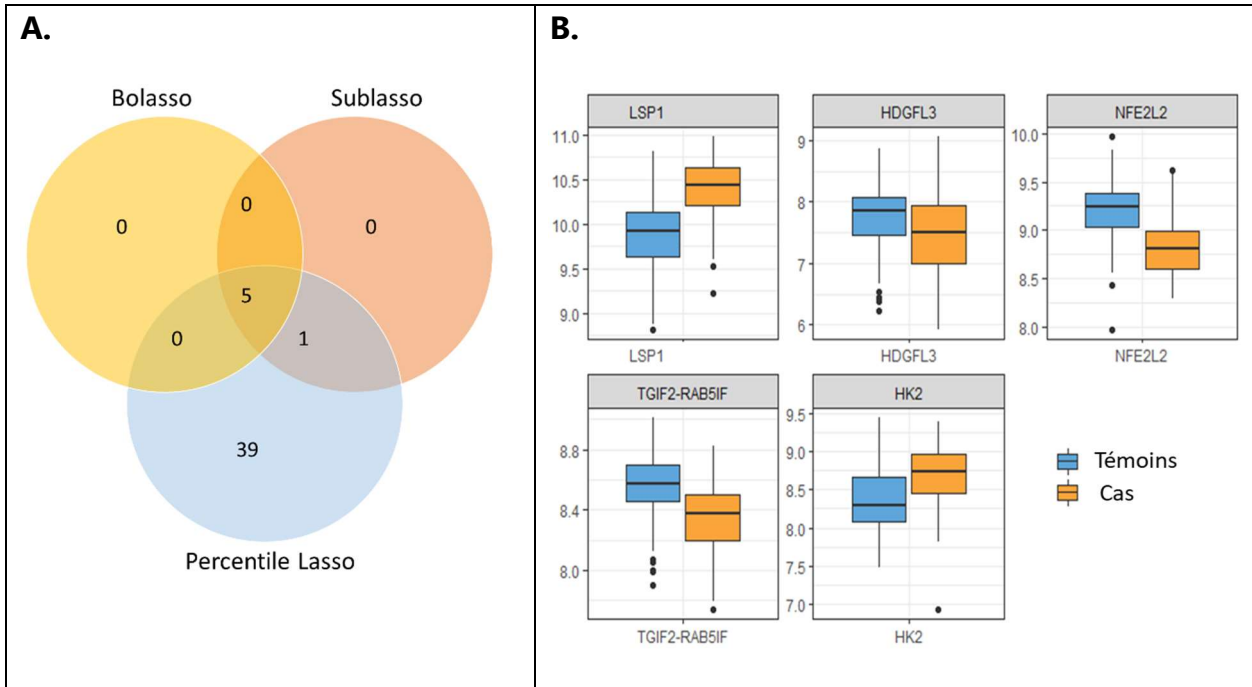


Figure 20. A. Diagramme de Venn représentant les gènes, identifiés par les approches Percentile Lasso ($\theta=95\%$), Bolasso et Sublasso (seuil 50 %) ; **B.** Box plots de l'expression des gènes associés aux pathologies cardiaques chez les patients présentant une pathologie cardiaque (cas) et les patients n'en présentant pas (témoins), gènes qui résultent de l'intersection des trois approches de stabilisation

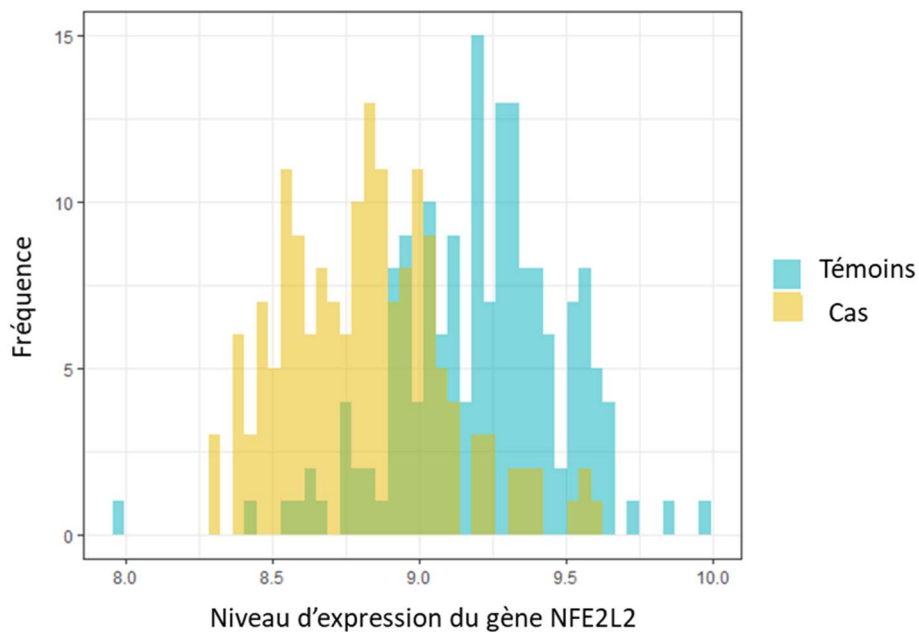


Figure 21. Histogramme de l'expression de NFE2L2 chez les patients présentant une pathologie cardiaque (cas) et les patients n'en présentant pas (témoins)

6.3.2.2 Estimation de l'effet des gènes

Le **Tableau 25** rapporte l'effet du gène NFE2L2 estimé à partir d'une régression logistique conditionnelle sur la pathologie cardiaque ajustée sur l'année du diagnostic du premier cancer, les doses cumulées d'anthracycline et les doses moyennes de radiations au cœur (modèle final), et montre qu'une augmentation de l'expression du gène NFE2L2 est significativement associée à une réduction du risque de pathologies cardiaques (OR=0,16 ; 95 % IC= 0,09-0,29, $p < 0,0001$). A noter que les effets des doses cumulées d'anthracyclines et des doses moyennes de radiations au cœur sont non significatifs dans cette analyse. Or dans un modèle n'incluant pas l'expression du gène NFE2L2, ces deux variables de traitement étaient significatives (**Annexe C 1**).

Aucune interaction significative n'a été observée entre l'expression du gène NFE2L2 et les doses cumulées d'anthracycline et les doses moyennes de radiations au cœur (**Tableau 26**). Même si les quatre autres biomarqueurs n'ont finalement pas été retenus, leurs interactions potentielles avec les variables liées aux anthracyclines et à la radiothérapie ont été étudiées et n'ont pas révélé d'interactions significatives (**Annexe C 2**).

Tableau 25. Effet de l'expression du gène NFE2L2 sur les pathologies cardiaques estimé à partir d'une régression logistique conditionnelle ajustée sur l'année de diagnostic du cancer et l'exposition aux anthracyclines et à la radiothérapie incluant le cœur

Variables	n / N	Odds Ratio (95 % IC)	p-value
Année de diagnostic du cancer			0,3341
< 1970	17 / 46	1,00	
[1970-1979]	57 / 100	1,29 (0,31 - 5,29)	
[1980-1989]	67 / 132	0,78 (0,17 - 3,64)	
[1990-2000]	24 / 52	2,66 (0,41 - 17,11)	
Dose cumulée d'anthracyclines (mg/m²)			0,2121
Aucune	57 / 132	1,00	
]0 - 250]	32 / 77	1,01 (0,29 - 3,52)	
]250 - 360]	25 / 48	2,96 (0,79 - 11,06)	
>360	51 / 73	3,56 (0,88 - 10,53)	
Dose moyenne de radiations au cœur (Gy)			0,5218
Aucune	51 / 107	1,00	
]0 – 5]	54 / 120	0,97 (0,38 - 2,53)	
]5 – 15]	10 / 24	0,65 (0,12 - 3,43)	
>15	50 / 79	1,94 (0,62 - 6,05)	
NFE2L2[⊥]		0,16 (0,09 - 0,29)	< 0,0001

Abréviations : n : nombre de cas ; N : nombre total de patients ; IC : intervalle de confiance ; [⊥] L'odd ratio a été estimé à 0,17 (95 % IC, 0,10 - 0,28) lorsque les doses cumulées d'anthracyclines et les doses moyennes d'irradiation du cœur ne sont pas incluses dans le modèle logistique conditionnel.

Tableau 26. Effet de l'expression gène NFE2L2 en fonction des niveaux de doses cumulées d'anthracyclines et des doses moyennes d'irradiation du cœur (termes d'interaction) à l'aide d'une régression logistique conditionnelle ajustée sur l'année de diagnostic du cancer et les variables liées au anthracycline et à l'irradiation

Gene	n/N	Odds Ratio (95 % IC)	p-value
Cumulative anthracyclines dose (mg/m²)			
NFE2L2			0,3288
Aucune	57 / 132	0,08 (0,03 – 0,24)	
1 – 250	32 / 77	3,23 (0,82 – 12,60)	
250 – 360	25 / 48	2,98 (0,76 – 11,71)	
> 360	51 / 73	2,16 (0,54 – 8,66)	
Doses moyennes de radiations au cœur (Gy)			
NFE2L2			0,6933
Aucune	51 / 107	0,22 (0,10 – 0,47)	
0 – 5	54 / 120	0,80 (0,28 – 2,35)	
5 – 15	10 / 24	0,46 (0,11 – 7,16)	
>15	50 / 79	1,06 (0,37 – 3,09)	

Abréviations : Gy : Gray, IC : Intervalle de Confiance, N : nombre de patients, n : nombre de cas

Des corrélations, cependant, ont été observées entre l'expression du gène NFE2L2 et les doses cumulées d'anthracyclines ($p= 0,0081$) et les doses moyennes de radiations au cœur ($p= 0,0089$) (**Tableau 27**). Les patients recevant de fortes doses d'anthracyclines ($>360 \text{ mg/m}^2$) ont montré une plus faible expression du gène NFE2L2 par rapport aux patients non exposés aux anthracyclines ($p= 0,0046$). De plus, les patients exposés à des radiations au cœur supérieures à 15 Gy ont moins exprimé le gène NFE2L2 comparativement aux patients n'ayant pas reçu de radiations au cœur ($p= 0,0703$) ou les patients ayant reçu moins de 5 Gy ($p= 0,0045$).

Tableau 27. Estimation de la corrélation entre l'expression du gène NFE2L2 et les doses cumulées d'anthracycline ou les doses moyennes de radiations au cœur.

	n/N	Moyenne (min-max)	Comparaison	p-value ajustée [¶]
Doses cumulées d'anthracyclines (mg/m²)				
Expression de NFE2LE				
Aucune	57 / 132	9,08 (8,30 – 9,97)		
0 – 250	32 / 77	8,99 (7,97 – 9,64)	Aucune vs. 0 – 250	0,1881
250 – 360	25 / 48	9,01 (8,39 – 9,66)	Aucune vs. 250 - 360	0,2753
> 360	51 / 73	8,92 (8,28 – 9,62)	Aucune vs. > 360	0,0046
			0 – 250 vs. 250 - 360	0,7264
			0 – 250 vs. > 360	0,2563
			250 – 360 vs. > 360	0,2423
Kruskal-Wallis $\chi^2 = 11,79$, p-value = 0,0081				
Doses moyennes de radiations au cœur (Gy)				
Expression de NFE2LE				
Aucune	51 / 107	9,02 (8,26 – 9,84)		
0 – 5	54 / 120	9,07 (8,30 – 9,97)	Aucune vs. 0 – 5	0,3769
5 – 15	10 / 24	9,04 (8,39 – 9,62)	Aucune vs. 5 – 15	0,8663
> 15	50 / 79	8,90 (7,97 – 9,64)	Aucune vs. > 15	0,0703
			0 – 5 vs. 5 – 15	0,7303
			0 – 5 vs. > 15	0,0045
			5 – 15 vs. > 15	0,2168
Kruskal-Wallis $\chi^2 = 11,59$, p-value = 0,0089				

Abréviations : n = nombre de cas ; N = nombre total de patients ; ¶ Test post-hoc de Dunn pour les comparaisons multiples. Les corrélations sont testées une par une.

6.3.2.3 Performances des modèles

Pour quantifier la valeur ajoutée du gène NFE2L2, nous avons comparé la statistique du rapport de vraisemblance des trois modèles décrits dans la section 6.2.5.4. Nous avons aussi comparé leur performance discriminante en comparant les AUC obtenus à partir d'une régression logistique conditionnelle. Les résultats sont présentés pour les deux approches dans le **Tableau 28**.

Tableau 28. Comparaison de différents modèles de régression logistique conditionnelle sur la base de la statistique du rapport de vraisemblance et de l'aire sous la courbe (AUC) pour le critère pathologie cardiaque

Modèle	-2 log vraisemblance	ddl	AUC (95 % IC)	AUC (95 % IC)	corrigé [†]
Régression logistique conditionnelle					
Modèle Nul (sans covariable)	228,7				
Modèle 1: Année de diagnostic du cancer + variables de traitement [†]	197,7	9	0,66 (0,62 – 0,73)	0,64 (0,60 – 0,68)	
Modèle 2: Année de diagnostic du cancer + NFE2L2	110,9	4	0,84 (0,79 – 0,89)	0,83 (0,79 – 0,86)	
Modèle 3: Année de diagnostic du cancer + variables de traitement + NFE2L2	102,9	10	0,85 (0,80 – 0,89)	0,83 (0,80 – 0,87)	

ddl : degrés de liberté ; †Les variables de traitement sont la dose cumulée d'anthracyclines (aucune, 0 à 250, 250 à 360, >360 mg/m²) et la dose moyenne d'irradiation cardiaque (aucune, 0 à 5, 5 à 15, >15 Gy). AUC aire sous la courbe correspond à une AUC non pondérée, CI intervalle de confiance calculé à partir de bootstrap †correspond à une AUC corrigée par bootstrap pour contrôler le surajustement.

Nous présentons les résultats en termes d'AUC obtenus à partir de l'approche logistique conditionnelle puisque les résultats produits par l'approche non conditionnelle sont similaires. A partir du modèle nul (sans covariable, -2log-vraisemblance= 228,7), l'ajout des variables de traitement à l'année de diagnostic du cancer améliore considérablement l'ajustement aux données (modèle 1, -2log-vraisemblance= 197,7, ddl= 9). Cette diminution est encore plus importante lorsque l'expression du gène NFE2L2 est ajouté à l'année de diagnostic du premier cancer (modèle 2, -2log-vraisemblance= 110,9, ddl= 4). Cela montre bien la valeur ajoutée de ce gène. Même si l'ajout des deux variables de traitement au modèle 2 (modèle 3) n'améliore pas significativement la qualité de l'ajustement (p= 0,2398, ddl= 6), ce dernier présente la meilleure performance globale (model 3, -2log-vraisemblance= 102,9, ddl= 10). L'information ajoutée par NFE2L2 est également confirmée par une

meilleure capacité discriminante avec un AUC plus élevé pour le modèle 3 (0,85, 95 % IC= 0,80-0,89) par rapport au modèle 1 (0,66, 95 % IC= 0,62-0,73). Ces AUC restent robustes si l'on considère une estimation corrigée de l'AUC pour contrôler un surajustement (dernière colonne du **Tableau 28**).

La validation interne par échantillonnage fractionné réalisée à partir de l'ajustement du modèle final (modèle 3) sur 500 jeu de données représentant 60 % du jeu de données cas-témoin a montré un bon AUC non corrigé sur l'échantillon des données de validation (0,84, 95 % IC= 0,77-0,89). Les boîtes à moustaches de la **Figure 22** montrent la distribution des AUC estimés sur l'échantillon d'entraînement, représentant 60 % du jeu de donnée cas-témoins apparié, et sur l'échantillon de validation, qui lui représente 40 % du jeu de donnée cas-témoins apparié.

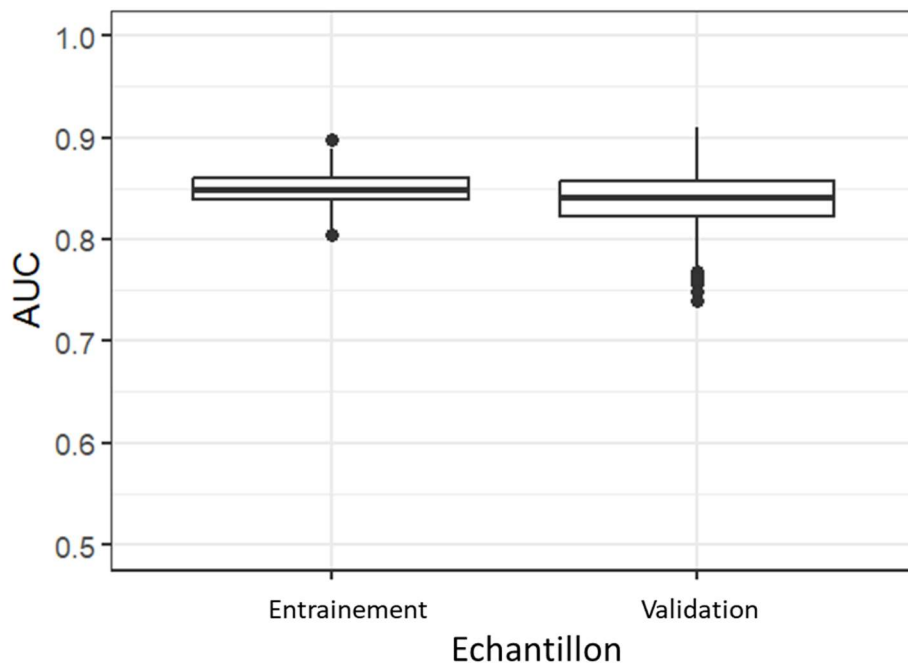


Figure 22. Boîtes à moustaches de l'AUC calculée sur les 500 répliques du jeu de données d'entraînement et du jeu de données de validation

Les analyses considérant non plus la dose moyenne de radiations au cœur mais le volume du cœur ayant reçu au moins 15 Gy (V15) donnent des résultats similaires (**Annexe C 3**)

6.3.2.4 Analyse de sensibilité

L'analyse de sensibilité a été menée sur un sous-échantillon de 132 cas ayant une pathologie cardiaque de grade ≥ 2 , et leurs 132 témoins appariés. L'expression du gène NFE2L2 est également significativement associé à la pathologie cardiaque de grade ≥ 2 (OR= 0,14, 95 % IC= 0,07 – 0,29) (**Tableau 29**), et les variables de traitements (doses cumulées d'anthracyclines et doses moyennes de radiations au cœur) ne sont pas significatives dans ce modèle. De plus, ce modèle incluant à la fois les variables de traitement et le gène NFE2L2 a une très bonne capacité discriminante (AUC non corrigé= 0,84, 95 % IC= 0,78 – 0,89).

Tableau 29. Effet de l'expression du gène NFE2L2 sur les pathologies cardiaques de grade ≥ 2 , estimé à partir d'une régression logistique conditionnelle ajustée sur l'année de diagnostic du cancer et l'exposition aux anthracyclines et à la radiothérapie au cœur

Variables	n/N	Odds Ratio (95 % IC)	p-value
Année de diagnostic du cancer			0,1470
< 1970	53/ 101	1,00	
[1970-1979]	18/ 34	1,2 (0,26 – 6,20)	
[1980-1989]	45/ 88	0,79 (0,13 – 4,72)	
[1990-2000]	16 / 41	5,21 (0,58 – 46,55)	
Dose cumulée d'anthracyclines (mg/m²)			0,2753
Aucune	42/ 97	1,00	
]0 - 250]	24/ 55	0,77 (0,20 - 3,08)	
]250 - 360]	22/ 43	2,63 (0,64 – 10,80)	
>360	44/ 69	2,94 (0,80 - 10,90)	
Dose moyenne de radiations au cœur (Gy)			0,6457
Aucune	42/ 88	1,00	
]0 – 5]	39/ 91	0,85 (0,28 - 2,58)	
]5 – 15]	7/ 19	0,58 (0,09 - 3,57)	
>15	44/ 66	1,66 (0,42 - 6,50)	
NFE2L2⁺		0,14 (0,07 - 0,29)	< 0,0001

Abréviations : N : Nombre total de patient, n : nombre de cas ; Gy : Gray, IC : Intervalle de confiance

6.3.3 Sélection de gènes associés à l'insuffisance cardiaque (critère secondaire)

À partir de l'étude cas-témoins appariée, un sous-échantillon de 148 patients, dont 74 cas diagnostiqués avec une insuffisance cardiaque comme premier évènement cardiaque et leurs témoins appariés, a été sélectionné. Les caractéristiques démographiques et cliniques de ce sous-échantillon de patients sont présentées en **Annexe C 4**.

6.3.3.1 Sélection des gènes

La procédure de sélection des gènes est identique à celle utilisée pour l'évènement pathologie cardiaque. La régression correspondant au 95^{ème} percentile dans l'approche Percentile Lasso sélectionne 23 gènes. En raison de la taille limitée de l'échantillon, le seuil de sélection a été fixé à 50 % pour les approches Bolasso et Sublasso qui sélectionnent toutes les deux le même gène NFE2L2 (**Tableau 30**).

Tableau 30. Nombre de gènes sélectionnés par les approches Percentile lasso, Bolasso et Sublasso avec ou sans contrainte

Approche	Sans contrainte**	Contrainte appliquée					
		$\theta=95\%$	40 %	50 %*	60 %	70 %	80 %
Percentile Lasso	41	23					
Bolasso	1 320		1	1	0	0	0
Sublasso	1 182		1	1	0	0	0

Abréviations : **Sans contrainte : correspond à $\theta = 0$ pour le Percentile Lasso et aucun seuil spécifié pour les approches Bolasso et Sublasso ; $\theta=95\%$ correspond à $\hat{\lambda}(95)=5,41$, * le même gène, NFE2L2, est sélectionné pour les deux approches Bolasso et Sublasso lors de l'application du seuil de 50 %.

L'intersection des trois approches de stabilisation est illustrée par un diagramme de Venn (**Figure 23. A**) et met en évidence l'unique gène, NFE2L2, qui constitue notre sélection finale. Ce gène a une fréquence de sélection de 52,6 % dans Bolasso et de 57,0 % dans Sublasso. Comme l'insuffisance cardiaque représentant 45 % des pathologies cardiaques étudiées, ce résultat n'est pas complètement surprenant.

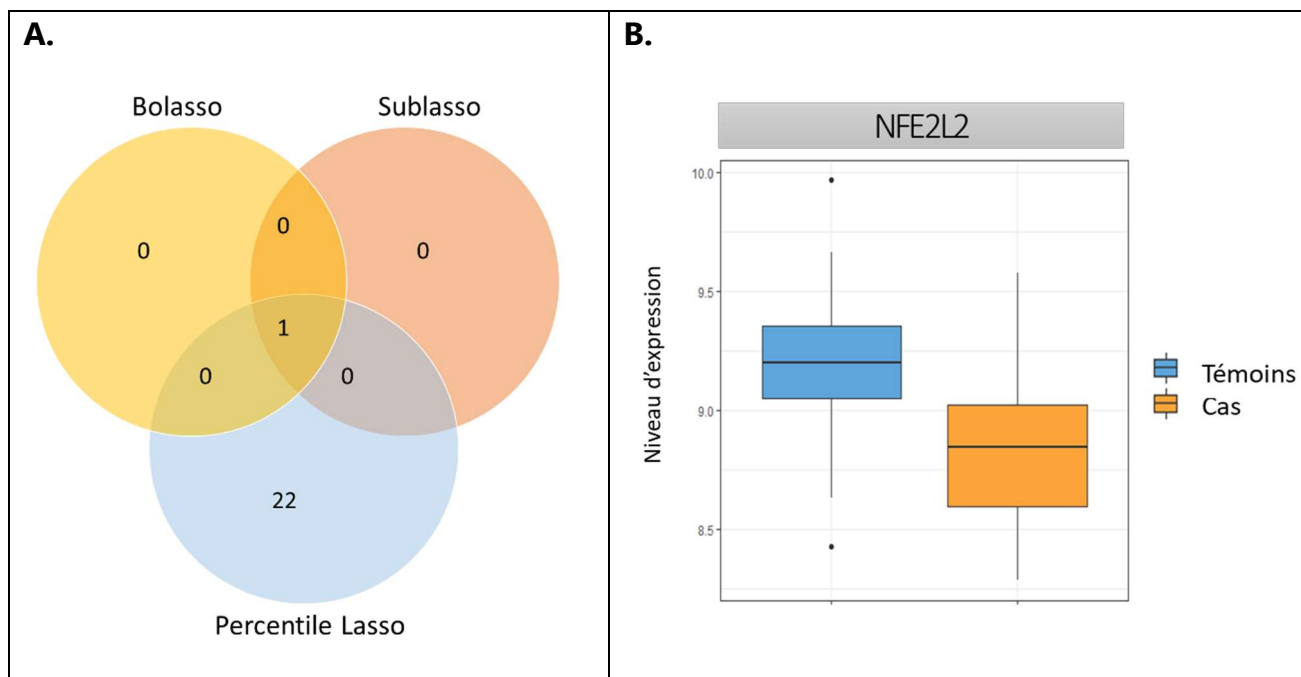


Figure 23. A. Diagramme de Venn représentant le gène identifié par les approches Percentile Lasso ($\theta = 95\%$), Bolasso et Sublasso (seuil 50%) ; **B.** Box plots de l'expression du gène NFE2L2 associés à l'insuffisance cardiaque chez les patients présentant une insuffisance cardiaque (cas) et les patients n'en présentant pas (témoins)

6.3.3.2 Estimation de l'effet des biomarqueurs sélectionnés

Le gène NFE2L2 est sous-exprimé chez les cas par rapport aux témoins (**Figure 23. B**), de sorte qu'une augmentation de l'expression du gène NFE2L2 est associée à une diminution du risque d'insuffisance cardiaque (OR= 0,11, 95 % IC= 0,03-0,37) (**Tableau 31**), indiquant un effet protecteur des forts niveaux d'expression du gène NFE2L2. De plus, l'expression du gène NFE2L2 est significativement corrélée aux doses cumulées d'anthracyclines ($p = 0,0002$), de sorte que les patients recevant de fortes doses d'anthracyclines ($> 360 \text{ mg/m}^2$) expriment moins le gène NFE2L2, par rapport à ceux non exposés aux anthracyclines ($p < 0,0001$) (**Tableau 32**). Aucune interaction entre l'expression du gène NFE2L2 et les doses de traitement reçues n'a été observée.

Tableau 31. Effet de l'expression du gène NFE2L2 sur l'insuffisance cardiaque estimé à partir d'une régression logistique conditionnelle ajustée sur l'année de diagnostic du cancer et l'exposition aux anthracyclines et à la radiothérapie incluant le cœur

Variables	n / N	Odds Ratio (95 % IC) ^{⊥⊥}	p-value
Année de diagnostic du cancer[†]			0,8241
< 1980	33 / 68	1,00	
[1980-2000]	41 / 80	1,22 (0,22 - 6,76)	
Dose cumulée d'anthracyclines (mg/m²)[†]			0,0004
Aucune	7 / 37	1,00	
]0 - 250]	13 / 33	12,97 (0,47 - 358,30)	
>250	54 / 78	53,62 (1,86 - 1548,03)	
Dose moyenne de radiations au cœur (Gy)[†]			0,9405
Aucune	27 / 56	1,00	
]0 - 15]	31 / 68	1,47 (0,30 - 7,10)	
>15	16 / 24	1,38 (0,12 - 16,30)	
NFE2L2^{⊥⊥}		0,11 (0,03 - 0,37)	<0,0001

Abréviations : n = nombre de cas ; N = nombre total de patients ; IC= intervalle de confiance: ⊥⊥ L'Odds Ratio a été estimé à 0,12 (IC à 95 %, 0,04 - 0,31) lorsque les doses de traitements ne sont pas incluses dans le modèle logistique conditionnel, † Certaines catégories ont été regroupées en raison de la petite taille de l'échantillon.

Tableau 32. Estimation de la corrélation entre l'expression du gène NFE2L2 et les doses cumulées d'anthracycline ou les doses moyennes de radiations du cœur. Les corrélations sont testées une par une.

	n/N	Moyenne (min-max)	Comparaison	p-value ajustée [†]
Doses cumulées d'anthracyclines (mg/m²)				
NFE2LE expression				
Aucune	7 / 37	9,21 (8,75 - 9,97)		
0 - 250	13 / 33	8,99 (8,42 - 9,60)	Aucune vs, 0 - 250	0,0124
250 - 360	18 / 33	9,04 (8,39 - 9,66)	Aucune vs, 250 - 360	0,0618
> 360	36 / 45	8,89 (8,28 - 9,62)	Aucune vs, > 360	< 0,0001
			0 - 250 vs, 250 - 360	0,4898
			0 - 250 vs, > 360	0,2680
			250 - 360 vs, > 360	0,0751
Kruskal-Wallis $\chi^2 = 19,61$ p-value = 0,0002				
Doses moyennes de radiations au cœur (Gy)				
NFE2LE expression				
Aucune	27 / 56	9,02 (8,29 - 9,58)		
0 - 5	28 / 60	9,07 (8,41 - 9,97)	Aucune vs, 0 - 5	0,6738
>5*	19 / 32	8,94 (8,38 - 9,64)	Aucune vs, > 5	0,1807
			0 - 5 vs, > 5	0,1610
Kruskal-Wallis $\chi^2 = 3,89$, p-value = 0,1432				

Abréviations : n = nombre de cas ; N = nombre total de patients ; † Test post-hoc de Dunn pour les comparaisons multiples. Les corrélations sont testées une par une.

6.3.3.3 Performances des modèles

Comme pour l'évènement pathologie cardiaque (critère principal), trois modèles ont été comparés afin de quantifier la valeur ajoutée du gène NFE2L2 par l'estimation de la statistique du rapport de vraisemblance, et d'évaluer les performances discriminantes à partir des AUC obtenus à partir d'une régression logistique conditionnelle. Des résultats similaires à ceux des pathologies cardiaques ont été observés (**Tableau 33**).

Tableau 33. Comparaison de différents modèles de régression logistique conditionnelle sur la base de la statistique du rapport de vraisemblance et de l'aire sous la courbe (AUC) pour le critère insuffisance cardiaque

Modèle	-2 log vraisemblance	ddl	AUC (95 % IC)	AUC corrigé [¶] (95 % IC)
Régression logistique conditionnelle				
Modèle Nul (sans covariable)	102,6			
Modèle 1: Année de diagnostic du cancer + variables de traitement [†]	66,9	5	0,77 (0,68 – 0,84)	0,74 (0,68 – 0,82)
Modèle 2: Année de diagnostic du cancer + NFE2L2	49,7	2	0,82 (0,75 – 0,88)	0,81 (0,75 – 0,88)
Modèle 3: Année de diagnostic du cancer + variables de traitement + NFE2L2	33,7	6	0,85 (0,80 – 0,90)	0,83 (0,78 – 0,88)

Abréviations : ddl : degrés de liberté ; [†] Les variables de traitement sont la dose cumulée d'anthracyclines (aucune, 0 à 250, >250 mg/m²) et la dose moyenne d'irradiation cardiaque (aucune, 0 à 5, 5 > Gy). Nous avons regroupé les deux dernières catégories en raison du faible nombre de patients dans la dernière catégorie. AUC aire sous la courbe correspond à une AUC non pondérée, CI intervalle de confiance calculé à partir de bootstrap [¶] correspond à une AUC corrigée par bootstrap pour contrôler le surajustement

Comme pour les pathologies cardiaques, le modèle 3 est le modèle qui ajuste le mieux les données (-2log-vraisemblance= 33,7, ddl= 6), ce qui met bien en évidence la valeur ajoutée du gène NFE2L2 dans le modèle. Le modèle 1 qui n'inclut pas de variable génétique a une bonne capacité discriminante (AUC= 0,77, 95 % IC= 0,68 – 0,84), et celle-ci est améliorée par l'ajout de l'expression du gène NFE2L2 à l'année de diagnostic de premier cancer (modèle 2, AUC= 0,82, 95 % IC= 0,75 – 0,88). Par ailleurs,

contrairement au critère principale (pathologies cardiaques), l'ajout des variables de traitements au modèle 2 a permis une augmentation de 3 points de sa capacité discriminante (modèle 3, AUC= 0,85, 95 % IC= 0.80 – 0.90). Ces AUC restent robustes si l'on considère une estimation corrigée de l'AUC pour contrôler un surajustement (dernière colonne du **Tableau 33**).

6.3.4 Etude *in-vitro*

L'analyse *in-vitro* a montré une diminution significative de l'expression du gène NFE2L2 pour des doses d'anthracyclines cliniquement pertinentes (10 µM). Le traitement des cardiomyocytes avec 1 µM d'anthracyclines pendant 24 heures n'a pas affecté de manière significative l'expression du gène NFE2L2 par rapport aux cardiomyocytes non exposés (différence moyenne de -0,09, 95 % IC= -0,37-0,19 ; p= 0,7586). Cependant, l'exposition à une concentration de 10 µM d'anthracyclines pendant 24 heures a diminué l'expression du gène NFE2L2 de 0,65 unité (95 % IC= 0,37- 0,93, p= 0,0022) et de 0,74 unité (95 % IC= 0,46-1,03, p= 0,0032), par rapport aux cardiomyocytes non exposés et à ceux exposés à 1 µM d'anthracyclines, respectivement (**Tableau 34**).

Tableau 34. Comparaison des niveaux d'expression du gène NFE2L2 avant et après exposition aux anthracyclines dans des cardiomyocytes de rats adultes

	Moyenne (min-max)	Comparaison	Différence moyenne (95 % IC)	p-value ajustée †
Témoins	1,00 (1,00 – 1,00)			
Ant 1 µM	1,09 (0,81 – 1,73)	Témoins vs. Ant 1 µM	-0,09 (-0,37 – 0,19)	0,7586
Ant 10 µM	0,35 (0,18 – 0,63)	Témoins vs. Ant 10 µM	0,65 (0,37 – 0,93)	0,0022
		Ant 1 µM vs. Ant 10 µM	0,74 (0,46 – 1,02)	0,0032
Kruskal-Wallis $\chi^2 = 13,97$, p-value = 0,0009				

Abréviations : Ant : Anthracyclines, IC : Intervalle de confiance, † Test post-hoc de Dunn pour les comparaisons multiples.

6.4 DISCUSSION

Ce chapitre rapporte, à travers une étude cas-témoin appariée nichée dans la cohorte FCCSS, l'identification du gène NFE2L2 dont l'expression est associée aux pathologies cardiaques de manière générale et plus spécifiquement à l'insuffisance cardiaque après le traitement d'un cancer de l'enfant. Ce gène est sous-exprimé chez les cas ayant développé une pathologie cardiaque (médiane= 8,81, Etendue interquartile, EI= 8,60 - 8,99) par rapport aux témoins (médiane= 9,24, EI= 9,03 - 9,39) présentant les mêmes caractéristiques cliniques : sexe, âge au premier diagnostic de cancer (± 3 ans) et type de diagnostic de cancer. Ce résultat a également été observé dans le sous-échantillon constitué de patients atteints d'insuffisance cardiaque (médiane= 8,49, EI= 8,60-9,02) et leurs témoins (médiane= 9,20, EI= 9,05-9,36). L'expression du gène NFE2L2 aurait donc un effet protecteur associé à la pathologie cardiaque (OR ajusté= 0,16 ; 95 % IC= 0,09-0,29) et à l'insuffisance cardiaque (OR ajusté= 0,11 ; 95 % IC= 0,03-0,37) pour une augmentation d'une unité de l'expression de NFE2L2. De plus, l'étude *in vitro* portant sur l'évolution des niveaux d'expression du gène NFE2L2 sur des cardiomyocytes de rats exposés à des anthracyclines, a montré une diminution de l'expression de ce gène lorsque les cardiomyocytes étaient exposés pendant 24h à des doses d'anthracyclines de 10 μ M par rapport aux cardiomyocytes non exposés. Chez les survivants d'un cancer de l'enfant, l'inclusion de l'expression du gène NFE2L2 dans un modèle comprenant les doses d'anthracycline cumulées et des doses moyennes de radiations du cœur a significativement amélioré sa capacité discriminante, atteignant une AUC de 0,85 (95 % IC= 0,80-0,89) à la fois pour la pathologie cardiaque et l'insuffisance cardiaque, et suggérant une meilleure prise en compte de la variabilité interindividuelle dans la survenue des pathologies cardiaques. A titre de comparaison, notre modèle sans NFE2L2 modélisant l'insuffisance cardiaque a montré un AUC (0,77 ; 95 % IC= 0,68 – 0,84) légèrement meilleure que celui du modèle développé par Chow et al. (2015) portant sur les survivants d'un cancer dans l'enfance (0,74, intervalle de confiance non

rapporté). Ce modèle, incluait des variables démographiques (sexe, âge au diagnostic du cancer) et des caractéristiques de traitement du cancer (doses d'anthracycline et de radiothérapie thoracique).

Dans la littérature (Chapitre 5), de nombreuses études génétiques antérieures sur la cardiotoxicité chez les survivants du cancer de l'enfant ont montré des résultats discordants, identifiant différents variants génétiques dans différents gènes. L'étude de l'expression des gènes, qui est plus sensible aux effets du traitement, est d'une importance majeure pour la compréhension des mécanismes associés à la cardiotoxicité (Joseph, 2017). À ce jour, une seule et très récente étude a examiné les données transcriptomiques, identifiant 36 gènes différentiellement exprimés chez les survivants d'un cancer dans l'enfance exposés aux anthracyclines (Singh et al., 2023). Ces gènes sont impliqués dans différentes voies biologiques canoniques comme la "fibrose hépatique/activation des cellules stellaires hépatiques" et la "signalisation de l'homéostasie du fer". Cependant, le gène NFE2L2 n'a pas été identifié dans cette étude. Deux raisons principales peuvent être avancées pour expliquer ce résultat : (i) l'utilisation de puces à ARN différentes et (ii) la méthodologie statistique utilisée. En ce qui concerne cette dernière, notre résultat découle d'une approche qui est adaptée aux données aux grandes dimensions, la régression Lasso, adaptée au design cas-témoins de l'étude. De plus, l'utilisation de trois approches de stabilisation, Percentile Lasso, Bolasso et Sublasso, dérivées de la régression logistique conditionnelle pénalisée Lasso, a permis l'obtention de résultats plus stables. La sélection des gènes, résultant de l'intersection de ces trois procédures de sélection, nous permet de minimiser le taux de faux positifs. Cette stratégie a abouti à l'identification de cinq gènes, dont le gène NFE2L2 qui apparaît comme un biomarqueur robuste aux différents seuils de sélection. D'un point de vue biologique, le gène NFE2L2 code pour le *nuclear factor erythroid 2-related factor 2* (NRF2), une protéine de la réponse antioxydante, impliquée dans diverses fonctions cellulaires telles que la défense antioxydante, l'inflammation, la

biogénèse des mitochondries, et l'homéostasie du fer, cette dernière étant l'une des principales voies identifiées dans l'étude de Singh et al. (2023). Cet aspect antioxydant suggère alors un lien mécanique potentiel avec la cardiotoxicité induite par le traitement (J. A. Roberts et al., 2023). En tant que régulateur principal des réponses antioxydantes, la protéine NRF2 cible directement environ 1% du génome humain et module plus de 240 gènes à travers diverses fonctions biologiques. Cette protéine protège aussi les tissus cardiaques contre le stress oxydatif (Q. M. Chen & Maltagliati, 2018; J. A. Roberts et al., 2023; Taverne et al., 2013), et son rôle dans l'homéostasie du fer et son effet anti-inflammatoire peuvent protéger contre la fibrose cardiaque précédant la cardiomyopathie et la carence en fer fréquente dans les pathologies cardiaques (J. A. Roberts et al., 2023; Savarese et al., 2023). En effet, le développement de la pathologie cardiaque est un phénomène biologique complexe qui peut être causé par l'exposition répétée des cellules du système cardiovasculaire à des substances toxiques telles que les médicaments pharmaceutiques dont les anthracyclines, ainsi qu'à des molécules endogènes telles que les dérivés réactifs de l'oxygène (DRO) qui peuvent agir comme des stimuli toxiques à des concentrations élevées (Roberts et al., 2023 ; Taverne et al., 2013).

En ce qui concerne le rapport entre l'expression du gène NFE2L2 et l'exposition aux anthracyclines et aux radiations au cœur, nous avons remarqué que les variables de traitement sont significativement associées à la pathologie cardiaque dans un modèle ajusté sur l'année de diagnostic du cancer mais ne sont plus significativement associées lorsque l'expression du gène NFE2L2 est ajoutée. Ceci peut s'expliquer par les importantes corrélations observées entre les doses cumulées d'anthracyclines et l'expression du gène NFE2L2. Ainsi, les patients recevant de fortes doses d'anthracyclines exprimaient moins le gène NFE2L2, par rapport à ceux non exposés aux anthracyclines. Cette tendance a aussi été observée dans le modèle *in-vitro* mené sur des cardiomyocytes de rats adultes exposés à des doses cliniquement pertinentes d'anthracyclines. Cela suggère une implication plus large de NFE2L2 dans la pathologie

cardiaque au-delà des effets induits par le traitement du cancer, ou que l'association de la dose de traitement avec la pathologie cardiaque peut être médiée par l'impact sur l'expression de NFE2L2. Cela démontre la double nature du transcriptome, qui peut être considéré comme le reflet d'un processus changeant ou comme un lien dans un chemin de causalité (Zeller et al., 2010).

Chez les souris et les cardiomyocytes isolés traités avec des anthracyclines, l'augmentation de l'expression de NFE2L2 protège contre le dysfonctionnement cardiaque, tandis que la déficience en NRF2 augmente les effets cardiotoxiques des anthracyclines (Li et al., 2014; X. Zhang et al., 2020). Notre étude a mis en évidence une diminution de l'expression du gène NFE2L2 dans les cardiomyocytes de rat exposés à des doses cliniquement pertinentes d'anthracyclines, suggérant que sa régulation à la baisse pourrait être associée à l'affaiblissement du tissu cardiaque. Toutefois, il convient de noter que le transcriptome de notre étude provient de cellules mononucléaires du sang périphérique, dont la réponse aux traitements peut différer de celle des cardiomyocytes. En outre, une étude sur l'athérosclérose a montré que le transcriptome des monocytes peut intégrer des facteurs génétiques et non génétiques, qui sont importants pour comprendre la physiopathologie de la maladie et l'évaluation des risques (Zeller et al., 2010). Plusieurs molécules induisant la production de NRF2 ont été testées sur des modèles animaux (souris et rats), et il a été prouvé qu'ils protégeaient contre les troubles cardiaques. La tanshinone IIA, un constituant bioactif majeur ayant une activité antioxydante, a été signalée comme protégeant contre la cardiotoxicité induite par les anthracyclines. Cette molécule a été testée en prétraitement chez des souris traitées par des anthracyclines afin de déterminer si elle induisait ou non la signalisation de NRF2. Il s'est avéré qu'elle augmentait de manière significative les niveaux de protéines NRF2 dans le tissu cardiaque par rapport aux témoins (Guo et al., 2018). L'étude de la tanshinone IIA (Guo et al., 2018) reste une piste intéressante pour de futures études cherchant à minimiser le risque de pathologies cardiaques après un traitement par anthracyclines d'un cancer de l'enfant.

Les limites de cette étude doivent, cependant, être discutées. L'évènement principal comprend plusieurs pathologies cardiaques, ayant chacune des physiopathologies différentes, ce qui nécessite des analyses séparées pour chaque pathologie, comme cela a été fait pour l'insuffisance cardiaque. Cependant, le faible nombre de patients avec une pathologie particulière n'a pas permis d'effectuer ces analyses. L'analyse principale est basée sur la survenue de pathologie cardiaque quel que soit le grade, mais l'analyse de sensibilité restreinte aux seuls grades 2 ou supérieurs a montré des résultats cohérents. Une autre limite est le fait que le sous-échantillon de la cohorte FCCSS utilisé dans l'étude cas-témoin peut ne pas être représentatif la cohorte entière, ce qui pourrait entraîner une sous-estimation de l'AUC pour un modèle ne comprenant que des variables cliniques. Il serait alors possible de pallier à cette sous-estimation en utilisant des méthodes pondérées pour le calcul de l'AUC. En l'absence de données externes, nous n'avons pas pu valider définitivement le rôle de NFE2L2 dans un échantillon indépendant de patients, même si la validation interne donne une capacité discriminante prometteuse au modèle incluant l'expression du gène NFE2L2.

En conclusion, cette étude suggère que les survivants d'un cancer dans l'enfance présentant une forte expression du gène NFE2L2 étaient moins susceptibles de souffrir d'une pathologie cardiaque. Ce résultat doit, cependant, être validé par un suivi longitudinal de l'expression du gène NFE2L2 avant et après le traitement du cancer de l'enfant, afin de proposer des traitements cardioprotecteurs aux patients en amont du développement d'une pathologie cardiaque (en tant que biomarqueur pronostique) ou lors du diagnostic précoce d'un dysfonctionnement cardiaque (en tant que biomarqueur diagnostique).

Chapitre 7. DISCUSSION GENERALE

Ce chapitre résume les résultats des trois principales parties de la thèse et fournit notre interprétation de ces résultats, les implications potentielles pour l'amélioration du suivi à long terme des survivants d'un cancer de l'enfant, ainsi que des propositions pour de futures recherches.

7.1 RESUME DES PRINCIPAUX RESULTATS

Dans la première partie de cette thèse, Chapitre 4, nous avons étudié les risques associés aux traitements cardiotoxiques dans la survenue de pathologies cardiaques tardives suite au traitement d'un cancer de l'enfant au sein de la cohorte FCCSS. Nous avons mis en évidence le rôle de l'irradiation à de faibles volumes cardiaques ainsi que l'interaction entre l'irradiation cardiaque et l'administration d'anthracyclines. Nous avons également évalué au sein de cette cohorte les risques d'hospitalisations pour causes cardiaques et l'impact des facteurs de risque liés aux traitements et aux comorbidités sur ces hospitalisations. Nous avons ensuite, à travers une revue de la littérature dans le Chapitre 5, fait l'état actuel des connaissances sur les facteurs génétiques associés aux pathologies cardiaques chez les survivants d'un cancer de l'enfant, et discuté l'importance de la prise en compte de ces facteurs dans la surveillance et la prise en charge des dysfonctionnements cardiaques. Dans le chapitre 6, nous avons cherché à identifier des biomarqueurs du transcriptome associés aux pathologies cardiaques dans une sous population de la cohorte FCCSS.

Les résultats présentés dans le Chapitre 4 confirment que le traitement du cancer de l'enfant par anthracyclines et/ou radiothérapie cardiaque est fortement associé à un risque élevé de pathologies cardiaques tardives, y compris d'insuffisance cardiaque. L'exposition des patients à des doses de radiations cardiaques de 5 à 20 Gy entraîne une augmentation du risque de développer des pathologies cardiaques quel que soit

le volume du cœur irradié. Nous avons également observé une augmentation du risque de pathologies cardiaques chez les survivants exposés à des doses de radiations au cœur dès 1 Gy, en comparaison aux patients non exposés aux radiations cardiaques. L'interaction entre l'administration d'anthracyclines et les doses moyennes de radiations reçues au cœur a été confirmée par notre étude. Ces résultats viennent renforcer les résultats précédemment observés dans la même cohorte mais dans une population plus restreinte (Haddy et al., 2016; Mansouri et al., 2019). Le risque de survenue de pathologies cardiaques associé aux fortes doses de radiations (>25 Gy) est similaire que les patients soient traités ou non par anthracyclines. Toutefois, le risque de base de pathologies cardiaques associé aux anthracyclines est plus élevé chez les patients qui y sont exposés.

En termes d'hospitalisations cardiaques, les survivants d'un cancer de l'enfant présentent un taux 4,6 fois (95 % IC= 4,0 - 5,2) plus élevé à celui attendu dans la population générale. Cela représente 45,8 hospitalisations supplémentaires (95 % IC= 38,7 - 53,6) pour 10000 personnes-années. Les femmes ont d'hospitalisation pour cause cardiaque plus important que les hommes. Et pour les deux genres, ce risque est plus élevé chez les survivants d'un cancer de l'enfant par rapport à la population générale. Les taux d'hospitalisations pour cause cardiaque suite au traitement d'un lymphome ou d'un neuroblastome sont plus élevés que pour les autres types de tumeurs primaires. De plus, l'excès de risque d'hospitalisation associé au diagnostic des lymphomes comme tumeur primaire est de 93,5 par rapport à la population générale. Ces résultats peuvent être expliqués par l'intensité des protocoles de traitements de ces types de cancers, particulièrement avant la prise de conscience des effets à long termes associés aux traitements anticancéreux. Les patients traités par des doses modérées d'anthracyclines (entre 250 et 360 mg/m²) ont un risque deux fois plus élevé d'être hospitalisés par rapport à ceux non exposés aux anthracyclines. Par ailleurs, les survivants ayant reçu une irradiation cardiaque d'au moins 30 Gy présentent un risque très élevé d'hospitalisation, même lorsque de petits volumes cardiaques (< 10%) sont

irradiés, et l'augmentation de ce risque est proportionnelle au volume du cœur irradié. Même si pour les doses d'anthracyclines $< 250 \text{ mg/m}^2$ et de radiations au cœur allant de 5 à 15 Gy, le risque d'hospitalisation cardiaque n'était pas significatif, une tendance similaire au risque associé à la survenue de pathologies cardiaques tous grades et grades ≥ 3 était observée. Cette différence peut s'expliquer par des différences dans la population étudiée. En effet, dans l'étude sur les hospitalisations cardiaques, seulement 83% des patients sont chaînés au SNDS donc inclus dans les analyses. De plus, seules les hospitalisations ayant eu lieu entre 2011 et 2018 ont pu être étudiées ce qui pourrait conduire à une sous-estimation des risques associés aux hospitalisations cardiaques.

La prise en compte des traitements cardiotoxiques a permis d'améliorer la prédiction de la survenue des pathologies cardiaques grâce au développement de modèles de prédiction (Chow et al., 2015) définissant des profils de patients à risque de cardiotoxicité. Cependant, la capacité prédictive de ces modèles reste limitée (Chow et al., 2015; Leerink et al., 2020). Une variabilité interindividuelle liée à la cardiotoxicité a été observée chez les survivants d'un cancer de l'enfant, soulignant la nécessité d'une meilleure compréhension de la susceptibilité génétique aux pathologies cardiaques induites par les traitements anticancéreux. Cette compréhension pourrait non seulement améliorer le suivi à long terme de ces patients, mais aussi le développement de stratégies adaptées au suivi et à prise en charge thérapeutiques des futurs patients. En réponse à ceci, nous avons proposé dans la seconde partie de cette thèse (Chapitre 5), une vue d'ensemble des résultats publiés sur les facteurs génétiques associés aux pathologies cardiaques chez les survivants d'un cancer dans l'enfance. Il en ressort, que sur les 19 études identifiées, seules les cardiomyopathies ont été étudiées, et plusieurs variants génétiques dans 63 gènes ont été retrouvés comme étant significativement associées à la cardiomyopathie. Ces études ont pour certaines cherché à mettre en évidence des interactions entre les gènes/variants et les doses cumulées d'anthracyclines. Cependant, aucune étude n'a exploré les interactions entre les gènes ou variants identifiés et l'exposition des patients à la radiothérapie. Le nombre élevé

de variants et de gènes identifiés à travers cette revue de la littérature, ce qui rend difficile le développement de stratégies de prévention définitives. L'hétérogénéité des résultats observés à travers les différentes études fait qu'à ce jour, il n'existe encore aucun consensus sur les gènes impliqués dans la cardiotoxicité liée aux traitements des cancers de l'enfant. Cette absence de consensus conduit à explorer de nouveaux domaines tels que le transcriptome. En effet, le transcriptome est sensible à l'exposition aux xénobiotiques, permettant d'identifier des biomarqueurs directement utilisables pour la détection de pathologies résultant de l'exposition à des traitements (Joseph, 2017). Dans ce contexte et dans la troisième partie de cette thèse (Chapitre 6), nous avons cherché à identifier des biomarqueurs associés à la survenue de pathologies cardiaques chez des patients traités pour un cancer dans l'enfance, à travers l'étude de l'expression des gènes. Nous avons donc mené une étude cas-témoins appariée, nichée dans la cohorte FCCSS, afin d'identifier des biomarqueurs transcriptomiques associés à la survenue de pathologies cardiaques dans cette population. Nous avons identifié le gène NFE2L2 dont l'expression est associée aux pathologies cardiaques de manière générale, et plus spécifiquement à l'insuffisance cardiaque après le traitement d'un cancer de l'enfant. Ce gène était sous-exprimé chez les patients ayant développé une pathologie cardiaque comparé aux témoins appariés, et une tendance similaire était observée lorsque nous nous sommes concentrés sur les insuffisances cardiaques. L'expression du gène NFE2L2 aurait donc un effet protecteur associé à la pathologie cardiaque et à l'insuffisance cardiaque. Toutefois, nous n'avons pas identifié d'interaction significative entre l'expression de ce gène et l'exposition aux radiations ou aux anthracyclines, ce qui pourrait être dû à un manque de puissance statistique de notre étude. Nous avons également mis en évidence une amélioration significative de la capacité prédictive du modèle contenant les doses d'anthracycline cumulées, des doses moyennes d'irradiation cardiaque et l'inclusion de l'expression du gène NFE2L2 pour la survenue de pathologie cardiaque et l'insuffisance cardiaque. Ce résultat suggère une meilleure prise en compte de la variabilité interindividuelle dans la

survenue des pathologies cardiaques et de l'insuffisance cardiaque. A titre de comparaison, le modèle sans NFE2L2 modélisant l'insuffisance cardiaque a montré un AUC légèrement meilleure que celui du modèle développé par Chow et al. (2015) portant sur les survivants d'un cancer dans l'enfance. Le modèle de Chow et al. incluait des variables démographiques (sexe, âge au diagnostic du cancer) et des caractéristiques de traitement du cancer (doses d'anthracycline et de radiothérapie cardiaque).

D'un point de vue biologique, le gène NFE2L2 code pour le *nuclear factor erythroid 2-related factor 2* (NRF2), une protéine de la réponse antioxydante, impliquée dans diverses fonctions cellulaires telles que la défense antioxydante, l'inflammation, et l'homéostasie du fer. Cet aspect antioxydant suggère alors un lien potentiel avec la cardiotoxicité induite par le traitement (J. A. Roberts et al., 2023). Il est important de noter que l'association entre les doses de traitements et la survenue de la pathologie cardiaque n'était plus significative après inclusion de l'expression du gène NFE2L2. Cela suggère que le gène NFE2L2 pourrait jouer un rôle plus étendu dans la pathologie cardiaque, au-delà des effets causés par le traitement du cancer, ou que l'association entre la dose de traitement et la pathologie cardiaque pourrait être médiée par son impact sur l'expression du gène NFE2L2. Cela illustre la dualité du transcriptome, qui peut être perçu à la fois comme le reflet d'un processus en évolution et comme un maillon dans une chaîne de causalité (Zeller et al., 2010).

7.2 POINTS FORTS ET LIMITES DE LA THESE

Les résultats et les conclusions de cette thèse sont basés sur l'utilisation d'une large cohorte de survivants à long terme d'un cancer de l'enfant : la FCCSS. Cette cohorte multicentrique, rétrospective à suivi prospectif, comporte des données démographiques, cliniques, et de traitements exhaustives. Par ailleurs, nous avons cherché à étudier un évènement de santé cliniquement pertinent, les pathologies cardiaques, qui sont la première cause de décès non cancéreux chez les survivants d'un

cancer de l'enfant. De plus, de tous les événements médicaux auto-déclarés inclus dans nos analyses ont été validés. Un autre point fort de cette thèse est la disponibilité des estimations de doses de radiations et des volumes cardiaques irradiés, et de prélèvements biologiques, qui nous ont permis de répondre aux objectifs fixés par la thèse. En outre, dans cette thèse, nous avons cherché à améliorer la surveillance cardiaque chez les survivants du cancer de l'enfant. Ainsi, le Chapitre 6 décrit la seconde étude recherchant des associations entre des biomarqueurs du transcriptome et les pathologies cardiaques. Cette recherche s'est faite en utilisant une méthode statistique (Lasso) pour l'analyse de données en grande dimension contrairement aux études identifiées dans notre revue de la littérature, qui ont utilisé des méthodes de sélection classiques et non adaptées au contexte de la grande dimension. Bien que les patients inclus dans la cohorte FCCSS aient pu être suivis réalisés au travers de plusieurs sources (chainage au SNDS, réponses aux questionnaires auto-administrés, consultations de suivi à long terme à Gustave Roussy), certains patients ont été perdus de vue, et n'ont pas pu être suivis au cours des dernières années, entraînant un biais d'attrition dans notre cohorte. Par ailleurs, certaines hospitalisations antérieures à 2011 ont été omises, en particulier pour les patients traités au cours des années précédentes ; cependant 83 % des survivants en vie en 2011 ont pu être chaînés aux données du SNDS.

Cette thèse rapporte l'identification d'un biomarqueur du transcriptome associé aux pathologies cardiaques, toutes pathologies confondues. Ces pathologies ont chacune des physiopathologies différentes, ce qui nécessite des analyses séparées pour chaque pathologie, comme cela a été fait pour l'insuffisance cardiaque. De plus, pour la majorité des cas, les prélèvements sanguins ont été effectués après le développement de la pathologie cardiaque, rendant difficile d'établir la nature du lien de causalité entre l'expression du gène NFE2L2 et la survenue d'une pathologie cardiaque. Par ailleurs, les analyses étant menées sur un échantillon de patients niché dans la cohorte FCCSS, celui-ci ne peut pas être représentatif de la cohorte, et peut entraîner une sous-estimation de l'AUC pour un modèle ne comprenant que des variables cliniques, qui

ont servi à l'appariement. Et finalement, en l'absence d'un échantillon de données externes, nous n'avons pas pu valider définitivement le rôle du gène identifié dans un autre échantillon de patients.

7.3 PERSPECTIVES FUTURES

7.3.1 La méthodologie

Concernant la sélection des biomarqueurs associés aux pathologies cardiaques, nous nous sommes principalement focalisés sur la pénalisation Lasso. Le choix de cette méthode est justifié de par son utilisation fréquente à des fins de sélection de variables dans le contexte des données en grande dimension. Cependant, d'autres méthodes pourraient être utilisées comme la procédure boosting basée sur des splines de lissage ou le Matched Forest. La procédure boosting basée sur les splines de lissage a été développée pour le modèle logistique conditionnel, et propose une sélection automatique de variables avec une prise en compte des effets non linéaires potentiels des covariables (Shi & Yin, 2018). Shomal Zadeh et al., (2020) proposent le Matched Forest, un algorithme de sélection des variables dans le cadre des études cas-témoins appariées, qui se base sur des arbres de régression, et donc propose un score d'importance des variables. Une étude de simulation comparant les performances de ces deux méthodes avec la régression logistique conditionnelle Lasso serait une perspective de recherche méthodologique intéressante.

D'autre part, il serait également intéressant de prendre en compte les interactions de l'expression des biomarqueurs avec les doses de traitements sur le développement des pathologies cardiaques (Rothwell, 2005). Plusieurs méthodes existent pour la prise en compte de l'effet d'interaction lors de la sélection de variables mais ne sont pas adaptées aux design cas-témoins appariés.

D'un point de vue épidémiologique, la performance discriminante du modèle que nous avons développé devrait être validée dans un échantillon externe afin de vérifier sa

généralisabilité à des données provenant d'une cohorte indépendante de la FCCSS. Il serait envisageable de valider nos résultats au sein de cohortes européennes et/ou américaines à travers des collaborations internationales, d'autant plus que certaines de ces cohortes, telles que l'étude américaine du Children's Oncology Group (COG), disposent désormais des données de transcriptome nécessaires.

D'autre part, il serait intéressant de mettre en place un dosage régulier de l'expression du gène NFE2L2 chez les patients traités pour un cancer dans l'enfance, et d'étudier l'évolution des niveaux d'expression de ce gène au cours du temps jusqu'au possible développement d'une pathologie cardiaque.

7.3.2 Recommandations sur le suivi des patients

Les résultats de cette thèse pourraient avoir des implications cliniques importantes pour l'amélioration des directives de suivi à long terme des survivants d'un cancer de l'enfant qui pour l'instant préconisent la surveillance des patients traités par des doses d'anthracyclines ≥ 100 mg/m², et/ou de radiations ≥ 15 Gy, et ne prennent pas en compte les prédispositions génétiques dans la stratification du risque de pathologie cardiaque à long terme (Ehrhardt et al., 2023).

7.3.2.1 Les traitements cardiotoxiques

Les protocoles de traitements reçus par les générations de survivants actuels pouvaient présenter des effets iatrogènes à long termes importants, et nos résultats fournissent des preuves supplémentaires pour orienter le dépistage de la pathologie cardiaque en fonction du risque.

Une question se pose quant aux seuils de dose d'irradiation cardiaque à partir desquels la surveillance des patients est recommandée. En effet, Ehrhardt et al., (2023) indiquent qu'aucune directive de surveillance des cardiomyopathies chez les survivants d'un cancer de l'enfant n'a été émise pour patients exposés à des doses de radiations au cœur inférieures à 15 Gy. Hors nos résultats montrent un effet des doses de radiations

au cœur comprises entre 1 et 5 Gy, sur l'augmentation du risque de pathologie cardiaque. De plus, nous avons pu mettre en évidence un effet important des doses de radiations < 15 Gy chez les patients traités par anthracyclines sur le risque d'insuffisance cardiaque. Bien que les risques associés à ces doses restent faibles, une surveillance adaptée de ces patients devrait tout de même être recommandée.

Nos résultats ont aussi mis en exergue la nécessité de la prise en compte des mesures doses-volumes pour les doses d'irradiation cardiaque entre 5 et 20 Gy dans le suivi des survivants d'un cancer dans l'enfance. Par ailleurs, des recommandations de suivi plus étroit peuvent être recommandée chez les survivants à haut risque, c'est-à-dire traités par des doses de radiations > 15 Gy et d'anthracyclines > 250 mg/m², et ayant des comorbidités à risque tel que le diabète sucré, et pathologies rénales chroniques. De plus, une surveillance particulière des survivants ayant développé une valvulopathie, une ischémie ou des troubles du rythme et de la conduction pourrait être recommandé, au vu de l'effet très important des antécédents de pathologies cardiaques sur le risque d'insuffisance cardiaque dans notre cohorte. Ces résultats devraient être par ailleurs confirmés dans une cohorte externe indépendante.

7.3.2.2 L'inclusion des données génétiques

Un certain nombre d'études ont montré que l'inclusion de variables génétiques dans des modèles statistiques intégrant des variables cliniques en améliorerait leur capacité à discriminer les patients ayant développé une cardiomyopathie de ceux n'en ayant pas développé (Visscher et al., 2012, 2013, 2015). Cependant, les recommandations internationales actuelles n'incluent pas de dépistage génétique (Armenian et al., 2015; J.-Z. Chen & Liang, 2023; Ehrhardt et al., 2023; Van Kalsbeek et al., 2021). Aucun consensus sur les gènes impliqués dans la cardiotoxicité liée aux traitements des cancers de l'enfant n'a été trouvé, cela rend donc difficile d'intégrer ces résultats génétiques dans les recommandations pour la surveillance des patients traités pour cancer de l'enfant. Des études sur les biomarqueurs transcriptomiques pourrait être

une solution du fait de leur sensibilité aux traitements, et donc de leur possible utilisation dans la détection d'une cardiotoxicité précoce. Le transcriptome pourrait alors fournir des éléments de recherche précieux pour le développement de traitements cardioprotecteurs.

Malgré les avancées thérapeutiques, ces patients sont toujours plus à risque de développer des problèmes de santé tardifs. Il serait extrêmement utile d'établir un ensemble de variants génétiques et de biomarqueurs, afin d'identifier les patients les plus à risque de développer une pathologie cardiaque, et d'adapter autant que possible leur exposition aux thérapies cardiotoxiques, ainsi que de mettre en place un suivi et une prévention à long terme.

7.4 CONCLUSION GENERALE

Les avancées thérapeutiques en oncologie pédiatrique ont considérablement amélioré la survie des patients. Toutefois, ces traitements sont associés à un risque élevé de cardiotoxicité. Cette thèse a exploré les mécanismes de cardiotoxicité liés à la survenue de pathologies cardiaques après traitement d'un cancer chez l'enfant. Elle a révélé un effet cardiotoxique des volumes irradiés < 10 % à des doses moyennes > 15 Gy, ainsi que des volumes > 10 % irradiés à des doses entre 5 et 20 Gy. En outre, une revue systématique de la littérature et une étude cas-témoins nichée dans la cohorte FCCSS ont permis d'identifier certains facteurs génétiques et biomarqueurs transcriptomiques, notamment le gène NFE2L2, dont l'expression élevée est associée à une réduction du risque de pathologies cardiaques. Ces résultats fournissent une base factuelle pour l'élaboration de recommandations de suivi clinique des survivants d'un cancer de l'enfant. Ils peuvent également être utilisés pour la stratification du risque de pathologies cardiaques à travers des facteurs de risque liés aux traitements, aux prédispositions génétiques (variants génétiques), et aux biomarqueurs d'expression génique.

La validation de ce biomarqueur pourrait contribuer à réorienter la pratique clinique vers une personnalisation des traitements et du suivi des survivants d'un cancer pédiatrique, et éventuellement guider le développement de stratégies thérapeutiques, y compris de potentiels traitements cardioprotecteurs. Cela permettrait de réduire les complications cardiaques tardives, qui constituent un fardeau pour les patients guéris de leur cancer dans l'enfance, et d'améliorer leur qualité de vie en accentuant le suivi à long terme des patients les plus à risque de développer une pathologie cardiaque.

REFERENCES BIBLIOGRAPHIQUES

- Aalen, O. (1978). Nonparametric Estimation of Partial Transition Probabilities in Multiple Decrement Models. *The Annals of Statistics*, 6(3), 534-545.
<https://doi.org/10.1214/aos/1176344198>
- Allen, C., Her, S., & Jaffray, D. A. (2017). Radiotherapy for Cancer: Present and Future. *Advanced Drug Delivery Reviews*, 109, 1-2.
<https://doi.org/10.1016/j.addr.2017.01.004>
- Aminkeng, F., Bhavsar, A. P., Visscher, H., Rassekh, S. R., Li, Y., Lee, J. W., Brunham, L. R., Caron, H. N., van Dalen, E. C., Kremer, L. C., van der Pal, H. J., Amstutz, U., Rieder, M. J., Bernstein, D., Carleton, B. C., Hayden, M. R., Ross, C. J. D., The Canadian Pharmacogenomics Network for Drug Safety Consortium, Hayden, M. R., ... Hand, J. (2015). A coding variant in RARG confers susceptibility to anthracycline-induced cardiotoxicity in childhood cancer. *Nature Genetics*, 47(9), 1079-1084.
<https://doi.org/10.1038/ng.3374>
- Aminkeng, F., Ross, C. J. D., Rassekh, S. R., Hwang, S., Rieder, M. J., Bhavsar, A. P., Smith, A., Sanatani, S., Gelmon, K. A., Bernstein, D., Hayden, M. R., Amstutz, U., & Carleton, B. C. (2016). Recommendations for genetic testing to reduce the incidence of anthracycline-induced cardiotoxicity. *British Journal of Clinical Pharmacology*, 82(3), 683-695. <https://doi.org/10.1111/bcp.13008>
- Amir-Behghadami, M., & Janati, A. (2020). Population, Intervention, Comparison,

Outcomes and Study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews. *Emergency Medicine Journal*, 37(6), 387-387.

<https://doi.org/10.1136/emered-2020-209567>

Armenian, S. H., Armstrong, G. T., Aune, G., Chow, E. J., Ehrhardt, M. J., Ky, B., Moslehi, J., Mulrooney, D. A., Nathan, P. C., Ryan, T. D., van der Pal, H. J., van Dalen, E. C., & Kremer, L. C. M. (2018). Cardiovascular Disease in Survivors of Childhood Cancer: Insights Into Epidemiology, Pathophysiology, and Prevention. *Journal of Clinical Oncology*, 36(21), 2135-2144. <https://doi.org/10.1200/JCO.2017.76.3920>

Armenian, S. H., Hudson, M. M., Mulder, R. L., Chen, M. H., Constine, L. S., Dwyer, M., Nathan, P. C., Tissing, W. J. E., Shankar, S., Sieswerda, E., Skinner, R., Steinberger, J., van Dalen, E. C., van der Pal, H., Wallace, W. H., Levitt, G., & Kremer, L. C. M. (2015). Recommendations for Cardiomyopathy Surveillance for Survivors of Childhood Cancer: A Report from the International Late Effects of Childhood Cancer Guideline Harmonization Group. *The Lancet. Oncology*, 16(3), e123-e136. [https://doi.org/10.1016/S1470-2045\(14\)70409-7](https://doi.org/10.1016/S1470-2045(14)70409-7)

Armstrong, G. T., Joshi, V. M., Ness, K. K., Marwick, T. H., Zhang, N., Srivastava, D., Griffin, B. P., Grimm, R. A., Thomas, J., Phelan, D., Collier, P., Krull, K. R., Mulrooney, D. A., Green, D. M., Hudson, M. M., Robison, L. L., & Plana, J. C. (2015). Comprehensive Echocardiographic Detection of Treatment-Related Cardiac Dysfunction in Adult Survivors of Childhood Cancer: Results From the St. Jude Lifetime Cohort Study. *Journal of the American College of Cardiology*, 65(23), 2511-2522.

<https://doi.org/10.1016/j.jacc.2015.04.013>

Armstrong, G. T., Plana, J. C., Zhang, N., Srivastava, D., Green, D. M., Ness, K. K., Daniel Donovan, F., Metzger, M. L., Arevalo, A., Durand, J.-B., Joshi, V., Hudson, M. M., Robison, L. L., & Flamm, S. D. (2012). Screening Adult Survivors of Childhood Cancer for Cardiomyopathy: Comparison of Echocardiography and Cardiac Magnetic Resonance Imaging. *Journal of Clinical Oncology*, *30*(23), 2876-2884.
<https://doi.org/10.1200/JCO.2011.40.3584>

Arruebo, M., Vilaboa, N., Sáez-Gutierrez, B., Lambea, J., Tres, A., Valladares, M., & González-Fernández, Á. (2011). Assessment of the Evolution of Cancer Treatment Therapies. *Cancers*, *3*(3), 3279-3330.
<https://doi.org/10.3390/cancers3033279>

Avalos, M., Pouyes, H., Grandvalet, Y., Orriols, L., & Lagarde, E. (2015). Sparse conditional logistic regression for analyzing large-scale matched data from epidemiological studies: A simple algorithm. *BMC Bioinformatics*, *16*(S6), S1.
<https://doi.org/10.1186/1471-2105-16-S6-S1>

Bansal, N., Amdani, S., Lipshultz, E. R., & Lipshultz, S. E. (2017). Chemotherapy-induced cardiotoxicity in children. *Expert Opinion on Drug Metabolism & Toxicology*, *13*(8), 817-832. <https://doi.org/10.1080/17425255.2017.1351547>

Bates, J. E., Howell, R. M., Liu, Q., Yasui, Y., Mulrooney, D. A., Dhakal, S., Smith, S. A., Leisenring, W. M., Indelicato, D. J., Gibson, T. M., Armstrong, G. T., Oeffinger, K. C., & Constine, L. S. (2019). Therapy-Related Cardiac Risk in Childhood Cancer

- Survivors: An Analysis of the Childhood Cancer Survivor Study. *Journal of Clinical Oncology*, 37(13), 1090-1101. <https://doi.org/10.1200/JCO.18.01764>
- Belhechmi, S., Bin, R. D., Rotolo, F., & Michiels, S. (2020). Accounting for grouped predictor variables or pathways in high-dimensional penalized Cox regression models. *BMC Bioinformatics*, 21(1), 277. <https://doi.org/10.1186/s12859-020-03618-y>
- Berg, T., Böhmer, J., Nwaru, B. I., Karason, K., & Jarfelt, M. (2022). Heart failure in childhood cancer survivors—A systematic review protocol. *Systematic Reviews*, 11, 54. <https://doi.org/10.1186/s13643-022-01929-0>
- Berkman, A. M., Hildebrandt, M. A. T., & Landstrom, A. P. (2021). The genetic underpinnings of anthracycline-induced cardiomyopathy predisposition. *Clinical Genetics*, 100(2), 132-143. <https://doi.org/10.1111/cge.13968>
- Blanco, J. G., Leisenring, W. M., Gonzalez-Covarrubias, V. M., Kawashima, T. I., Davies, S. M., Relling, M. V., Robison, L. L., Sklar, C. A., Stovall, M., & Bhatia, S. (2008). Genetic polymorphisms in the carbonyl reductase 3 gene CBR3 and the NAD(P)H:quinone oxidoreductase 1 gene NQO1 in patients who developed anthracycline-related congestive heart failure after childhood cancer. *Cancer*, 112(12), 2789-2795. <https://doi.org/10.1002/cncr.23534>
- Blanco, J. G., Sun, C.-L., Landier, W., Chen, L., Esparza-Duran, D., Leisenring, W., Mays, A., Friedman, D. L., Ginsberg, J. P., Hudson, M. M., Neglia, J. P., Oeffinger, K. C., Ritchey, A. K., Villaluna, D., Relling, M. V., & Bhatia, S. (2012). Anthracycline-

Related Cardiomyopathy After Childhood Cancer: Role of Polymorphisms in Carbonyl Reductase Genes—A Report From the Children’s Oncology Group. *Journal of Clinical Oncology*, 30(13), 1415-1421. <https://doi.org/10.1200/JCO.2011.34.8987>

Botta, L., Gatta, G., Capocaccia, R., Stiller, C., Cañete, A., Dal Maso, L., Innos, K., Mihor, A., Erdmann, F., Spix, C., Lacour, B., Marcos-Gragera, R., Murray, D., Rossi, S., Hackl, M., Van Eycken, E., Van Damme, N., Valerianova, Z., Sekerija, M., ... Huws, D. W. (2022). Long-term survival and cure fraction estimates for childhood cancer in Europe (EUROCARE-6): Results from a population-based study. *The Lancet Oncology*, 23(12), 1525-1536. [https://doi.org/10.1016/S1470-2045\(22\)00637-4](https://doi.org/10.1016/S1470-2045(22)00637-4)

Bottinor, W., Im, C., Doody, D. R., Armenian, S. H., Arynchyn, A., Hong, B., Howell, R. M., Jacobs, D. R., Ness, K. K., Oeffinger, K. C., Reiner, A. P., Armstrong, G. T., Yasui, Y., & Chow, E. J. (2024). Mortality After Major Cardiovascular Events in Survivors of Childhood Cancer. *Journal of the American College of Cardiology*, 83(8), 827-838. <https://doi.org/10.1016/j.jacc.2023.12.022>

Box, G. E. P., & Tidwell, P. W. (1962). Transformation of the Independent Variables. *Technometrics*, 4(4), 531-550. <https://doi.org/10.1080/00401706.1962.10490038>

Breslow, N. E., & Day, N. E. (1980). *Statistical methods in cancer research. 1: The analysis of case-control studies* (8. reimpr).

Breslow, N. E., & Day, N. E. (1991). *Statistical methods in cancer research. 2: The design*

and analysis of cohort studies (Reprinted).

Brouwer, C. A. J., Postma, A., Vonk, J. M., Zwart, N., Berg, M. P. van den, Bink-Boelkens, M. T. E., Dolsma, W. V., Smit, A. J., Vries, E. G. E. de, Tissing, W. J. E., & Gietema, J. A. (2011). Systolic and diastolic dysfunction in long-term adult survivors of childhood cancer. *European Journal of Cancer*, 47(16), 2453-2462.
<https://doi.org/10.1016/j.ejca.2011.05.023>

Bühlmann, P., & Van De Geer, S. (2011). *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Springer Berlin Heidelberg.
<https://doi.org/10.1007/978-3-642-20192-9>

Chaix, M.-A., Parmar, N., Kinnear, C., Lafreniere-Roula, M., Akinrinade, O., Yao, R., Miron, A., Lam, E., Meng, G., Christie, A., Manickaraj, A. K., Marjerrison, S., Dillenburg, R., Bassal, M., Loughheed, J., Zelcer, S., Rosenberg, H., Hodgson, D., Sender, L., ... Mital, S. (2020). Machine Learning Identifies Clinical and Genetic Factors Associated With Anthracycline Cardiotoxicity in Pediatric Cancer Survivors. *JACC: CardioOncology*, 2(5), 690-706.
<https://doi.org/10.1016/j.jaccao.2020.11.004>

Chen, J.-Z., & Liang, B. (2023). Comparison of American and European guidelines for cardio-oncology of heart failure. *Heart Failure Reviews*, 28(5), 1211-1220.
<https://doi.org/10.1007/s10741-023-10304-7>

Chen, Q. M., & Maltagliati, A. J. (2018). Nrf2 at the heart of oxidative stress and cardiac protection. *Physiological Genomics*, 50(2), 77-97.

<https://doi.org/10.1152/physiolgenomics.00041.2017>

Choi, J.-Y., Nowell, S. A., Blanco, J. G., & Ambrosone, C. B. (2006). The role of genetic variability in drug metabolism pathways in breast cancer prognosis. *Pharmacogenomics*, 7(4), 613-624. <https://doi.org/10.2217/14622416.7.4.613>

Chounta, S., Lemler, S., Haddy, N., Fresneau, B., Mansouri, I., Bentriou, M., Demoor-Goldschmidt, C., Diallo, I., Souchard, V., Do, T.-D., Veres, C., Surun, A., Doz, F., Llanas, D., Vu-Bezin, G., Rubino, C., de Vathaire, F., Letort, V., & Allodji, R. S. (2023). The risk of valvular heart disease in the French Childhood Cancer Survivors' Study: Contribution of dose-volume histogram parameters. *Radiotherapy and Oncology: Journal of the European Society for Therapeutic Radiology and Oncology*, 180, 109479. <https://doi.org/10.1016/j.radonc.2023.109479>

Chow, E. J., Chen, Y., Kremer, L. C., Breslow, N. E., Hudson, M. M., Armstrong, G. T., Border, W. L., Feijen, E. A. M., Green, D. M., Meacham, L. R., Meeske, K. A., Mulrooney, D. A., Ness, K. K., Oeffinger, K. C., Sklar, C. A., Stovall, M., van der Pal, H. J., Weathers, R. E., Robison, L. L., & Yasui, Y. (2015). Individual Prediction of Heart Failure Among Childhood Cancer Survivors. *Journal of Clinical Oncology*, 33(5), 394-402. <https://doi.org/10.1200/JCO.2014.56.1373>

Clemens, E., van der Kooi, A. L. F., Broer, L., van Dulmen-den Broeder, E., Visscher, H., Kremer, L., Tissing, W., Loonen, J., Ronckers, C. M., Pluijm, S. M. F., Neggers, S. J. C. M. M., Zolk, O., Langer, T., Zehnhoff-Dinnesen, A. am, Wilson, C. L., Hudson,

- M. M., Carleton, B., Laven, J. S. E., Uitterlinden, A. G., & van den Heuvel-Eibrink, M. M. (2018). The influence of genetic variation on late toxicities in childhood cancer survivors: A review. *Critical Reviews in Oncology/Hematology*, *126*, 154-167. <https://doi.org/10.1016/j.critrevonc.2018.04.001>
- Couture, L., Nash, J. A., & Turgeon, J. (2006). The ATP-Binding Cassette Transporters and Their Implication in Drug Disposition: A Special Look at the Heart. *Pharmacological Reviews*, *58*(2), 244-258. <https://doi.org/10.1124/pr.58.2.7>
- Cox, D. R. (1972). Regression Models and Life-Tables—Cox—1972—Journal of the Royal Statistical Society: Series B (Methodological)—Wiley Online Library. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187-202.
- de la Grange, P., Dutertre, M., Correa, M., & Auboeuf, D. (2007). A new advance in alternative splicing databases: From catalogue to detailed analysis of regulation of expression and function of human alternative splicing variants. *BMC Bioinformatics*, *8*, 180. <https://doi.org/10.1186/1471-2105-8-180>
- de Baat, E. C., Feijen, E. A. M., Reulen, R. C., Allodji, R. S., Bagnasco, F., Bardi, E., Belle, F. N., Byrne, J., van Dalen, E. C., Debiche, G., Diallo, I., Grabow, D., Hjorth, L., Jankovic, M., Kuehni, C. E., Levitt, G., Llanas, D., Loonen, J., Zaletel, L. Z., ... Kremer, L. C. M. (2023). Risk Factors for Heart Failure Among Pan-European Childhood Cancer Survivors: A PanCareSurFup and ProCardio Cohort and Nested Case-Control Study. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *41*(1), 96-106.

<https://doi.org/10.1200/JCO.21.02944>

Debela, D. T., Muzazu, S. G., Heraro, K. D., Ndalama, M. T., Mesele, B. W., Haile, D. C., Kitui, S. K., & Manyazewal, T. (2021). New approaches and procedures for cancer treatment: Current perspectives. *SAGE Open Medicine*, 9, 20503121211034366.

<https://doi.org/10.1177/20503121211034366>

Ducimetière, P. (2004). La fréquence des maladies cardio-vasculaires en France et en Europe. *Actualité et dossier en santé publique*, 47.

Ducimetière, P., Ruidavets, J. B., Montaye, M., Haas, B., Yarnell, J., & PRIME Study Group. (2001). Five-year incidence of angina pectoris and other forms of coronary heart disease in healthy men aged 50-59 in France and Northern Ireland: The Prospective Epidemiological Study of Myocardial Infarction (PRIME) Study. *International Journal of Epidemiology*, 30(5), 1057-1062.

<https://doi.org/10.1093/ije/30.5.1057>

Ehret, G. B., Munroe, P. B., Rice, K. M., Bochud, M., Johnson, A. D., Chasman, D. I., Smith, A. V., Tobin, M. D., Verwoert, G. C., Hwang, S.-J., Pihur, V., Vollenweider, P., O'Reilly, P. F., Amin, N., Bragg-Gresham, J. L., Teumer, A., Glazer, N. L., Launer, L., Zhao, J. H., ... Johnson, T. (2011). Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*, 478(7367), 103-109.

<https://doi.org/10.1038/nature10405>

Ehrhardt, M. J., Leerink, J. M., Mulder, R. L., Mavinkurve-Groothuis, A., Kok, W., Nohria, A., Nathan, P. C., Merks, R., De Baat, E., Asogwa, O. A., Skinner, R., Wallace, H.,

- Lieke Feijen, E. A. M., De Ville De Goyet, M., Prasad, M., Bárdi, E., Pavasovic, V., Van Der Pal, H., Fresneau, B., ... Armenian, S. H. (2023). Systematic review and updated recommendations for cardiomyopathy surveillance for survivors of childhood, adolescent, and young adult cancer from the International Late Effects of Childhood Cancer Guideline Harmonization Group. *The Lancet Oncology*, 24(3), e108-e120. [https://doi.org/10.1016/S1470-2045\(23\)00012-8](https://doi.org/10.1016/S1470-2045(23)00012-8)
- Erdmann, F., Frederiksen, L. E., Bonaventure, A., Mader, L., Hasle, H., Robison, L. L., & Winther, J. F. (2021). Childhood cancer: Survival, treatment modalities, late effects and improvements over time. *Cancer Epidemiology*, 71, 101733. <https://doi.org/10.1016/j.canep.2020.101733>
- Fallacara, A., Baldini, E., Manfredini, S., & Vertuani, S. (2018). Hyaluronic Acid in the Third Millennium. *Polymers*, 10(7), 701. <https://doi.org/10.3390/polym10070701>
- Fan, J., & Lv, J. (2010). A Selective Overview of Variable Selection in High Dimensional Feature Space. *Statistica Sinica*, 20(1), 101-148.
- Feijen, E. A. M., Leisenring, W. M., Stratton, K. L., Ness, K. K., Van Der Pal, H. J. H., Van Dalen, E. C., Armstrong, G. T., Aune, G. J., Green, D. M., Hudson, M. M., Loonen, J., Oeffinger, K. C., Robison, L. L., Yasui, Y., Kremer, L. C. M., & Chow, E. J. (2019). Derivation of Anthracycline and Anthraquinone Equivalence Ratios to Doxorubicin for Late-Onset Cardiotoxicity. *JAMA Oncology*, 5(6), 864. <https://doi.org/10.1001/jamaoncol.2018.6634>
- Fine, J. P., & Gray, R. J. (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/10.2307/2670170>
- Fogli, S., Nieri, P., & Cristina Breschi, M. (2004). The role of nitric oxide in anthracycline toxicity and prospects for pharmacologic prevention of cardiac damage. *The FASEB Journal*, 18(6), 664-675. <https://doi.org/10.1096/fj.03-0724rev>
- Fonoudi, H., Jouni, M., Cejas, R. B., Magdy, T., Blancard, M., Ge, N., Shah, D. A., Lyra, - Leite Davi M., Neupane, A., Gharib, M., Jiang, Z., Sapkota, Y., & BurrIDGE, P. W. (2024). Functional Validation of Doxorubicin-Induced Cardiotoxicity-Related Genes. *JACC: CardioOncology*, 6(1), 38-50. <https://doi.org/10.1016/j.jaccao.2023.11.008>
- Frelay, J., Ervik, M., Lam, F., Laversanne, M., Colombet, M., Meryl, L., Piñeros, M., Znaor, A., Soerjomataram, I., & Bray, F. (2024). *Global Cancer Observatory*. <https://gco.iarc.fr/>
- Geenen, M. M., Cardous-Ubbink, M. C., Kremer, L. C. M., van den Bos, C., van der Pal, H. J. H., Heinen, R. C., Jaspers, M. W. M., Koning, C. C. E., Oldenburger, F., Langeveld, N. E., Hart, A. A. M., Bakker, P. J. M., Caron, H. N., & van Leeuwen, F. E. (2007). Medical Assessment of Adverse Health Outcomes in Long-term Survivors of Childhood Cancer. *JAMA*, 297(24), 2705-2715. <https://doi.org/10.1001/jama.297.24.2705>
- Gooley, T. A., Leisenring, W., Crowley, J., & Storer, B. E. (1999). Estimation of failure probabilities in the presence of competing risks: New representations of old

estimators. *Statistics in Medicine*, 18(6), 695-706.

[https://doi.org/10.1002/\(SICI\)1097-0258\(19990330\)18:6<695::AID-](https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<695::AID-SIM60>3.0.CO;2-O)

[SIM60>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<695::AID-SIM60>3.0.CO;2-O)

Greenland, S., Schwartzbaum, J. A., & Finkle, W. D. (2000). Problems due to small samples and sparse data in conditional logistic regression analysis. *American Journal of Epidemiology*, 151(5), 531-539.
<https://doi.org/10.1093/oxfordjournals.aje.a010240>

Gröbner, S. N., Worst, B. C., Weischenfeldt, J., Buchhalter, I., Kleinheinz, K., Rudneva, V. A., Johann, P. D., Balasubramanian, G. P., Segura-Wang, M., Brabetz, S., Bender, S., Hutter, B., Sturm, D., Pfaff, E., Hübschmann, D., Zipprich, G., Heinold, M., Eils, J., Lawrenz, C., ... Pfister, S. M. (2018). The landscape of genomic alterations across childhood cancers. *Nature*, 555(7696), Article 7696.
<https://doi.org/10.1038/nature25480>

Guo, Z., Yan, M., Chen, L., Fang, P., Li, Z., Wan, Z., Cao, S., Hou, Z., Wei, S., Li, W., & Zhang, B. (2018). Nrf2-dependent antioxidant response mediated the protective effect of tanshinone IIA on doxorubicin-induced cardiotoxicity. *Experimental and Therapeutic Medicine*, 16(4), 3333-3344. <https://doi.org/10.3892/etm.2018.6614>

Haddy, N., Diallo, S., El-Fayech, C., Schwartz, B., Pein, F., Hawkins, M., Veres, C., Oberlin, O., Guibout, C., Pacquement, H., Munzer, M., N'Guyen, T. D., Bondiau, P.-Y., Berchery, D., Laprie, A., Scarabin, P.-Y., Jouven, X., Bridier, A., Koscielny, S., ... de Vathaire, F. (2016). Cardiac Diseases Following Childhood Cancer Treatment:

- Cohort Study. *Circulation*, 133(1), 31-38.
<https://doi.org/10.1161/CIRCULATIONAHA.115.016686>
- Harrell, F. E. (2015). *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-19425-7>
- He, L., Ma, Q., Wang, Y., Liu, X., Yuan, Y., Zhang, Y., Ou, W., Liu, L., Tan, X., & Wang, X. (2014). Association of Variants in KCNK17 Gene with Ischemic Stroke and Cerebral Hemorrhage in a Chinese Population. *Journal of Stroke and Cerebrovascular Diseases*, 23(9), 2322-2327.
<https://doi.org/10.1016/j.jstrokecerebrovasdis.2014.04.029>
- Heinze, G., & Puhr, R. (2010). Bias-reduced and separation-proof conditional logistic regression with small or sparse data sets. *Statistics in Medicine*, 29(7-8), 770-777.
<https://doi.org/10.1002/sim.3794>
- Hildebrandt, M. A. T., Reyes, M., Wu, X., Pu, X., Thompson, K. A., Ma, J., Landstrom, A. P., Morrison, A. C., & Ater, J. L. (2017). Hypertension Susceptibility Loci are Associated with Anthracycline-related Cardiotoxicity in Long-term Childhood Cancer Survivors. *Scientific Reports*, 7(1), 9698. <https://doi.org/10.1038/s41598-017-09517-2>
- Ho, K. K., Anderson, K. M., Kannel, W. B., Grossman, W., & Levy, D. (1993). Survival after the onset of congestive heart failure in Framingham Heart Study subjects. *Circulation*, 88(1), 107-115. <https://doi.org/10.1161/01.cir.88.1.107>

- Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *TECHNOMETRICS*, 12(1).
- Hoffmann, T. J., Ehret, G. B., Nandakumar, P., Ranatunga, D., Schaefer, C., Kwok, P.-Y., Iribarren, C., Chakravarti, A., & Risch, N. (2017). Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nature Genetics*, 49(1), 54-64. <https://doi.org/10.1038/ng.3715>
- Huang, K. M., Hu, S., & Sparreboom, A. (2018). Drug transporters and anthracycline-induced cardiotoxicity. *Pharmacogenomics*, 19(11), 883-888. <https://doi.org/10.2217/pgs-2018-0056>
- Hudson, M. M., Neglia, J. P., Woods, W. G., Sandlund, J. T., Pui, C.-H., Kun, L. E., Robison, L. L., & Green, D. M. (2012). Lessons from the past: Opportunities to improve childhood cancer survivor care through outcomes investigations of historical therapeutic approaches for pediatric hematological malignancies. *Pediatric Blood & Cancer*, 58(3), 334-343. <https://doi.org/10.1002/pbc.23385>
- Hudson, M. M., Ness, K. K., Gurney, J. G., Mulrooney, D. A., Chemaitilly, W., Krull, K. R., Green, D. M., Armstrong, G. T., Nottage, K. A., Jones, K. E., Sklar, C. A., Srivastava, D. K., & Robison, L. L. (2013). Clinical ascertainment of health outcomes among adults treated for childhood cancer. *JAMA*, 309(22), 2371-2381. <https://doi.org/10.1001/jama.2013.6296>
- Jaafar, Z., & Ismail, N. (2022). *PENALIZED REGRESSION METHOD IN HIGH*

DIMENSIONAL DATA. 100(8).

- Janes, H., & Pepe, M. S. (2008). Matching in Studies of Classification Accuracy: Implications for Analysis, Efficiency, and Assessment of Incremental Value—Janes—2008—Biometrics—Wiley Online Library. *International Biometric Society*, 64(1), pp 1-9. <https://doi.org/10.1111/j.l541-0420.2007.00823>
- Johnston, W. T., Erdmann, F., Newton, R., Steliarova-Foucher, E., Schüz, J., & Roman, E. (2021). Childhood cancer: Estimating regional and global incidence. *Cancer Epidemiology*, 71, 101662. <https://doi.org/10.1016/j.canep.2019.101662>
- Joseph, P. (2017). Transcriptomics in toxicology. *Food and chemical toxicology: an international journal published for the British Industrial Biological Research Association*, 109(Pt 1), 650-662. <https://doi.org/10.1016/j.fct.2017.07.031>
- Kellermayer, D., Smith, J. E., & Granzier, H. (2019). Titin mutations and muscle disease. *Pflugers Archiv: European journal of physiology*, 471(5), 673-682. <https://doi.org/10.1007/s00424-019-02272-5>
- Koleini, N., & Kardami, E. (2017). Autophagy and mitophagy in the context of doxorubicin-induced cardiotoxicity. *Oncotarget*, 8(28), 46663-46680. <https://doi.org/10.18632/oncotarget.16944>
- Krajinovic, M., Elbared, J., Drouin, S., Bertout, L., Rezgui, A., Ansari, M., Raboisson, M.-J., Lipshultz, S. E., Silverman, L. B., Sallan, S. E., Neuberg, D. S., Kutok, J. L., Laverdière, C., Sinnett, D., & Andelfinger, G. (2016). Polymorphisms of ABCC5 and NOS3 genes influence doxorubicin cardiotoxicity in survivors of childhood acute

lymphoblastic leukemia. *The Pharmacogenomics Journal*, 16(6), 530-535.

<https://doi.org/10.1038/tpj.2015.63>

Kremer, L. C. M., Mulder, R. L., Oeffinger, K. C., Bhatia, S., Landier, W., Levitt, G., Constine, L. S., Wallace, W. H., Caron, H. N., Armenian, S. H., Skinner, R., & Hudson, M. M. (2013). A Worldwide Collaboration To Harmonize Guidelines For The Long-Term Follow-Up Of Childhood And Young Adult Cancer Survivors: A Report From The International Late Effects Of Childhood Cancer Guideline Harmonization Group.

Pediatric blood & cancer, 60(4), 10.1002/pbc.24445.

<https://doi.org/10.1002/pbc.24445>

Kuo, C.-L., Duan, Y., & Grady, J. (2018). Unconditional or Conditional Logistic Regression Model for Age-Matched Case–Control Data? *Frontiers in Public Health*, 6, 57.

<https://doi.org/10.3389/fpubh.2018.00057>

Lacour, B., Goujon, S., Guissou, S., Guyot-Goubin, A., Desmée, S., Désandes, E., & Clavel, J. (2014). Childhood cancer survival in France, 2000–2008. *European Journal of Cancer Prevention*, 23(5), 449. <https://doi.org/10.1097/CEJ.0000000000000006>

Lam, C. G., Howard, S. C., Bouffet, E., & Pritchard-Jones, K. (2019). Science and health for all children with cancer. *Science (New York, N.Y.)*, 363(6432), 1182-1186.

<https://doi.org/10.1126/science.aaw4892>

Landier, W., Armenian, S., & Bhatia, S. (2015). Late Effects of Childhood Cancer and Its Treatment. *Pediatric Clinics of North America*, 62(1), 275-300.

<https://doi.org/10.1016/j.pcl.2014.09.017>

- Latouche, A., Allignol, A., Beyersmann, J., Labopin, M., & Fine, J. P. (2013). A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *Journal of Clinical Epidemiology*, epub ahead of print. <https://doi.org/10.1016/j.jclinepi.2012.09.017>
- L'Ecuyer, T., Allebban, Z., Thomas, R., & Vander Heide, R. (2004). Glutathione S-transferase overexpression protects against anthracycline-induced H9C2 cell death. *American Journal of Physiology. Heart and Circulatory Physiology*, 286(6), H2057-2064. <https://doi.org/10.1152/ajpheart.00778.2003>
- Leerink, J. M., de Baat, E. C., Feijen, E. A. M., Bellersen, L., van Dalen, E. C., Grotenhuis, H. B., Kapusta, L., Kok, W. E. M., Loonen, J., van der Pal, H. J. H., Pluijm, S. M. F., Teske, A. J., Mavinkurve-Groothuis, A. M. C., Merckx, R., & Kremer, L. C. M. (2020). Cardiac Disease in Childhood Cancer Survivors. *JACC: CardioOncology*, 2(3), 363-378. <https://doi.org/10.1016/j.jacc.2020.08.006>
- Leerink, J. M., Feijen, E. A. M., de Baat, E. C., Merckx, R., van der Pal, H. J. H., Tissing, W. J. E., Louwerens, M., van den Heuvel-Eibrink, M. M., Versluys, A. B., van Dalen, E. C., van der Heiden-van der Loo, M., Bresters, D., Ronckers, C. M., de Vries, A. C. H., Neggers, S., Kapusta, L., Loonen, J., Pinto, Y. M., Kremer, L. C. M., ... Kok, W. E. M. (2024). A Biomarker-Based Diagnostic Model for Cardiac Dysfunction in Childhood Cancer Survivors. *JACC: CardioOncology*, 6(2), 236-247. <https://doi.org/10.1016/j.jacc.2024.02.008>
- Li, S., Wang, W., Niu, T., Wang, H., Li, B., Shao, L., Lai, Y., Li, H., Janicki, J. S., Wang, X. L.,

- Tang, D., & Cui, T. (2014). Nrf2 Deficiency Exaggerates Doxorubicin-Induced Cardiotoxicity and Cardiac Dysfunction. *Oxidative Medicine and Cellular Longevity*, 2014, 748524. <https://doi.org/10.1155/2014/748524>
- Linschoten, M., Teske, A. J., Cramer, M. J., van der Wall, E., & Asselbergs, F. W. (2018). Chemotherapy-Related Cardiac Dysfunction: A Systematic Review of Genetic Variants Modulating Individual Risk. *Circulation: Genomic and Precision Medicine*, 11(1). <https://doi.org/10.1161/CIRCGEN.117.001753>
- Lipshultz, S. E., Adams, M. J., Colan, S. D., Constine, L. S., Herman, E. H., Hsu, D. T., Hudson, M. M., Kremer, L. C., Landy, D. C., Miller, T. L., Oeffinger, K. C., Rosenthal, D. N., Sable, C. A., Sallan, S. E., Singh, G. K., Steinberger, J., Cochran, T. R., Wilkinson, J. D., & American Heart Association Congenital Heart Defects Committee of the Council on Cardiovascular Disease in the Young, Council on Basic Cardiovascular Sciences, Council on Cardiovascular and Stroke Nursing, Council on Cardiovascular Radiology. (2013). Long-term cardiovascular toxicity in children, adolescents, and young adults who receive cancer therapy: Pathophysiology, course, monitoring, management, prevention, and research directions: a scientific statement from the American Heart Association. *Circulation*, 128(17), 1927-1995. <https://doi.org/10.1161/CIR.0b013e3182a88099>
- Lipshultz, S. E., Lipsitz, S. R., Kutok, J. L., Miller, T. L., Colan, S. D., Neuberg, D. S., Stevenson, K. E., Fleming, M. D., Sallan, S. E., Franco, V. I., Henkel, J. M., Asselin,

- B. L., Athale, U. H., Clavell, L. A., Michon, B., Laverdiere, C., Larsen, E., Kelly, K. M., & Silverman, L. B. (2013). Impact of hemochromatosis gene mutations on cardiac status in doxorubicin-treated survivors of childhood high-risk leukemia: HFE Mutations and Survivor Cardiac Status. *Cancer*, *119*(19), 3555-3562. <https://doi.org/10.1002/cncr.28256>
- Lipshultz, S. E., Lipsitz, S. R., Sallan, S. E., Dalton, V. M., Mone, S. M., Gelber, R. D., & Colan, S. D. (2005). Chronic Progressive Cardiac Dysfunction Years After Doxorubicin Therapy for Childhood Acute Lymphoblastic Leukemia. *Journal of Clinical Oncology*, *23*(12), 2629-2636. <https://doi.org/10.1200/JCO.2005.12.121>
- Lo, A. C., Liu, A., Liu, Q., Yasui, Y., Castellino, S. M., Kelly, K. M., Herrera, A. F., Friedberg, J. W., Friedman, D. L., Schwartz, C. L., Pei, Q., Kessel, S., Bergeron-Gravel, S., Dama, H., Roberts, K., Constine, L. S., & Hodgson, D. C. (2024). Late Cardiac Toxic Effects Associated With Treatment Protocols for Hodgkin Lymphoma in Children. *JAMA Network Open*, *7*(1), e2351062. <https://doi.org/10.1001/jamanetworkopen.2023.51062>
- Luo, L., Zhou, C., Hui, Y., Kurogi, K., Sakakibara, Y., Suiko, M., & Liu, M.-C. (2016). Human cytosolic sulfotransferase SULT1C4 mediates the sulfation of doxorubicin and epirubicin. *Drug Metabolism and Pharmacokinetics*, *31*(2), 163-166. <https://doi.org/10.1016/j.dmpk.2016.01.003>
- Lyon, A. R., Dent, S., Stanway, S., Earl, H., Brezden-Masley, C., Cohen-Solal, A., Tocchetti, C. G., Moslehi, J. J., Groarke, J. D., Bergler-Klein, J., Khoo, V., Tan, L. L., Anker, M.

- S., von Haehling, S., Maack, C., Pudil, R., Barac, A., Thavendiranathan, P., Ky, B., ... Lenihan, D. (2020). Baseline cardiovascular risk assessment in cancer patients scheduled to receive cardiotoxic cancer therapies: A position statement and new risk assessment tools from the Cardio-Oncology Study Group of the Heart Failure Association of the European Society of Cardiology in collaboration with the International Cardio-Oncology Society. *European Journal of Heart Failure*, 22(11), 1945-1960. <https://doi.org/10.1002/ejhf.1920>
- Malatkova, P., Maser, E., & Wsol, V. (2010). Human Carbonyl Reductases. *Current Drug Metabolism*, 11(8), 639-658.
- Mansouri, I., Allodji, R. S., Hill, C., El-Fayech, C., Pein, F., Diallo, S., Schwartz, B., Vu-Bezin, G., Veres, C., Souchard, V., Dumas, A., Bolle, S., Thomas-Teinturier, C., Pacquement, H., Munzer, M., Bondiau, P.-Y., Berchery, D., Fresneau, B., Oberlin, O., ... Haddy, N. (2019). The role of irradiated heart and left ventricular volumes in heart failure occurrence after childhood cancer. *European Journal of Heart Failure*, 21(4), 509-518. <https://doi.org/10.1002/ejhf.1376>
- Mattioli, R., Ilari, A., Colotti, B., Mosca, L., Fazi, F., & Colotti, G. (2023). Doxorubicin and other anthracyclines in cancers: Activity, chemoresistance and its overcoming. *Molecular Aspects of Medicine*, 93, 101205. <https://doi.org/10.1016/j.mam.2023.101205>
- Mazevet, M., Belhadef, A., Ribeiro, M., Dayde, D., Llach, A., Laudette, M., Belleville, T., Mateo, P., Gressette, M., Lefebvre, F., Chen, J., Bachelot-Loza, C., Rucker-Martin,

C., Lezoualch, F., Crozatier, B., Benitah, J.-P., Vozenin, M.-C., Fischmeister, R., Gomez, A.-M., ... Morel, E. (2023). EPAC1 inhibition protects the heart from doxorubicin-induced toxicity. *eLife*, 12, e83831. <https://doi.org/10.7554/eLife.83831>

Megias-Vericat, J. E., Martinez-Cuadron, D., Herrero, M. J., Alino, S. F., Poveda, J. L., Sanz, M. A., & Montesinos, P. (2018). Pharmacogenetics of Metabolic Genes of Anthracyclines in Acute Myeloid Leukemia. *Current Drug Metabolism*, 19(1), 55-74. <https://doi.org/10.2174/1389200218666171101124931>

Megías-Vericat, J. E., Montesinos, P., Herrero, M. J., Moscardó, F., Bosó, V., Rojas, L., Martínez-Cuadrón, D., Rodríguez-Veiga, R., Sendra, L., Cervera, J., Poveda, J. L., Sanz, M. Á., & Aliño, S. F. (2018). Impact of NADPH oxidase functional polymorphisms in acute myeloid leukemia induction chemotherapy. *The Pharmacogenomics Journal*, 18(2), 301-307. <https://doi.org/10.1038/tpj.2017.19>

Meinshausen, N., & Buehlmann, P. (2009). *Stability Selection* (arXiv:0809.2932). arXiv. <http://arxiv.org/abs/0809.2932>

Mladosevicova, B., Urbanova, D., Radvanska, E., Slavkovsky, P., & Simkova, I. (2012). Role of NT-proBNP in detection of myocardial damage in childhood leukemia survivors treated with and without anthracyclines. *Journal of Experimental & Clinical Cancer Research: CR*, 31(1), 86. <https://doi.org/10.1186/1756-9966-31-86>

Mulrooney, D. A., Yeazel, M. W., Kawashima, T., Mertens, A. C., Mitby, P., Stovall, M.,

- Donaldson, S. S., Green, D. M., Sklar, C. A., Robison, L. L., & Leisenring, W. M. (2009). Cardiac outcomes in a cohort of adult survivors of childhood and adolescent cancer: Retrospective analysis of the Childhood Cancer Survivor Study cohort. *The BMJ*, *339*, b4606. <https://doi.org/10.1136/bmj.b4606>
- Niska, J. R., Thorpe, C. S., Allen, S. M., Daniels, T. B., Rule, W. G., Schild, S. E., Vargas, C. E., & Mookadam, F. (2018). Radiation and the heart: Systematic review of dosimetry and cardiac endpoints. *Expert Review of Cardiovascular Therapy*, *16*(12), 931-950. <https://doi.org/10.1080/14779072.2018.1538785>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. <https://doi.org/10.1136/bmj.n71>
- Peters, U., Preisler-Adams, S., Hebeisen, A., Hahn, M., Seifert, E., Lanvers, C., Heinecke, A., Horst, J., Jürgens, H., & Lamprecht-Dinnesen, A. (2000). Glutathione S-transferase genetic polymorphisms and individual sensitivity to the ototoxic effect of cisplatin: *Anti-Cancer Drugs*, *11*(8), 639-643. <https://doi.org/10.1097/00001813-200009000-00007>
- Petrykey, K., Rezgui, A. M., Guern, M. L., Beaulieu, P., St-Onge, P., Drouin, S., Bertout, L., Wang, F., Baedke, J. L., Yasui, Y., Hudson, M. M., Raboisson, M.-J., Laverdière, C.,

- Sinnett, D., Andelfinger, G. U., & Krajinovic, M. (2021). Genetic factors in treatment-related cardiovascular complications in survivors of childhood acute lymphoblastic leukemia. *Pharmacogenomics*, 22(14), 885-901. <https://doi.org/10.2217/pgs-2021-0067>
- Podlesnikar, T., Berlot, B., Dolenc, J., Goričar, K., & Marinko, T. (2022). Radiotherapy-Induced Cardiotoxicity: The Role of Multimodality Cardiovascular Imaging. *Frontiers in Cardiovascular Medicine*, 9, 887705. <https://doi.org/10.3389/fcvm.2022.887705>
- Pommier, Y., Leo, E., Zhang, H., & Marchand, C. (2010). DNA Topoisomerases and Their Poisoning by Anticancer and Antibacterial Drugs. *Chemistry & biology*, 17(5), 421-433. <https://doi.org/10.1016/j.chembiol.2010.04.012>
- Ponikowski, P., Voors, A. A., Anker, S. D., Bueno, H., Cleland, J. G. F., Coats, A. J. S., Falk, V., González-Juanatey, J. R., Harjola, V.-P., Jankowska, E. A., Jessup, M., Linde, C., Nihoyannopoulos, P., Parissis, J. T., Pieske, B., Riley, J. P., Rosano, G. M. C., Ruilope, L. M., Ruschitzka, F., ... ESC Scientific Document Group. (2016). 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *European Heart Journal*, 37(27), 2129-2200. <https://doi.org/10.1093/eurheartj/ehw128>
- Pritchard-Jones, K., Pieters, R., Reaman, G. H., Hjorth, L., Downie, P., Calaminus, G.,

- Naafs-Wilstra, M. C., & Steliarova-Foucher, E. (2013). Sustaining innovation and improvement in the treatment of childhood cancer: Lessons from high-income countries. *The Lancet Oncology*, *14*(3), e95-e103. [https://doi.org/10.1016/S1470-2045\(13\)70010-X](https://doi.org/10.1016/S1470-2045(13)70010-X)
- Purdy, J. A. (2008). DOSE TO NORMAL TISSUES OUTSIDE THE RADIATION THERAPY PATIENT'S TREATED VOLUME: A REVIEW OF DIFFERENT RADIATION THERAPY TECHNIQUES. *Health Physics*, *95*(5), 666-676. <https://doi.org/10.1097/01.HP.0000326342.47348.06>
- Qian, J., Payabvash, S., Kemmling, A., Lev, M. H., Schwamm, L. H., & Betensky, R. A. (2014). Variable Selection and Prediction Using a Nested, Matched Case-Control Study: Application to Hospital Acquired Pneumonia in Stroke Patients. *Biometrics*, *70*(1), 153-163. <https://doi.org/10.1111/biom.12113>
- Qiu, Y., Jiang, P., & Huang, Y. (2023). Anthracycline-induced cardiotoxicity: Mechanisms, monitoring, and prevention. *Frontiers in Cardiovascular Medicine*, *10*, 1242596. <https://doi.org/10.3389/fcvm.2023.1242596>
- Rahmenführer, J., De Bin, R., Benner, A., Ambroggi, F., Lusa, L., Boulesteix, A.-L., Migliavacca, E., Binder, H., Michiels, S., Sauerbrei, W., McShane, L., & for topic group "High-dimensional data" (TG9) of the STRATOS initiative. (2023). Statistical analysis of high-dimensional biomedical data: A gentle introduction to analytical goals, common approaches and challenges. *BMC Medicine*, *21*(1), 182. <https://doi.org/10.1186/s12916-023-02858-y>

Registre National des Cancers de l'Enfant. (s. d.). *Taux d'incidence de 2014 à 2020, par groupe diagnostique—RNCE*. Consulté 7 juillet 2024, à l'adresse <https://rnce.inserm.fr/index.php/fr/statistiques/statistiques-d-incidence/taux-d-incidence-de-2014-a-2020-par-groupe-diagnostique>

Reid, S., & Tibshirani, R. (2014). Regularization Paths for Conditional Logistic Regression: The clogitL1 Package. *Journal of statistical software*, 58(12), 12.

Reulen, R. C., Winter, D. L., Frobisher, C., Lancashire, E. R., Stiller, C. A., Jenney, M. E., Skinner, R., Stevens, M. C., Hawkins, M. M., & British Childhood Cancer Survivor Study Steering Group, for the. (2010). Long-term Cause-Specific Mortality Among Survivors of Childhood Cancer. *JAMA*, 304(2), 172-179. <https://doi.org/10.1001/jama.2010.923>

Roberts, J. A., Rainbow, R. D., & Sharma, P. (2023). Mitigation of Cardiovascular Disease and Toxicity through NRF2 Signalling. *International Journal of Molecular Sciences*, 24(7), 6723. <https://doi.org/10.3390/ijms24076723>

Roberts, S., & Nowak, G. (2014). Stabilizing the lasso against cross-validation variability. *Computational Statistics & Data Analysis*, 70, 198-211. <https://doi.org/10.1016/j.csda.2013.09.008>

Robison, L. L., & Hudson, M. M. (2014). Survivors of Childhood/Adolescent Cancer: Life-long Risks and Responsibilities. *Nature reviews. Cancer*, 14(1), 61-70. <https://doi.org/10.1038/nrc3634>

Rocca, C., Pasqua, T., Cerra, M. C., & Angelone, T. (2020). Cardiac Damage in

- Anthracyclines Therapy: Focus on Oxidative Stress and Inflammation. *Antioxidants & Redox Signaling*, 32(15), 1081-1097. <https://doi.org/10.1089/ars.2020.8016>
- Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology* (3rd ed). Wolters Kluwer health - Lippincott Williams & Wilkins.
- Rothwell, P. M. (2005). Subgroup analysis in randomised controlled trials: Importance, indications, and interpretation. *The Lancet*, 365(9454), 176-186. [https://doi.org/10.1016/S0140-6736\(05\)17709-5](https://doi.org/10.1016/S0140-6736(05)17709-5)
- Sági, J. C., Egyed, B., Kelemen, A., Kutszegi, N., Hegyi, M., Gézsi, A., Herlitschke, M. A., Rzepiel, A., Fodor, L. E., Ottóffy, G., Kovács, G. T., Erdélyi, D. J., Szalai, C., & Semsei, Á. F. (2018). Possible roles of genetic variations in chemotherapy related cardiotoxicity in pediatric acute lymphoblastic leukemia and osteosarcoma. *BMC Cancer*, 18(1), 704. <https://doi.org/10.1186/s12885-018-4629-6>
- Samuel, Y., Babu, A., Karagkouni, F., Ismail, A., Choi, S., & Boussios, S. (2023). Cardiac Toxicities in Oncology: Elucidating the Dark Box in the Era of Precision Medicine. *Current Issues in Molecular Biology*, 45(10), 8337-8358. <https://doi.org/10.3390/cimb45100526>
- Sapkota, Y., Liu, Q., Li, N., Bhatt, N. S., Ehrhardt, M. J., Wilson, C. L., Wang, Z., Jefferies, J. L., Zhang, J., Armstrong, G. T., Hudson, M. M., Robison, L. L., Mulrooney, D. A., & Yasui, Y. (2022). Contribution of Genome-Wide Polygenic Score to Risk of Coronary Artery Disease in Childhood Cancer Survivors. *JACC: CardioOncology*,

4(2), 258-267. <https://doi.org/10.1016/j.jaccao.2022.04.003>

Sapkota, Y., Qin, N., Ehrhardt, M. J., Wang, Z., Chen, Y., Wilson, C. L., Estepp, J., Rai, P., Hankins, J. S., Burridge, P. W., Jefferies, J. L., Zhang, J., Hudson, M. M., Robison, L. L., Armstrong, G. T., Mulrooney, D. A., & Yasui, Y. (2021). Genetic Variants Associated with Therapy-Related Cardiomyopathy among Childhood Cancer Survivors of African Ancestry. *Cancer Research*, *81*(9), 2556-2565. <https://doi.org/10.1158/0008-5472.CAN-20-2675>

Satagopan, J. M., Ben-Porat, L., Berwick, M., Robson, M., Kutler, D., & Auerbach, A. D. (2004). A note on competing risks in survival data analysis. *British Journal of Cancer*, *91*(7), 1229-1235. <https://doi.org/10.1038/sj.bjc.6602102>

Savarese, G., von Haehling, Stephan, Butler, Javed, Cleland, John GF., Ponikowski, Piotr, & Anker, Stefan D. (2023). Iron deficiency and cardiovascular disease. *European Heart Journal*, *44*, 14-27. <https://doi.org/10.1093/eurheartj/ehac569>

Semsei, A. F., Erdelyi, D. J., Ungvari, I., Csagoly, E., Hegyi, M. Z., Kiszal, P. S., Lautner-Csorba, O., Szabolcs, J., Masat, P., Fekete, G., Falus, A., Szalai, C., & Kovacs, G. T. (2012). *ABCC1* polymorphisms in anthracycline-induced cardiotoxicity in childhood acute lymphoblastic leukaemia. *Cell Biology International*, *36*(1), 79-86. <https://doi.org/10.1042/CBI20110264>

Shi, H., & Yin, G. (2018). Boosting conditional logit model. *Journal of Choice Modelling*, *26*, 48-63. <https://doi.org/10.1016/j.jocm.2017.07.002>

Shomal Zadeh, N., Lin, S., & Runger, G. C. (2020). Matched Forest: Supervised learning

- for high-dimensional matched case-control studies. *Bioinformatics*, 36(5), 1570-1576. <https://doi.org/10.1093/bioinformatics/btz785>
- Siaravas, K. C., Katsouras, C. S., & Sioka, C. (2023). Radiation Treatment Mechanisms of Cardiotoxicity: A Systematic Review. *International Journal of Molecular Sciences*, 24(7), 6272. <https://doi.org/10.3390/ijms24076272>
- Singh, P., Shah, D. A., Jouni, M., Cejas, R. B., Crossman, D. K., Magdy, T., Qiu, S., Wang, X., Zhou, L., Sharafeldin, N., Hageman, L., McKenna, D. E., Armenian, S. H., Balis, F. M., Hawkins, D. S., Keller, F. G., Hudson, M. M., Neglia, J. P., Ritchey, A. K., ... Bhatia, S. (2023). Altered Peripheral Blood Gene Expression in Childhood Cancer Survivors With Anthracycline-Induced Cardiomyopathy – A COG-ALTE03N1 Report. *Journal of the American Heart Association*, 12(19), e029954. <https://doi.org/10.1161/JAHA.123.029954>
- Singh, P., Wang, X., Hageman, L., Chen, Y., Magdy, T., Landier, W., Ginsberg, J. P., Neglia, J. P., Sklar, C. A., Castellino, S. M., Dreyer, Z. E., Hudson, M. M., Robison, L. L., Blanco, J. G., Relling, M. V., Burrige, P., & Bhatia, S. (2020). Association of *GSTM1* null variant with anthracycline-related cardiomyopathy after childhood cancer—A Children’s Oncology Group ALTE03N1 report. *Cancer*, 126(17), 4051-4058. <https://doi.org/10.1002/cncr.32948>
- Sohani, Z. N., Meyre, D., de Souza, R. J., Joseph, P. G., Gandhi, M., Dennis, B. B., Norman, G., & Anand, S. S. (2015). Assessing the quality of published genetic association studies in meta-analyses: The quality of genetic studies (Q-Genie) tool. *BMC*

- Genetics*, 16(1), 50. <https://doi.org/10.1186/s12863-015-0211-2>
- Songbo, M., Lang, H., Xinyong, C., Bin, X., Ping, Z., & Liang, S. (2019). Oxidative stress injury in doxorubicin-induced cardiotoxicity. *Toxicology Letters*, 307, 41-48. <https://doi.org/10.1016/j.toxlet.2019.02.013>
- Steinmeier, T., Schulze Schleithoff, S., & Timmermann, B. (2019). Evolving Radiotherapy Techniques in Paediatric Oncology. *Clinical Oncology*, 31(3), 142-150. <https://doi.org/10.1016/j.clon.2018.12.005>
- Steliarova-Foucher, E., Colombet, M., Ries, L. A. G., Moreno, F., Dolya, A., Bray, F., Hesselning, P., Shin, H. Y., & Stiller, C. A. (2017). International incidence of childhood cancer, 2001–10: A population-based registry study. *The Lancet Oncology*, 18(6), 719-731. [https://doi.org/10.1016/S1470-2045\(17\)30186-9](https://doi.org/10.1016/S1470-2045(17)30186-9)
- Steliarova-Foucher, E., Stiller, C., Lacour, B., & Kaatsch, P. (2005). International Classification of Childhood Cancer, third edition. *Cancer*, 103(7), 1457-1467. <https://doi.org/10.1002/cncr.20910>
- Steyerberg, E. W. (2009). *Clinical Prediction Models*. Springer New York. <https://doi.org/10.1007/978-0-387-77244-8>
- Taverne, Y. J. H. J., Bogers, A. J. J. C., Duncker, D. J., & Merkus, D. (2013). Reactive Oxygen Species and the Cardiovascular System. *Oxidative Medicine and Cellular Longevity*, 2013, 1-15. <https://doi.org/10.1155/2013/862423>
- Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.

<https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>

Truong, Q. A., Massaro, J. M., Rogers, I. S., Mahabadi, A. A., Kriegel, M. F., Fox, C. S., O'Donnell, C. J., & Hoffmann, U. (2012). Reference values for normal pulmonary artery dimensions by noncontrast cardiac computed tomography: The Framingham Heart Study. *Circulation. Cardiovascular Imaging*, 5(1), 147-154. <https://doi.org/10.1161/CIRCIMAGING.111.968610>

Tukenova, M., Guibout, C., Oberlin, O., Doyon, F., Mousannif, A., Haddy, N., Guérin, S., Pacquement, H., Aouba, A., Hawkins, M., Winter, D., Bourhis, J., Lefkopoulos, D., Diallo, I., & de Vathaire, F. (2010). Role of Cancer Treatment in Long-Term Overall and Cardiovascular Mortality After Childhood Cancer. *Journal of Clinical Oncology*, 28(8), 1308-1315. <https://doi.org/10.1200/JCO.2008.20.2267>

US Department of Health and Human Services. (2010). *Common Terminology Criteria for Adverse Events (CTCAE). Version 4.03*. https://www.eortc.be/services/doc/ctc/ctcae_4.03_2010-06-14_quickreference_5x7.pdf

Van Kalsbeek, R. J., Van Der Pal, H. J. H., Kremer, L. C. M., Bardi, E., Brown, M. C., Effeney, R., Winther, J. F., Follin, C., Den Hartogh, J., Haupt, R., Hjorth, L., Kepak, T., Kepakova, K., Levitt, G., Loonen, J. J., Mangelschots, M., Muraca, M., Renard, M., Sabic, H., ... Mulder, R. L. (2021). European PanCareFollowUp Recommendations for surveillance of late effects of childhood, adolescent, and young adult cancer. *European Journal of Cancer*, 154, 316-328.

<https://doi.org/10.1016/j.ejca.2021.06.004>

van Nimwegen, F. A., Ntentas, G., Darby, S. C., Schaapveld, M., Hauptmann, M., Lugtenburg, P. J., Janus, C. P. M., Daniels, L., van Leeuwen, F. E., Cutter, D. J., & Aleman, B. M. P. (2017). Risk of heart failure in survivors of Hodgkin lymphoma: Effects of cardiac exposure to radiation and anthracyclines. *Blood*, *129*(16), 2257-2265. <https://doi.org/10.1182/blood-2016-09-740332>

Veres, C., Allodji, R. S., Llanas, D., Bezin, J. V., Chavaudra, J., Mège, J. P., Lefkopoulos, D., Quiniou, E., Deutsh, E., Vathaire, F. de, & Diallo, I. (2014). Retrospective Reconstructions of Active Bone Marrow Dose-Volume Histograms. *International Journal of Radiation Oncology, Biology, Physics*, *90*(5), 1216-1224. <https://doi.org/10.1016/j.ijrobp.2014.08.335>

Visscher, H., Rassekh, S. R., Sandor, G. S., Caron, H. N., van Dalen, E. C., Kremer, L. C., van der Pal, H. J., Rogers, P. C., Rieder, M. J., Carleton, B. C., Hayden, M. R., Ross, C. J., & CPNDS consortium. (2015). Genetic variants in SLC22A17 and SLC22A7 are associated with anthracycline-induced cardiotoxicity in children. *Pharmacogenomics*, *16*(10), 1065-1076. <https://doi.org/10.2217/pgs.15.61>

Visscher, H., Ross, C. J. D., Rassekh, S. R., Barhdadi, A., Dubé, M.-P., Al-Saloos, H., Sandor, G. S., Caron, H. N., van Dalen, E. C., Kremer, L. C., van der Pal, H. J., Brown, A. M. K., Rogers, P. C., Phillips, M. S., Rieder, M. J., Carleton, B. C., Hayden, M. R., & Canadian Pharmacogenomics Network for Drug Safety Consortium. (2012). Pharmacogenomic prediction of anthracycline-induced cardiotoxicity in

- children. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 30(13), 1422-1428. <https://doi.org/10.1200/JCO.2010.34.3467>
- Visscher, H., Ross, C. J. D., Rassekh, S. R., Sandor, G. S. S., Caron, H. N., van Dalen, E. C., Kremer, L. C., van der Pal, H. J., Rogers, P. C., Rieder, M. J., Carleton, B. C., Hayden, M. R., & CPNDS Consortium. (2013). Validation of variants in SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. *Pediatric Blood & Cancer*, 60(8), 1375-1381. <https://doi.org/10.1002/pbc.24505>
- Vũ Bezin, J., Allodji, R. S., Mège, J.-P., Beldjoudi, G., Saunier, F., Chavaudra, J., Deutsch, E., de Vathaire, F., Bernier, V., Carrie, C., Lefkopoulos, D., & Diallo, I. (2017). A review of uncertainties in radiotherapy dose reconstruction and their impacts on dose-response relationships. *Journal of Radiological Protection: Official Journal of the Society for Radiological Protection*, 37(1), R1-R18. <https://doi.org/10.1088/1361-6498/aa575d>
- Vuong, J. T., Stein-Merlob, A. F., Cheng, R. K., & Yang, E. H. (2022). Novel Therapeutics for Anthracycline Induced Cardiotoxicity. *Frontiers in Cardiovascular Medicine*, 9. <https://doi.org/10.3389/fcvm.2022.863314>
- Wain, L. V., Vaez, A., Jansen, R., Joehanes, R., van der Most, P. J., Erzurumluoglu, A. M., O'Reilly, P. F., Cabrera, C. P., Warren, H. R., Rose, L. M., Verwoert, G. C., Hottenga, J.-J., Strawbridge, R. J., Esko, T., Arking, D. E., Hwang, S.-J., Guo, X., Kutalik, Z., Trompet, S., ... Ehret, G. B. (2017). Novel Blood Pressure Locus and Gene

Discovery Using Genome-Wide Association Study and Expression Data Sets From Blood and the Kidney. *Hypertension (Dallas, Tex.: 1979)*, 70(3), e4-e19. <https://doi.org/10.1161/HYPERTENSIONAHA.117.09438>

Wang, K., & Tepper, J. E. (2021). Radiation therapy-associated toxicity: Etiology, management, and prevention. *CA: A Cancer Journal for Clinicians*, 71(5), 437-454. <https://doi.org/10.3322/caac.21689>

Wang, X., Liu, W., Sun, C.-L., Armenian, S. H., Hakonarson, H., Hageman, L., Ding, Y., Landier, W., Blanco, J. G., Chen, L., Quiñones, A., Ferguson, D., Winick, N., Ginsberg, J. P., Keller, F., Neglia, J. P., Desai, S., Sklar, C. A., Castellino, S. M., ... Bhatia, S. (2014). Hyaluronan synthase 3 variant and anthracycline-related cardiomyopathy: A report from the children's oncology group. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 32(7), 647-653. <https://doi.org/10.1200/JCO.2013.50.3557>

Wang, X., Singh, P., Zhou, L., Sharafeldin, N., Landier, W., Hageman, L., Burridge, P., Yasui, Y., Sapkota, Y., Blanco, J. G., Oeffinger, K. C., Hudson, M. M., Chow, E. J., Armenian, S. H., Neglia, J. P., Ritchey, A. K., Hawkins, D. S., Ginsberg, J. P., Robison, L. L., ... Bhatia, S. (2022). Genome-Wide Association Study Identifies ROBO2 as a Novel Susceptibility Gene for Anthracycline-Related Cardiomyopathy in Childhood Cancer Survivors. *Journal of Clinical Oncology*, JCO.22.01527. <https://doi.org/10.1200/JCO.22.01527>

Wang, X., Sun, C.-L., Quiñones-Lombraña, A., Singh, P., Landier, W., Hageman, L.,

- Mather, M., Rotter, J. I., Taylor, K. D., Chen, Y.-D. I., Armenian, S. H., Winick, N., Ginsberg, J. P., Neglia, J. P., Oeffinger, K. C., Castellino, S. M., Dreyer, Z. E., Hudson, M. M., Robison, L. L., ... Bhatia, S. (2016). *CELF4* Variant and Anthracycline-Related Cardiomyopathy: A Children's Oncology Group Genome-Wide Association Study. *Journal of Clinical Oncology*, *34*(8), 863-870. <https://doi.org/10.1200/JCO.2015.63.4550>
- Xiang, S. Y., Ouyang, K., Yung, B. S., Miyamoto, S., Smrcka, A. V., Chen, J., & Brown, J. H. (2013). PLC ϵ , PKD1, and SSH1L Transduce RhoA Signaling to Protect Mitochondria from Oxidative Stress in the Heart. *Science Signaling*, *6*(306), ra108-ra108. <https://doi.org/10.1126/scisignal.2004405>
- Xu, H., Qian, J., Paynter, N. P., Zhang, X., Whitcomb, B. W., Tworoger, S. S., Rexrode, K. M., Hankinson, S. E., & Balasubramanian, R. (2019). Estimating the receiver operating characteristic curve in matched case control studies: Estimating the ROC curve in matched case control studies. *Statistics in Medicine*, *38*(3), 437-451. <https://doi.org/10.1002/sim.7986>
- Yang, X., Li, G., Yang, T., Guan, M., An, N., Yang, F., Dai, Q., Zhong, C., Luo, C., Gao, Y., Das, S., Xing, Y., & Shang, H. (2020). Possible Susceptibility Genes for Intervention against Chemotherapy-Induced Cardiotoxicity. *Oxidative Medicine and Cellular Longevity*, *2020*, 1-30. <https://doi.org/10.1155/2020/4894625>
- Zeller, T., Wild, P., Szymczak, S., Rotival, M., Schillert, A., Castagne, R., Maouche, S., Germain, M., Lackner, K., Rossmann, H., Eleftheriadis, M., Sinning, C. R., Schnabel,

- R. B., Lubos, E., Mennerich, D., Rust, W., Perret, C., Proust, C., Nicaud, V., ... Cambien, F. (2010). Genetics and Beyond – The Transcriptome of Human Monocytes and Disease Susceptibility. *PLoS ONE*, 5(5), e10693. <https://doi.org/10.1371/journal.pone.0010693>
- Zhang, S., Liu, X., Bawa-Khalfe, T., Lu, L.-S., Lyu, Y. L., Liu, L. F., & Yeh, E. T. H. (2012). Identification of the molecular basis of doxorubicin-induced cardiotoxicity. *Nature Medicine*, 18(11), 1639-1642. <https://doi.org/10.1038/nm.2919>
- Zhang, X., Hu, C., Kong, C.-Y., Song, P., Wu, H.-M., Xu, S.-C., Yuan, Y.-P., Deng, W., Ma, Z.-G., & Tang, Q.-Z. (2020). FNDC5 alleviates oxidative stress and cardiomyocyte apoptosis in doxorubicin-induced cardiotoxicity via activating AKT. *Cell Death & Differentiation*, 27(2), Article 2. <https://doi.org/10.1038/s41418-019-0372-z>
- Zhang, Z. (2017). Survival analysis in the presence of competing risks. *Annals of Translational Medicine*, 5(3), 47. <https://doi.org/10.21037/atm.2016.08.62>
- Zou, H., & Hastie, T. (2005). Regularization and Variable Selection Via the Elastic Net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), 301-320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

ANNEXES

ANNEXE A. ANNEXES EN LIEN AVEC LES RISQUES DE PATHOLOGIES CARDIAQUES ET D'HOSPITALISATION CARDIAQUE

Annexe A. 1. Effet des doses cumulées d'anthracyclines et des doses moyennes de radiations cardiaques sur le risque de décès non cardiaque dans la cohorte FCCSS

Caractéristiques	N	Décès avant pathologie cardiaque (n= 1166)	
		csHR (95 % IC)	p-value
Âge au diagnostic du premier cancer (années)		1,0 (1,0-1,1)	<0,0001
Genre			0,0581
Masculin	732	Ref	
Féminin	571	0,9 (0,8-1)	
Année de diagnostic du premier cancer			0,2195
Avant1970]	264	Ref	
[1970-1980[409	1,2 (1-1,4)	
[1980-1990[370	1,2 (1-1,5)	
[1990 – 2000[260	1,2 (1-1,5)	
Doses cumulées d'anthracyclines (mg/m²)			<0,0001
Aucune	1010	Ref	
]0-250]	135	0,6 (0,5-0,7)	
]250-360]	67	0,7 (0,6-1)	
>360	91	0,8 (0,7-1)	
Doses moyennes de radiations au cœur (Gy)			<0,0001
Aucune	263	Ref	
]0-1]	451	3,1 (2,6-3,7)	
]1-5]	201	2,8 (2,3-3,5)	
]5-15]	153	3,1 (2,5-3,8)	
]15-25]	131	3,2 (2,6-4,1)	
]25-35]	43	3,1 (2,2-4,3)	
>35	33	4,2 (2,8-6,2)	

Abréviations : ; N : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence.* Les csHR ont été estimés à partir d'un modèle de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), e le genre, et l'année calendaire du premier cancer

Annexe A. 2. Effet des pourcentages de volume du cœur recevant des doses spécifiques de radiothérapie sur le risque de décès non cardiaque dans la cohorte FCCSS

Décès avant pathologie cardiaque			
Caractéristiques	N	csHR (95 % IC)	p-value
Volume (%) du cœur ayant reçu au moins 5 Gy			<0,0001
Pas de RT	201	Ref	
0%	518	3,2 (2,7-3,8)	
]0-10] %	39	1,8 (1,3-2,5)	
]10-50] %	106	3,2 (2,5-4,1)	
>50 %	276	3,4 (2,8-4,1)	
Volume (%) du cœur ayant reçu au moins 15 Gy			<0,0001
Pas de RT	201	Ref	
0%	598	3,1 (2,6-3,7)	
]0-10] %	56	2,5 (1,9-3,4)	
]10-50] %	99	3,3 (2,6-4,3)	
>50 %	186	3,5 (2,9-4,3)	
Volume (%) du cœur ayant reçu au moins 20 Gy			<0,0001
Pas de RT	201	Ref	
0%	633	3,1 (2,6-3,6)	
]0-10] %	73	2,8 (2,1-3,7)	
]10-50] %	102	3,4 (2,7-4,4)	
>50 %	131	3,7 (3-4,7)	
Volume (%) du cœur ayant reçu au moins 30 Gy			<0,0001
Pas de RT	201	Ref	
0%	738	3,1 (2,6-3,6)	
]0-10] %	80	3,1 (2,4-4,1)	
]10-50] %	62	4,5 (3,4-6)	
>50 %	59	3,7 (2,7-5)	
Volume (%) du cœur ayant reçu 0,1 Gy quand V5 = 0%			<0,0001
Pas de RT	201	Ref	

0%	56	1,75 (1,29-2,36)
]0-10] %	8	2,29 (1,13-4,66)
]10-50] %	12	2,26 (1,25-4,07)
>50 %	442	3,8 (3,19-4,54)
Volume (%) du cœur ayant reçu 5 Gy quand V20 = 0%		<0,0001
Pas de RT	201	Ref
0%	518	3,3 (2,7-3,9)
]0-10] %	36	1,8 (1,2-2,6)
]10-50] %	46	3,8 (2,7-5,2)
>50 %	33	2,8 (2-4,1)

Abréviations ; N : nombre de décès; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence ; V5 : Volume du cœur ayant reçu au moins 5Gy ; V20 : Volume du cœur ayant reçu au moins 20Gy ; RT : Radiothérapie.

Toutes les sections de ce tableau proviennent de modèle distincts, exclusifs l'un de l'autre. Les csHR ont été estimés à partir de modèles de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), le genre, l'année calendaire du premier cancer, et la dose cumulée d'anthracyclines

Annexe A. 3. Effet et incidence du traitement par anthracyclines et des doses moyennes de radiations cardiaque sur le risque de décès non cardiaque dans la cohorte FCCSS

Administration d'anthracycline	Dose moyenne de radiations au cœur, Gy	N	Incidence à 40 ans de suivi (95 % IC)	csHR (95 % IC)
Non	Aucune	155	10,0 (8,1 – 12,1)	Ref
]0-1]	332	32,9 (29,5 – 36,4)	3,6 (2,9-4,4)
]1-5]	143	24,1 (20,1 – 28,2)	2,9 (2,3-3,7)
]5-15]	118	30,2 (25,2 – 35,4)	3,3 (2,6-4,2)
]15-25]	86	34,4 (27,9 – 41,1)	3,8 (2,9-5,0)
	> 25†	53	28,9 (22,2 – 35,8)	3,6 (2,6-4,9)
				11,2 (6,3 – 17,7)
Oui	Aucune	77	17,5 (13,4 – 22,0)	0,9 (0,7-1,2)
]0-1]	74	25,3 (17,5 – 33,9)	2,0 (1,5-2,6)
]1-5]	39	20,7 (12,0 – 30,9)	2,6 (1,8-3,7)
]5-15]	21	14,1 (9,2 – 20,1)	2,2 (1,4-3,4)
]15-25]	26	30,9 (8,8 – 56,7)	1,8 (1,2-2,8)
	> 25†	16	10,0 (8,1 – 12,1)	3,2 (1,9-5,3)

Abréviations ; N : nombre total de survivants ; n : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence. †Certaines catégories de doses ont été regroupées à cause du faible nombre d'évènement dans les catégories précédentes.

Les csHR ont été estimés à partir d'un modèle de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), le genre, et l'année calendaire du premier cancer

Annexe A. 4. Effet des doses cumulées d'anthracyclines et des doses moyennes de radiations cardiaques sur le risque de décès sans d'avoir développé une insuffisance cardiaque dans la cohorte FCCSS

Décès avant pathologie cardiaque (n= 1166)			
Caractéristiques	N	csHR (95 % IC)	p-value
Âge au diagnostic du premier cancer (années)		1,0 (1,0-1,1)	<0,0001
Genre			0,0441
Masculin	691	Ref	
Féminin	533	0,9 (0,8-1,0)	
Année de diagnostic du premier cancer			0,1635
Avant1970]	253	Ref	
[1970-1980[393	1,2 (1,0-1,4)	
[1980-1990[341	1,2 (1,0-1,5)	
[1990 – 2000[237	1,2 (1,0-1,5)	
Doses cumulées d'anthracyclines (mg/m²)			<0,0001
Aucune	956	Ref	
]0-250]	120	0,6 (0,5-0,7)	
]250-360]	60	0,7 (0,6-0,9)	
>360	88	0,8 (0,6-1)	
Doses moyennes de radiations au cœur (Gy)			<0,0001
Aucune	236	Ref	
]0-1]	412	3,1 (2,6-3,6)	
]1-5]	187	2,8 (2,3-3,4)	
]5-15]	145	3,1 (2,5-3,8)	
]15-25]	125	3,5 (2,8-4,3)	
]25-35]	54	3,8 (2,8-5,2)	
>35	39	4,8 (3,4-6,8)	

Abréviations : N : nombre d'évènement ; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence.* Les csHR ont été estimés à partir d'un modèle de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), e le genre, et l'année calendaire du premier cancer

Annexe A. 5. Effet des pourcentages de volume du cœur recevant des doses spécifiques de radiothérapie sur le risque de décès sans d'avoir développé une insuffisance cardiaque dans la cohorte FCCSS

Caractéristiques	N	Décès avant pathologie cardiaque	
		csHR (95 % IC)	p-value
Volume (%) du cœur ayant reçu au moins 5 Gy			<0,0001
Pas de RT	205		
0%	527	3,2 (2,7-3,7)	
]0-10] %	40	1,8 (1,2-2,5)	
]10-50] %	109	3,2 (2,5-4)	
>50 %	317	3,7 (3,1-4,4)	
Volume (%) du cœur ayant reçu au moins 15 Gy			<0,0001
Pas de RT	205		
0%	610	3,1 (2,6-3,6)	
]0-10] %	57	2,5 (1,9-3,4)	
]10-50] %	104	3,4 (2,7-4,3)	
>50 %	222	3,9 (3,2-4,8)	
Volume (%) du cœur ayant reçu au moins 20 Gy			<0,0001
Pas de RT	205		
0%	646	3 (2,6-3,6)	
]0-10] %	74	2,8 (2,1-3,6)	
]10-50] %	110	3,6 (2,8-4,6)	
>50 %	163	4,3 (3,5-5,3)	
Volume (%) du cœur ayant reçu au moins 30 Gy			<0,0001
Pas de RT	205		
0%	761	3,1 (2,6-3,6)	
]0-10] %	85	3,2 (2,5-4,2)	
]10-50] %	65	4,5 (3,4-6)	
>50 %	82	4,5 (3,5-5,9)	
Volume (%) du cœur ayant reçu 0,1 Gy quand V5 = 0%			<0,0001
Pas de RT	205		
0%	58	1,78 (1,33-2,4)	
]0-10] %	8	2,26 (1,11-4,6)	

]10-50] %	12	2,22 (1,24-4)	
>50 %	449	3,76 (3,16-4,48)	
Volume (%) du cœur ayant reçu 5 Gy quand V20 = 0%			<0,0001
Pas de RT	205		
0%	527	3,2 (2,7-3,8)	
]0-10] %	37	1,8 (1,3-2,6)	
]10-50] %	47	3,7 (2,7-5,1)	
>50 %	35	2,9 (2-4,1)	

Abréviations ; N : nombre de décès; IC : Intervalle de confiance ; csHR : Hazard Ratio cause spécifique ; Gy : Gray, Ref : Référence ; V5 : Volume du cœur ayant reçu au moins 5Gy ; V20 : Volume du cœur ayant reçu au moins 20Gy ; RT : Radiothérapie.

Toutes les sections de ce tableau proviennent de modèle distincts, exclusifs l'un de l'autre. Les csHR ont été estimés à partir de modèles de Cox cause spécifique utilisant le délai de suivi comme échelle de temps, ajusté pour l'âge (en années), le genre, l'année calendaire du premier cancer, et la dose cumulée d'anthracyclines

Annexe A. 6. Tableau de la 10e classification internationale des maladies (CIM-10) correspondant aux hospitalisations pour cause cardiaque

Pathologie	ICD-10 code	N	%
Pathologies cardiaques ischémiques			
Angine de poitrine	I20	53	8.2
Infarctus du myocarde	I21-22	41	6.3
Autres pathologies ischémiques	I24	4	0.6
Autres pathologies ischémiques chroniques	I25	80	12.4
Péri-, myo- and endocardites			
Pericardite	I30-32	51	7.9
Endocardite	I33, I38-39	12	1.9
Myocardite	I40-41; I51.4	2	0.3
Valvulopathie (non rhumatismale)			
Maladie valvulaire mitrale	I34	12	1.9
Valvulopathie aortique	I35	39	6.0
Autres valvulopathies	I36-37	0	0.0
Cardiomyopathie et arythmie			
Cardiomyopathie	I42-43	106	16.4
Troubles de la conduction	I44-45	17	2.6
Arrêt cardiaque	I46	10	1.5
Troubles du rythme cardiaque	I47-49, R00	68	10.5
Insuffisance cardiaque			
Insuffisance cardiaque	I50	152	23.5
Total		647	100

Abbreviations: N: Nombre de patients

ANNEXE B. ANNEXES EN LIEN AVEC LA REVUE SYSTEMATIQUE

La procédure de sélection a été traduite dans les bases de données bibliographiques électroniques comme suit :

Annexe B 1. Stratégie de recherche NCBI/PubMed et Scopus

Stratégie de recherche NCBI PubMed :

#1 (((((((((((((((Cancer) OR (Oncology)) OR (neoplasms)) OR (malignant)) OR (tumor)) OR (carcinoma)) OR (leukemia)) OR (retinoblastoma)) OR (sarcoma)) OR (osteosarcoma)) OR (Wilmstumor)) OR (nephroblastoma)) OR (neuroblastoma)) OR (brain neoplasms)) OR (CNS)) OR (central nervous system)) OR (lymphoma)) OR (Hodgkin disease)) OR (Hodgkin Diseases)))) AND (((Cancer survivors) OR (« long-term survival »)) OR (« followed up »)) OR (Survivor*))

#2 (((Radiotherapy) OR (Radiations)) OR (« Dose volume »)) OR (((((((((((Chemotherapy) OR (Antimetabolites)) OR (alkylating)) OR (Topoisomerase)) OR (« Vinca Alkaloid »)) OR (« Tyrosine kinase »)) OR (Anthracyclines)) OR (Daunorubicin)) OR (Adriamycin)) OR (Doxorubicin)) OR (Mitoxantrone)) OR (Antineoplastic agent)))

#3 (((((((((((((((Heart diseases) OR (Heart Disorders)) OR (Cardiac Diseases)) OR (Cardiac Disorders)) OR (cardiovascular)) OR (Cardiotoxicit*)) OR (Cardiomyopath*)) OR (Heart failure)) OR (arrhythmia)) OR (Conduction disorders)) OR (Valvular diseases)) OR (heart valve disease)) OR (ischemia)) OR (myocardial infarction)) OR (Coronary artery disease)) OR (Ischemic heart diseases)) OR (Angina)) OR (Heart attack)) OR (Pericarditis)) OR (Cardiac toxicity)))) AND (((((((((((genetics) OR (Genotype)) OR (gene expression)) OR (Transcript*)) OR « SNPs »)) OR « single nucleotid polymorphism »)) OR « copy number variations »)) OR « Sequencing »)) OR « ma seq »)) OR « GWAS »)) OR « genetic association studies »)) OR « Biomarkers »)) OR « genome-wide association study »))

#4 (((Cohort Stud*)) OR (Observational Stud*)) OR (Case control studi*)) OR (Prospective Stud*)) OR (Retrospective Stud*))

#1 AND #2 AND #3 AND #4

Stratégie de recherche Scopus :

#1 ((TITLE-ABS-KEY (((((((((((((((cancer) OR (oncology)) OR (neoplasms)) OR (malignant)) OR (tumor)) OR (carcinoma)) OR (leukemia)) OR (retinoblastoma)) OR (sarcoma)) OR (osteosarcoma)) OR («Wilms tumor»)) OR (nephroblastoma)) OR (neuroblastoma)) OR («brain neoplasms»)) OR (CNS)) OR («central nervous system»)) OR (lymphoma)) OR («hodgkin disease»)))) AND (ALL (((«long-term survival»)) OR («Followed up»)) OR (survivor*))))

#2 ((ALL (((radiotherapy) OR (radiations)) OR («Dose-volume»))) OR (ALL (((((((((((chemotherapy) OR (antimetabolites)) OR (alkylating)) OR (topoisomerase)) OR («Vinca Alkaloid»)) OR («Tyrosine kinase»)) OR (anthracyclines)) OR (daunorubicin)) OR (adriamycin)) OR (doxorubicin)) OR (mitoxantrone)) OR («antineoplastic agent»))))

#3 ((TITLE-ABS (((((((((((((((«heart diseases»)) OR («heart disorders»)) OR («cardiac diseases»)) OR («cardiac disorders»)) OR (cardiovascular)) OR (cardiotoxicit*)) OR (cardiomyopath*)) OR («heart failure»)) OR (arythmia)) OR («conduction disorders»)) OR («valvular diseases»)) OR («heart valve disease»)) OR (ischemia)) OR («myocardial infarction»)) OR («coronary artery disease»)) OR («ischemic heart diseases»)) OR (angina)) OR («heart attack»)) OR (pericarditis)) OR («cardiac toxicity»)))) AND (TITLE-ABS (((((((((((((((«genome-wide association study»)) OR (polymorphism)) OR (genetic)) OR (genotype)) OR («gene expression»)) OR (transcription)) OR (transcriptome)) OR (snps)) OR («single nucleotid polymorphism»)) OR («copy number variations»)) OR (sequencing)) OR («ma seq»)) OR (gwas)) OR (biomarkers))))

#4 (ALL (cohort* OR «cohort stud*» OR «observational stud*» OR «case-control stud*» OR «prospective stud*» OR «Retrospective Stud*» OR «longitudinal stud*»))

#5 (KEY (mouse, AND mice, «Animal Model», «Animal Tissue», AND animal, AND non-human)) AND (LIMIT-TO (SUBJAREA, «MEDI») OR LIMIT-TO (SUBJAREA, »BIOC») OR LIMIT-TO (SUBJAREA, »PHAR») OR LIMIT-TO (SUBJAR-EA, «COMP»)) AND (LIMIT-TO (DOCTYPE, »ar») OR LIMIT-TO (DOCTYPE, »re»)) AND (LIMIT-TO (LANGUAGE, »English»))

#1 AND #2 AND #3 AND #4 AND NOT #5

Annexe B 2. Références des articles exclus après lecture du texte et raisons d'exclusions

Références	Raison d'exclusion
Wojnowski L, Kulle B, Schirmer M, Schlüter G, Schmidt A, Rosenberger A, et al. NAD(P)H oxidase and multidrug resistance protein genetic polymorphisms are associated with doxorubicin-induced cardiotoxicity. <i>Circulation</i> . 13 déc 2005 ;112(24) :3754-62.	Cancer chez l'adulte
Rajić V, Aplenc R, Debeljak M, Prestor VV, Karas-Kuzelicki N, Mlinaric-Rascan I, et al. Influence of the polymorphism in candidate genes on late cardiac damage in patients treated due to acute leukemia in childhood. <i>Leuk Lymphoma</i> . Oct 2009 ;50(10) :1693-8	Small number of Childhood Cancer Survivor
Dube S, Rajan A, Yalamanchili S, Thomas A, Abbott L, Benz P, et al. Single nucleotide polymorphisms in sarcomeric protein genes : Association with chemotherapy induced cardiomyopathy. <i>Internet Journal of Cardiology</i> . 2011 ;10(1).	Cancer chez l'adulte
Volkan-Salanci B, Aksoy H, Kiratli PÖ, Tülümen E, Güler N, Öksüzoglu B, et al. The relationship between changes in functional cardiac parameters following anthracycline therapy and carbonyl reductase 3 and glutathione S transferase Pi polymorphisms. <i>J Chemother</i> . Oct 2012 ;24(5) :285-91.	Pas de survivants d'un cancer de l'enfant
Cascales A, Pastor-Quirante F, Sánchez-Vega B, Luengo-Gil G, Corral J, Ortuño-Pacheco G, et al. Association of anthracycline-related cardiac histological lesions with NADPH oxidase functional polymorphisms. <i>Oncologist</i> . 2013 ;18(4) :446-53.	Autopsie
Schneider BP, Shen F, Gardner L, Radovich M, Li L, Miller KD, et al. Genome-Wide Association Study for Anthracycline-Induced Congestive Heart Failure. <i>Clin Cancer Res</i> . 1 janv 2017 ;23(1) :43-51.	Cancer chez l'adulte
Skitch A, Mital S, Mertens L, Liu P, Kantor P, Grosse-Wortmann L, et al. Novel approaches to the prediction, diagnosis and treatment of cardiac late effects in survivors of childhood cancer : a multi-centre observational study. <i>BMC Cancer</i> . 3 août 2017 ;17(1) :519.	Pas d'association génétique
Megias-Vericat JE, Montesinos P, Herrero MJ, Moscardó F, Bosó V, Rojas L, et al. Impact of NADPH oxidase functional polymorphisms in acute myeloid leukemia induction chemotherapy. <i>Pharmacogenomics J</i> . avr 2018 ;18(2) :301-7.	Cancer chez l'adulte
Ruiz-Pinto S, Pita G, Martín M, Alonso-Gordoa T, Barnes DR, Alonso MR, et al. Exome array analysis identifies ETFB as a novel susceptibility gene for anthracycline-induced cardiotoxicity in cancer patients. <i>Breast Cancer Res Treat</i> . 1 janv 2018 ;167(1) :249-56.	Cancer chez l'adulte
Garcia-Pavia P, Kim Y, Restrepo-Cordoba MA, Lunde IG, Wakimoto H, Smith AM, et al. Genetic Variants Associated With Cancer Therapy-Induced Cardiomyopathy. <i>Circulation</i> . 2 juill 2019 ;140(1) :31-41.	Nombre faible de patients
Goel S, Liu J, Guo H, Barry W, Bell R, Murray B, et al. Decline in Left Ventricular Ejection Fraction Following Anthracyclines Predicts Trastuzumab Cardiotoxicity. <i>JACC Heart Fail</i> . Sept 2019 ;7(9) :795-804.	Cancer chez l'adulte
Hellmann F, Völler S, Kirschke M, Jamieson D, André N, Bisogno G, et al. Genetic Polymorphisms Affecting Cardiac Biomarker Concentrations in Children with Cancer : an Analysis from the « European Paediatric Oncology Off-patents Medicines Consortium » (EPOC) Trial. <i>Eur J Drug Metab</i>	Cardiotoxicité en phase aiguë et en tant qu'effet tardif

Pharmacokinet. Juin 2020 ;45(3) :413-22.	
McOwan TN, Craig LA, Tripdayonis A, Karavendzas K, Cheung MM, Porrello ER, et al. Evaluating anthracycline cardiotoxicity associated single nucleotide polymorphisms in a paediatric cohort with early onset cardiomyopathy. Cardio-Oncology. 21 mai 2020 ;6(1) :5.	Nombre faible de patients
Norton N, Crook JE, Wang L, Olson JE, Kachergus JM, Serie DJ, et al. AssociabeforeGenetic Variants at TRPC6 With Chemotherapy-Related Heart Failure. Front Cardiovasc Med. 2020 ;7 :142.	Cancer chez l'adulte
Vaitiekus D, Muckiene G, Vaitiekiene A, Sereikaite L, Inciuraite R, Insodaite R, et al. HFE Gene Variants' Impact on Anthracycline-Based Chemotherapy-Induced Subclinical Cardiotoxicity. Cardiovasc Toxicol. Janv 2021 ;21(1) :59-66.	Cancer chez l'adulte
Loucks CM, Yan K, Tanoshima R, Ross CJD, Rassekh SR, Carleton BC. Pharmacogenetic testing to guide therapeutic decision-making and improve outcomes for children undergoing anthracycline-based chemotherapy. Basic Clin Pharmacol Toxicol. Janv 2022 ;130 Suppl 1 :95-9.	Étude de cas
Sapkota Y, Liu Q, Li N, Bhatt NS, Ehrhardt MJ, Wilson CL, et al. Contribution of Genome-Wide Polygenic Score to Risk of Coronary Artery Disease in Childhood Cancer Survivors. JACC CardioOncol. Juin 2022 ;4(2) :258-67.	Pas d'association génétique
Zolk O, von dem Knesebeck A, Graf N, Simon T, Hero B, Abdul-Khaliq H, et al. Cardiovascular Health Status And Genetic Risk In Survivors of Childhood Neuroblastoma and Nephroblastoma Treated With Doxorubicin : Protocol of the Pharmacogenetic Part of the LESS-Anthra Cross-Sectional Cohort Study. JMIR Res Protoc. 17 févr 2022 ;11(2) :e27898.	Étude non finalisée
Vargas-Neri JL, Carleton B, Ross CJ, Medeiros M, Castañeda-Hernández G, Clark P. Pharmacogenomic study of anthracycline-induced cardiotoxicity in Mexican pediatric patients. Pharmacogenomics. Avr 2022 ;23(5) :291-301.	Nombre faible de patients

Annexe B 3. Méthodes statistiques et procédure de sélection des 19 études incluses

Etude	Méthode de génotypage	Analyses statistiques	Variables d'appariement ou d'ajustement	Procédure de sélection	Correction pour multiplicité des tests	Évaluation de performance	Validation des résultats
Sági et al., 2018	Gène candidat	Régression logistique Régression logistique	Ajustement : âge au diagnostic, sexe, protocoles de chimiothérapie, groupes à risque et dose cumulée d'anthracycline.		Benjamini-Hochberg FDR, $p \leq 8,90 \times 10^{-3}$		
Blanco et al., 2012	Gène candidat	Régression logistique conditionnelle	Appariement : diagnostic de cancer, année du diagnostic (± 5 ans), race/ethnicité et durée du suivi Ajustement : dose cumulée d'ANT, sexe, âge au diagnostic du cancer		Correction de Bonferroni pour les comparaisons multiples ($p < 0,025$) lors de l'analyse de l'association avec les deux SNP.		
Visscher et al., 2013	Gène candidat	Régression logistique sur un unique SNP Régression logistique	Ajustement : La dose cumulée d'anthracycline, l'âge au traitement, le sexe et la radiothérapie du cœur sont inclus comme covariables. incluant de multiples variants génétiques et/ou variables cliniques ont été formés dans la cohorte originale	Régression par étapes avec sélection préalable : les associations avec cardiotoxicité induite par les anthracyclines à $p < 0,01$ ont été retenues dans le modèle final.		Calcul de la courbe ROC et de la statistique (aire sous la courbe, AUC)	Testé dans la cohorte de réplication (patients néerlandais-EKZ et canadiens-CPNDS combinés)
Singh et al., 2020	Gène candidat	Régression logistique conditionnelle	Appariement : diagnostic de cancer, année du diagnostic (± 5 ans), race/ethnie et durée du suivi Ajustement : dose d'anthracycline, irradiation thoracique, âge au moment du diagnostic du cancer primaire et sexe.				Analyse de l'expression génétique des cardiomyocytes et du sang périphérique

Blanco et al., 2008	Gène candidat	Régression logistique conditionnelle	<p>Appariement : âge au diagnostic (± 5 ans), race/ ethnie, dose d'anthracycline, irradiation du cœur (appariement pour 81 % des témoins ; ce facteur a été inclus dans tous les modèles pour tenir compte de l'appariement incomplet) et durée du suivi pour les témoins</p> <p>Ajustement : sexe, antécédents de tabagisme, première récurrence de la maladie initiale, cœur dans le faisceau d'irradiation et antécédents familiaux de cardiopathie.</p> <p>Appariement : âge au moment du diagnostic, sexe, dose d'anthracycline, irradiation thoracique et type de cancer.</p> <p>Ajustement : durée du suivi, âge au diagnostic, sexe, race, hypertension, dose d'anthracycline, irradiation thoracique et siège du cancer.</p>			
Hildebrandt et al., 2017	Gène candidat	Régression logistique	<p>Appariement : âge au diagnostic, sexe, dose d'anthracycline, irradiation thoracique et type de cancer.</p> <p>Ajustement : durée du suivi, âge au diagnostic, sexe, race, hypertension, dose d'anthracycline, irradiation thoracique et siège du cancer.</p>	Gène associé à la cardiotoxicité à $p \leq 0,05$	L'expression des gènes identifiés a été analysée dans les cardiomyocytes	
Wang et al., 2022	GWAS	Modèle de Cox	Ajustement : 10 composantes principales basées sur le génotype, l'âge au diagnostic du cancer primaire, le sexe, l'irradiation thoracique, la dose d'anthracycline et la présence de facteurs de risque cardiovasculaire (et les interactions SNP-anthracycline pour l'étape 2).	Gène associé à la cardiotoxicité à $p \leq 5 \times 10^{-8}$	Réplication sur un modèle cas-témoins apparié pour vérifier les interactions SNP et SNP-anthracycline significatives identifiées lors de la phase de découverte.	
Aminkeng et al., 2015	GWAS	Régression logistique	Ajustement : âge au début du traitement, dose cumulée d'anthracycline, type de tumeur et radiothérapie cardiaque.	Association avec l'ACT : $p < 1 \times 10^{-5}$ pour la cohorte de découverte et $p < 0,006$ pour la cohorte de réplication, et $p < 0,05$ dans la	Correction pour tests multiples $p < 1 \times 10^{-5}$ pour la	Les marqueurs qui ont atteint le seuil de signification dans la cohorte de découverte ont été testés pour la réplication (patients européens-néerlandais) et les SNP significatifs après correction de

				population non européenne indépendante.	cohorte de découverte ; Correction de Bonferroni $p < 0,006$ pour la cohorte de réplification.		Bonferroni dans la cohorte de réplification ont été testés dans la population non européenne, et dans l'analyse combinée de tous les patients d'ascendance européenne dans les cohortes de découverte et de réplification.
Visscher et al., 2015	GWAS	Régression logistique	Ajustement : dose cumulée d'anthracycline, âge au début du traitement, sexe, radiothérapie, et les deux premières composantes principales, et SLC28A3 (rs7853758) et UGT1A6 (rs17863783).	Analyse par paliers pour identifier les SNP associés à l'ACT à $p < 0,01$ dans la cohorte de découverte et qui restent associés dans la cohorte de réplification à $p < 0,05$.	Corrigé de Bonferroni $p < 1,7 \times 10^{-5}$	Calcul de la courbe ROC et de la statistique (aire sous la courbe, AUC)	Testé dans la cohorte de réplification (patients néerlandais-EKZ et canadiens-CPNDS combinés))
		Régression logistique	Incluant de multiples variants génétiques et/ou variables cliniques ont été formés dans la cohorte originale	Pour les SNP non identifiés dans l'analyse par paliers, une valeur $p < 0,005$ dans la cohorte combinée a été considérée comme une preuve suggestive.			
Wang et al., 2016	GWAS	Régression logistique conditionnelle	Appariement : diagnostic de cancer, année du diagnostic (± 5 ans), race/ethnie et durée du suivi Ajustement : dose cumulée d'ANT, exposition à la RT, sexe et âge au diagnostic (étape 1), et interactions SNP-anthracycline (étape 2).	Association avec la cardiomyopathie : $p < 0,004$	Correction des tests multiples : $p < 7,77 \times 10^{-5}$		Réplification sur un modèle de cas uniquement, afin de vérifier les interactions SNP-anthracycline significatives identifiées lors de la phase de découverte. Analyse de l'expression génétique sur des échantillons de cœur
Wang et al., 2014	GWAS	Régression logistique conditionnelle	Appariement : diagnostic de cancer, année du diagnostic (± 5 ans), race/ethnicité et durée du suivi Ajustement : dose cumulée d'ANT, dose totale de RT, sexe, âge au moment du cancer et interactions SNP-anthracycline.	Association avec la cardiomyopathie : $p < 5 \times 10^{-6}$	Correction pour tests multiples $p < 5 \times 10^{-6}$		Réplification sur un modèle de cas uniquement, afin de vérifier les interactions SNP-anthracycline significatives identifiées lors de la phase de découverte. Analyse de l'expression génétique sur des échantillons de cœur
Visscher et al., 2012	GWAS	Régression logistique	Ajustement : La dose cumulée, l'âge, le sexe et la radiothérapie du cœur sont inclus comme covariables ainsi que les deux premières composantes principales dans l'hypothèse d'un modèle additif.	Analyse par paliers d'un seul marqueur pour identifier les variants	Le seuil de signification global tenant compte des tests multiples a	Une analyse ROC a été réalisée pour	Le SNP le plus important a été évalué dans la cohorte Dutch-EKZ et dans les trois cohortes combinées. Les SNP ayant un $p < 0,01$ dans le test SNP

				associés à la cardiotoxicité induite par les anthracyclines dans la cohorte de découverte ($p < 0,01$) qui restent associées dans la cohorte de réplification canadienne ($p < 0,01$).	été calculé à $1,5 \times 10^{-4}$ à l'aide de la correction simpleM	chaque échantillon du modèle de régression multi-marqueurs.	unique pour la cohorte canadienne combinée ont été considérés pour une inclusion potentielle dans un modèle de prédiction du risque multimarqueurs.	
				En utilisant la régression logistique par étapes avec une sélection en amont incluant les covariables, les SNP ont été retenus avec $p < 0,01$ dans le modèle multi-SNP final.				
Chaix et al., 2020	WES	Burden CMC. SKAT. SKAT-O.	Ajustement : sexe, diagnostic de cancer, âge au début de l'anthracycline, utilisation de dexaméthosone, irradiation thoracique, durée du suivi depuis la première dose d'anthracycline, et les deux premières composantes principales déduisant l'appartenance ethnique.	Association de gènes avec la cardiotoxicité $p < 0,001$ par au moins 2 méthodes de regroupement de gènes et gènes dans des voies biologiquement pertinentes avec une association de $p < 0,001$ par au moins 1 méthode.	Correction de Bonferroni : $2,87 \times 10^{-6}$	de $p <$	Modèle de forêt aléatoire sur la cohorte de découverte. AUC ROC, sensibilité, spécificité, valeur prédictive positive, valeur prédictive négative, taux de faux positifs, taux de faux négatifs et calcul du taux de mauvaise classification.	Réplification sur un modèle cas-témoins 1:1 apparié selon la propension, comprenant un sous-ensemble de patients récemment inscrits à l'étude PCS2.

Krajinovic et al., 2016	Gène candidat	Analyse de variant ou test de Student		Soit une association a résisté à la correction des tests multiples (FDR $\leq 0,05$) avec au moins un des paramètres testés, soit l'association a été préférentiellement observée chez les patients à haut risque ayant reçu des doses cumulées de doxorubicine plus élevées.	Correction des tests multiples (FDR $\leq 0,05$)	Les résultats ont été validés sur un ensemble de 44 patients atteints de LAL et traités avec le protocole DFCI 95-01.
Semsei A et al., 2012	Gène candidat	Linear regression	Ajustement : Sexe, âge au moment du diagnostic, centre clinique, dose totale d'antracyclines, dexrazoxane, protocole de chimiothérapie.	Association des gènes avec la FRVG p< 0,005	Correction de Bonferroni p< 0,005	
Lipshultz S et al., 2013	Gène candidat	t-tests or Wilcoxon rank-sum tests	Les scores Z pour la masse du ventricule gauche, l'épaisseur de la paroi postérieure en fin de systole et en fin de diastole et les dimensions du ventricule gauche ont été ajustés en fonction de la surface corporelle, et les scores Z pour le raccourcissement fractionnel du ventricule gauche ont été ajustés en fonction de l'âge au moment de l'échocardiographie.	0.05		
Petrykey et al., 2021	WES	Statistique de Wald et régression linéaire pour variants communs	Ajustement : âge au moment du diagnostic, temps écoulé depuis la fin du traitement, sexe, protocole de traitement et scores risque-traitement.	Variants communs dans les gènes impliqués dans le fonctionnement du système cardiaque p< 0.0006	La procédure de Benjamini-Hochberg pour le FDR était p< 0,05 p-value de Bonferroni corrigée p < 0,0006	Analyses des variants rares les plus importants chez les survivants de la LAL d'origine européenne inclus dans la cohorte SJLIFE
Sapkota et al., 2022	WGS	Régression linéaire Régression logistique SKAT-O pour variants rares	Ajustement : sexe, âge au moment du diagnostic du cancer, âge au moment du dernier suivi, dose cumulée d'antracycline, dose moyenne d'irradiation du cœur et les 20 premières composantes principales basées sur le génotype.	P< 5 x 10 ⁻⁸ P<0.05 significatif p< 0.05	Ajustements de tests multiples avec une valeur seuil de FDR <0,05	Les variants qui ont atteint les seuils de signification statistique dans l'analyse de découverte ont été testées dans deux cohortes de réplification comprenant 301 survivants d'origine africaine de l'étude SJLIFE (pour la FEVG et la CCD) et 4020 survivants supplémentaires d'origine européenne de l'étude CCSS (pour les données autodéclarées sur la FEVG et la ICC).

Sapkota et al., 2021	WGS	Régression linéaire pour variants communs	Ajustement : sexe, âge au moment du diagnostic du cancer, âge au moment du dernier suivi, dose cumulée d'anthracycline, dose moyenne d'irradiation du cœur et les 10 premières composantes principales.	$P < 5 \times 10^{-8}$ significatif au niveau du génome entier	Les résultats ont été reproduits sur 1645 patients de l'étude SJLIFE, survivants d'ascendance européenne exposés à des thérapies cardiotoxiques (radiothérapie cardiaque et anthracyclines).
		SKAT-O pour variants rares et à fréquence faible		$p < 7,3 \times 10^{-8}$	
		Régression logistique pour les variants communs, de faible fréquence et rares	Appariement sur les déciles de PS. Calcul du PS : en régressant chaque SNP sur le sexe, l'âge au moment du diagnostic du cancer, l'âge au moment du dernier suivi, la dose cumulée d'anthracycline, la dose moyenne d'irradiation cardiaque et les 10 premières composantes principales.		

Abréviations : SNP, Single-nucleotide polymorphism ; ICC, Insuffisance Cardiaque Congestive CMC, Combined Multivariate and Collapsing ; SKAT, Sequence kernel association test ; SKAT-O, sequence kernel association test-optimized ; SJLIFE, St. Jude Lifetime ; DFCl, Dana-Farber Cancer Institute ; PCS2 : Preventing Cardiac Sequelae in Pediatric Cancer Survivors ; CCSS, Childhood Cancer survivor Study ; FDR, False discovery rate ou Taux de fausses découvertes ; PS, Score de propension ;

Annexe B 4. Liste des 61 gènes et 2 locus identifiés dans les 19 études incluses dans la revue systématique présentés dans l'ordre alphabétique

ABCA1 ;ABCC1 ; ABCC2 ; ABCC5 ; ABCC9 ; ABCG2 ; AKR1C3 ; ATG4C ; ATP2B1 ; ATP2B2 ; CABYR ; CBLB ; CBR1 ; CBR3 ; CDIP1 ; CELF4 ; CHAD ; CYP3A5 ; DYNAP ; ELAC2 ; FAM83C ; GJB7 ; GSTM1 ; HAS3 ; HFE ; JAGN1 ; KCNK17 ; KDM4B ; LRRFIP2 ; MCU ; NOD2 ; NOS1 ; NOS3 ; NOXO1 ; NQO1 ; NUGGC ; ODF3 ; OR1D2 ; PIK3R2 ; PLCE1 ; PSM1 ; RARG ; RGS3 ; ROBO2 ; SAC3D1 ; SEC62 ; SIM2 ; SIMC1 ; SLC22A17 ; SLC22A6 ; SLC22A7 ; SLC28A3 ; SULT2B1 ; TGM3 ; TTN ; UACA ; UGT1A6 ; USP42 ; NF267 ; ZNF71 ; ZNF827 ; Locus 1p13. 2 ; Locus 15q25. 3.

ANNEXE C. ANNEXES EN LIEN AVEC L'ÉTUDE DES BIOMARQUEURS ASSOCIÉS AUX PATHOLOGIES CARDIAQUES

Annexe C 1. Effets des traitements anticancéreux sur les pathologies cardiaques et l'insuffisance cardiaque dans l'étude cas-témoins appariée estimés à partir d'une régression logistique conditionnelle

Variables	n/N	Odds Ratio (95 % IC)	p-value
Pathologies cardiaques			
Année de diagnostic du cancer			0.8262
< 1970	17 / 46	1.00	
[1970-1979]	57 / 100	1.53 (0.60 – 3.87)	
[1980-1989]	67 / 132	1.42 (0.46 – 3.62)	
[1990-2000]	24 / 52	1.57 (0.51 – 4.85)	
Doses cumulées d'anthracyclines (mg/m²)			0.0008
Aucune	57 / 132	1.00	
]0 – 250]	32 / 77	1.03 (0.49 – 2.17)	
]250 – 360]	25 / 48	2.04 (0.88 – 4.76)	
>360	51 / 73	3.85 (1.74 – 8.52)	
Doses moyennes de radiations au cœur (Gy)			0.0037
None	51 / 107	1.00	
]0 – 5]	54 / 120	0.86 (0.45 – 1.63)	
]5 – 15]	10 / 24	0.74 (0.26 – 2.12)	
>15	50 / 79	2.81 (1.33 – 5.92)	
Insuffisance cardiaque			
Année de diagnostic du cancer[†]			0.6171
< 1980	33 / 68	1.00	
[1980-2000]	41 / 80	1.32 (0.44 – 4.00)	
Doses cumulées d'anthracyclines (mg/m²)[†]			< 0.0001
None	7 / 37	1.00	
]0 - 250]	13 / 33	9.40 (1.10 – 80.58)	
>250	54 / 78	40.27 (4.70 – 344.77)	
Doses moyennes de radiations au cœur (Gy)[†]			0.4086
None	27 / 56	1.00	
]0 – 15]	31 / 68	1.32 (0.44 – 3.94)	
>15	16 / 24	1.86 (0.63 – 12.96)	

Annexe C 2. Effet de l'expression des gènes *NFE2L2*, *TGIF2-RAB5IF*, *HK2*, *HDGFL3* et *LSP1* en fonction des doses cumulées d'anthracyclines et des doses moyennes de radiations au cœur (termes d'interaction) estimé à travers d'une régression logistique conditionnelle ajustée sur l'année de diagnostic du cancer pour l'évènement pathologies cardiaques. Les termes d'interaction sont testés un par un.

Gene	n/N	Odds Ratio (95 % IC)	p-value
Cumulative anthracyclines dose (mg/m²)			
LSP1			0,8315
Aucune	57 / 132	7,78 (3,17 - 19,09)	
1 - 250	32 / 77	9,17 (3,49 - 25,16)	
250 - 360	25 / 48	8,49 (3,35 - 21,86)	
> 360	51 / 73	8,60 (3,38 - 22,35)	
HDGFL3			0,8219
Aucune	57 / 132	0,54 (0,34 - 0,86)	
1 - 250	32 / 77	1,92 (1,03 - 3,62)	
250 - 360	25 / 48	1,72 (0,85 - 3,57)	
> 360	51 / 73	1,62 (0,86 - 3,08)	
NFE2L2			0,3288
Aucune	57 / 132	0,08 (0,03 - 0,24)	
1 - 250	32 / 77	3,23 (0,82 - 12,60)	
250 - 360	25 / 48	2,98 (0,76 - 11,71)	
> 360	51 / 73	2,16 (0,54 - 8,66)	
TGIF2 - RAB5IF			0,6357
Aucune	57 / 132	0,28 (0,15 - 0,52)	
1 - 250	32 / 77	1,39 (0,56 - 3,55)	
250 - 360	25 / 48	2,18 (0,85 - 5,69)	
> 360	51 / 73	1,36 (0,54 - 3,56)	
HK2			0,6037
Aucune	57 / 132	3,30 (1,92 - 5,68)	
1 - 250	32 / 77	4,42 (2,38 - 8,35)	
250 - 360	25 / 48	3,94 (2,15 - 7,45)	
> 360	51 / 73	3,99 (2,23 - 7,18)	
Doses moyennes de radiations au cœur (Gy)			
LSP1 - MIR7847			0,5521
Aucune	51 / 107	6,19 (2,63 - 14,57)	
0 - 5	54 / 120	7,21 (2,98 - 17,53)	
> 5 [†]	60 / 103	8,06 (3,15 - 21,24)	
HDGFL3			0,9576
Aucune	51 / 107	0,61 (0,40 - 0,93)	
0 - 5	54 / 120	1,55 (0,93 - 2,62)	
5 - 15	10 / 24	1,69 (0,77 - 4,13)	
> 15	50 / 79	1,46 (0,83 - 2,58)	
NFE2L2			0,6933
Aucune	51 / 107	0,22 (0,10 - 0,47)	
0 - 5	54 / 120	0,80 (0,28 - 2,35)	

5 – 15	10 / 24	0,46 (0,11 – 7,16)	
>15	50 / 79	1,06 (0,37 – 3,09)	
TGIF2 - RAB5IF			0,4234
Aucune	51 / 107	0,26 (0,13 – 0,52)	
0 – 5	54 / 120	1,86 (0,79 – 4,36)	
5 – 15	10 / 24	0,50 (0,14 – 6,51)	
>15	50 / 79	1,51 (0,62 – 3,72)	
HK2			0,8768
Aucune	51 / 107	2,59 (1,53 – 4,38)	
1 – 5	54 / 120	3,74 (2,11 – 6,66)	
5 – 15	10 / 24	3,37 (1,78 – 6,91)	
> 15	50 / 79	3,76 (2,07 – 6,92)	

Annexe C 3. Effet de l'expression du gène NFE2L2 sur les pathologies cardiaques estimé à partir d'une régression logistique conditionnelle ajustée sur l'année de diagnostic du cancer et l'exposition à l'anthracycline et la métrique dose-volume de la radiothérapie (V15)

Model	-2 log vraisemblance	ddl	AUC (95 % IC)	AUC corrigé^T (95 % IC)
Modèle Nul (sans covariables)	228.7			
Modèle 1: Année de diagnostic du cancer + variables de traitement	197.7	9	0.67 (0.63 – 0.74)	0.66 (0.62 – 0.69)
Modèle 2: Année de diagnostic du cancer + NFE2L2	110.9	4	0.84 (0.79 – 0.89)	0.83 (0.79 – 0.86)
Modèle 3: Année de diagnostic du cancer + variables de traitement + NFE2L2	102.9	10	0.85 (0.81 – 0.89)	0.83 (0.80 – 0.87)

Annexe C 4. Caractéristiques démographiques et cliniques des patients inclus dans l'étude de contrôle des cas appariés pour l'évènement insuffisance cardiaque (n=148)

	Cas (n=74)	Témoins (n=74)
Genre, n (%)*		
Féminin	32 (43,2)	32 (43,2)
Masculin	52 (56,8)	52 (56,8)
Âge au diagnostic du premier cancer (années), n (%)*		
Médiane (EI)	4 (1 – 7)	4 (2 - 7)
[0 – 5[44 (59,5)	45 (60,8)
[5 – 10[21 (28,4)	21 (28,4)
[10 – 15[7 (9,5)	6 (8,1)
[15 – 22[2 (2,7)	2 (2,7)
Type de tumeur, n (%)*		
Lymphome	17 (23,0)	17 (23,0)
Neuroblastome	14 (18,9)	14 (18,9)
Sarcome osseux	14 (18,9)	14 (18,9)
Néphroblastome	10 (13,5)	10 (13,5)
Sarcome des tissus mous	11 (14,9)	11 (14,9)
Tumeur gonadique	3 (4,1)	3 (4,1)
Autres	5 (6,8)	5 (6,8)
Suivi après diagnostic du cancer (années) , n (%)*		
Médiane (EI)	18 (12 – 23)	35 (25 – 42)
[0 – 10[*	15 (20,3)	0 (0,0)
[10 – 20[26 (35,1)	5 (6,8)
[20 – 30[24 (32,4)	20 (27,0)
[30 – 40[9 (12,2)	26 (35,1)
[40 – 50[0 (0,0)	14 (18,9)
≥ 50	0 (0,0)	9 (12,2)
Année de diagnostic du premier cancer, n (%)		
Avant 1970	0 (0,0)	15 (20,3)
[1970, 1979[33 (44,6)	20 (27,0)
[1980, 1989]	33 (44,6)	24 (32,4)
[1990, 2000]	8 (10,8)	15 (20,3)
Anthracyclines, n (%)		
Non	7 (9,5)	30 (40,5)
Oui	67 (90,5)	44 (59,5)
Doses cumulées d'anthracyclines (mg/m²), n (%)		
Médiane (min - max)	362,3 (106,2-	280,4 (84,4-546,1)
Aucune	6 (8,1)	29 (39,2)
]0 – 250]	14 (18,9)	21 (28,4)
]250 – 360]	20 (27,0)	15 (20,3)
>360	34 (46,0)	9 (12,2)
Radiations au cœur, n (%)		
Non	27 (36,5)	29 (39,2)
Oui	47 (63,5)	45 (60,8)
Doses moyennes de radiations au cœur (Gy), n (%)		
Médiane (min - max)	0,2 (0,0 – 31,7)	0,3 (0,0 – 41,6)

Aucune	27 (36,5)	29 (39,2)
]0 – 5]	28 (37,8)	32 (43,2)
]5 – 15]	3 (4,1)	5 (6,8)
>15	16 (21,6)	8 (10,8)

PUBLICATIONS

Influence of Genetic Biomarkers on Cardiac Diseases in Childhood Cancer Survivors: a Systematic Review

Naïla Aba ^{1,2}, Claire Ducos ^{1,2}, Eric Morel ³, Chiraz El Fayech ^{2,4}, Brice Fresneau ^{1,2,4,5}, Florent de Vathaire ^{1,2,5}, Gwénaél Le Teuff* ^{6,7}, Nadia Haddy* ^{1,2,5}

¹ Université Paris-Saclay, UVSQ, Inserm, CESP, 94807, Villejuif, France

² CESP, INSERM U 1018, Cancer and Radiation team, Gustave Roussy, Villejuif, France

³ Inserm, U769, Univ. Paris-Sud 11, IFR141, Labex Lermite, Châtenay-Malabry, France

⁴ Department of Childhood and Adolescent Oncology, Gustave Roussy, Villejuif, France

⁵ Department of Research, Gustave Roussy, Villejuif, France

⁶ Bureau de Biostatistique et d'Epidémiologie, Gustave Roussy, Villejuif, France

⁷ Université Paris-Saclay, CESP, INSERM U1018 Oncostat, labeled Ligue Contre le Cancer, Villejuif, France.

*These authors contributed equally to this work

Corresponding author: Naïla Aba

Radiation Epidemiology Group / CESP - Unit 1018 INSERM, Gustave Roussy, B2M, 114, rue Édouard Vaillant 94805 Villejuif Cedex, E-mail: nailamyriam.aba@gustaveroussy.fr

Funding support and author disclosures: Naïla Aba received a doctoral grant from the Paris Sud-Paris Saclay University. This work has been supported by the Fondation ARC pour la recherche sur le cancer [ARCDOC42023010005942].

The authors have reported that they have no relationships relevant to the contents of this paper to disclose.

Conflict of interest: The authors declare no conflict of interest.

Abstract

Childhood cancer survivors (CCS) often suffer from cardiac disease (CD) after treatment that included anthracycline and radiotherapy involving the heart. However, the variability in CD occurrence cannot be explained solely by these treatments, suggesting the existence of genetic predisposition. We conducted a systematic review searching on Medline-PubMed and Scopus, to identify studies reporting associations between genetic factors and CD in CCS. We included studies published up to 11 April 2023, with no lower limit, and assessed the quality of genetic associations by the Q genie tool. As a result, 19 studies were included (14 case-control and five cohorts), revealing several genes and variants associated with cardiomyopathy, among which, *SLC28A3*-rs7853758 and *RARG*-rs2229774 variants gave the most consistent findings. This review highlights the necessity to establish a set of clinically useful genes and variants to identify patients most at risk of developing cardiomyopathy, and to implement monitoring and prevention strategies.

Keywords: Childhood Cancer Survivors, Anthracyclines, Radiotherapy, Cardiomyopathy, Genomics, Genetic association, Systematic Review

1. Introduction

Five-year survival after childhood cancer treatment has improved significantly over the past five decades (1), with a rate exceeding 80% in the 2000s (2). Unfortunately, these therapeutic advances are associated with multiple secondary diseases at a relatively young age in childhood cancer survivors (CCS) (2,3). Cardiac diseases (CD) are the major non-malignant cause of death in this population (2–7), and CD risk increases in CCS treated with anthracyclines (8) and radiotherapy (RT) involving the heart (4,8,9). These CD can manifest as cardiomyopathy and heart failure (10–12), but also as coronary artery disease, valvular disease, arrhythmias and conduction disorders, and pericardial disease (4,10).

The mechanisms underlying radiation and anthracycline treatment-induced cardiotoxicities are multifactorial and include mitochondrial dysfunction, oxidative stress, apoptosis, and dysregulation of autophagy (13,14). Several other risk factors associated with CD have been identified, such as a younger age of exposure to cancer treatment, female sex, comorbidities, and African ethnicity (15). However, the ability of models including these factors to predict patients at risk of cardiotoxicity remains limited (6,16). Indeed, variability in the occurrence of cardiomyopathy has been observed in similarly treated CCS, suggesting the existence of inter-individual predispositions (3). Some studies have investigated the role of genetic factors mainly on cardiomyopathy occurrence after treatment of childhood cancer with anthracyclines and/or RT involving the heart. These studies used different sequencing approaches, such as candidate genes, focusing on one or several specific variants, genome-wide association studies (GWAS), whole-exome sequencing (WES), and whole-genome sequencing (WGS), and aim to identify genes and variants associated with cancer treatment-related cardiomyopathy.

Identifying genetic biomarkers is essential to improve the prediction of CD risk after childhood cancer treatment. This improvement could target patients at higher CD risk and intensify their cardiac monitoring. Moreover, this enhancement could also allow the identification of high-risk individuals as soon as childhood cancer is diagnosed, in order to prevent cardiotoxicity by adapting therapies (3,17).

26 We performed a systematic review to summarize the published findings on the genetic factors
27 associated with CD in CCS and report the potential interactions of genes/variants with cancer
28 treatments.

29 2. Materials and Methods

30 2.1. Search strategy and selection criteria

31 This systematic review was registered on PROSPERO CRD N°2023388689. Literature research was
32 performed until April 2023 with no lower limit to identify studies reporting genetic associations
33 between genes/variants and CD in CCS, according to the PICOS (Population, Intervention, Comparison,
34 Outcome, Study design) framework: Population – patients treated for childhood cancer before 21
35 years; Intervention – patients treated with chemotherapy and/or radiotherapy; Comparator – not
36 applicable for this review; Outcomes – papers that included genetic data involved in CD risk in patients
37 treated for childhood cancer; Study – case-control and cohort studies (**Appendix Table A**).

38 The first part of the search strategy involved searching in the PubMed/MEDLINE and Scopus electronic
39 databases. The second part is a manual search from references cited in the eligible studies identified
40 in the first part to include omitted studies. Details of the complete search strategy are provided in
41 **Appendix A**.

42 Two independent reviewers (NA and CD) screened the titles and abstracts to determine whether a full-
43 text review should be performed based on inclusion criteria (**Appendix B**). One investigator (NA)
44 screened full-text articles to assess compliance with the selection criteria. If there was any uncertainty
45 about the inclusion of an article, it was reviewed by the second investigator (CD), and the decision was
46 made by mutual agreement.

47 2.2. Data collection, quality control, and risk bias evaluation

48 Information from eligible studies was recorded from the main text and supplementary material using
49 a data extraction form (**Appendix Table C**) comprising 14 items relating to study design, sequencing

50 method, sample size, country of origin and/or ethnicity, cancer type, methods and duration of follow-
51 up, cancer treatment information, cancer treatment or cancer diagnosis area, age at cancer diagnosis,
52 cardiac outcome definition, case and controls definition, gene selection method, statistical analysis,
53 selected genes and/or variants and their association risks with CD.

54 Study quality was evaluated using the Q-Genie tool for quality assessment of genetic association
55 studies (18), which consists of a total of 11 characteristics rated from "Poor" (score=1) to "Excellent"
56 (score=7). The quality score is the sum of these ratings, with scores ≤ 35 indicating poor quality study,
57 >35 and ≤ 45 moderate quality study, and >45 good quality study (18).

58 Data reporting followed Preferred Reporting Items for Systematic Reviews and Meta-Analysis
59 (PRISMA) (19) guidelines.

60 3. Results

61 3.1. Study, patient data description

62 The search strategy identified 978 studies, of which 56 were duplicates. Of the remaining 922 studies,
63 884 records did not meet inclusion criteria after screening titles and abstracts (main exclusion criteria:
64 study out of scope, no genetic findings), leading to 38 studies. Finally, 19 studies were selected on the
65 full-text screening (20–38) (**Appendix Table D1**) and the 19 other studies were excluded mainly
66 because they were related to adult patients (n=8) or a small number of patients (n=4) (**Appendix Table**
67 **D2**). The identification and screening process results are detailed in a PRISMA flow diagram (**Figure 1**).
68 The 19 included studies were classified as moderate (n=4), and good quality (n=15) according to the
69 Q-Genie tool (**Appendix Table E**).

70 **Table 1** reports the main characteristics of the 19 included studies published between 2008 and 2022.
71 These studies involved patients diagnosed before the age of 21 with various types of malignant tumors:
72 acute lymphocytic leukemia (ALL, n=4) (33–35,37), ALL and osteosarcoma (n=1) (31), and unspecified
73 tumors (n=14) (20–30,32,36,38). All studies reported administration of anthracyclines to patients, and

74 14 studies (74%) reported treatment with RT involving the heart. For the former, cumulative doses of
75 anthracyclines were specified for each patient. For the latter, the exposure was mainly categorized as
76 "received thoracic radiation: yes/no" (20–23,25,27,28,30,32,38), while five studies considered the
77 heart radiation doses in the analyses, using metrics such as cumulative dose (24,26) and mean dose
78 (29,36). Five studies reported the use of Dexrazoxane, a medication used in the management and
79 treatment of anthracycline-induced cardiotoxicity (30,33–35,37) (Details in **Supplementary Table1**).

80 3.2. Genetic sequencing

81 Genetic sequencing was performed on various biological specimens taken from peripheral blood (n=9)
82 (27,29–32,34–37), buccal cells/saliva (n=1) (24), and blood or buccal cells/saliva (n=6) (21,22,25–
83 27,38). Three studies did not specify the biological sample used for the genotyping (20,23,33). Different
84 genotyping methods were used: candidate gene analyses (n=9) (21,24,25,28,31–35), GWAS (n=6)
85 (20,22,23,26,27,38), WGS (n=2) (29,36), and WES (n=2) (30,37) (Details in **Supplementary Table 2**).

86 3.3. Cardiac disease outcomes

87 Although no restriction in the search strategy was performed about the type of CD outcome, all the
88 included studies investigated cardiomyopathy defined either by diagnosis or by monitoring
89 echocardiographic parameters. Cardiomyopathy diagnosis was designated by the authors using
90 different denominations: cardiomyopathy (n=4), cancer treatments induced cardiomyopathy (n=1),
91 anthracycline-induced cardiotoxicity (n=4), or clinical cardiotoxicity (n=4). Some authors used
92 congestive heart failure (CHF) as the outcome (n=1), which is a consequence of cardiomyopathy. The
93 monitoring of echocardiographic parameters revealing a decrease in LVEF and/ or in left ventricular
94 fractional shortening (LVFS) (n=5) constituted another outcome (29,33–35,37). For simplicity, the
95 wording cardiomyopathy is used in the rest of the paper to cover the different denominations reported
96 above (Details of outcomes per study are given in **Supplementary Table 3**).

97 3.4. Design and statistical methods

98 Among the 19 selected studies, 14 were case-control studies (20–28,30–32,38), including six matched
99 studies, with case numbers ranging from 38 to 277. The remaining five were cohort studies (29,33–
100 37), with patient numbers ranging from 184 to 1870. Sixty-three percent (n=12) of the studies
101 replicated their findings using independent data from an external cohort (n=4) (20–23) or from patients
102 of the overall cohort not included in the discovery set (26,27,29,30,33,36–38). Only eight studies
103 investigated the effect of variants according to cumulative anthracycline doses. More details on the
104 statistical methods used are reported in **Supplementary Table 4**.

105 3.5. Summary of genetic findings

106 Regarding the heterogeneity in outcomes and sequencing methods in the 19 included studies, we
107 reported the results of genetic associations according to study design (case-control and cohort), which
108 is strongly related to study outcome and the sequencing methods. **Table 2** reports the main results,
109 and **Table 3** summarizes the effect of variants according to cumulative doses of anthracyclines.

110 3.5.1. Case-control studies (n=14)

111 Among the 14 case-control studies, six used a gene candidate approach, six a GWAS, and one a WES
112 approach.

113 a) Candidate gene

114 The number of variants tested ranges from two to 70 per study. In survivors of ALL or osteosarcoma,
115 carrying the *ABCC2*-rs3740066-GG genotype or the *SLC22A6*-rs6591722-AA genotype increased the
116 risk of cardiomyopathy 5 to 10 years after cancer diagnosis compared to other genotypes ($p=7.11 \times 10^{-4}$,
117 $p=1.71 \times 10^{-3}$, respectively). A similar association was observed in patients carrying the T allele in the
118 *NQO1*-rs1043470 variant ($p=5.82 \times 10^{-3}$) (31). The *CYP3A5*-rs4646450-TT and the *SLC28A3*-rs7853758-
119 AA genotypes were associated respectively with a 6.6-fold (95% CI: 1.7–25.7, $p=7.00 \times 10^{-3}$) and 9.8-
120 fold (95% CI: 1.7 - 56.0, $p=0.01$) increased risk of cardiomyopathy (31).

121 In survivors treated for heterogeneous types of cancers, the *CBR3*-rs1056892-GG variant genotype
122 showed an increased risk of developing cardiomyopathy (OR=1.79, 95% CI: 1.08 - 2.96), p=0.02) (25),
123 and also showed an interaction effect with exposition to low-to-moderate doses of anthracycline (1 to
124 250 mg/m²), increasing by 3.3-fold increased risk of cardiomyopathy compared to GA and AA
125 genotypes (25). For both *CBR3*-rs1056892 and *CBR1*-rs9024 variants, carrying the GG genotype along
126 with receiving low-to-moderate doses of anthracycline increased the risk of cardiomyopathy by 3.6-
127 fold compared to AA and GA genotypes (25). Carrying the A allele in the *SLC28A3*- rs7853758 variant
128 was associated with a lower risk of cardiomyopathy (OR=0.36, 95% CI: 0.22–0.60) (21), and the variant
129 *UGT1A6*-rs17863783 was significantly associated with a 6.2-fold (95% CI: 2.5-15.4; p=1.1 x 10⁻⁴)
130 increased risk of severe cardiomyopathy (21). The *GSTM1* null variant was more prevalent among cases
131 with cardiomyopathy than controls (60.0% vs. 38.0%, p=0.005): individuals with *GSTM1* null genotype
132 had 2.7-times greater risk of developing cardiomyopathy (p=0.007) compared to those with *GSTM1*
133 positive genotype (28). The analysis of *NQO1*-rs1800566 (*NQO1**2 polymorphism) and *CBR3*-
134 rs1056892 (*V244M* polymorphism) variants revealed that they were not associated with the
135 development of CHF (24). The *PLCE1*-rs932764-G and *ATP2B1*-rs17249754-G variants, known for their
136 role in hypertension, were significantly associated with a decreasing risk of cardiomyopathy (p=0.006
137 and 0.04, respectively) (32).

138 b) GWAS

139 Patients carrying *ROBO2*-rs17736312-AA genotype and exposed to anthracyclines >250 mg/m²
140 demonstrated a 2.2-fold (95% CI: 1.2–4.0, p=0.009) higher risk of CHF compared to patients carrying
141 at least one G allele and exposed to less than 250 mg/m² of anthracyclines (38). However, in the same
142 study, no variant reached the threshold for genome-wide significance in their association with
143 cardiomyopathy (38). In the *RARG*-rs2229774 variant, carrying at least one A allele was associated with
144 an increased risk of cardiomyopathy following anthracycline exposition compared to the GG genotype
145 (p=5.9 × 10⁻¹⁰) (23). Variants *SLC22A17*-rs4982753 and *SLC22A7*-rs4149178 were associated with

146 decreasing risk of cardiomyopathy in Canadian CCS ($p=0.0078$ and 0.0034 , respectively) and were
147 validated in a replication cohort ($p = 0.0071$ and 0.047 , respectively). Additionally, the variant *SULT2B1*-
148 rs10426628 was associated with a 1.9-fold increased risk of cardiomyopathy in the combined cohort
149 (95% CI: 1.2–2.5, $p=3.2 \times 10^{-4}$) (22). The *CELF4*-rs1786814-GG genotype was associated with an
150 increased risk of cardiomyopathy compared to GA and AA genotypes ($p=0.006$), and when exposed to
151 doses lower or equal to 300 mg/m^2 , patients carrying GG genotype were at 5.8-fold higher risk of
152 developing cardiomyopathy compared to GA and AA genotypes (27). In addition, exposition to
153 anthracycline doses $>250 \text{ mg/m}^2$ and carrying at least one A allele in the *HAS3*-rs2232228 variant were
154 at increased risk of cardiomyopathy compared to the GG genotype (AA: $p=0.003$ and AG: $p=0.02$) (26).
155 Moreover, carrying the A allele in the *SLC28A3*-rs7853758 variant was associated with a decreased risk
156 effect on cardiomyopathy (OR=0.31, 95% CI: 0.16–0.60) (20).

157 c) WES

158 Twenty-eight genes carrying at least one rare to low-frequency variant were identified as associated
159 with protective effects against cardiomyopathy. Among the controls, 89.6% harbored rare-to-low-
160 frequency variants, while only 42.6% of patients with cardiomyopathy were carriers. The *ZNF827* and
161 the *PI3KR2* genes showed a more significant variant burden in controls compared to patients
162 experiencing cardiomyopathy (30).

163 3.5.2. Cohort studies

164 Among the five case-control studies, three used a gene candidate approach, two a WGS, and one a
165 WES. The outcomes for these five studies were LVEF or LVFS; four studies investigated cohorts of ALL
166 survivors, and only one worked on a cohort of CCS treated for heterogeneous first malignancies.

167 a) Candidate gene

168 In ALL survivors, the *ABCC5*-rs7627754-TT genotype was associated with a 12% reduction LVEF and 8%
169 reduction of LVFS. However, prior to cancer treatment exposure, carriers of the TT genotype already
170 had lower LVEF and LVFS values compared to other genotypes (33). A protective effect of the *NOS3*-

171 rs1799983 variant was observed in patients carrying the TT genotype, who presented a higher LVEF
172 than individuals with at least one G allele (61.4% vs 59%, $p=0.02$). This protective effect was more
173 significant in high-risk patients who did not receive dexrazoxane before doxorubicin (64.1% vs. 56.8%,
174 $p=0.002$) but was not observed in patients who received that treatment (33). Carrying TT genotypes
175 for the *ABCC1*-rs3743527 variant was associated with a reduced LVFS at the end of treatment (34.0%),
176 compared to CC (39.5%) or CT (39.3%) genotypes ($p=0.001$). In addition, patients with at least one T
177 allele in the *ABCC1*-rs246221 variant showed a lower LVFS than patients with homozygous C alleles
178 (Allele frequency: TC: 38.4%, TT: 38.5%, CC: 40.7%, $p=0.027$) (34). After anthracycline administration,
179 the mean LVFS was significantly lower in patients with a combination of *ABCC1*-rs3743527-TT and
180 *ABCC1*-rs246221-TC or TT genotypes compared to patients carrying other genotypes ($p=0.001$) (34).
181 Compared to a population of healthy children, high-risk ALL patients with the C282Y and/or H63D
182 allelic variants in the *HFE* gene had significantly lower LV function, mass, and wall thickness two years
183 after their diagnosis (35).

184 b) WES

185 Common variants analysis revealed that the genotypes *TTN*-rs72648998-CC and *TTN*-rs3829747-CC
186 were associated with an increased LVFS ($p=0.0002$ and 0.0004 , respectively) and LVEF ($p=0.0003$ and
187 0.0003 , respectively), and that *TTN*-rs2303838-CC genotype was associated with an increased LVEF
188 ($p=0.0005$) (37). Rare variants *NOS1*-rs76090928, *ABCG2*-rs199473672, *ABCG2*-rs45605536, *AKR1C3*-
189 rs200981816, and *AKR1C3*-rs34186955 were associated with a protective effect on LVEF, and a
190 threatening effect was observed for the *CBR1*-rs2835266 variant (37). In the high-ALL risk group, the
191 *ABCC5* gene had a protective effect on LVEF and LVFS ($p=0.0008$ and $p=0.0014$, respectively). In
192 addition, in survivors treated according to the high-ALL risk protocols, the rare variant enrichment in
193 the *ABCC5* gene was associated with a protective effect on LVEF and LVFS (37). An association between
194 LVFS and rare variants enrichment in the *NOD2* and *ZNF267* genes ($p=1.39 \times 10^{-6}$ and 3.51×10^{-6} ,

195 respectively) was detected. An association was also observed between rare variant enrichment in
196 *NOD2* and LVEF ($p=2 \times 10^{-6}$) (37).

197 a) WGS

198 The *KCNK17*-rs2815063-A allele was associated with a reduced LVEF in patients diagnosed with a
199 cardiomyopathy grade ≥ 3 after adjusting for cardiac medication ($p=7.5 \times 10^{-6}$). After classification into
200 therapy-based risk groups, the association between *KCNK17*-rs2815063-A and LVEF reduction was
201 stronger in the high and moderate cardiomyopathy risk groups ($p=9.6 \times 10^{-4}$ and $p=1.4 \times 10^{-4}$,
202 respectively), compared to the low-risk group ($p=0.002$). This association was also replicated in
203 individuals with African ancestry ($p=0.004$) (36). In a secondary case-control analysis, the same author
204 observed a 1.4-fold increased risk of CHF (95% CI: 1.02–1.92, $p=0.004$) in individuals carrying the
205 *KCNK17*-rs2815063 variant with the A allele. This polymorphism also revealed an elevated risk of
206 cardiac events with OR=1.6 (95% CI: 1.19–2.09) in grade ≥ 2 cardiomyopathy and OR=1.8 (95% CI: 1.17–
207 2.77) in grade ≥ 3 cardiomyopathy (36).

208 In CCS with African ancestry, 31 sets of rare/ low-frequency variants and two loci (1p13.2 and 15q25.3)
209 in the common variant analysis (rs6689879 and rs9788776, respectively) were significantly associated
210 with LVEF (29). The rs6689879-C and rs9788776-G alleles were significantly associated with a reduced
211 LVEF ($p=2.8 \times 10^{-8}$ and $p=3.5 \times 10^{-8}$, respectively), and also showed a reduction effect in patients
212 exposed to anthracyclines only. However, in patients exposed to RT only, the rs9788776-G allele
213 showed a reduction of 8.1% of the LVEF ($p=5.4 \times 10^{-4}$) (29). These two variants were not associated
214 with LVEF in patients not exposed to cardiotoxic treatments (anthracyclines and chest-RT).

215 4. Discussion

216 This systematic review identified 19 studies investigating the associations between genetics and
217 cardiomyopathy in CCS, fourteen of which (76%) conducted case-controls, while five (26%) conducted
218 cohort studies. Several genetic variants in 63 genes were identified as significantly associated with
219 cardiomyopathy, using diverse sequencing methods. Among these variants, *SLC28A3*-rs7853758 and

220 *RARG*-rs2229774 represent the most consistent findings, showing strong and significant associations
221 with cardiomyopathy following childhood cancer treatment, including anthracycline; these
222 associations have been successfully replicated in patients from external cohorts (20,21,23). Our review
223 also identified numerous other genes and variants that require further validation, and highlighted
224 interactions between genes/variants and cumulative doses of anthracyclines, as well as ethnic and
225 gender-specific genes/variants. However, too few studies have investigated the excess risk of RT doses
226 to the heart on the development of CD in CCS, and furthermore, no study has investigated the
227 existence of interactions between identified genes/variants and RT.

228 The different genes/variants identified in this systematic review were implicated in various biological
229 mechanisms, first among them, the genes involved in drug transport, accumulation, and/or
230 elimination. Different variants in the Solute carrier (SLC) transporters: *SLC28A3*, *SLC22A17*, *SLC22A7*,
231 and *SLC22A6* were identified as associated with cardiomyopathy in CCS. In particular, the rs7853758
232 variant in the *SLC28A3* gene, which contributes to anthracyclines transport across cell membranes and
233 their elimination (39,40). A candidate gene study on early breast cancer patients that focused on the
234 *SLC28A3* gene found no cardiomyopathy association after anthracycline treatment (41), suggesting a
235 potential specificity of this gene to CCS. An evidence association with an increased risk of
236 cardiomyopathy has also been observed for the *SULT2B1* gene, which contributes to the elimination
237 of drugs such as anthracyclines and their toxic metabolites (42). The rs17863783 variant in the *UGT1A6*
238 gene was also reported to be associated with a greater risk of cardiomyopathy. This variant encodes
239 for an enzyme that eliminates various drugs, and is known to impair the ability of this enzyme to
240 effectively eliminate anthracyclines (14). Also a cardiotoxicity effect was observed for genes from the
241 ABC transporters group of genes, specifically subfamilies B and C, which transport xenobiotics like
242 anthracyclines, and whose inhibition disrupts the distribution and metabolism of antineoplastic
243 agents, leading to increased plasma and tissue concentrations of anthracyclines (43).

244 Protein and enzyme regulation is also an important mechanism discussed in this review, including the
245 *RARG* gene and the *CELF4* gene. The former is the second most promising finding in this review; this
246 gene is involved in the transcriptional regulation of an enzyme (Topoisomerase-2 β) that interacts with
247 doxorubicin in cardiac tissues, leading to cardiomyocyte apoptosis and heart failure in mice (14). The
248 latter regulates the mRNA splicing of a protein (Troponin 2) that enables cardiac muscle contraction
249 (14).

250 An important biological mechanism reported in the included studies is oxidative stress induced by
251 anthracyclines, resulting in an increased formation of reactive oxygen species (ROS), leading to a
252 reduction of cardiac contractility, tumor cell apoptosis, and cardiomyocyte damage (44–47). Different
253 genes involved in this mechanism were reported, such as the *PLCE1* gene, which encodes for a
254 cardioprotective protein known for its effect against oxidative stress; *NQO1*, which participates in the
255 anthracyclines reduction and prevents the formation of ROS (45); *NOS1* and *NOS3*, which catalyzes the
256 nitric oxide production that contributes to anthracycline cardiac damage and the production of ROS
257 (14,45); or *HAS3*, which encodes for the synthesis of an extracellular matrix known to capture, and
258 thus reduce, ROS (48). Also, the ROS generation induces the expression of *GSTM1* to cancel cellular
259 oxidative stress caused by anthracycline exposure (45,47), which explains how the complete deletion
260 of the *GSTM1* gene (*GSTM1* null variant) is associated with cardiomyopathy after anthracycline
261 treatment.

262 Other genes, not related to any of the biological mechanisms mentioned above, were also identified
263 in this review, like the *HFE* gene involved in iron metabolism (14) and the *TTN* gene, whose protective
264 effect contradicts a priori knowledge on its role in heart disease (49). In addition, two ethnically specific
265 variants were also identified in individuals with African ancestry - rs9788776 and rs6689879 - with
266 findings not replicated in the cohort of patients with European ancestry because of the monomorphic
267 nature of rs9788776 in this population and of the insignificant association of rs6689879 in these
268 patients (29).

269 One previous review (17) investigated genetic predispositions to anthracycline-induced cardiotoxicity,
270 but was not specific to pediatric cancers which are histopathologically and molecularly different from
271 adult cancers. It is, therefore, imperative to focus on CCS in order to build a complete list of genes
272 involved in pediatric malignancies and cancer-treatment late effects, and to develop more effective
273 and less toxic treatments (50,51). In this sense, Aminkeng, et al. (52) proposed recommendations for
274 using pharmacogenomic testing to reduce cardiomyopathy incidence in patients treated for cancer,
275 but did not specifically address CCS. The review of Berkman et al., (18) was focused on anthracycline-
276 induced cardiotoxicity, and did not address the effects of RT involving the heart. Although in our review
277 we sought to introduce the impact of RT involving the heart in CD development, no study has looked
278 for interactions between genes identified and this cancer treatment even though the risks of CD
279 associated with RT involving the heart are now well known. Moreover, research data on this topic are
280 now available with the help of tools that estimate the dose-volume distribution of radiation in the
281 whole heart and the substructures of the heart (4,9,53). However, recent findings have shown that
282 their mechanisms are not involved in the same genetic pathways as anthracyclines (54); it would be
283 valuable to carry out analyses on the genes involved in these biological pathways.

284 Future directions must focus on interactions between treatment-related risk factors, radiation doses
285 and volume metrics - considering the different cardiac substructure - anthracycline and other
286 chemotherapy doses with genetic predisposition to CD.

287 As described in the current review, the high number of identified variants and genes makes it difficult
288 to draw firm conclusions. Focusing on rare or ultra-rare genetic variants might reveal a more important
289 effect than common variants in CCS. Also, opening up new fields of research, such as genes expression,
290 could be an interesting: the dynamic nature of the transcriptome, responsive to xenobiotics, allows
291 the identification of biomarkers that can directly be used for detection of pathologies resulting from
292 exposure to treatments (55). It is also important not to focus solely on cardiomyopathies, but to study
293 all cardiac diseases, which may have different pathophysiological mechanisms involving different

294 genes. Studies in these different fields require the inclusion of thousands of CCS and can, therefore,
295 only be carried out through international collaboration. This would produce the robust results that
296 could be considered in the international guidelines of patients treated for childhood cancer. Indeed,
297 current CCS follow-up guidelines do not include genetic screening (56–58), even though a number of
298 studies have shown the benefits of including genetic variables in statistical models comprising clinical
299 variables by improving the discriminative accuracy of the models, and allowing a better distinction
300 between patients who developed cardiomyopathy and those who did not (20–22).

301 In conclusion, this systematic review presents an overview of the genetic variants predisposing to
302 cardiomyopathy, enabling us to gain a better understanding of the underlying biological mechanisms,
303 and also to generate new hypotheses and more effective investigation of research fields. Indeed, it is
304 crucial to prevent late effects, such as cardiac disease, and to improve CCS management. Despite
305 therapeutic advances, CCS are still at increased risk of developing late health concerns. It would be
306 extremely useful to establish a set of gene variants and biomarkers, in order to identify patients most
307 at risk of developing CD, and to adapt as much as possible their exposition to cardiotoxic therapies, as
308 well as to implement monitoring and prevention.

References

1. Robison LL, Hudson MM. Survivors of childhood and adolescent cancer: life-long risks and responsibilities. *Nat Rev Cancer*. janv 2014;14(1):61-70.
2. Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. *CA: A Cancer Journal for Clinicians*. 2014;64(2):83-103.
3. Clemens E, van der Kooi ALF, Broer L, van Dulmen-den Broeder E, Visscher H, Kremer L, et al. The influence of genetic variation on late toxicities in childhood cancer survivors: A review. *Critical Reviews in Oncology/Hematology*. juin 2018;126:154-67.
4. Haddy N, Diallo S, El-Fayech C, Schwartz B, Pein F, Hawkins M, et al. Cardiac Diseases Following Childhood Cancer Treatment: Cohort Study. *Circulation*. 5 janv 2016;133(1):31-8.
5. Mulrooney DA, Hyun G, Ness KK, Ehrhardt MJ, Yasui Y, Duprez D, et al. Major cardiac events for adult survivors of childhood cancer diagnosed between 1970 and 1999: report from the Childhood Cancer Survivor Study cohort. *BMJ*. 15 janv 2020;l6794.
6. Chow EJ, Chen Y, Kremer LC, Breslow NE, Hudson MM, Armstrong GT, et al. Individual Prediction of Heart Failure Among Childhood Cancer Survivors. *JCO*. 10 févr 2015;33(5):394-402.
7. Reulen RC, Winter DL, Frobisher C, Lancashire ER, Stiller CA, Jenney ME, et al. Long-term Cause-Specific Mortality Among Survivors of Childhood Cancer. *JAMA*. 14 juill 2010;304(2):172-9.
8. Geenen MM, Cardous-Ubbink MC, Kremer LCM, van den Bos C, van der Pal HJH, Heinen RC, et al. Medical Assessment of Adverse Health Outcomes in Long-term Survivors of Childhood Cancer. *JAMA*. 27 juin 2007;297(24):2705-15.
9. Mansouri I, Allodji RS, Hill C, El-Fayech C, Pein F, Diallo S, et al. The role of irradiated heart and left ventricular volumes in heart failure occurrence after childhood cancer. *Eur J Heart Fail*. avr 2019;21(4):509-18.
10. Mulrooney DA, Yeazel MW, Kawashima T, Mertens AC, Mitby P, Stovall M, et al. Cardiac outcomes in a cohort of adult survivors of childhood and adolescent cancer: retrospective analysis of the Childhood Cancer Survivor Study cohort. *BMJ*. 8 déc 2009;339:b4606.
11. Lipshultz SE, Lipsitz SR, Sallan SE, Dalton VM, Mone SM, Gelber RD, et al. Chronic Progressive Cardiac Dysfunction Years After Doxorubicin Therapy for Childhood Acute Lymphoblastic Leukemia. *JCO*. 20 avr 2005;23(12):2629-36.
12. Armstrong GT, Plana JC, Zhang N, Srivastava D, Green DM, Ness KK, et al. Screening Adult Survivors of Childhood Cancer for Cardiomyopathy: Comparison of Echocardiography and Cardiac Magnetic Resonance Imaging. *J Clin Oncol*. 10 août 2012;30(23):2876-84.
13. Siaravas KC, Katsouras CS, Sioka C. Radiation Treatment Mechanisms of Cardiotoxicity: A Systematic Review. *Int J Mol Sci*. 27 mars 2023;24(7):6272.
14. Linschoten M, Teske AJ, Cramer MJ, van der Wall E, Asselbergs FW. Chemotherapy-Related Cardiac Dysfunction: A Systematic Review of Genetic Variants Modulating Individual Risk. *Circ*:

Genomic and Precision Medicine [Internet]. janv 2018 [cité 6 déc 2022];11(1). Disponible sur: <https://www.ahajournals.org/doi/10.1161/CIRCGEN.117.001753>

15. Bansal N, Amdani S, Lipshultz ER, Lipshultz SE. Chemotherapy-induced cardiotoxicity in children. *Expert Opinion on Drug Metabolism & Toxicology*. 3 août 2017;13(8):817-32.
16. Leerink JM, de Baat EC, Feijen EAM, Bellersen L, van Dalen EC, Grotenhuis HB, et al. Cardiac Disease in Childhood Cancer Survivors. *JACC CardioOncol*. 15 sept 2020;2(3):363-78.
17. Berkman AM, Hildebrandt MAT, Landstrom AP. The genetic underpinnings of anthracycline-induced cardiomyopathy predisposition. *Clinical Genetics*. 2021;100(2):132-43.
18. Sohani ZN, Meyre D, de Souza RJ, Joseph PG, Gandhi M, Dennis BB, et al. Assessing the quality of published genetic association studies in meta-analyses: the quality of genetic studies (Q-Genie) tool. *BMC Genet*. déc 2015;16(1):50.
19. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews | The BMJ [Internet]. [cité 21 oct 2022]. Disponible sur: <https://www.bmj.com/content/372/bmj.n71>
20. Visscher H, Ross CJD, Rassekh SR, Barhdadi A, Dubé MP, Al-Saloos H, et al. Pharmacogenomic prediction of anthracycline-induced cardiotoxicity in children. *J Clin Oncol*. 1 mai 2012;30(13):1422-8.
21. Visscher H, Ross CJD, Rassekh SR, Sandor GSS, Caron HN, van Dalen EC, et al. Validation of variants in SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. *Pediatr Blood Cancer*. août 2013;60(8):1375-81.
22. Visscher H, Rassekh SR, Sandor GS, Caron HN, van Dalen EC, Kremer LC, et al. Genetic variants in SLC22A17 and SLC22A7 are associated with anthracycline-induced cardiotoxicity in children. *Pharmacogenomics*. 2015;16(10):1065-76.
23. Aminkeng F, Bhavsar AP, Visscher H, Rassekh SR, Li Y, Lee JW, et al. A coding variant in RARG confers susceptibility to anthracycline-induced cardiotoxicity in childhood cancer. *Nat Genet*. sept 2015;47(9):1079-84.
24. Blanco JG, Leisenring WM, Gonzalez-Covarrubias VM, Kawashima TI, Davies SM, Relling MV, et al. Genetic polymorphisms in the carbonyl reductase 3 gene CBR3 and the NAD(P)H:quinone oxidoreductase 1 gene NQO1 in patients who developed anthracycline-related congestive heart failure after childhood cancer. *Cancer*. 15 juin 2008;112(12):2789-95.
25. Blanco JG, Sun CL, Landier W, Chen L, Esparza-Duran D, Leisenring W, et al. Anthracycline-Related Cardiomyopathy After Childhood Cancer: Role of Polymorphisms in Carbonyl Reductase Genes—A Report From the Children’s Oncology Group. *JCO*. 1 mai 2012;30(13):1415-21.
26. Wang X, Liu W, Sun CL, Armenian SH, Hakonarson H, Hageman L, et al. Hyaluronan synthase 3 variant and anthracycline-related cardiomyopathy: a report from the children’s oncology group. *J Clin Oncol*. 1 mars 2014;32(7):647-53.
27. Wang X, Sun CL, Quiñones-Lombraña A, Singh P, Landier W, Hageman L, et al. *CELF4* Variant and Anthracycline-Related Cardiomyopathy: A Children’s Oncology Group Genome-Wide Association Study. *JCO*. 10 mars 2016;34(8):863-70.

28. Singh P, Wang X, Hageman L, Chen Y, Magdy T, Landier W, et al. Association of *GSTM1* null variant with anthracycline-related cardiomyopathy after childhood cancer—A Children's Oncology Group ALTE03N1 report. *Cancer*. sept 2020;126(17):4051-8.
29. Sapkota Y, Qin N, Ehrhardt MJ, Wang Z, Chen Y, Wilson CL, et al. Genetic Variants Associated with Therapy-Related Cardiomyopathy among Childhood Cancer Survivors of African Ancestry. *Cancer Research*. 1 mai 2021;81(9):2556-65.
30. Chaix MA, Parmar N, Kinnear C, Lafreniere-Roula M, Akinrinade O, Yao R, et al. Machine Learning Identifies Clinical and Genetic Factors Associated With Anthracycline Cardiotoxicity in Pediatric Cancer Survivors. *JACC: CardioOncology*. déc 2020;2(5):690-706.
31. Sági JC, Egyed B, Kelemen A, Kutszegi N, Hegyi M, Gézsi A, et al. Possible roles of genetic variations in chemotherapy related cardiotoxicity in pediatric acute lymphoblastic leukemia and osteosarcoma. *BMC Cancer*. déc 2018;18(1):704.
32. Hildebrandt MAT, Reyes M, Wu X, Pu X, Thompson KA, Ma J, et al. Hypertension Susceptibility Loci are Associated with Anthracycline-related Cardiotoxicity in Long-term Childhood Cancer Survivors. *Sci Rep*. 29 août 2017;7:9698.
33. Krajcinovic M, Elbared J, Drouin S, Bertout L, Rezgui A, Ansari M, et al. Polymorphisms of *ABCC5* and *NOS3* genes influence doxorubicin cardiotoxicity in survivors of childhood acute lymphoblastic leukemia. *Pharmacogenomics J*. nov 2016;16(6):530-5.
34. Semsei AF, Erdelyi DJ, Ungvari I, Csagoly E, Hegyi MZ, Kiszal PS, et al. *ABCC1* polymorphisms in anthracycline-induced cardiotoxicity in childhood acute lymphoblastic leukaemia. *Cell Biol Int*. 1 janv 2012;36(1):79-86.
35. Lipshultz SE, Lipsitz SR, Kutok JL, Miller TL, Colan SD, Neuberg DS, et al. Impact of hemochromatosis gene mutations on cardiac status in doxorubicin-treated survivors of childhood high-risk leukemia: HFE Mutations and Survivor Cardiac Status. *Cancer*. 1 oct 2013;119(19):3555-62.
36. Sapkota Y, Ehrhardt MJ, Qin N, Wang Z, Liu Q, Qiu W, et al. A Novel Locus on 6p21.2 for Cancer Treatment-Induced Cardiac Dysfunction Among Childhood Cancer Survivors. *JNCI: Journal of the National Cancer Institute*. 1 août 2022;114(8):1109-16.
37. Petrykey K, Rezgui AM, Guern ML, Beaulieu P, St-Onge P, Drouin S, et al. Genetic factors in treatment-related cardiovascular complications in survivors of childhood acute lymphoblastic leukemia. *Pharmacogenomics*. 22(14):885-901.
38. Wang X, Singh P, Zhou L, Sharafeldin N, Landier W, Hageman L, et al. Genome-Wide Association Study Identifies *ROBO2* as a Novel Susceptibility Gene for Anthracycline-Related Cardiomyopathy in Childhood Cancer Survivors. *JCO*. 12 déc 2022;JCO.22.01527.
39. Pizzagalli MD, Bensimon A, Superti-Furga G. A guide to plasma membrane solute carrier proteins. *FEBS J*. mai 2021;288(9):2784-835.
40. Huang KM, Hu S, Sparreboom A. Drug transporters and anthracycline-induced cardiotoxicity. *Pharmacogenomics*. 1 juill 2018;19(11):883-8.

41. Vulsteke C, Pfeil AM, Maggen C, Schwenkglens M, Pettengell R, Szucs TD, et al. Clinical and genetic risk factors for epirubicin-induced cardiac toxicity in early breast cancer patients. *Breast Cancer Res Treat.* juill 2015;152(1):67-76.
42. Luo L, Zhou C, Hui Y, Kurogi K, Sakakibara Y, Suiko M, et al. Human cytosolic sulfotransferase SULT1C4 mediates the sulfation of doxorubicin and epirubicin. *Drug Metabolism and Pharmacokinetics.* 1 avr 2016;31(2):163-6.
43. Couture L, Nash JA, Turgeon J. The ATP-Binding Cassette Transporters and Their Implication in Drug Disposition: A Special Look at the Heart. *Pharmacol Rev.* juin 2006;58(2):244-58.
44. Fogli S, Nieri P, Cristina Breschi M. The role of nitric oxide in anthracycline toxicity and prospects for pharmacologic prevention of cardiac damage. *The FASEB Journal.* 2004;18(6):664-75.
45. Megias-Vericat JE, Martinez-Cuadron D, Herrero MJ, Alino SF, Poveda JL, Sanz MA, et al. Pharmacogenetics of Metabolic Genes of Anthracyclines in Acute Myeloid Leukemia. *Curr Drug Metab.* 2018;19(1):55-74.
46. Koleini N, Kardami E. Autophagy and mitophagy in the context of doxorubicin-induced cardiotoxicity. *Oncotarget.* 7 avr 2017;8(28):46663-80.
47. L'Ecuyer T, Allebban Z, Thomas R, Vander Heide R. Glutathione S-transferase overexpression protects against anthracycline-induced H9C2 cell death. *Am J Physiol Heart Circ Physiol.* juin 2004;286(6):H2057-2064.
48. Mendoza G, Álvarez AI, Pulido MM, Molina AJ, Merino G, Real R, et al. Inhibitory effects of different antioxidants on hyaluronan depolymerization. *Carbohydrate Research.* 15 janv 2007;342(1):96-102.
49. Kellermayer D, Smith JE, Granzier H. Titin mutations and muscle disease. *Pflugers Arch.* mai 2019;471(5):673-82.
50. Gröbner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, et al. The landscape of genomic alterations across childhood cancers. *Nature.* mars 2018;555(7696):321-7.
51. Downing JR, Wilson RK, Zhang J, Mardis ER, Pui CH, Ding L, et al. The Pediatric Cancer Genome Project. *Nat Genet.* 29 mai 2012;44(6):619-22.
52. Aminkeng F, Ross CJD, Rassekh SR, Hwang S, Rieder MJ, Bhavsar AP, et al. Recommendations for genetic testing to reduce the incidence of anthracycline-induced cardiotoxicity. *Br J Clin Pharmacol.* sept 2016;82(3):683-95.
53. Diallo I, Lamon A, Shamsaldin A, Grimaud E, de Vathaire F, Chavaudra J. Estimation of the radiation dose delivered to any point outside the target volume per patient treated with external beam radiotherapy. *Radiother Oncol.* mars 1996;38(3):269-71.
54. Wang K, Tepper JE. Radiation therapy-associated toxicity: Etiology, management, and prevention. *CA: A Cancer Journal for Clinicians.* 2021;71(5):437-54.
55. Joseph P. Transcriptomics in toxicology. *Food Chem Toxicol.* nov 2017;109(Pt 1):650-62.

56. Chen JZ, Liang B. Comparison of American and European guidelines for cardio-oncology of heart failure. *Heart Fail Rev.* 1 sept 2023;28(5):1211-20.
57. Van Kalsbeek RJ, Van Der Pal HJH, Kremer LCM, Bardi E, Brown MC, Effenev R, et al. European PanCareFollowUp Recommendations for surveillance of late effects of childhood, adolescent, and young adult cancer. *European Journal of Cancer.* sept 2021;154:316-28.
58. Armenian SH, Hudson MM, Mulder RL, Chen MH, Constone LS, Dwyer M, et al. Recommendations for Cardiomyopathy Surveillance for Survivors of Childhood Cancer: A Report from the International Late Effects of Childhood Cancer Guideline Harmonization Group. *Lancet Oncol.* mars 2015;16(3):e123-36.

Table 1. Characteristics of the 19 included studies

Study	Country/ethnicity	Tumor type	Age at cancer diagnosis/treatment (years)	Follow-up (years)	Treatments (anthracyclines /radiotherapy)	Sequencing method	Study design	Cohort size (case vs. controls)
Sági et al., <i>BMC Cancer</i> , 2018	Hungary	Acute Lymphoblastic Leukemia and osteosarcoma	Median (range) ALL: 5.2 (0 - 18) OSC: 13.2 (5 -18)	NA	Anthracyclines (all patients)	Candidate Gene	Case-control study	661 (20 vs. 641)
Blanco et al., <i>JCO</i> , 2012	North America	Any cancer type	Median (range) Cases: 7.3 (0–20.7) Controls: 7.6 (0 - 21.1)	Median (range) 9.7 (0.1 - 40.3)	Anthracyclines and/ or radiotherapy	Candidate Gene	Matched case-control study	487 (170 vs. 317)
Visscher et al., <i>Pediatr Blood Cancer</i> , 2013	Netherlands and Canada	Any cancer type	Median (range) Cases: 9.9 (0.5–17.0) Controls: 8.16 (0.5 - 17.7)	Median (range) 13.8 (0.4 - 31.6)	Anthracyclines +/- radiotherapy	Candidate Gene	Case-control study	344 (78 vs. 266)
Singh et al., <i>Cancer</i> , 2020	Asian, Black, Hispanic, White	Any cancer type	Median (range) Cases: 7.8 (3.8 - 11.5) Controls: 9.6 (3.3 - 14.8)	Median (range) 9.3 (1.3 - 17.2)	Anthracyclines +/- or radiotherapy	Candidate Gene	Matched case-control study	167 (75 vs. 92)
Blanco et al., <i>Cancer</i> , 2008	North America: Black, White, Others	Any cancer type	Mean (SD) Cases: 10.3 (6.5) Controls: 9.1 (5.8)	≥ 5 years after diagnosis	Anthracyclines +/- radiotherapy	Candidate Gene	Nested matched case-control study	140 (29 vs. 111)
Hildebrandt et al., <i>Nature</i> , 2017	White, Hispanic, Black and other	Any cancer type	Mean (SD) Cases: 9.2 (4.7) Controls: 9.3 (5.7)	Mean (SD) Cases: 21.2(11.2) Controls: 15.7(7.6)	Anthracyclines +/- radiotherapy	Candidate Gene	Matched case-control study	108 (46 vs. 62)
Wang et al., <i>JCO</i> , 2022	North America: Non-Hispanic white,	Any cancer type	Median (range) Cases: 7 (0 - 20) Controls: 11 (0 -20)	Median (range) Cases: 22 (0 - 39) Controls: 31 (14 - 45)	Anthracyclines +/- radiotherapy	GWAS	Case-control study	1866 (126 - 1740)
Aminkeng et al., <i>Nat Genet.</i> 2015	Canada: European and non-European ancestry	Any cancer type	Median (IQR) Cases: 8.5 (4 -14) Controls: 5 (2 - 10)	NA	Anthracyclines +/- radiotherapy	GWAS	Case-control study	376 (54 vs. 322)
Visscher et al., <i>Pharmacogenomics</i> , 2015	Netherlands and Canada	Any cancer type	Median (range) Cases: 9.9 (0.5–17.0) Controls: 4.9 (0.1-17.7)	Median (range) 13.8 (0.4 - 31.6)	Anthracyclines +/- radiotherapy	GWAS	Case-control study	335 (78 vs. 257)
Wang et al., <i>JCO</i> , 2016	Non-Hispanic whites	Any cancer type	Median (range) Cases: 7.5 (0 - 20) Controls: 8.3 (5 - 8)	Median (range) 11.7 (0.1 - 41)	Anthracyclines and/ or radiotherapy	GWAS	Matched case-control study	331 (112 vs. 219)

Wang et al., JCO, 2014	Non-Hispanic whites	Any cancer type	Median (range) Cases: 6.9 (0 - 20.2) Controls: 6.3 (0 - 20.6)	Median (range) 10.87 (0.1 - 41.0)	Anthracyclines and/ or radiotherapy	GWAS	Matched case-control study	287 (93 vs. 194)
Visscher et al., JCO, 2012	Canada: European and non-European ancestry	Any cancer type	Median (range) Cases: 6.1 (0.04–17.6) Controls: 3.8 (0.05 - 16.9)	Median (range) 8.2 (0.1 - 18.6)	Anthracyclines +/- radiotherapy	GWAS	Case-control study	156 (38 vs. 118)
Chaix et al., JACC: Cardiooncology, 2020.	North America	Any cancer type	Median (IQR) Cases: 4.0 (2.0 - 7.0) Controls: 6.0 (2.0 - 10.0)	Median (IQR) Cases: 9.0 (6.0-12.3) Controls: 8.5 (5.0-12.3)	Anthracyclines +/- radiotherapy	WES	Nested case-control study	289 (183 vs. 106)
Krajinovic et al., The Pharmacogenomics Journal, 2016	Canada: Caucasian	Acute Lymphoblastic Leukemia	Median (range) 5.0 (1.0 - 18.0)	Median (range) 8.00 (1.00 - 18)	Anthracyclines +/- radiotherapy	Candidate Gene	Cohort study	251
Semsei A., Cell Biol. Int, 2012	Hungary	Acute Lymphoblastic Leukemia	Mean (range) 5.7 (1.2 - 18.0)	Median (range) 6.3 (2.4 - 13.7)	Anthracyclines (all patients)	Candidate Gene	Cohort study	235
Lipshultz S., Cancer, 2013	Missing	High risk Acute Lymphoblastic Leukemia	59% : <10 years old 41% : ≥10 years old	Median (range) 7.0 (1.1 - 17.1)	Anthracyclines	Candidate Gene	Cohort study	184
Petrykey et al., Pharmacogenomics, 2021	Canada: European ancestry	Acute Lymphoblastic Leukemia	Median (range) 4 (0 - 18)	Median (range) 13 (3 - 24)	Anthracyclines (all patients)	WES	Cohort study	233
Sapkota et al., JNCI J Natl Cancer Inst, 2022	USA : European ancestry; African ancestry;	Any cancer type	Median (range) Cases: 10.5 (0.0 - 20.5) Controls: 6.9 (0.0 - 23.6)	NA	Anthracyclines and/ or radiotherapy	WGS	Cohort study	1870 (227 - 1643)
Sapkota et al., AACR, 2021	USA: African ancestry	Any cancer type	Median (range) 8.8 (0.1-19.7)	NA	Anthracyclines and/ or radiotherapy	WGS	Cohort study	246 (40 vs. 206)

Abbreviations: SD, Standard Deviation; IQR, Interquartile Range; NA, Not Available.

Table 1. Main genetic associations' findings of the 19 included studies

Study	Sequencing method	Outcome	Gene/ region	Variant	Effect allele	OR (95% CI)	p-values
Sági et al., <i>BMC Cancer</i> , 2018	Candidate gene (70 SNP in 26 genes)	Cardiotoxicity	CYP3A5	rs4646450	TT	7.25 (1.83 - 28.78)	4.81 x 10 ⁻³
			SLC28A3	rs7853758	AA	9.84 (1.73 - 56.02)	1.00 x 10 ⁻²
		LVFS decrease 5 to 10 years after diagnosis	ABCC2	rs3740066	GG/ GA/ AA	39.1 vs. 40.6 vs. 42.4 ^c	7.11 x 10 ⁻⁴
			NQO1	rs1043470	TT/ CC + CT	38.5 vs. 40.6 ^c	5.82 x 10 ⁻³
		SLC22A6	rs6591722	AA/ TA + TT	37.5 vs. 40.6 ^c	1.71 x 10 ⁻³	
Blanco et al., <i>JCO</i> , 2012	Candidate gene (2 SNP in 2 genes)	Cardiomyopathy	CBR3	rs1056892	GG	1.79 (1.08 - 2.96)	0.02
			CBR1	rs9024	GG	0.81 (0.45 - 1.47)	0.49
			Combine CBR3 and CBR1	Combine: rs1056892/ rs9024	GG + GG	1.53 (0.93 - 2.51)	0.09
Visscher et al., <i>Pediatr Blood Cancer</i> , 2013	Candidate gene (23 SNP)	ACT	SLC28A3	rs7853758	A	0.36 (0.22 - 0.60)	1.6 x 10 ⁻⁵
			SLC28A3	rs885004	A	0.34 (0.20 - 0.60)	3.0 x 10 ⁻⁵
Singh et al., <i>Cancer</i> , 2020	Gene candidate (1 gene)	Cardiomyopathy	GSTM1	Null/ Positive		2.52 (1.2 - 5.2)	0.01
Blanco et al., <i>Cancer</i> , 2008	Candidate gene (2 SNP in 2 genes)	CHF	CBR3	rs1056892	GG/ AA	5.63 (0.80 - 39.57)	0.083
					GA/ AA	3.66 (0.64 - 21.07)	0.15
			NQO1	rs1800566	CC/ TT	1.26 (0.14 - 11.39)	0.84
					CT/ TT	0.65 (0.06 - 6.77)	0.72
Hildebrandt et al., <i>Nature</i> , 2017	Candidate gene (12 SNP)	Cardiotoxicity	PLCE1	rs932764	G	0.48 (0.27–0.85)	0.012
			ATP2B1	rs17249754	G	0.33 (0.12 - 0.92)	0.034
			PLCE1	rs932764	G	0.36 (0.18 - 0.76)	0.0068
			ATP2B1	rs17249754	G	0.26 (0.07 - 0.96)	0.04
Aminkeng et al., <i>Nat Genet.</i> , 2015	GWAS	ACT	RARG	rs2229774		4.7 (2.7–8.3)	5.9 x 10 ⁻⁸
			RARG	rs2229774	AA + AG/ GG	5.2 (3.0 - 9.0)	5.9 x 10 ⁻¹⁰
Visscher et al., <i>Pharmacogenomics</i> , 2015	GWAS	ACT	SLC22A17	rs4982753	A/ G	0.52 (0.31 - 0.85)	0.0078
			SLC22A7	rs4149178	G/ A	0.41 (0.21 - 0.77)	0.0034
Wang et al., <i>JCO.</i> , 2016	GWAS	Cardiomyopathy	CELF4	rs1786814	GA/ AA	3.18 (0.5 - 18.9)	0.2
					GG/ AA	6.73 (1.1 - 41.1)	0.04
					GG/AA + GA	2.26 (1.2 - 4.0)	0.006
Wang et al., <i>JCO</i> , 2014	GWAS	Cardiomyopathy	HAS3	rs2232228	GA/ GG	2.0 (0.8 - 4.6)	0.12
			HAS3	rs2232228	AA/ GG	1.8 (0.7 - 4.7)	0.2

Visscher et al., JCO, 2012	GWAS	ACT	SLC28A3	rs7853758	A	0.31 (0.16 - 0.60)	1.0 x 10 ⁻⁴
Chaix et al., JACC : Cardiooncology, 2020 ^a	WES	Cardiotoxicity	CEL4			0.18 (0.06 - 0.56)	<0.001
			ZNF71			0.05 (0.00 - 0.92)	<0.001
			PIK3R2			0.05 (0.00 - 0.92)	<0.001
Krajinovic et al., The Pharmacogenomics Journal, 2016	Candidate gene (33 common SNP in 12 genes)	LVEF	ABCC5	rs7627754	TT/ AA + AT	- 12% ^e	<0.0005
		LVFS	ABCC5	rs7627754	TT/ AA + AT	- 8% ^e	0.001
Semsei A et al., Cell Biol. Int, 2012	Candidate gene (9 SNP in 1 gene)	LVFS at the last follow-up time	ABCC1	rs246221	TT/ CC/ TC	38.5%/ 40.7%/ 38.4% ^c	0.027
Lipshultz S et al., Cancer. 2013	Candidate gene (2 SNP in 1 gene)	LVSF 4 to 7 years after diagnosis	HFE	Combined H63D/C282Y	Carriers	-0.93% (0.27) ^d	0.003
Petrykey et al., Pharmacogenomics, 2021	WES ^b	LVEF	TTN	rs72648998	CC/ CT + TT	4.78 (2.4 - 7.3) ^d	0.0003
			TTN	rs3829747	CC/ CT + TT	3.33 (1.5 - 5.2) ^d	0.0003
			TTN	rs2303838	CC/ CT + TT	2.59 (0.9 - 4.2) ^d	0.0005
		LVFS	TTN	rs72648998	CC/ CT + TT	3.35 (1.6 - 5.1) ^d	0.0002
		LVEF	NOS1	rs76090928			0.0013
			ABCG2	rs199473672			0.0026
Sapkota et al., JNCI , 2022	WGS	CCD	KCNK17	rs2815063	A	1.57 (1.19 - 2.09)	0.002
		LVEF	KCNK17	rs2815063	A	- 0.016 (0.003) ^d	2.1 x 10 ⁻⁸
				rs115537302	A	- 0.019 (0.004) ^d	8.7 x 10 ⁻⁸
				rs115057884	C	- 0.019 (0.004) ^d	9.0 x 10 ⁻⁸
				rs116181858	G	- 0.019 (0.004) ^d	9.6 x 10 ⁻⁸
				rs114566901	C	- 0.019 (0.004) ^d	1.7 x 10 ⁻⁷
Sapkota et al., AACR, 2021	WGS	LVEF	1p13. 2	rs6689879	C/ T	-0.042 (0.007) ^d	2.8 x 10 ⁻⁸
			15q25. 3	rs9788776	G/ A	-0.059 (0.01) ^d	3.5 x10 ⁻⁸
		Cardiomyopathy (grade 2-4)	1p13. 2	rs6689879	C/ T	3.71 (1.87 - 7.37)	1.8 x 10 ⁻⁴
			15q25. 3	rs9788776	G/ A	5.24 (2.30 - 11.92)	7.9 x 10 ⁻⁵

Abbreviations: OR, Odds ratio; CI, confidence interval, SE, Standard error; ACT, Anthracycline-induced cardiotoxicity; CHF, Congestive heart failure; CCD, Cancer treatment-induced cardiac dysfunction; LVEF, Left-ventricular ejection fraction; LVFS, Left-ventricular fractional shortening;

^a28 genes significantly associated with case-control status but we only reported the 3 genes considered as the more important by the authors; ^bWith a focus on genetic variants in genes selected for doxorubicin action pathways, in the functioning of the cardiac system, and in the mitochondrial function; ^cAllele frequency; ^dBeta (SE) or Beta (95%CI); ^eEffect size measured by echo-cardiac parameters

Table 3. Genes and variant effect according to cumulative anthracycline doses on cardiomyopathy

Study	Outcome	Interaction or stratification	Categories definitions	Genes /variants	Effect allele	Cat 1	Cat 2	Cat 3
						OR (95% CI)	OR (95% CI)	OR (95% CI)
Blanco et al., JCO, 2012	Cardiomyopathy	Interaction SNP * cumulative anthracycline doses (3 categories)	Cat 1: no exposure Cat 2: 1 to 250 mg/m ² Cat 3: > 250 mg/m ²	CBR3 / rs1056892	GA + AA	Ref	1.66 (0.49 - 5.69)	18.92 (6.13 - 58.45)
					GG	0.86 (0.22-3.39)	5.48 (1.81 - 16.63)	25.91 (7.67 - 87.57)
				CBR1/ rs9024	GA + AA	Ref	0.68 (0.13-3.59)	24.61 (5.4-112)
					GG	0.35 (0.08-1.54)	2.25 (0.62-8.21)	11.17 (2.9-12.59)
				CBR3 / rs1056892 and CBR1/ rs9024	GA + AA and GA + AA	Ref	1.51 (0.47-4.81)	17.40 (5.91-51.22)
GG and GG	0.53 (0.10-2.67)	5.39 (1.79-16.17)	17.94 (5.64-57.08)					
Singh et al., Cancer, 2020	Cardiomyopathy	Stratification on cumulative anthracycline doses	Cat 1: ≤ 250 mg/m ² Cat 2: > 250 mg/m ²	GSTM1	Null / Positive	2.3 (0.7 – 6.7)	2.6 (1.2 – 5.1)	
Wang et al., JCO, 2022	Cardiomyopathy	Interaction SNP * cumulative anthracycline doses (2 categories)	Cat 1: ≤ 250 mg/m ² Cat 2: > 250 mg/m ²	ROBO2/ rs17736312	GG + AG	Ref	0.14 (1.1 – 2.9)	
					AA	1.8 (1.1 – 2.9)	2.2 (1.2 – 4.0)	
Wang et al., JCO, 2016	Cardiomyopathy	Interaction SNP * cumulative anthracycline doses (2 categories)	Cat 1: ≤ 300 mg/m ² Cat 2: > 300 mg/m ²	CELF4/ rs1786814	GA + AA*	Ref	1.75 (0.6 – 5.0)	
					GG	1.03 (0.5 – 2.2)	10.16 (3.8 – 27.3)	
Wang et al., JCO, 2014	Cardiomyopathy	Interaction SNP * cumulative anthracycline doses (continuous)	Cat 1: cumulative anthracyclines dose > 450 mg/m ²	HAS3/ rs2232228	AA	56.5 (NA – NA)		
					GG	0.6 (NA – NA)		
			Cat 1: Dose > 250 mg/m ²	HAS3/ rs2232228	GG	Ref		
		GA			4.7 (1.4 – 16.2)			
		AA			8.9 (2.1 – 37.5)			
		Interaction SNP * cumulative anthracycline doses (2 categories)	Cat 1: 0 -250 mg/m ² Cat 2: > 250 mg/m ²	HAS3/ rs2232228	GG	Ref	1.1 (0.3 – 4.8)	
GA	0.5 (0.2 – 1.8)				5.2 (1.6 – 17.4)			
AA	0.2 (0.1 – 0.8)				9.9 (2.4 – 40.9)			

Abbreviations: ACT, Anthracyclines-Induced Cardiotoxicity; LVFS, Left-ventricular fractional shortening; LVEF, Left-ventricular ejection fraction; Abbreviations: OR, Odds ratio; CI, confidence interval;

Supplementary Table 1: Details of treatments of the 19 included studies

Study	Anthracycline	Radiotherapy	Other treatments
Sági et al., BMC Cancer, 2018	All patients were treated with anthracyclines, Cumulative anthracycline exposure calculated using doxorubicin equivalent. ALL low-risk and medium-risk groups of: doses between 180 and 240 mg/m ² ; ALL high-risk group and relapsed patients: doses between 240 and 380 mg/m ² . Osteosarcoma standard-risk group: 360 mg/m ² high-risk group: 180 mg/m ² ..	NA	Treatment protocols: ALL BFM (Berlin–Frankfurt–Münster) 88, ALL BFM 90, ALL BFM 95, ALL IC-BFM (ALL Intercontinental) 2002 or ALL IC-BFM 2009; Interfant 98 or Interfant 2006. COSS (German-Austrian-Swiss osteosarcoma study group) -86 and COSS-96 protocols. 29% of the patients received 12 Gy cranial radiotherapy.
Blanco et al., JCO, 2012	Cumulative anthracycline exposure calculated using a factor that reflects the drug’s cardiotoxic potential (mg/m ²). Cumulative anthracycline exposure was treated as a categorical variable (0, 1- 100, 101 - 150, 151 - 200, 201- 250, 251 - 300, > 300).	Radiation therapy to the chest including the heart: (Yes/No) variable.	NA
Visscher et al., Pediatr Blood Cancer, 2013	All Canadian patients received anthracyclines. Cumulative anthracycline doses were calculated using doxorubicin equivalents. Cumulative anthracycline exposure was treated as a continuous variable.	Radiotherapy to the heart: (Yes/No) variable.	NA
Singh et al., Cancer, 2020	Cumulative anthracycline exposure calculated using a factor that reflects the drug’s cardiotoxic potential (mg/m ²). Cumulative anthracycline exposure was treated as a continuous variable.	Radiation therapy to the chest including the heart: (Yes/No) variable.	NA
Blanco et al., Cancer. 2008	Cumulative doses of anthracyclines.	Radiation to the heart: total cumulative dose in cGy to the heart.	NA
Hildebrandt et al., Nature, 2017	Cumulative anthracyclines dose were standardized to doxorubicin equivalents.	NA	All patients who received anthracycline-based chemotherapy were also treated with other agents as indicated by cancer diagnosis. Some patients were treated with cardiac medication.
Wang et al., JCO, 2022	Anthracycline exposure included exposure to doxorubicin, daunorubicin, idarubicin, epirubicin, or mitoxantrone. Cumulative anthracycline exposure calculated using a factor that reflects the drug’s cardiotoxic potential (mg/m ²).	Radiation to the chest with heart in the field was captured as a yes/no variable.	NA
Aminkeng et al., Nat Genet. 2015	All patients received anthracyclines; Cumulative anthracycline doses were calculated using doxorubicin equivalents.	Radiation therapy included significant radiation exposure to the heart or surrounding tissue: mantle and mediastinal radiation, whole-	NA

	Cumulative anthracycline exposure was treated as a continuous variable.	lung radiation, whole-abdomen or upper abdominal radiation, left-side flank radiation, and total-body irradiation: : (Yes/No) variable.	
Visscher et al., Pharmacogenomics, 2015	All Canadian patients received anthracyclines. Cumulative anthracycline doses were calculated using doxorubicin equivalents. Cumulative anthracycline exposure was treated as a continuous variable.	Radiotherapy to the heart: (Yes/No) variable.	NA
Wang et al., JCO, 2016	Cumulative anthracycline exposure calculated using a factor that reflects the drug's cardiotoxic potential (mg/m ²) Cumulative anthracycline exposure was treated as a continuous variable.	Radiation therapy to the chest including the heart: (Yes/No) variable.	NA
Wang et al., JCO, 2014	Cumulative anthracycline exposure calculated using a factor that reflects the drug's cardiotoxic potential (mg/m ²) Cumulative anthracycline exposure was treated as a continuous variable.	Total radiation therapy to the chest including the heart: continuous variable.	NA
Visscher et al., Journal of Clinical Oncology, 2012	All Canadian patients received doxorubicin or daunorubicin. 23/96 patients in the Dutch-EKZ cohort received other anthracyclines instead. Cumulative anthracycline doses were calculated using doxorubicin equivalents. Cumulative anthracycline exposure was treated as a continuous variable.	Radiotherapy to the heart: (Yes/No) variable.	NA
Chaix et al., JACC: Cardiooncology, 2020.	The anthracycline cumulative dose was measured in doxorubicin equivalents (mg/m ²). Cases received low cumulative anthracycline dose (≤250 mg/m ²) Controls received high dose anthracycline (>250 mg/m ²).	Radiation therapy to the chest including the heart: (Yes/No) variable.	Dexrazoxane
Krajinovic et al., The Pharmacogenomics Journal, 2016	All patients received at least two 30 mg/m ² doses of doxorubicin. High-risk patients continued to receive one weekly doxorubicin dose during the consolidation phase, for a total cumulative dose of 300 mg/m ² for the 95-01 and 00-01 protocols, or 360 mg/m ² for the 87-01 and 91-01 protocols.	NA	DFCI consortium 87-01, 91-01, 95-01, or 00-01. For 95-01 protocol, high-risk patients were randomized to receive (or not) a 300 mg/m ² dose of Dexrazoxane during induction and consolidation therapy. For 00-01 protocol, all high-risk patients received dexrazoxane before doxorubicin administration.
Semsei A., Cell Biol. Int, 2012	Cumulative anthracyclines doses in mg/m ²	NA	Chemotherapy protocol: ALL BFM 90 or ALL BFM 95. The chemotherapy regimen included repeated doses of intravenous vincristine, L-asparaginase, daunorubicin, doxorubicin, methotrexate, cyclophosphamide, cytosine arabinoside, oral

			<p>prednisone, dexamethasone, mercaptopurine, thioguanine, intrathecal methotrexate, intrathecal prednisone and ifosfamide (the latter only in patients with high-risk leukaemia). The protocol in the low-risk and medium-risk arms differed only in the number of anthracycline doses, the cumulative anthracycline doses being between 180 and 300 mg/m².</p> <p>The high-risk arm differed considerably from the low-risk and medium-risk arms in terms of the applied therapy dosage.</p> <p>Administration of dexrazoxane (Yes / No).</p>
Lipshultz S., Cancer, 2013	<p>Total planned cumulative doxorubicin dosage: 300–360 mg/m².</p> <p>Cumulative DOX dose in mg/m².</p>	NA	<p>ALL Consortium Protocol (1991).</p> <p>Administration of dexrazoxane (Yes / No).</p>
Petrykey et al., Pharmacogenomics, 2021	<p>All patients were given two 30 mg/m² doses of doxorubicin. HR patients continued to receive one weekly doxorubicin dose during the consolidation phase, for a total cumulative dose of 294.52 mg/m².</p>	NA	<p>HR patients treated with 95-01 protocol were randomized to receive (or not) dexrazoxane and all HR patients treated with 00-01 and 00-05 protocols, received dexrazoxane before doxorubicin administration.</p>
Sapkota et al., JNCI J Natl Cancer Inst, 2022	<p>Cumulative anthracycline exposure calculated using doxorubicin equivalent (mg/m²).</p>	Mean heart radiation dose in cGy.	NA
Sapkota et al., AACR, 2021	<p>Cumulative anthracycline exposure calculated using doxorubicin equivalent (mg/m²).</p>	Average radiation dose in Gy.	NA

Abbreviations: BFM: Berlin-Frankfurt-Münster; DFCI: Dana-Farber Cancer Institute; Gy, Gray; ALL, Acute lymphocytic leukemia, NA, Not Available.

Supplementary Table 2: Sequencing techniques of the 19 included studies

Study	biological specimens	Technique
Sági et al., BMC Cancer, 2018	Blood	Based on the scientific literature, 70 single nucleotide polymorphisms (SNPs) in 26 genes were selected and genotyped. Genotyping 63 of the SNPs was conducted using TaqMan® OpenArray™ Genotyping following the manufacturer's instructions at the Department of Medical Chemistry, Molecular Biology and Pathobiochemistry, Semmelweis University. Other 7 SNPs were genotyped using KBioscience Competitive Allele-Specific Polymerase chain reaction-on-Demand prevalidated on 7900HT Fast Real-Time PCR System. The genotyping was unsuccessful in the case of three SNPs.
Blanco et al., JCO, 2012	Blood, saliva or buccal swabs	The CBR3 V244M and CBR1 1096GA polymorphisms (rs1056892, rs9024) were analyzed using validated assays for allelic discrimination with specific fluorescent probes. Laboratory personnel were blinded to case-control status.
Visscher et al., Pediatr Blood Cancer, 2013	blood, saliva, or buccal swabs	Twenty-three SNPs were selected that previously showed evidence of association ($p < 0.01$) with ACT. DNA samples were genotyped using a custom 96-plex Illumina Veracode GoldenGate SNP genotyping assay. This assay included an additional 63 non-study SNPs used for quality control purposes only. All SNP genotypes were manually clustered using Illumina GenomeStudio software. One SNP in FMO2 (rs2020870) could not be reliably clustered and was therefore removed from further analyses.
Singh et al., Cancer, 2020	Blood	DNA and RNA were isolated using the Gentra Puregene Blood Kit and the PAXgene Blood RNA kit (both from Qiagen Inc), respectively. The concentration was measured on Nanodrop ND-1000 Spectrophotometers (Thermo Fisher Scientific). DNA integrity was analyzed by using the Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific). RNA quality was checked using a Bioanalyzer Nanochip (Agilent Technologies), and samples with RNA integrity numbers >7 were used for microarray analysis. GSTM1 deletion was examined by multiplex PCR using the B2M gene as control. Illumina HumanHT-12 v4.0 Expression BeadChips (Illumina, Inc) were used for mRNA transcription. Expression of the GSTM1 gene was measured using qrt-PCR. For RNA sequencing of blood, libraries were prepared using the TruSeq RNA Sample Preparation Kit, according to the protocols recommended by the manufacturer. Each library was paired-end sequenced (2×50 bp) by using the TruSeq SBS Kit v4-HS, on a HiSeq2000 platform. For RNA sequencing of cardiomyocytes, gene expression levels and exon use were estimated using the featureCounts function in the Subread software.
Blanco et al., Cancer. 2008	Buccal cell	The NQO1*2 polymorphism [rs1800566] was examined with a validated PCR-restriction fragment length polymorphism technique. The CBR3 V244M polymorphism (rs1056892) was analyzed with a validated assay for allelic discrimination with specific fluorescent probes.
Hildebrandt et al., Nature, 2017	Blood	Twelve index genetic variants were selected for genotyping based on previous evidence of significant association with hypertension from a GWAS involving $>200\ 000$ individuals in the general population. TaqMan Genotyping Assays were available for 11 of the 12 loci. Genotyping was performed according to standard protocols on the ABI 7900HT platform. All assays were performed blinded to the cardiotoxicity status of the individuals. RNA sequencing of cardiomyocytes: RNAseq was performed in MD Anderson's Sequencing and Microarray Facility using the Illumina TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Gold. The 13 strand-specific libraries were pooled and divided across two lanes for sequencing with the Illumina HiSeq. 3000. Demultiplexed reads of ~ 100 bp length were generated for each sample.

Wang et al., JCO, 2022	Blood or saliva in the replication set	<p>Germline DNA from 5,739 CCSS participants was genotyped on HumanOmni5Exome arrays. Imputation on the basis of the 1000 Genomes Project release v.3 reference haplotypes using IMPUTE v.2.3.0 yielded 26 135 904 high-quality SNPs and small insertions or deletions.</p> <p>In the replication cohort, genotyping was performed on the Juno system according to manufacturer's instructions. End point fluorescence values were measured on the BioMark HD system, and the Fluidigm SNP Genotyping Analysis software program was used to generate genotyping calls for each sample.</p> <p>RNA concentration was measured using Nanodrop ND-1000. RNA quality was checked on Bioanalyzer Nanochip and samples with RNA integrity number >7 were submitted to the Genomic Services Laboratory at HudsonAlpha Institute for Biotechnology, Huntsville, AL. Poly-adenylated RNAs were isolated using NEBNext Magnetic Oligo d(T)25 beads. Libraries were prepared using the TruSeq RNA Sample Preparation Kit (Illumina Inc).</p>
Aminkeng et al., Nat Genet. 2015	Not Specified	<p>Discovery cohort and genotyped on the Illumina Infinium HumanOmniExpress panel (~740,000 markers), genotypes were called with the Illumina GenomeStudio software package, and SNPs were clustered using the Illumina 740K cluster file.</p>
Visscher et al., Pharmacogenomics, 2015	blood, saliva or buccal swabs	<p>DNA samples were genotyped for 4536 SNPs using a customized Illumina GoldenGate SNP genotyping assay, designed to capture the genetic variation of over 300 key drug biotransformation genes.</p> <p>In addition, tagSNPs were included that were identified using the ldSelect algorithm to select a maximally informative set of tagSNPs to assay the candidate genes.</p> <p>In addition, the main SNP panel was supplemented with a custom 96-SNP Illumina Veracode GoldenGate genotyping assay to include both functional and tagSNPs in genes involved in the metabolism of anthracyclines into alcohol metabolites, as well as other SNPs possibly related to ACT that were not included in the main panel.</p> <p>All SNP genotypes were manually clustered using Illumina GenomeStudio software.</p>
Wang et al., JCO, 2016	Blood, saliva or buccal swabs	<p>Genotyping was performed on the Illumina HumanOmniExpress-12 v1.0 DNA analysis bead-chip. The final data set included 583 748 autosomal SNPs.</p> <p>Significant SNP(s) identified in discovery were genotyped in the replication set by using Sequenom iPLEX SNP chemistry on a MassArray system.</p> <p>Presence of SNP(s) of interest in cardiac DNA was investigated with TaqMan genotyping assays per the manufacturer's guidelines. Splicing isoforms were amplified by nested PCR.</p>
Wang et al., JCO, 2014	Blood, saliva or buccal swabs	<p>Genotyping was performed on the Illumina IBC cardiovascular SNP array. The final data set retained 34 912 autosomal SNPs.</p> <p>Significant SNP(s) identified in the discovery stage were genotyped by using Sequenom iPLEX SNP chemistry on a MassARRAY system.</p> <p>Significant SNPs validated in the replication stage were genotyped with TaqMan genotyping assays. Total RNA from heart tissue samples was reverse transcribed and amplified by using one-step QuantiTect SYBR Green rt-PCR kits. Relative hyaluronan synthase 3 (HAS3) mRNA levels were obtained after normalization to reference gene (ACTB).</p>
Visscher et al., Journal of Clinical Oncology, 2012	Not Specified	<p>DNA samples from the two Canadian cohorts were genotyped for 2 977 SNPs using a customized Illumina GoldenGate SNP genotyping assay, designed to capture the genetic variation of 220 key drug biotransformation genes.</p> <p>After quality control, 1 931 SNPs remained for analysis. DNA samples from the Dutch-EKZ replication cohort were genotyped for rs7853758 by TaqMan SNP genotyping.</p>
Chaix et al., JACC: Cardiooncology, 2020.	Blood	<p>WES was performed (average: 100 x depth) by using the Illumina HiSeq X platform. High quality paired-end reads (2 x 150bp) were mapped to the human genome reference sequence by using the bwa mem aligner, version 0.7.8, and variants were called using the Genome Analysis Toolkit, version 3.8.0.</p>
Krajinovic et al., The Pharmacogenomics Journal, 2016	Not Specified	<p>Thirty-three common polymorphisms (with minor allele frequency, MAF, higher than 5%) in 12 genes (ABCC1, ABCC2, ABCC5, ABCB1 ABCG2, NFKB1, NQO1, NOS3, MLH1, MLH2, GSTM1 and GSTT1) involved in DOX metabolic and functional pathways⁷ were genotyped.</p>

		Genotyping of ABCC5 and NOS3 gene polymorphisms was not described in the text.
Semsei A., Cell Biol. Int, 2012	Blood in patients in remission; bone marrow smears, neonatal Guthrie spots, stored buffy coats in patients who were deceased before sample collection	SNPs (single nucleotide polymorphisms) were selected prioritized on the basis of their estimated functionality in this order: non-synonymous SNPs, SNPs in promoter and 3'-UTR (3' untranslated region) region, synonymous SNPs and intronic SNPs. The goal was to cover every haplotype block in the gene defined by the Haploview 4.1 software with 1 or 2 SNPs. The ABCC1 rs45511401 genotypes were determined by multiplex single base extension using a SNaPshot Multiplex Kit followed by mini-sequencing on an ABI 310 genetic analyser. All other ABCC1 SNPs were genotyped using the GenomeLab SNPstream genotyping platform according to the manufacturer's instructions.
Lipshultz S., Cancer, 2013	Blood	Isolated DNA was prepared using NucleoSpin DNA isolation kits (BD Biosciences Clontech) in all samples at the time of mononuclear cell fraction isolation. The two most common HFE alleles associated with hemochromatosis, were detected by directly-sequencing genomic DNA with a commercial, clinically-validated kit (Biotage Inc.; PyroMark HFE Cat #40-0053) run on a Pyrosequencer instrument (PSQ HS 96, Biotage Inc.) or by Sequenom and Taqman genotyping assays performed at the Harvard Partners Center for Genetics and Genomics.
Petrykey et al., Pharmacogenomics, 2021	Blood	Whole exomes were captured in solution with Agilent's SureSelect Human All Exon 50 Mb kits and sequenced on either Life Technologies SOLiD System 4.0 (mean coverage = 40x) or Illumina HiSeq 2500 platform (mean coverage = 113.1x) at SJUHC integrated clinical genomic center in pediatrics. Reads were aligned to the hg19 reference genome using SOLiD LifeScope software for the SOLiD samples and BWA-MEM for the samples sequenced on the Illumina system. PICARD was used to mark PCR duplicates and collect sequencing quality control metrics. Variant calling was performed using the Haplotype Caller and quality score recalibration was performed using Variant Recalibrator, both implemented in the Genome Analysis Tool Kit (GATK). Variants were selected based on the variant quality score (VQSOR = PASS) and the minimum depth of coverage (DP ≥10). The final germline variants were annotated by ANNOVAR. Only missense, nonsense and splicing common and rare variants with predicted functional impact (Sift [<0.1] and PolyPhen2 [≥ 0.85]) were considered. Variants were defined as rare (minor allele frequency $<5\%$) and common (minor allele frequency $\geq 5\%$) according to the reported frequency for European populations in the 1000 Genomes and ESP6500 datasets.
Sapkota et al., JNCI J Natl Cancer Inst, 2022	Blood	DNA concentration was fluorometrically measured using the Quant-it dsDNA Assay Kit and DNA integrity verified visually by agarose gel electrophoresis. Unique samples were barcoded and WGS was performed at the HudsonAlpha Institute for Biotechnology Genomic Services Laboratory using the Illumina HiSeq X10 sequencers to generate 360 million paired-ends reads, each 150 base pairs (bp) in length, for each sample.
Sapkota et al., AACR, 2021	Blood	DNA concentration was fluorometrically measured using the Quant-it dsDNA Assay Kit (Life Technologies cat#Q33130) and DNA integrity verified visually by agarose gel electrophoresis (E-gel, Life Technologies, cat#G8008-01). Unique samples were barcoded and WGS was performed at the HudsonAlpha Institute for Biotechnology Genomic Services Laboratory using the Illumina HiSeq X10 sequencers to generate 360 million paired-ends reads, each 150 base pairs (bp) in length, for each sample. Paired-end reads were aligned to the GRCh38 human reference using the Burrows-Wheeler Aligner (BWA-ALN v0.7.12) and data was cleaned following the best practice guide by the Genome Analysis Toolkit (GATK v3.4.0). 9.3 million common [minor allele frequency (MAF) ≥ 0.05] variants ; 10.2 million rare/low frequency (MAF < 0.05 and minor allele count ≥ 3) variants.

Abbreviations: ACT, Anthracycline-induced cardiotoxicity; SNP, Single Nucleotide Polymorphism; PCR, Polymerase Chain Reaction; rt-PCR, reverse-transcriptase PCR; qrt-PCR, quantitative real-time PCR; NA, Not Available.

Supplementary Table 3. Outcome definitions of the 19 included studies

Study	Study design	Definition
Sági et al., BMC Cancer, 2018	case-control	<p>LVEF and LVFS data were analyzed in follow-up categories: 1) at the diagnosis (used as a control); 2) in acute phase: during the intensive chemotherapy phase; 3) during oral maintenance chemotherapy; 4) at the end of the treatment, which is after the oral maintenance chemotherapy period completed 2 or 3 years after the diagnosis; 5) from the end of the treatment until 5 years after the diagnosis; 6) 5-10 years after the diagnosis; 7) 10–15 years after the diagnosis; 8) more than 15 years after the diagnosis;</p> <p>The worst heart function of each patient was used to define patients for the case-control type study;</p> <p>The alteration of LVFS was computed and analyzed as dichotomous variable, which was defined as the difference between the LVFS value at diagnosis and at the end of the treatment. In this study, patients with decreased LVFS were compared to those with increased LVFS;</p> <p>Cases definition: survivors with a LVFS ≤ 28% at any time point during the follow-up;</p> <p>Controls definition: survivors who received the same chemotherapy but never had a LVFS ≤ 28%.</p>
Blanco et al., JCO, 2012	matched case-control	<p>Cardiomyopathy definition: cardiac compromise by presenting with symptoms (eg, dyspnea, orthopnea, fatigue) and/or signs (eg, edema, hepatomegaly, rales) of cardiac decompensation or, in the absence of symptoms/signs, if they had LVEF ≤ 40% and/or LVFS ≤ 28%;</p> <p>Cases definition: CCS who developed cardiomyopathy after completion of cancer therapy and were alive at study participation;</p> <p>Controls definition: CCS without cardiomyopathy, who had no symptoms or signs of cardiac compromise and had normal echocardiographic features or no clinical indication for echocardiographic examination because of lack of exposure to anthracyclines or chest radiation.</p>
Visscher et al., Pediatr Blood Cancer, 2013	case-control	<p>ACT definition: early- or late-onset LV dysfunction measured by echocardiogram (LVFS) and/or symptoms requiring intervention based on CTCAE v3;</p> <p>Cases definition: CCS with a LVFS ≤ 26% at any time during or after anthracycline therapy;</p> <p>Controls definition: CCS with a LVFS ≥ 30% during and after therapy, with follow-up of more than 5 years after completion of anthracycline therapy;</p> <p>To exclude transient acute cardiotoxicity, only echocardiograms obtained 21 days or more after a dose were used.</p>
Singh et al., Cancer, 2020	matched case-control	<p>Cardiomyopathy definition: cardiac compromise by presenting with symptoms (eg, dyspnea, orthopnea, fatigue) and/or signs (eg, edema, hepatomegaly, rales) of cardiac;</p> <p>Cases definition: anthracycline-exposed CCS who developed cardiomyopathy;</p> <p>Controls definition: anthracycline-exposed CCS with no signs or symptoms of cardiomyopathy.</p>
Blanco et al., Cancer, 2008	nested matched case-control	<p>CHF ascertained through self-report by the study participants, validated by a telephone interview, during which the signs and symptoms of CHF as well as the use of medications for management of the CHF were reviewed. The responses from the script were reviewed independently by 2 physicians, and the information was used to validate CHF;</p> <p>Cases: CCS who survived at least 5 years after cancer diagnosis presenting CHF;</p> <p>Controls: CCS who survived at least 5 years after cancer diagnosis and did not present signs of CHF.</p>
Hildebrandt et al., Nature, 2017	matched case-control	<p>Cardiotoxicity not defined;</p> <p>Cases definition: CCS either with a LVEF 45 – 50% and presenting symptoms and other echocardiogram findings considered by a cardiologist to warrant cardiac medications; or with a LVEF ≤ 45% and/or LVFS ≤ 25% on at least two echocardiograms.</p> <p>All patients included because they were started on cardiac medications had chart review to assure that medications were</p>

		started for cardiac dysfunction and not hypertension without echocardiogram abnormalities. Patients who had a single low LVEF or LVFS that returned to LVEF > 55% without the use of cardiac medications were not included as cases; Controls definitions: CCS with LVEF >55% and LVFS > 28% with at least two echocardiograms obtained more than 5 years off treatment. A patient with a single discrepant low LVEF or LVFS with subsequent normal echocardiograms not on medications were considered a control.
Wang et al., JCO, 2022	case-control	Cardiomyopathy, CHF, heart transplantation, and medications were classified and graded using the CTCAE; Only those outcomes graded as severe (grade 3; self-reported cardiomyopathy or CHF, plus medications), life-threatening (grade 4; requiring heart transplantation), or fatal (grade 5) were included.
Aminkeng et al., Nat Genet. 2015	case-control	ACT definition: early- or late-onset LV dysfunction measured by echocardiogram (LVFS) and/or symptoms (dyspnea, orthopnea and/or fatigue) and/or signs (edema, hepatomegaly and/or rales) requiring intervention based on CTCAE v3; Cases definition: CCS with a grade ≥ 2 ACT, LVFS $\leq 24\%$ for asymptomatic ACT; Controls definition: CCS with no signs or symptoms of cardiac compromise at study participation (grade 0) and LVFS $\geq 30\%$ with ≥ 5 years of follow-up after the end of anthracycline treatment; To exclude transient acute cardiotoxicity, only echocardiograms obtained 21 days or more after a dose were used.
Visscher et al., Pharmacogenomics, 2015	case-control	ACT definition: early- or late-onset LV dysfunction measured by echocardiogram (LVFS) and/or symptoms requiring intervention based on CTCAE v3; Cases definition: CCS with a LVFS $\leq 26\%$ at any time during or after anthracycline therapy; Controls definition: CCS with a LVFS $\geq 30\%$ during and after therapy, with follow-up of more than 5 years after completion of anthracycline therapy; To exclude transient acute cardiotoxicity, only echocardiograms obtained 21 days or more after a dose were used.
Wang et al., JCO, 2016	matched case-control	Cardiomyopathy definition: cardiac compromise by presenting with symptoms (eg, dyspnea, orthopnea, fatigue) and/or signs (eg, edema, hepatomegaly, rales) of cardiac decompensation or, in the absence of symptoms/signs, if they had LVEF $\leq 40\%$ and/or LVFS $\leq 28\%$; Cases definition: CCS who developed cardiomyopathy after completion of cancer therapy and were alive at study participation; Controls definition: CCS without cardiomyopathy.
Wang et al., JCO, 2014	matched case-control	Cardiomyopathy definition: cardiac compromise by presenting with symptoms (eg, dyspnea, orthopnea, fatigue) and/or signs (eg, edema, hepatomegaly, rales) of cardiac decompensation or, in the absence of symptoms/signs, if they had LVEF $\leq 40\%$ and/or LVFS $\leq 28\%$; Cases definition: CCS who developed cardiomyopathy after completion of cancer therapy and were alive at study participation; Controls definition: CCS without cardiomyopathy.
Visscher et al., Journal of Clinical Oncology, 2012	case-control	ACT definition: early- or late-onset LV dysfunction measured by echocardiogram (LVFS) and/or symptoms requiring intervention based on CTCAE v3.20; Cases definition: CCS with a LVFS $\leq 26\%$ at any time during or after anthracycline therapy; Controls definition: CCS with a LVFS $\geq 30\%$ during and after therapy, with follow-up of more than 5 years after completion of anthracycline therapy; To exclude transient acute cardiotoxicity, only echocardiograms obtained 21 days or more after a dose were used.
Chaix et al., JACC: Cardiooncology, 2020.	nested case-control	Cardiotoxicity definition: LVEF $\leq 50\%$ or $>10\%$ LVEF decline to $\leq 55\%$ from a previous echocardiogram during follow-up; Cases definition: patients who received low cumulative anthracycline dose (≤ 250 mg/m ²) but developed either: 1) clinically defined cardiotoxicity; or 2) low LVEF of $\leq 55\%$ based on American Society of Echocardiography guidelines; Controls definition: patients with preserved cardiac function (LVEF of $>55\%$) despite high dose anthracycline (>250 mg/m ²).

Krajinovic et al., The Pharmacogenomics Journal, 2016	cohort	Only LVFS and LVEF were analyzed, as these two parameters were significantly associated with genotypes in the ALL group. LV volumes were used for LVEF calculation and were estimated from diastolic and systolic dimensions using Teichholz' formula. Measurements obtained at ≥ 3 years after diagnosis were taken for genotype–phenotype comparison corresponding to the time between diagnosis and echocardiography of the ALL cohort.
Semsei A., Cell Biol. Int, 2012	Cohort	Patients were followed-up by echocardiography to assess LV function by measuring left ventricular end-diastolic-dimension and left ventricular end-systolic dimension. LVFS was calculated from these 2 data; LVFS was analysed at 3 time-points, first, at the time of diagnosis, secondly at the end of the treatment, that is, at a median of 2 years after diagnosis, while the third data-point was determined at the time of the latest follow-up.
Lipshultz S., Cancer, 2013	Cohort	Echocardiograms were obtained at diagnosis, after doxorubicin therapy, and every 2 years thereafter at local treatment sites and centrally re-measured at a single facility by study staff blinded to treatment status; Left ventricular status was assessed with LV endsystolic and end-diastolic dimensions; LV mass; LV end-systolic and end-diastolic posterior wall thicknesses; LV thickness-to-dimension ratio; and LV fractional shortening, an index of LV systolic performance influenced by heart rate, LV preload, LV afterload, and LV contractility; Echocardiographic measurements were standardized with Z-scores. Z-scores were calculated from the difference between LV outcome values in patients and known values in healthy children, divided by the standard deviation of a distribution of values in healthy children.
Petrykey et al., Pharmacogenomics, 2021	cohort	Echocardiographic parameters for the quantitation of LV morphology and function included the LV end-diastolic diameter, LVFS) and LVEF. Two measurements of the LVEF were studied – LVEF M-mode and LVEF 2D (Simpson).
Sapkota et al., JNCI J Natl Cancer Inst, 2022	cohort and case-control	LVEF as a quantitative trait, according to the American Society of Echocardiography guidelines; When multiple EF measurements were available, we used the lowest LVEF of each survivor to reflect the most severe form of CCD known for the survivor; In SJLIFE patients, CCD was clinically assessed based on LVEF and pharmacologic treatment. Using modifications of the National CTCAE v4.03, CCD was classified as moderate (grade 2: resting LVEF < 50%-40% or 10%-19% absolute drop from baseline), severe or disabling (grade 3: resting LVEF < 39%-20%, or >20% absolute drop from baseline, or medication initiated), life-threatening (grade 4: resting LVEF < 20%; refractory or poorly controlled heart failure; intervention such as ventricular assist device, intravenous vasopressor support; or heart transplant indicated), or fatal (grade 5: death); In CCSS, participants completed a multi-item questionnaire at baseline and follow-up surveys. Using the CTCAE, CHF was graded as moderate (grade 2: self-reported CHF not requiring medication), severe or disabling (grade 3: cardiomyopathy of CHF requiring medication), life-threatening (grade 4: cardiac transplantation), or fatal (grade 5: death as a result of heart failure).
Sapkota et al., AACR, 2021	cohort	LVEF as a quantitative trait, according to the American Society of Echocardiography guidelines; cardiomyopathy was clinically assessed and graded using a modified version of the National Cancer Institute's CTCAE v4.03, and classified as mild (grade 1), moderate (grade 2), severe or disabling (grade 3), life- threatening (grade 4), or fatal (grade 5).

Abbreviations: ACT, Anthracyclines-Induced Cardiotoxicity; LV: Left-ventricle; LVFS, Left-Ventricular Fractional Shortening; LVEF, Left-ventricular Ejection Fraction; CTCAE, Common Terminology Criteria for Adverse Events; CHF, Congestive Heart Failure; CCD, Cancer treatment–induced cardiac dysfunction; NA, Not Available.

Supplementary Table 4. Statistical methods and selection procedure of the 19 included studies

Study	Sequencing method	Statistical Analysis	Variables for Matching / adjustment	Selection procedure	Correction for multiplicity	Performance evaluation	Validation of the results
Sági et al., <i>BMC Cancer</i> , 2018	Candidate gene	Logistic regression for case-control analysis	Adjustment: age at the time of diagnosis, gender, chemotherapy protocols, risk groups, and cumulative dose of anthracycline	NA	Benjamini-Hochberg FDR, $p \leq 8.90 \times 10^{-3}$	NA	NA
		Linear regression for follow-up analysis					
Blanco et al., <i>JCO</i> , 2012	Candidate gene	Conditional logistic regression	Matching: cancer diagnosis, year of diagnosis (± 5 years), race/ethnicity, and duration of follow-up for controls to exceed time from cancer diagnosis to cardiomyopathy for index patient case Adjustment: cumulative ANT dose, sex, age at cancer	NA	Bonferroni correction for multiple comparisons ($p < 0.025$) when analyzing the association with the two SNPs.	NA	NA
Visscher et al., <i>Pediatr Blood Cancer</i> , 2013	Candidate Gene	Single SNP Logistic regression	Adjustment: Cumulative anthracycline dose, age at start of treatment, gender and radiation therapy to the heart included as covariates	NA	NA	ROC curve and -statistic (area under the curve, AUC) calculation	Tested in the replication cohort (Dutch-EKZ and Canadian-CPNDS patients combined)
		Logistic regression	including multiple genetic variants and/or clinical variables were trained in the original cohort	Step-wise regression with forward selection: association with ACT at $p < 0.01$ were retained in the final model			
Singh et al., <i>Cancer</i> , 2020	Candidate gene	Conditional logistic regression	Matching: cancer diagnosis, year of diagnosis (± 5 years), race/ethnicity, and duration of follow-up for controls to exceed time from cancer diagnosis to cardiomyopathy for index patient case Adjustment: anthracycline dose, chest radiation, age at diagnosis of primary cancer, and sex	NA	NA	NA	Gene expression analyses on cardiomyocytes and peripheral blood

Blanco et al., Cancer, 2008	Candidate gene	Conditional logistic regression	Matching: age at diagnosis (± 5 years), race/ethnicity, anthracycline dose, radiation to the heart (matched for 81% of controls; this factor was included in all models to adjust for incomplete matching) and duration of follow-up for controls to exceed time from cancer diagnosis to cardiomyopathy for index patient case Adjustment: sex, smoking history, first recurrence of original disease, heart in radiation beam, and family history of heart disease	NA	NA	NA	NA
Hildebrandt et al., Nature, 2017	Candidate gene	Logistic regression	Matching; age at diagnosis, gender, anthracycline dose, chest radiation, and cancer type Adjustment: follow-up time, age at diagnosis, gender, race, hypertension, anthracycline dose, chest radiation, and cancer site	Gene associated with cardiotoxicity at $p \leq 0.05$	NA	NA	Expression of the identified genes was analyzed in cardiomyocytes
Wang et al., JCO, 2022	GWAS	Cox regression	Adjustment: 10 genotype-based principal components, age at diagnosis of primary cancer, sex, chest radiation, anthracycline dose, and presence of cardiovascular risk factors (and and SNP-anthracycline interactions for step 2)	Gene associated with cardiotoxicity at $p \leq 5 \times 10^{-8}$	NA	NA	Replication on matched case-control design to verify significant SNP and SNP-anthracycline interactions identified in the discovery stage
Aminkeng et al., Nat Genet., 2015	GWAS	Logistic regression	Adjustment: age at the start of treatment, cumulative anthracycline dose, tumor type and cardiac radiation therapy	Association with ACT: $p < 1 \times 10^{-5}$ for the discovery cohort and $p < 0.006$ on the replication cohort, and $p < 0.05$ in the independent non-European population	Multiple testing correction $p < 1 \times 10^{-5}$ for the discovery cohort; Bonferroni correction $p < 0.006$	NA	Markers that reached significance threshold in the discovery cohort were tested for replication (European-Dutch patients) and the SNPs significant after Bonferroni correction in the replication

					for the replication cohort		cohort were tested in non-European population, and in combined analysis of all patients with European ancestry in discovery and replication cohorts
Visscher et al., Pharmacogenomics, 2015	GWAS	Logistic regression	Adjustment: cumulative anthracycline dose, age at start of treatment, gender, radiation therapy, and the first two principal components, and SLC28A3 (rs7853758) and UGT1A6 (rs17863783)	Tiered analysis to identify SNPs associated with ACT at $p < 0.01$ in the discovery cohort that remained associated in the replication cohort at $p < 0.05$	Bonferroni corrected $p < 1.7 \times 10^{-5}$	ROC curve and -statistic (area under the curve, AUC) calculation	Tested in the replication cohort (Dutch-EKZ and Canadian-CPNDS patients combined)
		Logistic regression	including multiple genetic variants and/or clinical variables were trained in the original cohort	For SNPs not identified in the tiered analysis, $p < 0.005$ in the combined cohort was considered suggestive evidence			
Wang et al., JCO., 2016	GWAS	Conditional logistic regression	Matching: cancer diagnosis, year of diagnosis (± 5 years), race/ethnicity, and duration of follow-up for controls to exceed time from cancer diagnosis to cardiomyopathy for index patient case Adjustment: cumulative ANT dose, RT exposition, sex, and age at cancer (step 1), and SNP-anthracycline interactions (Step 2)	Association with cardiomyopathy: $p < 0.004$	Multiple testing correction: $p < 7.77 \times 10^{-5}$	NA	Replication on case-only design, to verify significant SNP-anthracycline interactions identified in the discovery stage Gene expression analysis on heart samples
Wang et al., JCO, 2014	GWAS	Conditional logistic regression	Matching: cancer diagnosis, year of diagnosis (± 5 years), race/ethnicity, and duration of follow-up for controls to exceed time from cancer diagnosis to cardiomyopathy for index patient case Adjustment: cumulative ANT dose, total RT dose, sex, age at cancer and SNP-anthracycline interactions	Association with cardiomyopathy: $p < 5 \times 10^{-6}$	Multiple testing correction $p < 5 \times 10^{-6}$	NA	Replication on case-only design, to verify significant SNP-anthracycline interactions identified in the discovery stage Gene expression analysis on heart samples
Visscher et al., JCO, 2012	GWAS	Logistic regression	Adjustment: Cumulative dose, age, sex, and radiation	Tiered single-marker analysis to identify variants	The overall significance	ROC analysis was performed for each sample from the	The top SNP was further evaluated in the Dutch-EKZ

			therapy to the heart included as covariates as well as the first two principal components assuming an additive model	associated with ACT in the discovery cohort ($p < 0.01$) that remained associated in the Canadian replication cohort ($p < 0.01$) Using step-wise logistic regression with forward selection including covariates, SNPs were retained with $p < 0.01$ in the final multi-SNP model	threshold accounting for multiple testing was calculated at 1.5×10^{-4} using the simpleM correction	multimarker regression model.	cohort and in all three cohorts combined. SNPs with $p < 0.01$ in the single-SNP test for the combined Canadian cohort were considered for potential inclusion in a multimarker risk prediction model
Chaix et al., JACC : Cardiooncology , 2020	WES	Burden CMC. SKAT. SKAT-O. Logistic regression for associated genes	Adjustment: sex, cancer diagnosis, age at the start of anthracycline, use of dexrazoxane, chest radiation, duration of follow-up from first anthracycline dose, and the first 2 principal components inferring ethnicity	Gene association with cardiotoxicity $p < 0.001$ by at least 2 gene-collapsing methods and genes in biologically relevant pathways with an association of $p < 0.001$ by at least 1 method	Bonferroni correction: $p < 2.87 \times 10^{-6}$	Random forest model on the discovery cohort. ROC AUC, sensitivity, specificity, positive predictive value, negative predictive value, false positive rate, false negative rate and misclassification rate calculation	Replication on 1:1 propensity-matched case-control design, comprising a subset of patients with recent enrollment in the PCS2 study
Krajinovic et al., The Pharmacogenomics Journal, 2016	Candidate gene	Analysis of variance or t-test	NA	Either an association sustained multiple testing correction ($FDR \leq 0.05$) with at least one of the tested parameters, or association was preferentially seen in high-risk patients who received higher cumulative doxorubicin doses	multiple testing correction ($FDR \leq 0.05$)	NA	Findings were validated in set of 44 ALL patients who underwent treatment with DFCI 95-01 protocol
Semsei A et al., Cell Biol. Int, 2012	Candidate gene	Linear regression	Adjustment: Gender, age at diagnosis, clinical centre, total anthracyclines dose, dexrazoxane, chemotherapy protocol.	Gene association with LVFS $p < 0.005$	Bonferroni correction $p < 0.005$	NA	NA
Lipshultz S et al., Cancer. 2013	Candidate gene	t-tests or Wilcoxon rank-sum tests	The Z-scores for LV mass, LV end-systolic and end diastolic posterior wall thicknesses, and LV dimensions were adjusted for body-surface area, and Z-scores for LV fractional shortening were adjusted for age at echocardiography.	0.05	NA	NA	NA

Petrykey et al., Pharmacogenomics, 2021	WES	Wald statistic and linear regression for common variants	Adjustment: age at the time of diagnosis, time since the end of treatment; sex; treatment protocol; and a risk-treatment scores	Common variants in genes involved in cardiac system functioning p < 0.0006	Benjamini–Hochberg procedure for FDR was p < 0.05 Bonferroni p-value corrected p < 0.0006	NA	Analyses of top-ranking rare variants in ALL survivors of European ancestry enrolled in the SJLIFE cohort
Sapkota et al., JNCI, 2022	WGS	Linear regression	Adjustment: sex, age at cancer diagnosis, age at last follow-up, cumulative anthracycline dose, average heart radiation dose, and the top 20 genotype-based principal components	p < 5 x 10 ⁻⁸	NA	NA	Variants that met the statistical significance thresholds in the discovery analysis were tested in 2 replication cohorts including 301 survivors of African ancestry from the SJLIFE (for LVEF and CCD) and an additional 4020 of European ancestry from the CCSS (for self-reported CHF)
		Logistic regression		P < 0.05 considered statistically significant	NA	NA	
		SKAT-O for rare variants	NA	p < 0.05	Multiple test adjustments with a cutoff value of FDR < 0.05		
Sapkota et al., AACR, 2021	WGS	Linear regression for common variants	Adjustment: sex, age at cancer diagnosis, age at last follow-up, cumulative anthracycline dose, average heart radiation dose and the top 10 principal components	p < 5 × 10 ⁻⁸ considered significant at the genome-wide level	NA	NA	Findings were replicated on 1645 patients from the SJLIFE survivors of European ancestry exposed to cardiotoxic therapies (radiotherapy to the heart and anthracyclines)
		SKAT-O for rare and low-frequency variants	NA	p < 7.3 × 10 ⁻⁸	NA		
		Logistic regression for both common, low-frequency and rare variants	Matching on PS deciles. PS calculation: by regressing each SNP on all the sex, age at cancer diagnosis, age at last follow-up, cumulative anthracycline dose, average heart radiation dose and the top 10 principal components	NA	NA		

Abbreviations: ACT, Anthracycline-induced cardiotoxicity; SNP, Single-nucleotide polymorphism; CHF, Congestive heart failure; CCD: Cancer treatment–induced cardiac dysfunction; CMC, Combined Multivariate and Collapsing; SKAT, Sequence kernel association test; SKAT-O, sequence kernel association test-optimized; ALL, Acute lymphocytic leukemia; SJLIFE, St. Jude Lifetime; DFCI, Dana-Farber Cancer Institute; PCS2: Preventing Cardiac Sequelae in Pediatric Cancer Survivors; CCSS, Childhood Cancer survivor Study; FDR, False discovery rate; PS, Propensity Score; NA, Not Available

Supplementary Table 5. Subgroup analyses of the 19 selected studies

Study	Subgroup of Analysis	Cardiac outcome	Gene/ region	Variant	Effect allele	Effect size : OR (95% CI)	p-values
Sági et al., <i>BMC Cancer</i> , 2018	Treated for ALL	LVFS decrease from diagnosis to last follow-up	NQO1	rs1043470	CC /CT + TT	0.44 (0.24–0.81)	8.9 x 10 ⁻³
		Cardiotoxicity	CYP3A5	rs4646450	TT	6.56 (1.68 - 25.71)	7.0 x 10 ⁻³
	Males group	Cardiotoxicity	CYP3A5	rs4646450	TT	13.45 (2.26 - 80.1)	4.0 x 10 ⁻³
	Intermediate risk patients	Cardiotoxicity	CYP3A5	rs4646450	TT	23.34 (4.46 - 122.07)	2.0 x 10 ⁻⁴
	Recidive occurred	Cardiotoxicity	CYP3A5	rs4646450	TT	48.00 (2.96 – 778.53)	6.0 x 10 ⁻³
	Treated for ALL and OSC	Cardiotoxicity	CYP3A5	rs4646450	TT	6.44 (1.66 - 24.97)	0.007
	Treated for ALL	LVFS decrease: 2 - 5 years after diagnosis	ABCC1	rs35626	GG/ GT + TT	41.0 vs. 39.0 ^a	7.9 x 10 ⁻³
		LVFS decrease: 5 - 10 years after diagnosis	ABCC2	rs3740066	GG / GA/ AA	39.5 vs. 40.8 vs. 42.9 ^a	4.5 x 10 ⁻³
		LVFS decrease: 10- 15 years after diagnosis	ABCB1	rs9282564	AA/ AG + GG	41.5 vs. 37.9 ^a	2.5 x 10 ⁻³
		SLC28A3	rs7853758	GG/ GA + AA	41.3 vs. 38.4 ^a	4.80 x 10 ⁻³	
		SLC28A3	rs885004	GG/ GA + AA	41.3 vs. 38.0 ^a	2.50 x 10 ⁻³	
Visscher et al., <i>Pediatr Blood Cancer</i> , 2013	European ancestry	ACT grade 2 - 4	UGT1A6	rs17863783	A	6.22 (2.52 - 15.4)	1.1 x 10 ⁻⁴
	Received anthracyclines , and heart radiations	ACT	UGT1A6	rs17863783	A	7.98 (1.85–34.4)	0.0062
	Male group		SULT2B1	rs10426377	A	0.35 (0.20–0.64)	2.1 x 10 ⁻⁴
Singh et al., <i>Cancer</i> , 2020	Non-Hispanic white	Cardiomyopathy	GSTM1	Null / Positive		0.67 vs. 1.33 ^c	0.49
Aminkeng et al., <i>Nat Genet.</i> , 2015	European ancestry	ACT	RARG	rs2229774	A	5.4 (2.9–10.3)	7.8 x 10 ⁻⁸
	Canadian from European ancestry		RARG	rs2229774	A	7.0 (2.9–17.0)	5.0 x 10 ⁻⁶
	Dutch from European ancestry		RARG	rs2229774	A	4.1 (1.5 – 11.5)	0.0043
Visscher et al., <i>Pharmacogenomics</i> , 2015	Exposed only to doxorubicin or daunorubicin	ACT	SLC22A17	rs4982753	A/G	0.39 (0.19 – 0.81)	0.0071
			SLC22A7	rs4149178	G/A	0.39 (0.14–1.05)	0.0034
			SLC22A17	rs4982753	A/G	0.50 (0.33–0.75)	4.4 x 10 ⁻⁴
			SLC22A7	rs4149178	G/A	0.45 (0.26–0.75)	0.0013
			SULT2B1	rs10426628	A	1.92 (1.19–2.50)	3.2 x 10 ⁻⁴
			ABCC9	rs11046217	C	2.67 (0.24–0.79)	9.9 x 10 ⁻⁴
			ABCA1	rs3887137	A	2.28 (1.32–3.77)	9.5 x 10 ⁻⁴
	Exposed to anthracyclines	ACT	SLC22A17	rs4982753	A/G	0.53	6.7 x 10 ⁻⁴
			SLC22A7	rs4149178	G/A	0.45	9.4 x 10 ⁻⁴
Wang et al., <i>JCO.</i> , 2016	All racial / ethnic backgrounds	High-dose anthracyclines (Yes/ No)	CELF4	rs1786814	GG/ AA + GA	5.09 (1.03 -25.23)	0.046
Wang et al., <i>JCO</i> , 2014	All racial / ethnic backgrounds	High-dose anthracyclines (Yes/ No)	HAS3	rs2232228	GA/ GG	3.6 (0.9 - 15.0)	0.07
			HAS3	rs2232228	AA/ GG	4.5 (1.1 - 18.7)	0.04
			HAS3	rs2232228	GA/ GG	2.6 (1.0 - 6.9)	0.05
			HAS3	rs2232228	AA/ GG	3.7 (1.3 - 10.2)	0.01

Krajinovic et al., The Pharmacogenomics Journal, 2016	High-risk ALL	LVEF	NOS3	rs1799983	TT/ GG + TG	62.2% vs 58.3% ^a	0.02
	High-risk + No dexrazoxane before DOX	LVEF	NOS3	rs1799983	TT/ GG + TG	64.1% vs 56.8% ^a	0.002
Petrykey et al., Pharmacogenomics, 2021	Standard Risk ALL	LVEF	CBR1	rs2835266			0.0015
		LVFS	ABCG2	rs199473672			0.0027
		LVFS	CBR1	rs2835266			0.0037
		LVEF	ABCG2	rs45605536			0.0019
		LVEF	AKR1C3	rs200981816			0.0067
	High Risk ALL	LVEF	ABCC5	chr3:183681203			0.0008
		LVFS	ABCC5	chr3:183681203			0.0014
Sapkota et al., JNCI , 2022	European ancestry grade ≥ 3	CCD / CHF	KCNK17	rs2815063	A	1.57 (1.19 to 2.09)	0.002
	European ancestry exposed to high therapy based risk	CCD / CHF	KCNK17	rs2815063	A	1.80 (1.17 - 2.77)	0.008
	European ancestry exposed to moderate therapy based risk	CCD / CHF	KCNK17	rs2815063	A	1.57 (1.03 - 2.40)	0.035
	European ancestry exposed to moderate therapy based risk	CCD / CHF	KCNK17	rs2815063	A	1.73 (1.08 - 2.79)	0.024
	European ancestry, exposed to doxorubicin only	CCD / CHF	KCNK17	rs2815063	A	2.97 (1.37 - 6.43)	0.006
	European ancestry	LVEF	KCNK17	rs2815063	A	- 0.016 (0.003) ^b	2.1×10^{-8}
			KCNK17	rs115537302	A	-0.019 (0.004) ^b	8.7×10^{-8}
			KCNK17	rs115057884	C	-0.019 (0.004) ^b	9.0×10^{-8}
			KCNK17	rs116181858	G	-0.019 (0.004) ^b	9.6×10^{-8}
			KCNK17	rs114566901	C	-0.019 (0.004) ^b	1.7×10^{-7}
	Exposed to high therapy based risk	LVEF	KCNK17	rs2815063	A	-0.017 (0.005) ^b	9.6×10^{-4}
	African ancestry	LVEF	KCNK17	rs2815063	A	- 0.015 (0.005) ^b	0.004
	Exposed to anthracyclines and radiotherapy to the heart	LVEF	SLC22A7	rs4149178	G	0.008 (0.003) ^b	7.9×10^{-3}
Exposed to anthracyclines only	CCD	RARG	rs2229774		2.08 (0.36)	0.041	
Exposed to anthracyclines only ($\geq 250\text{mg}/\text{m}^2$)	LVEF	HAS3	rs2232228	AA	-0.32% ^d	0.04	
Sapkota et al., AACR, 2021	European ancestry	Cardiomyopathy (grade 2-4)	1p13. 2	rs6689879	C / T	1.24 (0.98 - 1.56)	0.071
		Cardiomyopathy (grade 3-4)				1.31 (0.88 - 1.94)	0.19
		LVEF	1p13. 2	rs6689879	C / T	-0.004 (0.002)	0.042
			15q25. 3	Monomorphic		-	-
	Not exposed to anthracyclines neither to radiotherapy to the heart	LVEF	1p13. 2	rs6689879	C / T	-0.027 (0.018)	0.15
			15q25. 3	rs9788776	G / A	-0.04 (0.036)	0.25
			1p13. 2	rs6689879	C / T	0.007 (0.007)	0.28
	Exposed to anthracyclines only	LVEF	1p13. 2	rs6689879	C / T	-0.03	0.044
			15q25. 3	rs9788776	G / A	-0.044	0.022
	Exposed to radiotherapy only	LVEF	1p13. 2	rs6689879	C / T	-0.024	0.14
15q25. 3			rs9788776	G / A	-0.081	5.4×10^{-4}	

Abbreviations: ACT, Anthracyclines-Induced Cardiotoxicity; LVEF, Left-ventricular ejection fraction; LVFS, Left-ventricular fractional shortening; CHF, Congestive Heart Failure; ALL, Acute Lymphoblastic Leukemia; OSC, Osteosarcoma; CCD, Cancer treatment–induced cardiac dysfunction; ^aAllele frequency; ^bBeta (SE) or Beta (95%CI); ^cMean expression Cases vs. Controls; ^dEffect size measured by echo-cardiac parameters.

Genetic biomarkers related to cardiac disease occurrence in childhood cancer survivors: a matched case-control study nested in the French Childhood Cancer Survivors Study

Running Head : Genetic Biomarkers in Cardiac Disease in Childhood Cancer Survivors

Naïla Aba, PhD^{1,2}, Gwénaél Le Teuff, PhD^{2,3,4}, Brice Fresneau, MD, PhD^{1,2,5}, Serge Koscielny, PhD⁴, Shaima Belhechmi, PhD^{2,3}, Boris Schwartz, MSc^{1,2}, Chiraz El-Fayech, MD⁵, Carole Rubino, PhD^{1,2}, Rodrigue S. Allodji, PhD^{1,2}, Eric Morel, PhD⁶, Delphine Dayde, PhD⁶, Pierre De la Grange, PhD⁷, Ariane Jolly, MSc⁷, Noémie Pata-Merci, MSc⁸, Camille Cordero, MD⁹, Isabelle Aerts, MD⁹, François Pein, MD¹⁰, Giau Vu-Bezin, MSc^{1,2}, Florent De Vathaire, PhD^{1,2}, Nadia Haddy, PhD^{1,2}

Affiliations:

1. CESP, INSERM U 1018, Radiation Epidemiology team, Villejuif, France.
2. Université Paris-Saclay, UVSQ, Inserm, CESP, 94807, Villejuif, France
3. Université Paris-Saclay, CESP, INSERM U1018 Oncostat, labeled Ligue Contre le Cancer, Villejuif, France.
4. Bureau de Biostatistique et d'Epidémiologie, Gustave Roussy, Villejuif, France.
5. Department of Childhood and Adolescent Oncology, Gustave Roussy, Villejuif, France.
6. Université Paris-Saclay, Inserm, UMR-S 1180, Orsay, France.
7. GenoSplice, Paris, France.
8. Institut Gustave Roussy.
9. PSL research university, Institut Curie, SIREDO oncology center, Paris, France.
10. Institut de Cancérologie de l'Ouest ICO Centre René Gauducheau, Saint-Herblain (Nantes), France.

Acknowledgments: The authors acknowledge The Paris-Saclay University, The French Ligue Against Cancer (“Ligue Nationale contre le cancer), The “Fondation ARC”, The INCA (Institut National de Cancer) and the PHRC-K Program

Corresponding author: Nadia Haddy (nadia.haddy@gustaveroussy.fr)

Prior presentations: This work has been presented at:

- the French-speaking conference on EPIdémiologie CLINique- EPICLIN 2023, Nancy, France, May 10-12, 2023
- the International Society of Radiation Epidemiology and Dosimetry – ISORED 2023, Sitges, Spain, May 16-18, 2023
- the International Society for Clinical Biostatistics- ISCB conference 2023, Milan, Italy, 27 – 31 August 2023

Abstract

PURPOSE: Cardiac diseases (CD) are the most common non-cancerous long-term adverse effects in childhood cancer survivors (CCS), leading to long-term morbidity and early mortality. While cardiotoxic treatments like anthracyclines and heart-directed radiotherapy contribute to this risk, individual differences in CD incidence remain unexplained. This study aims to identify genetic expression biomarkers associated with CD occurrence after childhood cancer treatment.

PATIENTS AND METHODS: We conducted a matched case-control study nested in the French Childhood Cancer Survivor Study (FCCSS) cohort. The research involved 165 CCS with a validated CD (cases) matched to 165 CD-free control. The expression of 8,739 genes was investigated to select genes associated with CD and heart failure (HF). Three stabilization approaches derived for the conditional logistic regression with the Lasso procedure (Percentile lasso, Bolasso, and Sublasso) were used, and the intersection of the three selected gene sets formed the final selection. The interactions between cancer treatment doses and selected genes were investigated.

RESULTS: One promising gene, NFE2L2, constituted the final selection, and its expression was lower in cases than in controls (CD: Odds Ratio=0.16 (0.09-0.29), HF: OR=0.11 (0.03-0.37)). No interaction between treatment doses and NFE2L2 expression levels was found in our study. Incorporating NFE2L2 gene expression into prognostic models improved discrimination between cases and controls compared to models based solely on clinical and treatment variables (CD: AUC= 0.85 vs. 0.66, HF: AUC= 0.87 vs. 0.77).

CONCLUSION: Using high-dimensional data selection methods enabled us to identify the gene NFE2L2, associated with CD and HF in CCS. Further research is needed to validate this finding and achieve a better understanding of the biological mechanisms leading to cardiac toxicities, and so to develop risk-adapted treatment and surveillance strategies.

Keywords: childhood cancer survivors, cardiac diseases, cardiomyopathy, transcriptome, gene expression, case-control, conditional logistic regression, lasso

1 **Introduction**

2 Annually, approximately 385,500 children and young adults are diagnosed with cancer worldwide¹.
3 Significant progress in treatments has drastically increased the 5-year overall survival rate for these
4 young patients, to over 80% in the modern era^{2,3}. However, the future of childhood cancer survivors
5 (CCS) raises important questions regarding long-term adverse effects. Among these, cardiac diseases
6 (CDs) emerge as the leading non-cancer cause of death among CCS, placing them at a sevenfold higher
7 risk of dying from cardiovascular disease compared to the general population⁴. Furthermore, CCS face
8 a significantly increased risk of mortality after a major cardiovascular event⁵.

9 Nearly 20% of CCS treated with anthracyclines experience cardiotoxic effects, with a risk of
10 cardiovascular events 3.5 times higher than those untreated⁶. Higher doses, particularly those exceeding
11 360 mg/m², significantly increase CD risk. Additionally, radiotherapy is associated with developing late-
12 onset CDs⁷. The incidence of CDs is higher in patients receiving high radiation doses^{8,9} with a
13 significantly increased risk for a dose greater than 15 Gray (Gy)^{10,11}.

14 Although these treatments are strongly associated with CDs, they do not entirely explain the individual
15 variability of the incidence of the pathologies related to CDs in CCS treated similarly¹², suggesting a
16 potential role of genetic predispositions^{9,13}. Several studies have also investigated genetic biomarkers,
17 focusing on (i) genetic characteristics of anthracycline-induced cardiotoxicity¹⁴ and (ii) prognostic
18 markers associated with this cardiotoxicity¹⁵⁻²². Although several genetic variants were identified in
19 association with CDs, no consensus has been reached. Unlike the genome, the transcriptome is highly
20 dynamic and sensitive to xenobiotic exposures, including anticancer treatments that can significantly
21 affect gene expression profiling²³. The gene expression study may provide new insights into the
22 mechanisms of cardiotoxicity in patients treated with anthracyclines and radiotherapy involving the
23 heart. To date, only one study has analyzed transcriptomic data of 104 patients identifying genes
24 differentially expressed between CCS with cardiomyopathy and controls²⁴.

25 This study aims to identify potential transcriptomic biomarkers associated with the occurrence of CDs
26 after childhood cancer treatment, using statistical methods developed for high-dimensional data.

27 **Methods**

28 *Study design and population*

29 This study is based on data from the French Childhood Cancer Survivors Study (FCCSS), a retrospective
30 cohort of 5-year CCS (n=7,670) diagnosed with solid cancer or lymphoma before the age of 21 between
31 1946 and 2000 in five French cancer centers^{25, 26} (**Table S1**). Clinical and treatment details were
32 extracted from medical records, including cumulative anthracycline doses. Heart radiation doses were
33 retrospectively estimated, and dose-volume metrics were reconstituted²⁷. Further data collection details
34 are in prior studies^{27,28}. CD diagnosis, including heart failure (HF), valvular disease, ischemic pericardial
35 disease, and arrhythmia diseases, was validated based on criteria from the European Society of
36 Cardiology²⁹ and the Framingham studies³⁰. CDs assessment details are in **Appendix 1**. CDs were
37 graded using the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE)
38 version 4.03. Biological samples were collected to establish a biobank linked to the FCCSS, and
39 lymphocyte cells were obtained for 1,571 patients (**Appendix 2**). Of these, 165 CCS with transcriptomic
40 data who developed a CD (cases) were matched to 165 CD-free controls on gender, age at cancer
41 diagnosis (± 3 years), first tumor type, and follow-up duration (each control has a follow-up higher or
42 equal to that of its case) (See flowchart **Figure 1**). The study's protocol was approved by a regional
43 ethics committee and by the national committee protecting data confidentiality (Commission Nationale
44 de l'Informatique et des Libertés), following French regulations. Written informed consent was obtained
45 from all patients, parents, or guardians, in line with national ethics standards.

46 *RNA extraction*

47 Peripheral blood mononuclear cells were isolated from peripheral blood, and cultures of lymphocyte
48 cell lines were made. RNA extraction was performed on lymphocyte cells of the 165 cases and their
49 controls using the TRIzol Reagent method (ThermoFischer, Invitrogen). To determine RNA
50 concentration, a Nanodrop spectrophotometer (www.nanodrop.com) was used, and the quality of RNA
51 preparations was further assessed using Lab-on-a-chip Bioanalyser 2100 technology (Agilent
52 technologies). The concentration of all RNAs was adjusted and verified with a second measurement on
53 a Nanodrop spectrophotometer.

54 ***Gene expression profiling***

55 The transcriptome of the 330 patients was obtained for over 33,000 genes. Analysis of the Affymetrix
56 Human ClariomD Array was conducted by GenoSplice (Paris, France). Only the 8,739 genes that passed
57 the quality control were included in the analyses. **Appendix 3** and **Figure S1** describe the different steps
58 involved in gene expression profiling.

59 ***Clinical endpoints***

60 The primary endpoint was the occurrence of the first event related to CDs, whatever the grade, defining
61 the cases. The secondary endpoint was the occurrence of HF. A sensitivity analysis was performed by
62 restricting to patients with a first grade ≥ 2 event related to CDs.

63 ***Statistical analysis***

64 To identify potential gene expression biomarkers associated with the clinical endpoints in the setting of
65 high dimensional data and considering the nature of the design (case-control matching), the Lasso (Least
66 Absolute Shrinkage and Selection Operator) selection procedure adapted to a conditional logistic
67 regression (CLR) was used^{31, 32}. Standardized gene expression biomarkers were included in CLR with
68 Lasso and without covariates adjustment. Three stabilization methods were considered to address the
69 instability of the optimal penalty term (λ) estimation in the Lasso and thus on the biomarkers selection³³,
70 ³⁴ (**Figure S2**). Percentile Lasso uses the θ th percentile of the distribution of the penalization parameter
71 $\hat{\lambda}(\theta)$ estimated from several runs of CLR with Lasso as the optimal penalty term. Bolasso and Sublasso,
72 respectively, generate bootstraps and sub-samples from the original dataset, applying CLR with Lasso
73 to each and selecting genes with a selection frequency higher or equal to a pre-defined cut-off (Details
74 in **Appendix 4**). Different cut-offs (50, 60, and 70%) were investigated for instability study. The final
75 gene selection is defined as the intersection of the sets of genes resulting from the Percentile lasso with
76 $\theta= 95\%$ (recommended parameter), and the Bolasso and Sublasso using a cut-off of 70%. The selected
77 genes' effects were also estimated in a CLR adjusted on the year of cancer diagnosis (before 1970, [1970
78 – 1979], [1980 – 1989], [1990 – 2000]), cumulative anthracycline doses (none,]0 – 250],]250 – 360],
79 >360 mg/m²) and mean heart radiation (MHR) doses (none,]0 - 5],]5 -15], >15 Gy). The gene
80 expression biomarkers were analyzed as continuous standardized variables since no deviation of the log-
81 linearity hypothesis was observed through the Box-Tidwell test (data not shown). The correlation and

82 interaction between the selected genes and the two treatment variables were investigated. The added
83 value of the selected genes was quantified through the likelihood ratio statistic, and the discriminative
84 performance was evaluated using the bias-corrected Area Under the Curve (AUC) with its bootstrapped
85 95% confidence interval calculated from a CLR³⁵ (Appendix 4). A split-sample internal validation was
86 performed to assess the discriminant performance. The clogitL1 R package was used for the CLR with
87 Lasso³². For Percentile Lasso, the distribution of the penalization parameter was estimated from 1000
88 runs. For both Bolasso and Sublasso (from a subsample of 60% of the original data), the resampling
89 number is set to 1000. A p-value of 0.05 was considered statistically significant. All analyses were
90 conducted using R (<http://www.r-project.org/>).

91 *In-vitro analysis*

92 We performed an analysis to measure the effect of anthracyclines on the expression of the identified
93 genes in cardiomyocytes derived from adult rat ventricles using RT-qPCR. Six samples were exposed
94 to three concentrations of anthracycline: 0 μ M (control), 1 μ M, and 10 μ M for 24 hours. EPAC1
95 expression was used as a positive control due to its known dose-dependent response to anthracyclines³⁶.
96 Gene expression levels across the three anthracycline concentrations were described and compared
97 using the non-parametric Kruskal-Wallis test.

98 **Results**

99 *Patients' characteristics*

100 The median age at cancer diagnosis was 6 years (IQR, 2–11 years) for both cases and control patients
101 (**Table 1**). The main types of malignancy were lymphoma (35.8%), neuroblastoma (16.4%), bone
102 sarcoma (13.9%), nephroblastoma (12.1%), and soft tissue sarcoma (10.9%). HF was the most common
103 CD (44.8%), followed by valvular heart diseases (24.8%), arrhythmia (19.4%), pericardial disease
104 (6.7%), and ischemia (4.2%). Cases received significantly higher cumulative anthracycline doses, with
105 30.9% receiving doses greater than 360 mg/m² compared to 13.3% of controls (p=0.0004). Additionally,
106 cases were exposed to higher MHR doses, with 30.3% receiving doses greater than 15 Gy compared to
107 17.6% of controls (p=0.0410).

108 ***Gene selection associated with cardiac diseases (primary endpoint)***

109 The gene selection by the three stabilization approaches highlighted the significant reduction in gene
110 numbers when constraints were applied (**Table S2**). The Percentile Lasso approach selected 45 genes
111 with a 95th percentile constraint. The Bolasso and Sublasso approaches, respectively, identified five and
112 six genes for a selection frequency greater than or equal to 50% (**Table S3**). The intersection of the gene
113 sets resulting from the three approaches (**Figure 2**) revealed five differentially expressed genes
114 (NFE2L2, TGIF2-RAB5IF, HK2, HDGFL3, and LSP1) between cases and controls (**Figure 3**). The
115 HDGFL3, NFE2L2, and TGIF2-RAB5IF genes were under-expressed in cases compared to controls,
116 while the genes LSP1 and HK2 were overexpressed in cases compared to controls. When a cut-off of
117 70% was applied to Bolasso and Sublasso, the intersection was defined by the NFE2L2 gene alone
118 (**Table S3, Figure S3**). **Table 2** reports the effect of NFE2L2 estimated by a CLR on CD occurrence
119 adjusted on the year of cancer diagnosis, cumulative anthracycline doses, and MHR doses and shows
120 that an increase in NFE2L2 expression was significantly associated with a reduced CD risk (OR= 0.16,
121 95% CI 0.09-0.29). Correlations were observed between NFE2L2 gene expression and cumulative
122 anthracycline doses (p=0.0081) and with MHR doses (p=0.0089) (**Table S4**). Patients receiving high-
123 anthracycline doses (>360 mg/m²) showed a lower NFE2L2 gene expression compared to patients not
124 exposed to anthracyclines (p=0.0046). Additionally, patients exposed to radiation to the heart with doses
125 higher than 15 Gy expressed the gene NFE2L2 less than patients who did not receive radiation to the
126 heart (p=0.0703) and patients who received less than 5 Gy (p=0.0045) (**Table S4**). No significant
127 interactions were observed between NFE2L2 expression and anthracycline doses or MHR exposure.
128 While the other four biomarkers were finally not selected, their potential interactions with treatment
129 variables (anthracyclines and radiotherapy) were investigated, revealing no significant interactions
130 (**Table S5**). To quantify the added value of NFE2L2, we compared the likelihood (L) ratio statistic of
131 three models: model 1 included the year of cancer diagnosis, cumulative anthracycline doses, and MHR
132 doses; model 2 included the year of cancer diagnosis and NFE2L2, and model 3 (final model) included
133 the year of cancer diagnosis, cumulative anthracycline doses, MHR doses and NFE2L2 (**Table 3**).
134 Adding only NFE2L2 to the year of cancer diagnosis (model 2, -2logL= 110.9, degrees of freedom=4)
135 produces a non-significant likelihood ratio test (p=0.2398) as compared to the final model (-2log

136 L=102.9, df=10), contrary to the addition of treatment variables to year of diagnosis (model 1, -2log
137 L=197.7, df=9). This indicates strong evidence of the value added of this gene. Moreover, the
138 information added by NFE2L2 increased the discriminant ability from AUC=0.66 (95%CI 0.62-0.73,
139 model 1) to 0.85 (95%CI 0.80–0.89, model 3). These AUCs remain robust when a corrected estimation
140 of AUC is considered to control for overfitting (**Table 3**). An internal validation evaluating the
141 performance of model 3 shows a good AUC on the validation dataset (0.85, 95%CI 0.76-0.91) (**Figure**
142 **S4**). The sensitivity analysis restricted only to grade 2 or higher showed similar results (data not shown).

143 ***Gene selection associated with heart failure (secondary endpoint)***

144 From the matched case-control study, a subsample of 148 patients, including 74 cases diagnosed with
145 HF as the first cardiac event and their matched controls, were selected (**Table S6**). Using a selection
146 procedure similar to the primary endpoint but with a reduced threshold for the Bolasso and Sublasso
147 approaches (50%) due to the limited sample size, NFE2L2 remains the only gene associated with HF
148 (**Table S2-S3, Figure S5**). NFE2L2 was under-expressed in cases compared to controls (**Figure S6**)
149 and significantly correlated with the cumulative anthracycline doses ($p=0.0002$), so that patients
150 receiving high-anthracycline doses (>360 mg/m²) expressed the NFE2L2 gene less, compared to those
151 not exposed to anthracyclines ($p<0.0001$) (**Table S5**). As for CDs, an increasing NFE2L2 expression is
152 associated with a decreased HF risk (OR=0.11, 95% CI 0.03-0.37) (**Table 2**). There was no interaction
153 between NFE2L2 gene expression and treatment doses received (data not shown). Similar results to
154 CDs' outcome for quantifying the added value of NFE2L2 and its discriminant ability are also observed
155 (**Table 3**).

156 ***In vitro-validation***

157 The *in-vitro* analysis showed a significant decrease in NFE2L2 gene expression for clinically relevant
158 anthracycline doses (10 μ M). Treatment of cardiomyocytes with 1 μ M of anthracyclines for 24 hours
159 did not significantly affect the NFE2L2 gene expression as compared to non-exposed cardiomyocytes
160 (mean difference of -0.09, 95%CI -0.37-0.19; $p=0.7586$). However, exposure to a concentration of
161 10 μ M of anthracycline for 24 hours decreased NFE2L2 gene expression by 0.65 units (95%CI 0.37-

162 0.93, $p=0.0022$) and 0.74 units (95%CI 0.46-1.03, $p=0.0032$), compared to the non-exposed
163 cardiomyocyte and those exposed to 1 μ M of anthracycline, respectively (**Figure S7 and Table S7**).

164 **Discussion**

165 This study reports the independent protective effect of NFE2L2 gene expression associated with CD
166 (adjusted OR=0.16, 95%CI 0.09-0.29) and with HF (adjusted OR=0.11, 95%CI 0.03-0.37) for a one-
167 unit increase in NFE2L2. Including NFE2L2 gene expression in a model comprising the doses of
168 cumulative anthracycline and of the mean heart radiation significantly improved its discriminative
169 ability, reaching an AUC of 0.85 (95%CI 0.80-0.89) for both CD and HF, and suggesting a better
170 consideration of the inter-individual variability in CD occurrence. For comparison, our model without
171 NFE2L2 for modelling HF showed an AUC (0.77 95%CI 0.68-0.84) slightly better than that of the
172 model developed by Chow et al.¹² in a previous CCS study (0.74), which included demographic
173 variables (sex, age at cancer diagnosis) and cancer treatment characteristics (anthracycline and chest
174 radiotherapy doses).

175 Previous genetic studies on cardiotoxicity in childhood cancer survivors have shown inconsistent results
176 with different genetic variants^{16, 17, 22}. Exploring gene expression, which is more responsive to treatment
177 effects, is of major importance in the understanding of the mechanisms associated with cardiotoxicity²³.
178 To date, only one study has investigated transcriptomic data, identifying 36 genes differentially
179 expressed in anthracycline-exposed CCS implicated in different pathways including iron homeostasis
180 signaling²⁴. However, no overlapping genes were identified with our study. Two main reasons can be
181 given to explain this: (i) the use of different microarray and (ii) the methodology used. For the latter,
182 our finding results from the use of three approaches (Percentile Lasso, Bolasso, and Sublasso) derived
183 from the CLR with Lasso³² for analyzing data of a matched case-control study in the setting of high-
184 dimensional data. The gene selection, resulting in the intersection of these three selection procedures,
185 enables us to minimize the false positive rate. This strategy has identified five genes, with NFE2L2
186 emerging as a consistent biomarker across varying thresholds. NFE2L2 encodes for the nuclear factor
187 erythroid 2-related factor 2 (NRF2), an antioxidant-response protein, suggesting a potential mechanistic
188 link to treatment-induced cardiotoxicity³⁷. The NRF2 is involved in several cellular functions, including

189 iron homeostasis, one of the main pathways identified in the study of Singh et al.²⁴. As a master regulator
190 of antioxidant responses, NRF2 modulates over 240 genes throughout various biological functions³⁷,
191 and protects cardiac tissues against oxidative stress³⁷⁻³⁹. Its role in iron homeostasis and its anti-
192 inflammatory effect may protect against cardiac fibrosis preceding cardiomyopathy³⁷ and the iron
193 deficiency common in CDs^{37,40}. Stimulation of NFE2L2 can significantly impact the heart's response to
194 toxic substances, including anthracyclines³⁷. In mice and isolated cardiomyocytes treated with
195 anthracyclines, increased NFE2L2 expression protected against cardiac dysfunction^{41, 42}, while
196 NRF2 deficiency increased anthracycline cardiotoxic effects^{41,42}. Our study highlighted a decrease in
197 NFE2L2 gene expression in rat cardiomyocytes exposed to clinically relevant doses of anthracyclines,
198 suggesting that its downregulation may be associated with the weakening of cardiac tissue. However, it
199 should be noted that the transcriptome in our study is derived from peripheral blood mononuclear cells,
200 whose response to treatments may differ from that of cardiomyocytes. Furthermore, a study on
201 atherosclerosis has shown that the monocyte transcriptome may integrate genetic and non-genetic
202 factors, which are relevant for understanding disease pathophysiology and risk assessment⁴³. In a model
203 including cancer treatment doses adjusted for the year of cancer diagnosis, treatment variables were
204 significantly associated with CD. However, when adding NFE2L2 expression, these variables were no
205 longer significant. This suggests a wider involvement of NFE2L2 in CD beyond cancer treatment-
206 induced effects, or that the association of treatment dose with CD may be mediated by the impact on
207 NFE2L2 expression. This demonstrates the dual nature of the transcriptome, which can be viewed as a
208 reflection of a changing process or a link in a causal path⁴³.

209 Despite the importance of our findings, some limitations must be acknowledged. The primary outcome
210 comprises various CDs with different pathophysiologies, requiring separate analyses for each CD, as
211 done for HF. However, the small number of patients in the different categories limited such subgroup
212 analyses. The main analysis is based on the occurrence of CD whatever the grade, but the sensitivity
213 analysis restricted only to grade 2 or higher showed consistent results. Additionally, the FCCSS cohort
214 subsample used in our study may not fully represent the entire cohort, thus potentially underestimating
215 the AUC for a model including only clinical variables. As external data were unavailable, we could not

216 definitively validate the role of NFE2L2 externally, even though the internal-external validation results
217 in a promising discriminant ability of the model including the NFE2L2 gene expression

218 Our findings indicate that CCS presenting a high expression of the NFE2L2 gene were less likely to
219 have a CD. This result should be validated through a longitudinal monitoring of NFE2L2 gene
220 expression before and after childhood cancer treatment, to provide cardioprotective treatments to
221 patients upstream of the development of CD (as a prognostic biomarker) or during early diagnosis of
222 cardiac dysfunction (as a diagnostic biomarker).

223

224 **References**

- 225 1. Steliarova-Foucher E, Colombet M, Ries LAG, et al: International incidence of childhood cancer,
226 2001–10: a population-based registry study. *Lancet Oncol* 18:719–731, 2017
- 227 2. Leerink JM, de Baat EC, Feijen EAM, et al: Cardiac Disease in Childhood Cancer Survivors. *JACC*
228 *CardioOncol* 2:363–378, 2020
- 229 3. Hudson MM, Ness KK, Gurney JG, et al: Clinical ascertainment of health outcomes among adults
230 treated for childhood cancer. *JAMA* 309:2371–2381, 2013
- 231 4. Armenian SH, Armstrong GT, Aune G, et al: Cardiovascular Disease in Survivors of Childhood
232 Cancer: Insights Into Epidemiology, Pathophysiology, and Prevention. *JCO* 36:2135–2144, 2018
- 233 5. Bottinor W, Im C, Doody DR, et al: Mortality After Major Cardiovascular Events in Survivors of
234 Childhood Cancer. *Journal of the American College of Cardiology* 83:827–838, 2024
- 235 6. Geenen MM, Cardous-Ubbink MC, Kremer LCM, et al: Medical Assessment of Adverse Health
236 Outcomes in Long-term Survivors of Childhood Cancer. *JAMA* 297:2705–2715, 2007
- 237 7. Haddy N, Diallo S, El-Fayech C, et al: Cardiac Diseases Following Childhood Cancer Treatment:
238 Cohort Study. *Circulation* 133:31–38, 2016
- 239 8. Yeh ETH, Tong AT, Lenihan DJ, et al: Cardiovascular complications of cancer therapy: diagnosis,
240 pathogenesis, and management. *Circulation* 109:3122–3131, 2004
- 241 9. Lipshultz SE, Adams MJ: Cardiotoxicity After Childhood Cancer: Beginning With the End in Mind.
242 *JCO* 28:1276–1281, 2010
- 243 10. Niska JR, Thorpe CS, Allen SM, et al: Radiation and the heart: systematic review of dosimetry and
244 cardiac endpoints. *Expert Review of Cardiovascular Therapy* 16:931–950, 2018

- 245 **11.** Mulrooney DA, Yeazel MW, Kawashima T, et al: Cardiac outcomes in a cohort of adult survivors
246 of childhood and adolescent cancer: retrospective analysis of the Childhood Cancer Survivor Study
247 cohort. *BMJ* 339:b4606, 2009
- 248 **12.** Chow EJ, Chen Y, Kremer LC, et al: Individual Prediction of Heart Failure Among Childhood
249 Cancer Survivors. *JCO* 33:394–402, 2015
- 250 **13.** Clemens E, van der Kooi ALF, Broer L, et al: The influence of genetic variation on late toxicities in
251 childhood cancer survivors: A review. *Critical Reviews in Oncology/Hematology* 126:154–167, 2018
- 252 **14.** Hildebrandt MAT, Reyes M, Wu X, et al: Hypertension Susceptibility Loci are Associated with
253 Anthracycline-related Cardiotoxicity in Long-term Childhood Cancer Survivors. *Sci Rep* 7:9698, 2017
- 254 **15.** Visscher H, Rassekh SR, Sandor GS, et al: Genetic variants in *SLC22A17* and *SLC22A7* are
255 associated with anthracycline-induced cardiotoxicity in children. *Pharmacogenomics* 16:1065–1076,
256 2015
- 257 **16.** Visscher H, Ross CJD, Rassekh SR, et al: Pharmacogenomic prediction of anthracycline-induced
258 cardiotoxicity in children. *J Clin Oncol* 30:1422–1428, 2012
- 259 **17.** Visscher H, Ross CJD, Rassekh SR, et al: Validation of variants in *SLC28A3* and *UGT1A6* as
260 genetic markers predictive of anthracycline-induced cardiotoxicity in children. *Pediatr Blood Cancer*
261 60:1375–1381, 2013
- 262 **18.** Wang X, Liu W, Sun C-L, et al: Hyaluronan synthase 3 variant and anthracycline-related
263 cardiomyopathy: a report from the children’s oncology group. *J Clin Oncol* 32:647–653, 2014
- 264 **19.** Wang X, Sun C-L, Quiñones-Lombraña A, et al: *CELF4* Variant and Anthracycline-Related
265 Cardiomyopathy: A Children’s Oncology Group Genome-Wide Association Study. *JCO* 34:863–870,
266 2016

- 267 **20.** Chaix M-A, Parmar N, Kinnear C, et al: Machine Learning Identifies Clinical and Genetic Factors
268 Associated With Anthracycline Cardiotoxicity in Pediatric Cancer Survivors. *JACC: CardioOncology*
269 2:690–706, 2020
- 270 **21.** Petrykey K, Rezgui AM, Guern ML, et al: Genetic factors in treatment-related cardiovascular
271 complications in survivors of childhood acute lymphoblastic leukemia. *Pharmacogenomics* 22:885–901,
272 2021
- 273 **22.** Aminkeng F, Bhavsar AP, Visscher H, et al: A coding variant in RARG confers susceptibility to
274 anthracycline-induced cardiotoxicity in childhood cancer. *Nat Genet* 47:1079–1084, 2015
- 275 **23.** Joseph P: Transcriptomics in toxicology. *Food Chem Toxicol* 109:650–662, 2017
- 276 **24.** Singh P, Shah DA, Jouni M, et al: Altered Peripheral Blood Gene Expression in Childhood Cancer
277 Survivors With Anthracycline-Induced Cardiomyopathy – A COG-ALTE03N1 Report. *JAHA*
278 12:e029954, 2023
- 279 **25.** Bejarano-Quisoboni D, Pelletier-Fleury N, Allodji RS, et al: Long-term hospitalisations in survivors
280 of paediatric solid tumours in France. *Sci Rep* 12:18068, 2022
- 281 **26.** Chounta S, Lemler S, Haddy N, et al: The risk of valvular heart disease in the French Childhood
282 Cancer Survivors' Study: Contribution of dose-volume histogram parameters. *Radiother Oncol*
283 180:109479, 2023
- 284 **27.** Mansouri I, Allodji RS, Hill C, et al: The role of irradiated heart and left ventricular volumes in heart
285 failure occurrence after childhood cancer. *Eur J Heart Fail* 21:509–518, 2019
- 286 **28.** Charrier T, Haddy N, Schwartz B, et al: Increased Cardiac Risk After a Second Malignant Neoplasm
287 Among Childhood Cancer Survivors. *JACC CardioOncol* 5:792–803, 2023
- 288 **29.** European Society of Cardiology [Internet][cited 2022 May 19] Available from:
289 <https://www.escardio.org/>, <https://www.escardio.org>

- 290 **30.** Truong QA, Massaro JM, Rogers IS, et al: Reference values for normal pulmonary artery dimensions
291 by noncontrast cardiac computed tomography: the Framingham Heart Study. *Circ Cardiovasc Imaging*
292 5:147–154, 2012
- 293 **31.** Tibshirani R: Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical*
294 *Society: Series B (Methodological)* 58:267–288, 1996
- 295 **32.** Reid S, Tibshirani R: Regularization Paths for Conditional Logistic Regression: The clogitL1
296 Package. *J Stat Softw* 58:12, 2014
- 297 **33.** Roberts S, Nowak G: Stabilizing the lasso against cross-validation variability. *Computational*
298 *Statistics & Data Analysis* 70:198–211, 2014
- 299 **34.** Xu H, Caramanis C, Mannor S: Sparse Algorithms Are Not Stable: A No-Free-Lunch Theorem.
300 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34:187–193, 2012
- 301 **35.** Janes H, Pepe MS: Adjusting for Covariates in Studies of Diagnostic, Screening, or Prognostic
302 Markers: An Old Concept in a New Setting. *American Journal of Epidemiology* 168:89–97, 2008
- 303 **36.** Mazevet M, Belhadeif A, Ribeiro M, et al: EPAC1 inhibition protects the heart from doxorubicin-
304 induced toxicity. *Elife* 12:e83831, 2023
- 305 **37.** Roberts JA, Rainbow RD, Sharma P: Mitigation of Cardiovascular Disease and Toxicity through
306 NRF2 Signalling. *Int J Mol Sci* 24:6723, 2023
- 307 **38.** Taverne YJHJ, Bogers AJJC, Duncker DJ, et al: Reactive Oxygen Species and the Cardiovascular
308 System. *Oxidative Medicine and Cellular Longevity* 2013:1–15, 2013
- 309 **39.** Chen QM, Maltagliati AJ: Nrf2 at the heart of oxidative stress and cardiac protection. *Physiol*
310 *Genomics* 50:77–97, 2018
- 311 **40.** Savarese G, von Haehling, Stephan, Butler, Javed, et al: Iron deficiency and cardiovascular disease.
312 *European Heart Journal* 44:14–27, 2023

- 313 **41.** Zhang X, Hu C, Kong C-Y, et al: FNDC5 alleviates oxidative stress and cardiomyocyte apoptosis
314 in doxorubicin-induced cardiotoxicity via activating AKT. *Cell Death Differ* 27:540–555, 2020
- 315 **42.** Li S, Wang W, Niu T, et al: Nrf2 Deficiency Exaggerates Doxorubicin-Induced Cardiotoxicity and
316 Cardiac Dysfunction. *Oxid Med Cell Longev* 2014:748524, 2014
- 317 **43.** Zeller T, Wild P, Szymczak S, et al: Genetics and Beyond – The Transcriptome of Human
318 Monocytes and Disease Susceptibility. *PLoS ONE* 5:e10693, 2010
- 319

Tables and Figures

Figure 1. Flowchart of patient's inclusion in the study

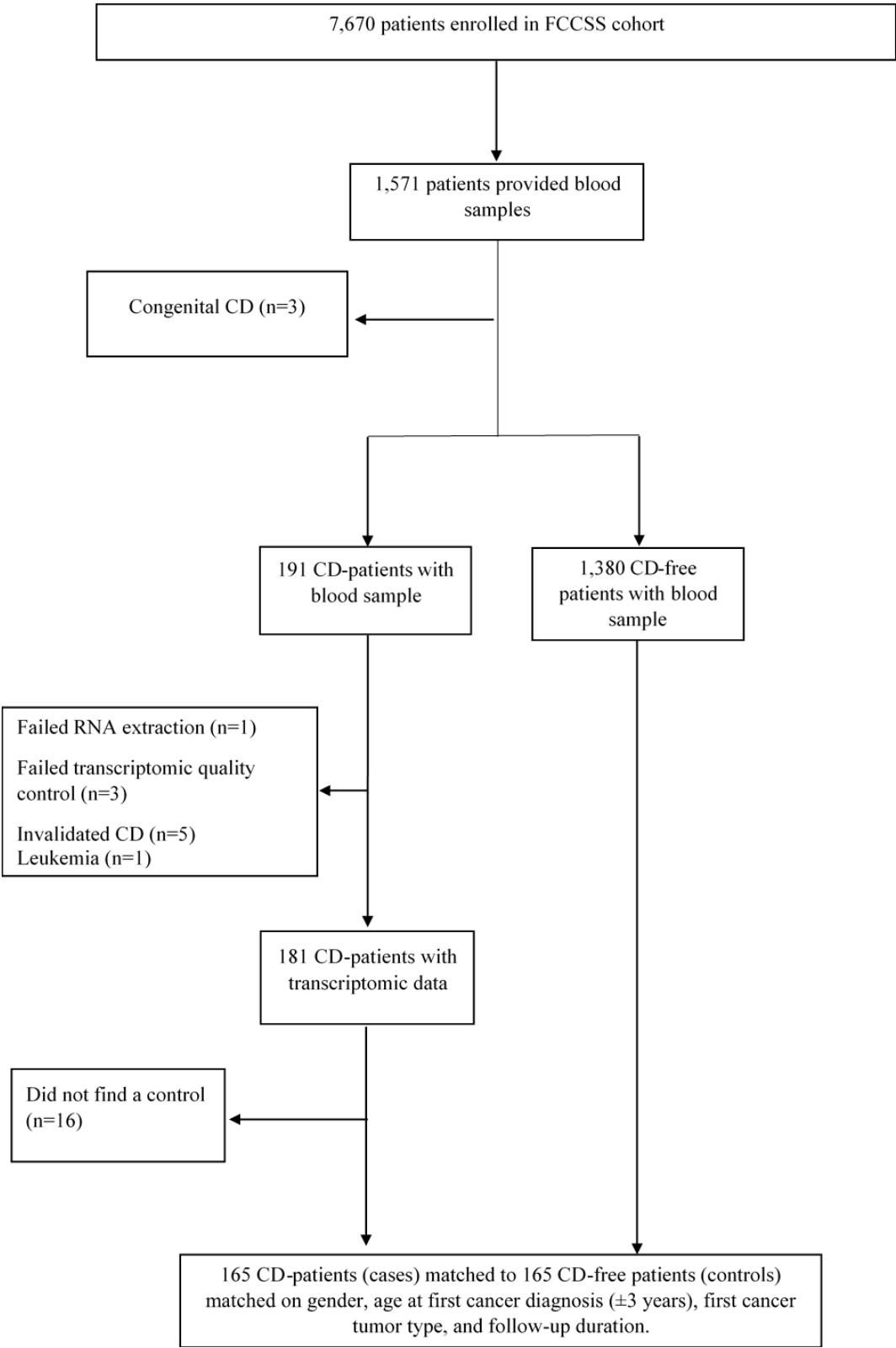


Figure 2. Venn diagram representing the genes, identified by the Percentile lasso, Bolasso and Sublasso for cardiac disease outcome. A threshold of 50% is used for the latter two approaches.

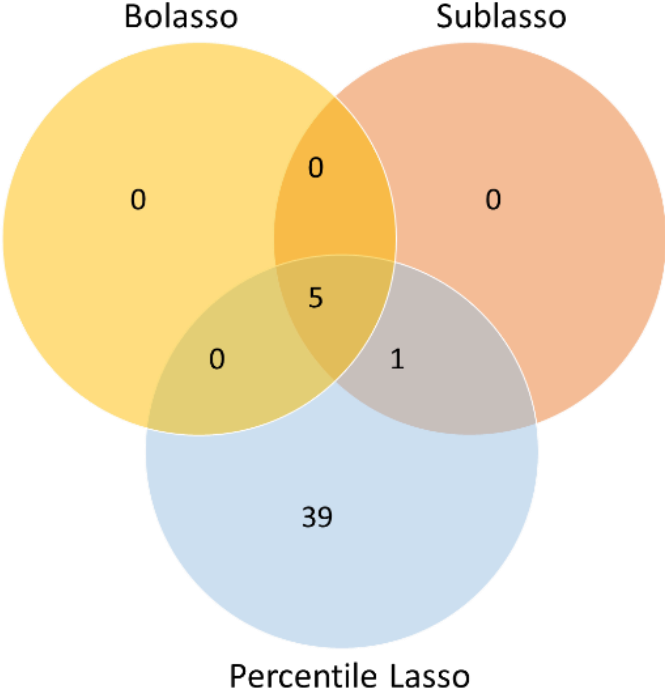
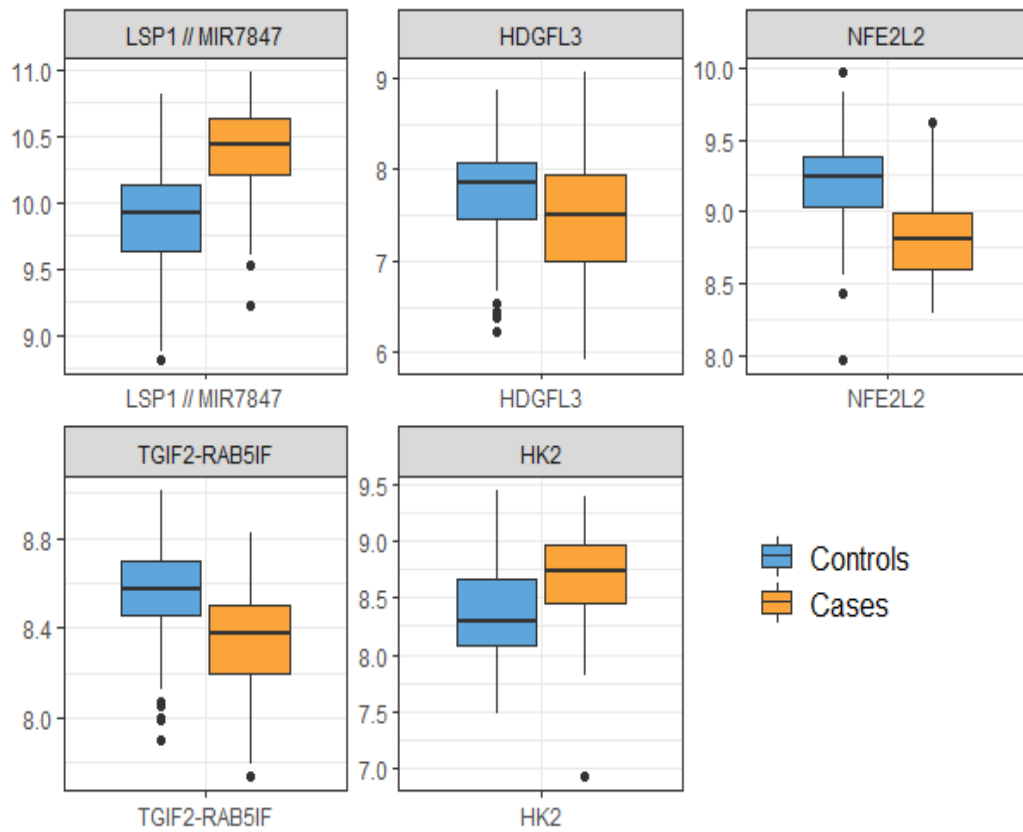


Figure 3. Box plots of the genes expression defining the intersection set of the Percentile lasso, Bolasso and Sublasso* in cases and controls associated to cardiac disease outcome.



*: For Bolasso and Sublasso, a threshold of 50% is used for the gene selection.

Table 1. Demographic and clinical characteristic of patients included in the matched-case control study for cardiac disease outcome (n=330)

	Cases (n=165)	Controls (n=165)	p-value
Gender, n (%)*			
Female	78 (47.3)	78 (47.3)	
Age at primary cancer diagnosis (years)*			
Median (IQR)	6 (2 – 11)	6 (2 – 11)	
[0 – 5[72 (43.6)	73 (44.2)	
[5 – 10[44 (26.7)	43 (26.1)	
[10 – 15[36 (21.8)	41 (24.8)	
[15 – 22[13 (7.9)	8 (4.8)	
Type of malignancy, n (%)*			
Lymphoma	59 (35.8)	59 (35.8)	
Neuroblastoma	27 (16.4)	27 (16.4)	
Bone sarcoma	23 (13.9)	23 (13.9)	
Nephroblastoma	20 (12.1)	20 (12.1)	
Soft tissue sarcoma	18 (10.9)	18 (10.9)	
Gonad	6 (3.6)	6 (3.6)	
Others	12 (7.3)	12 (7.3)	
Cardiac disease, n (%)			
Heart Failure	74 (44.8)		
Valvular Heart Diseases	41 (24.8)		
Arrhythmia	32 (19.4)		
Pericardial disease	11 (6.7)		
Ischemia	7 (4.2)		
Follow-up after cancer diagnosis (years)*			
Median (range)	20 (0 - 54)	33 (10 – 65)	
[0 – 10]*	30 (18.2)	0 (0.0)	
[10 – 20[48 (29.1)	15 (9.1)	
[20 – 30[50 (30.3)	46 (27.9)	
[30 – 40[32 (19.4)	59 (35.8)	
[40 – 50[4 (2.4)	32 (19.4)	
≥ 50	1 (0.6)	13 (7.9)	
Year of primary cancer diagnosis, n (%)			0.7383
Before 1970	17 (10.3)	29 (17.6)	
[1970, 1979[57 (34.5)	43 (26.0)	
[1980, 1989]	67 (40.6)	65 (39.4)	
[1990, 2000]	24 (14.5)	28 (17.0)	
Anthracyclines n (%)			
No	57 (34.5)	75 (45.5)	
Yes	108 (65.5)	90 (54.5)	
Cumulative anthracycline doses (mg/m²)			0.0004
Median (range)	355.1 (100.0 – 1207.5)	248.4 (20.8 – 546.1)	
None	57 (34.5)	75 (45.5)	
]0 – 250]	32 (19.4)	45 (27.3)	
]250 – 360]	25 (15.2)	23 (13.9)	
>360	51 (30.9)	22 (13.3)	
Heart Radiations			
No	51 (30.9)	56 (33.9)	
Yes	114 (69.1)	109 (66.1)	
Mean heart radiation doses (Gy)			0.0410

Median (range)	7.8 (0.0 – 43.2)	2.0 (0.0 – 43.1)
None	51 (30.9)	56 (33.9)
]0 – 5]	54 (32.7)	66 (40.0)
]5 – 15]	10 (6.1)	14 (8.5)
>15	50 (30.3)	29 (17.6)

IQR, Interquartile range; Gy, Gray; p-values were obtained using Stuart-Kendall Tau-c one sided exact test.

* Variables used for the matching between cases and controls; ‡ In cases, the follow-up starts from the cancer diagnosis and ends at CD development, which can occur before the first 5 years of follow-up

Table 2. Effect of NFE2L2 gene expression on cardiac diseases (main endpoint) and heart failure (secondary endpoint) estimated from a conditional logistic regression adjusted on the year of cancer diagnosis and exposition to anthracycline and radiotherapy

Variables	n / N	Odds Ratio (95% Confidence Interval)	p-value
Cardiac Diseases (primary outcome)			
Year of cancer diagnosis			0.3341
< 1970	17 / 46	1.00	
[1970-1979]	57 / 100	1.29 (0.31 - 5.29)	
[1980-1989]	67 / 132	0.78 (0.17 - 3.64)	
[1990-2000]	24 / 52	2.66 (0.41 - 17.11)	
Cumulative anthracycline dose (mg/m²)			0.2121
None	57 / 132	1.00	
]0 - 250]	32 / 77	1.01 (0.29 - 3.52)	
]250 - 360]	25 / 48	2.96 (0.79 - 11.06)	
>360	51 / 73	3.56 (0.88 - 10.53)	
Mean heart radiation dose (Gy)			0.5218
None	51 / 107	1.00	
]0 – 5]	54 / 120	0.97 (0.38 - 2.53)	
]5 – 15]	10 / 24	0.65 (0.12 - 3.43)	
>15	50 / 79	1.94 (0.62 - 6.05)	
NFE2L2[‡]		0.16 (0.09 - 0.29)	< 0.0001
Heart Failure (secondary outcome)			
Year of cancer diagnosis[¶]			0.8241
< 1980	33 / 68	1.00	
[1980-2000]	41 / 80	1.22 (0.22 - 6.76)	
Cumulative anthracycline dose (mg/m²)[¶]			0.0004
None	7 / 37	1.00	
]0 - 250]	13 / 33	12.97 (0.47 - 358.30)	
>250	54 / 78	53.62 (1.86 - 1548.03)	
Mean heart radiation dose (Gy)[¶]			0.9405
None	27 / 56	1.00	
]0 – 15]	31 / 68	1.47 (0.30 - 7.10)	
>15	16 / 24	1.38 (0.12 - 16.30)	
NFE2L2^{‡¶}		0.11 (0.03 - 0.37)	<0.0001

n = number of cases; N = total number of patients;

[‡] Odds ratio was estimated to 0.17 (95% CI, 0.10 - 0.28) when cumulative anthracycline dose and mean heart radiation dose are not included in the conditional logistic model, ^{‡‡} Odds ratio was estimated to 0.12 (95% CI, 0.04 - 0.31) when cumulative anthracycline dose and mean heart radiation dose are not included in the conditional logistic model, [¶] Some categories were collapsed due to the small sample size.

Results of the models which include cumulative anthracycline doses, MHR doses, and adjusted on the year of cancer diagnosis, without NFE2L2 gene expression, are presented in **Table S8**.

Table 3. Comparison of different conditional logistic regression models based on log-likelihood statistic and area under the curve (AUC) for cardiac diseases (top part) and heart failure (bottom part) outcomes

Model	-2 log Likelihood	df	AUC (95% CI)	corrected AUC¶ (95% CI)
Cardiac diseases (primary outcome)				
Null model (no covariate)	228.7			
Model 1: year of cancer diagnosis + treatment variables†	197.7	9	0.66 (0.62 – 0.73)	0.64 (0.60 – 0.68)
Model 2: year of cancer diagnosis + NFE2L2	110.9	4	0.84 (0.79 – 0.89)	0.83 (0.79 – 0.86)
Model 3: year of cancer diagnosis + treatment variables + NFE2L2	102.9	10	0.85 (0.80 – 0.89)	0.83 (0.80 – 0.87)
Heart Failure (secondary outcome)				
Null model (no covariate)	102.6			
Model 1: year of cancer diagnosis + treatment variables††	67.2	5	0.77 (0.68 – 0.84)	0.74 (0.68 – 0.82)
Model 2: year of cancer diagnosis + NFE2L2	49.7	2	0.82 (0.75 – 0.88)	0.81 (0.75 – 0.88)
Model 3: year of cancer diagnosis + treatment variables + NFE2L2	34.1	6	0.85 (0.80 – 0.90)	0.83 (0.78 – 0.89)

df: degrees of freedom

† Treatment variables are the cumulative dose of anthracyclines (none, 0 to 250, 250 to 360, >360 mg/m²) and the mean heart radiation dose (none, 0 to 5, 5 to 15, >15 Gy).

†† Treatment variables are the cumulative dose of anthracyclines (none, 0 to 250, >250 mg/m²) and the mean heart radiation dose (none, 0 to 5, 5> Gy). We collapsed the two last categories due to the small patients' number in the last category.

AUC Area under the Curve corresponds to an unweighted AUC, CI confidence interval calculated from bootstrap

¶ corresponds to a bootstrapped-corrected AUC for controlling the overfit (See appendix 3 Statistical consideration in Supplementary for more information)

Abbreviations

FCCSS: French Childhood Cancer Survivor Study; CCS: Childhood Cancer Survivors; CDs: Cardiac Diseases; HF: Heart Failure; MHR: mean heart radiation; RT-qPCR: Quantitative reverse transcription polymerase chain reaction; OR: Odds Ratio; CI: Confidence interval, CLR: Conditional Logistic Regression, df: degrees of freedom

Supplementary material

Appendix 1: Cardiac disease assessment details

Cases were defined as individuals who developed cardiac diseases (CD), including heart failure, valvular disease, ischemic pericardial disease, and arrhythmia diseases, at least five years after the childhood cancer diagnosis. CDs were identified from self-questionnaires, medical records, long-term follow-up consultation reports, and the National Health Insurance Inter-Scheme Information System database (SNIIR-AM), which contains medical data on all hospitalized patients in France. A questionnaire confirming the CD was completed by the general practitioner and/or a cardiologist for all patients, including B-type natriuretic peptide (N-terminal proBNP), echocardiographic data, the date of diagnosis, and which drug and/or interventional treatment was given. A copy of medical results was asked for when appropriate. Each diagnosis was validated based on criteria from the European Society of Cardiology and the Framingham studies ¹. The Common Terminology Criteria for Adverse Events (CTCAE version 4.03) was used to grade all CDs ².

Appendix 2: Biological Samples Collection

A biobank attached to the FCCSS cohort was constructed in order to study the biological impacts of the iatrogenic effects of cancer treatments. Biological samples were collected from adult patients during routine visits or in medical laboratories. Blood samples were available for 1,571 patients from the FCCSS, on which we constructed our matched case-control study. For most CD cases (n=160), blood sampling was done after the diagnosis of CD, and for the five other cases, blood sampling took place less than one year before the CD diagnosis. Finally, for all 165 cases, the median time between diagnosis of cardiac pathology and blood sampling was 12.5 years (Interquartile range, 6.1 – 16.6). Furthermore, the median age at blood sampling for cases (39.7, (IQR, 34.4 - 45.8)) was close to that of controls (41.7, (IQR, 34.6 - 46.7)).

Appendix 3: Gene Expression Profiling

Affymetrix Human ClariomD Array analysis was done by GenoSplice (Paris, France). Array data were normalized using quantile normalization using APT command lines (Affymetrix). For quality control, “pos_vs_neg_auc”, “all_probeset_rle_mean”, “%P” metrics, bacterial spikes, and polyA-control RNAs expression were checked. Background correction was made using the antigenomic probes, and probe selection was made as described previously³⁻⁵. Only probes targeting exons annotated from FAST DB transcripts (v2016_1) were selected in order to focus on well-annotated genes whose mRNA sequences are in public databases⁶. Among these selected probes, bad-quality probes (e.g., probes identified by Affymetrix as “cross-hybridizing”), and probes with too low an intensity of signal compared to antigenomic background probes with the same GC content were removed from the analysis. Only probes with a Detection Above Background (DABG) p -value ≤ 0.05 in at least half of the arrays were considered for statistical analyses. Only genes expressed in at least one compared condition were analyzed (i.e., case and control samples). The DABG p -value had to be ≤ 0.05 for at least half of the gene probes to be considered as expressed. Additionally, 3,247 genes insufficiently expressed for biological confirmation were excluded, resulting in 8,739 genes to include in the analysis (**Figure S1**).

Appendix 4: Statistical Considerations

The Lasso procedure

To identify potential biomarkers associated with the clinical endpoints in the high dimensional data setting, the Lasso (Least Absolute Shrinkage and Selection Operator) selection procedure is often used, allowing both selection of biomarkers and estimation of their effects ⁷. This process involves two main steps: first, the calculation of a tuning parameter, known as the regularization or penalization parameter, denoted λ , using k -fold cross-validation (high values of λ result in a small selection of biomarkers since the coefficients of the less important biomarkers are set to 0). The second step involves maximizing the penalized log-likelihood function ⁷. Several runs of the Lasso algorithm, including the same variables, give different penalization parameters since the k -fold cross-validation is different ($k = 5$ is used per default), leading to unstable biomarkers selection.

Conditional Logistic Regression with Lasso Penalties

Related to the study design, i.e., the matched case-control study, the Lasso procedure is adapted to the likelihood of a conditional logistic regression (CLR) ⁸.

The stabilization approaches

Three resampling-based approaches were employed to address the instability problem in the Lasso approach on the optimal λ estimation and subsequent biomarker selection.

- The Percentile Lasso approach involves selecting $\hat{\lambda}(\theta)$ corresponding to the θ^{th} percentile of the distribution of the penalization parameter (λ) derived from m replications of the CLR Lasso. Each execution of the Lasso procedure results in a single calculation of the optimal λ due to the fold assignment within the cross-validation process ⁹.
- Bolasso entails generating m bootstrap samples from the original data set, on which the CLR Lasso is applied ¹⁰.
- Sublasso is similar to Bolasso but involves generating m sub-samples of the original data set, on which the CLR Lasso is applied ¹¹.

In Bolasso and Sublasso approaches, a gene is retained if its selection frequency through $m=1000$ repetitions exceeds or equals a predetermined threshold (50%, 60%, 70%, or 80%). For Percentile Lasso, θ is set to 95%, and a subset of 60% of the original data set is used for the Sublasso. The CLR Lasso selection procedure is conducted using the clogitL1 R package with default parameters ⁸.

Discriminatory performance

The model's discriminative performance for a binary outcome is commonly estimated using the Area Under the Curve (AUC) metric. In the setting of a case-control design, different approaches are possible to estimate this quantity. They differ in their ways of estimating the linear predictor.

Let Y represent the binary outcome variable (1 = CD-case: 0 = CD-free control), $X = (X_1, \dots, X_p)$ denotes the selected biomarkers, year of cancer diagnosis and/or treatment variables, and $Z = (Z_1, \dots, Z_q)$ denotes the matching variables. The matching variables are, here, age at primary cancer diagnosis, type of malignancy, and follow-up after diagnosis.

- First approach: unconditional logistic regression including the selected biomarker, the year of cancer diagnosis, and/or treatment variables (X) adjusted for matching variables (Z) is fitted. We then estimate the regression coefficients related to the intercept ($\hat{\alpha}$), related to the covariates X ($\hat{\beta}$), and related to the matching variable Z ($\hat{\gamma}$). The AUC is then calculated using the estimated linear predictor $\hat{\alpha} + \hat{\beta}X + \hat{\gamma}Z$ ¹²⁻¹⁴.
- Second approach: a conditional logistic regression including the selected biomarkers and year of cancer diagnosis and/or treatment variables (X) is fitted. Only the coefficients $\hat{\beta}$ related to X are estimated ($\hat{\alpha} = \hat{\gamma} = 0$). The AUC is then estimated using the linear predictor $\hat{\beta}X$ ¹².

A bootstrap sampling technique was used to estimate a confidence interval of AUC. Specifically, 500 bootstrap samples were generated by sampling the matched sets with replacement. From the distribution of the AUCs estimated from the 500 bootstrap samples, the 0.025 and 0.975 percentiles were used as the 95% confidence interval limits (95% CI).

Note that these two approaches produce an unweighted estimate of AUC, i.e., not considering a possible unrepresentativeness of patients from the case-control compared to the parent cohort. Some methods have been developed to consider a weighted AUC for the nested design framework, such as weighted logistic regression, weighted conditional logistic regression, the two-stage procedure, and the simultaneous procedure¹².

For the analysis, the second approach was used to estimate the unweighted AUC of different models in a matched case-control design framework. These models differ in terms of the covariates X: (1) year of cancer diagnosis and treatment variables: categorized cumulative dose of anthracyclines (none,]0 – 250],]250 – 360], >360 mg/m²), and categorized mean heart radiation doses (none,]0 - 5],]5 -15], >15 Gy), (2) year of cancer diagnosis and the selected biomarkers alone and (3) year of cancer diagnosis, both treatment variables and selected biomarkers. This last model corresponds to the final model. Similar results were produced with the first approach (data not shown).

Bias-corrected AUC calculation

To estimate a bias-corrected AUC for both unconditional and conditional logistic regressions described above, a bootstrapping technique was used to estimate the bias due to overfitting termed optimism. When the AUC-optimism is estimated, it can be subtracted from the AUC derived from the original sample to obtain a bias-corrected AUC.

From the original data of size n , a bootstrapped sample of size n is constructed. The model is then fitted on the bootstrapped sample and applied without changing to the original sample. The AUC from the bootstrap sample ($AUC_{Bootstrap}$) minus the one computed on the original sample ($AUC_{Apparent}$) is an estimate of optimism. $optimism = AUC_{Bootstrap} - AUC_{Apparent}$

This process is repeated for B=500 bootstrap replications to obtain an average optimism, which is subtracted from the final model estimated AUC to obtain the bias-corrected AUC¹⁵.

$$AUC_{bias\ corrected} = AUC_{original} - mean(optimisms)$$

The 95% CI of this bias-corrected AUC is obtained by applying the 0.025 and 0.975 percentiles to the distribution of the corrected AUCs.

$$AUC_{bias\ corrected(B)} = AUC_{original} - optimism_{(B)}$$

For the analysis, conditional logistic regressions was used to estimate the unweighted bias-corrected AUC.

References:

1. Truong QA, Massaro JM, Rogers IS, et al: Reference values for normal pulmonary artery dimensions by noncontrast cardiac computed tomography: the Framingham Heart Study. *Circ Cardiovasc Imaging* 5:147–154, 2012
2. US Department of Health and Human Services: Common Terminology Criteria for Adverse Events (CTCAE). Version 4.03. [Internet], 2010 Available from: https://www.eortc.be/services/doc/ctc/ctcae_4.03_2010-06-14_quickreference_5x7.pdf
3. de la Grange P, Jolly A, Courageux C, et al: Genes coding for transcription factors involved in stem cell maintenance are repressed by TGF- β and downstream of Slug/Snail2 in COPD bronchial epithelial progenitors. *Mol Biol Rep* 48:6729–6738, 2021
4. Weiss E, de la Grange P, Defaye M, et al: Characterization of Blood Immune Cells in Patients With Decompensated Cirrhosis Including ACLF. *Front Immunol* 11:619039, 2020
5. Gorombeï P, Guidez F, Ganesan S, et al: BCL-2 Inhibitor ABT-737 Effectively Targets Leukemia-Initiating Cells with Differential Regulation of Relevant Genes Leading to Extended Survival in a NRAS/BCL-2 Mouse Model of High Risk-Myelodysplastic Syndrome. *Int J Mol Sci* 22:10658, 2021
6. de la Grange P, Dutertre M, Correa M, et al: A new advance in alternative splicing databases: from catalogue to detailed analysis of regulation of expression and function of human alternative splicing variants. *BMC Bioinformatics* 8:180, 2007
7. Tibshirani R: Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 58:267–288, 1996
8. Reid S, Tibshirani R: Regularization Paths for Conditional Logistic Regression: The clogitL1 Package. *J Stat Softw* 58:12, 2014
9. Roberts S, Nowak G: Stabilizing the lasso against cross-validation variability. *Computational Statistics & Data Analysis* 70:198–211, 2014

- 10.** Bach FR: Bolasso: model consistent Lasso estimation through the bootstrap [Internet], in Proceedings of the 25th international conference on Machine learning - ICML '08. Helsinki, Finland, ACM Press, 2008, pp 33–40[cited 2022 Jul 28] Available from: <http://portal.acm.org/citation.cfm?doid=1390156.1390161>
- 11.** Meinshausen N, Buehlmann P: Stability Selection [Internet], 2009[cited 2023 May 2] Available from: <http://arxiv.org/abs/0809.2932>
- 12.** Xu H, Qian J, Paynter NP, et al: Estimating the receiver operating characteristic curve in matched case control studies: Estimating the ROC curve in matched case control studies. *Statistics in Medicine* 38:437–451, 2019
- 13.** Pepe MS, Fan J, Seymour CW: Estimating the ROC Curve in Studies that Match Controls to Cases on Covariates. *Acad Radiol* 20:863–873, 2013
- 14.** Janes H, Pepe MS: Adjusting for Covariates in Studies of Diagnostic, Screening, or Prognostic Markers: An Old Concept in a New Setting. *American Journal of Epidemiology* 168:89–97, 2008
- 15.** Harrell , FE: *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis* [Internet]. Cham, Springer International Publishing, 2015[cited 2022 Oct 5] Available from: <http://link.springer.com/10.1007/978-3-319-19425-7>

Supplementary Figures

Figure S1. Flowchart of genes included in the study

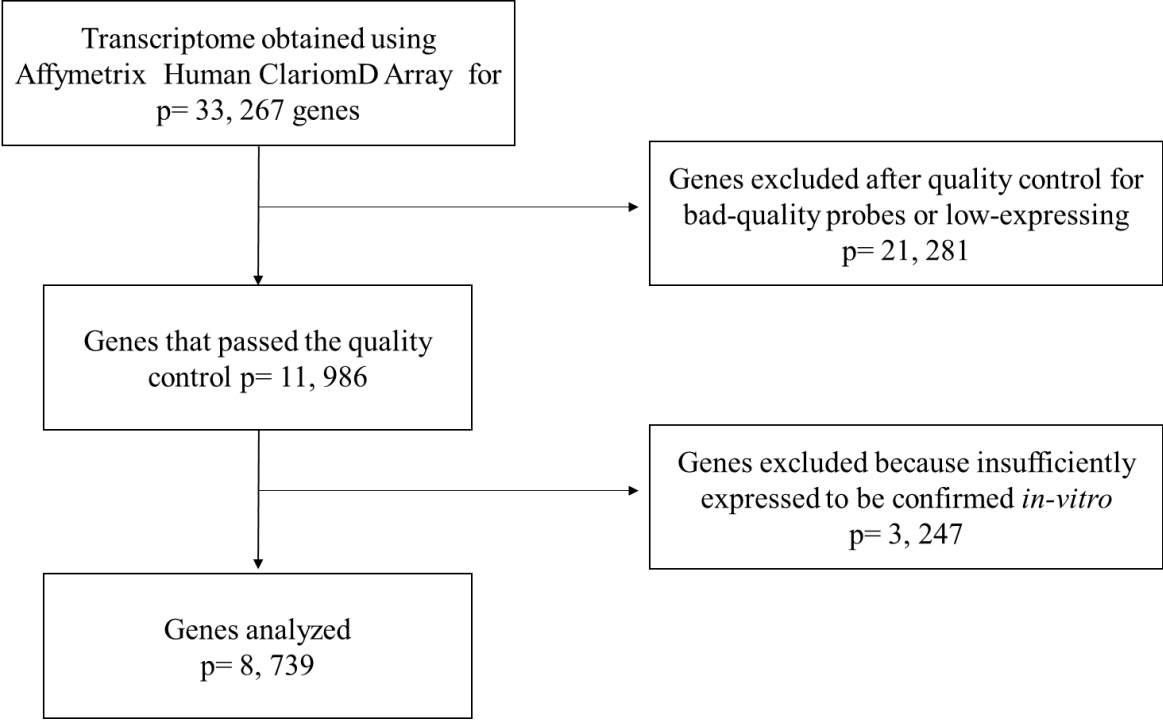
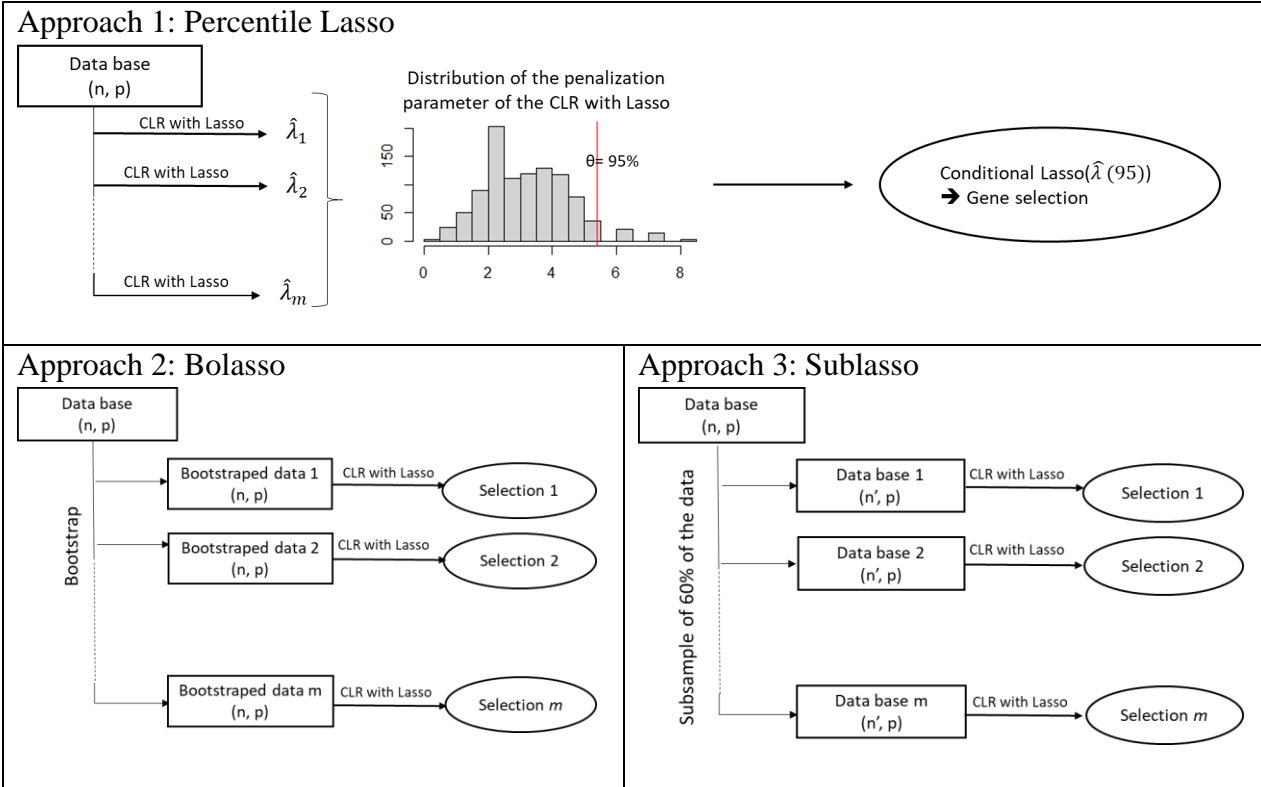


Figure S2. Description of the three stabilization approaches using condition logistic regression with lasso



n: sample size, n': subsample size, p: biomarkers number, λ : penalization parameter of Lasso, m: represents the number of runs for Percentile Lasso and resampling samples for Bolasso and Sublasso, CLR: conditional logistic regression.

The red vertical line in approach 1 (Percentile Lasso) corresponds to the 95th percentile of the penalization parameter distribution.

Figure S3. Histogram of NFE2L2 expression in cases and controls for cardiac disease outcome

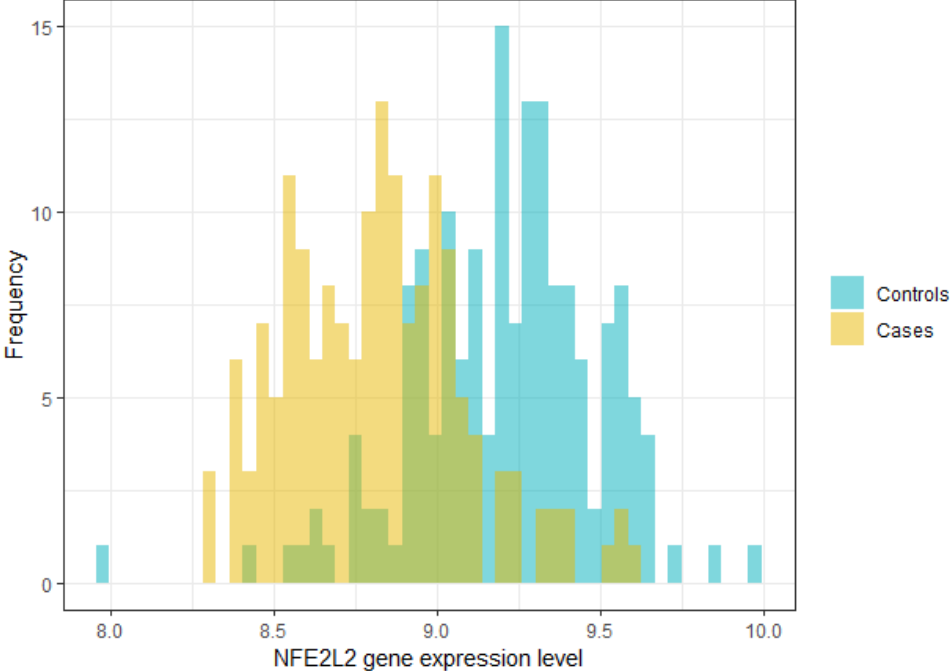
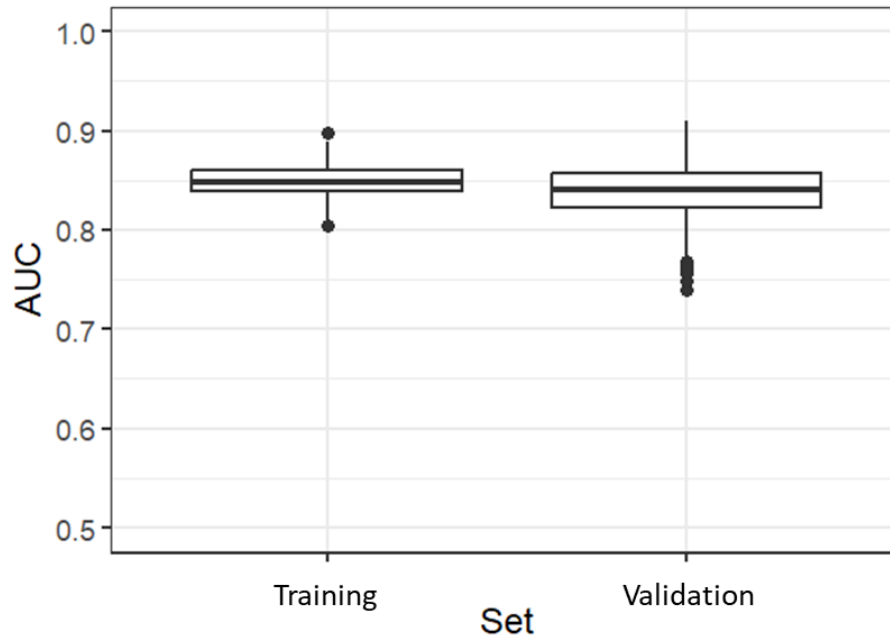


Figure S4. Box plots of the unweighted covariate-adjusted AUC^\perp calculated on the 500 replications of the training set (representing 60% of the matched case-control) and the validation set (representing 40%) for the cardiac disease outcome.



The box delineates the interquartile range (IQR) and contains a horizontal line corresponding to the median; outside the box, Tukey-style whiskers extend to a maximum of $1.5 \times \text{IQR}$ beyond the box. AUC is 0.85 (95%CI 0.81-0.88) and 0.84 (95%CI 0.77-0.89) for the training and validation sets, respectively

$^\perp$ Unweighted covariate-adjusted AUC was computed by the CLR model, including the year of cancer diagnosis, cumulative anthracycline doses, mean heart radiation doses, and NFE2L2 gene expression.

Figure S5. Venn diagram representing the genes identified by the Percentile Lasso, Bolasso, and Sublasso for heart failure outcome. A threshold of 50% is used for the latter two approaches. See Tables S2-S3 for additional information.

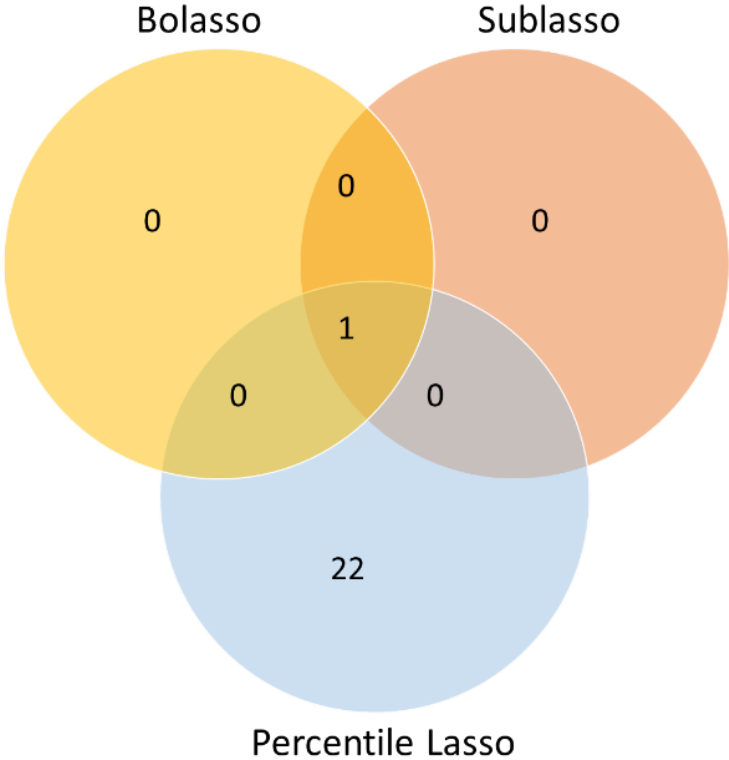
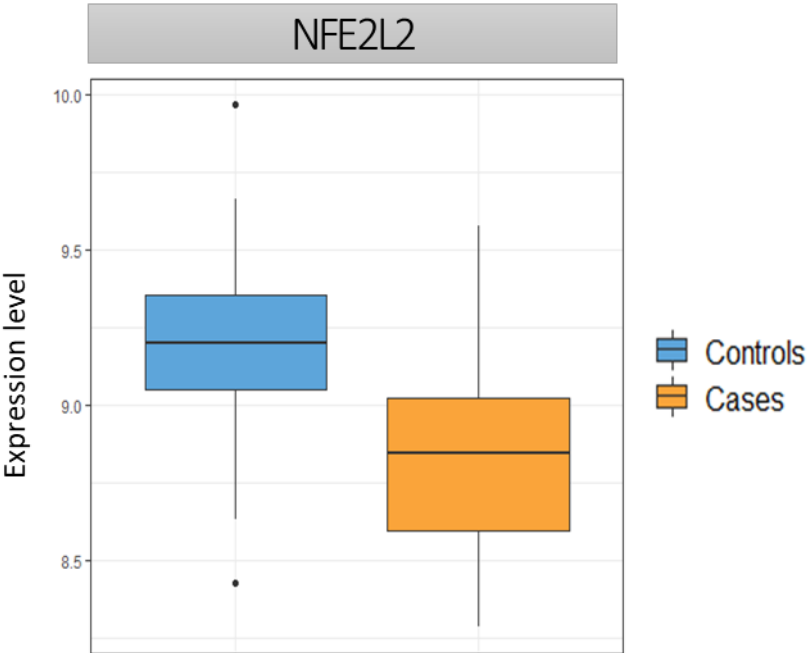
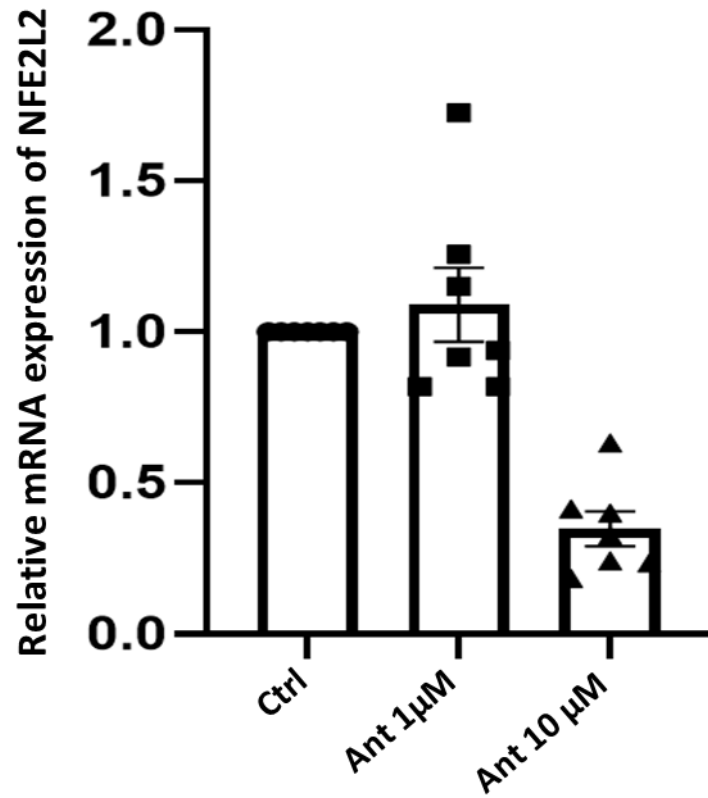


Figure S6. Box plots of the gene expression NFE2L2 for the heart failure outcome



The median of NFE2L2 expression is 9.20 (IQR, 9.05-9.36) and 8.85 (8.60-9.02) in controls and cases, respectively.

Figure S7. Expression levels of the NFE2L2 gene in six samples from adult rat ventricular cardiomyocytes before (controls) and after exposure to anthracycline (doses: 1 μ M and 10 μ M)



The expression level of the gene NFE2L2 for each cardiomyocyte are represented by dots, squares or triangles, respectively for the controls (0 μ M), 1 μ M and 10 μ M of anthracyclines.

Supplementary Tables

Table S1. Demographic and clinical characteristics of patients included in the FCCSS cohort (n=7670)

	Cases (n=495)		Controls (n=7,175)	
Gender, n (%)*				
Female	236	(47.7)	3,233	(45.1)
Age at primary cancer diagnosis (years)*				
Median (IQR)	6	(3 - 12)	5	(2 - 11)
[0 - 5[208	(42.0)	3518	(49.0)
[5 - 10[125	(25.3)	1553	(21.6)
[10 - 15[117	(23.6)	1506	(21.0)
[15 - 22[45	(9.1)	598	(8.3)
Type of malignancy, n (%)*				
Lymphoma	167	(33.7)	1,112	(15.5)
Neuroblastoma	83	(16.8)	1,045	(14.6)
Bone sarcoma	71	(14.3)	964	(13.4)
Nephroblastoma	53	(10.7)	648	(9.0)
Soft tissue sarcoma	51	(10.3)	768	(10.7)
Gonad	22	(4.4)	417	(5.8)
Others	48	(9.8)	2,221	(31.0)
Cardiac disease, n (%)				
Heart Failure	248	(50.1)		
Valvular Heart Diseases	120	(24.2)		
Arrhythmia	56	(11.3)		
Pericardial disease	28	(5.7)		
Ischemia	23	(4.6)		
Others	20	(4.0)		
Follow-up after cancer diagnosis*[‡] (years)				
Median (IQR)	6	(3 - 12)	5	(2 - 11)
[0 - 10[*	83	(16.8)	440	(6.1)
[10 - 20[131	(26.5)	588	(8.2)
[20 - 30[140	(28.3)	2578	(35.9)
[30 - 40[101	(20.4)	2100	(29.3)
[40 - 50[31	(6.3)	1095	(15.3)
≥ 50	9	(1.8)	374	(5.2)
Year of primary cancer diagnosis, n (%)				
Before 1970	87	(17.6)	669	(9.3)
[1970, 1979[193	(39.0)	1483	(20.7)
[1980, 1989]	158	(31.9)	2340	(32.6)
[1990, 2000]	57	(11.5)	2683	(37.4)
Anthracyclines, n (%)				
No	229	(46.2)	4831	67.3
Yes	266	(53.7)	2344	32.7
Anthracycline doses (mg/m²)				
Median (range)	352.3	(46.3-1705.1)	217.9	(2.6-3595.4)
None	229	(46.3)	4831	(67.3)
]0 - 250]	76	(15.4)	1390	(19.4)
]250 - 360]	67	(13.5)	463	(6.5)
>360	123	(24.8)	591	(6.8)
Heart Radiations				
Missing	1	(0.2)	174	(2.4)
No	127	(25.7)	3614	(50.4)
Yes	367	(74.1)	3387	(47.2)

Mean heart radiation doses (Gy)				
Median (range)	12.3	(0.0- 49.1)	1.3	(0.0- 61.2)
None	127	(25.7)	3614	(51.6)
]0 – 5]	149	(30.2)	2245	(32.1)
]5 – 15]	44	(8.9)	540	(7.7)
>15	174	(35.2)	602	(8.7)
Missing	1	(0.0)	174	(2.4)

IQR: Interquartile range, Gy: Gray; Cases are patients included in the FCCSS who developed a cardiac disease (CD) after a childhood cancer diagnosis. Controls are patients included in the FCCSS who did not develop a CD.

*: variables used for the matching between cases and controls; ‡ In cases, follow-up begins at cancer diagnosis and ends at CD development, which can occur before five years of follow-up.

Table S2. Number of genes selected by the Percentile lasso, Bolasso, and Sublasso approaches with or without constraint for the cardiac disease outcome (top part) and heart failure (bottom part)

Approach	Without Constraint [†]	Constraint					
		$\theta=95\%$ [‡]	40%	50%	60%	70% [*]	80% [*]
Cardiac Diseases (primary outcome)							
Percentile Lasso	61	45					
Bolasso	1525			5	2	1	0
Sublasso	1311			6	2	1	1
Heart Failure (secondary outcome)							
Approach	Without Constraint [†]	Constraint					
		$\theta=95\%$	40%	50%	60%	70%	80%
Percentile Lasso	41	23					
Bolasso	1320		1	1	0	0	0
Sublasso	1182		1	1	0	0	0

[†] Without constraint, corresponds to $\theta = 0$ for the Percentile Lasso and no cut-off specified for Bolasso and Sublasso approaches, [‡] $\theta=95\%$ corresponds to $\hat{\lambda}(95) = 5.41$, ^{*} The same gene, NFE2L2, is selected for both Bolasso and Sublasso approaches when applying the threshold of 70%, ^{*} Only the Sublasso approach selects one gene, NFE2L2, when applying a threshold of 80%

Table S3. Selection frequency of the genes selected by the Bolasso and Sublasso approaches with a cut-off set of 50% from 1000 replications for cardiac disease outcome (primary endpoint, top part) and heart failure (bottom part). These genes were also selected for the Percentile lasso approach (See Figure 2).

Selected genes	Selection frequency	
	Bolasso, n (%)	Sublasso, n (%)
Cardiac Diseases (primary outcome)		
NFE2L2	757 (75.7)	805 (80.5)
TGIF2 - RAB5IF	622 (62.2)	662 (66.2)
HK2	577 (57.7)	565 (56.5)
HDGFL3	535 (53.5)	527 (52.7)
LSP1 - MIR7847	508 (50.8)	551 (55.1)
RGS1*		527 (52.7)
Heart failure (secondary outcome)		
NFE2L2	526 (52.6)	570 (57.0)

*: As the selection frequency of the RGS1 gene is lower than 50% by using the Bolasso approach, we did not report its frequency.

Table S4. Estimates of correlation between NFE2L2 and cumulative anthracycline doses or mean heart radiation doses using linear regression. The correlations are tested one by one.

	n/N	Mean (min-max)	Comparison	Adjusted p-value [¶]
Cardiac Diseases (primary outcome)				
Cumulative anthracycline doses (mg/m²)				
NFE2LE expression				
None	57 / 132	9.08 (8.30 – 9.97)		
0 – 250	32 / 77	8.99 (7.97 – 9.64)	None vs. 0 – 250	0.1881
250 – 360	25 / 48	9.01 (8.39 – 9.66)	None vs. 250 - 360	0.2753
> 360	51 / 73	8.92 (8.28 – 9.62)	None vs. > 360	0.0046
			0 – 250 vs. 250 - 360	0.7264
			0 – 250 vs. > 360	0.2563
			250 – 360 vs. > 360	0.2423
Kruskal-Wallis $\chi^2= 11.79$, p-value = 0.0081				
Mean heart radiation doses (Gy)				
NFE2LE expression				
None	51 / 107	9.02 (8.26 – 9.84)		
0 – 5	54 / 120	9.07 (8.30 – 9.97)	None vs. 0 – 5	0.3769
5 – 15	10 / 24	9.04 (8.39 – 9.62)	None vs. 5 – 15	0.8663
> 15	50 / 79	8.90 (7.97 – 9.64)	None vs. > 15	0.0703
			0 – 5 vs. 5 – 15	0.7303
			0 – 5 vs. > 15	0.0045
			5 – 15 vs. > 15	0.2168
Kruskal-Wallis $\chi^2= 11.59$, p-value = 0.0089				
Heart failure (secondary outcome)				
Cumulative anthracycline doses (mg/m²)				
NFE2LE expression				
None	7 / 37	9.21 (8.75 – 9.97)		
0 – 250	13 / 33	8.99 (8.42 – 9.60)	None vs. 0 – 250	0.0124
250 – 360	18 / 33	9.04 (8.39 – 9.66)	None vs. 250 - 360	0.0618
> 360	36 / 45	8.89 (8.28 – 9.62)	None vs. > 360	< 0.0001
			0 – 250 vs. 250 - 360	0.4898
			0 – 250 vs. > 360	0.2680
			250 – 360 vs. > 360	0.0751
Kruskal-Wallis $\chi^2= 19.61$ p-value = 0.0002				
Mean heart radiation doses (Gy)				
NFE2LE expression				
None	27 / 56	9.02 (8.29 – 9.58)		
0 – 5	28 / 60	9.07 (8.41 – 9.97)	None vs. 0 – 5	0.6738
>5 [†]	19 / 32	8.94 (8.38 – 9.64)	None vs. > 5	0.1807
			0 – 5 vs. > 5	0.1610
Kruskal-Wallis $\chi^2= 3.89$, p-value = 0.1432				

n = number of cases; N = total number of patients; [¶] Dunn's Test post-hoc for multiple comparisons; [†]Some categories were collapsed due to the small sample size.

Table S5. Effect of NFE2L2, TGIF2-RAB5IF, HK2, HDGFL3, and LSP1-MIR7847 across dose levels of cumulative anthracycline dose and mean heart radiation dose (interaction terms) using conditional logistic regression adjusted on the year of cancer diagnosis for cardiac disease outcome. The interaction terms are tested one by one.

Gene	n/N	Odds Ratio (95% CI)	p-value
Cumulative anthracyclines dose (mg/m²)			
LSP1			0.8315
None	57 / 132	7.78 (3.17 - 19.09)	
1 – 250	32 / 77	9.17 (3.49 – 25.16)	
250 – 360	25 / 48	8.49 (3.35 – 21.86)	
> 360	51 / 73	8.60 (3.38 – 22.35)	
HDGFL3			0.8219
None	57 / 132	0.54 (0.34 - 0.86)	
1 – 250	32 / 77	1.92 (1.03 - 3.62)	
250 – 360	25 / 48	1.72 (0.85 - 3.57)	
> 360	51 / 73	1.62 (0.86 - 3.08)	
NFE2L2			0.3288
None	57 / 132	0.08 (0.03 – 0.24)	
1 – 250	32 / 77	3.23 (0.82 – 12.60)	
250 – 360	25 / 48	2.98 (0.76 – 11.71)	
> 360	51 / 73	2.16 (0.54 – 8.66)	
TGIF2 - RAB5IF			0.6357
None	57 / 132	0.28 (0.15 – 0.52)	
1 – 250	32 / 77	1.39 (0.56 – 3.55)	
250 – 360	25 / 48	2.18 (0.85 – 5.69)	
> 360	51 / 73	1.36 (0.54 – 3.56)	
HK2			0.6037
None	57 / 132	3.30 (1.92 – 5.68)	
1 – 250	32 / 77	4.42 (2.38 – 8.35)	
250 – 360	25 / 48	3.94 (2.15 – 7.45)	
> 360	51 / 73	3.99 (2.23 – 7.18)	
Mean heart radiation doses (Gy)			
LSP1 - MIR7847			0.5521
None	51 / 107	6.19 (2.63 – 14.57)	
0 – 5	54 / 120	7.21 (2.98 – 17.53)	
> 5 ¶	60 / 103	8.06 (3.15 – 21.24)	
HDGFL3			0.9576
None	51 / 107	0.61 (0.40 – 0.93)	
0 – 5	54 / 120	1.55 (0.93 – 2.62)	

5 – 15	10 / 24	1.69 (0.77 – 4.13)	
>15	50 / 79	1.46 (0.83 – 2.58)	
NFE2L2			0.6933
None	51 / 107	0.22 (0.10 – 0.47)	
0 – 5	54 / 120	0.80 (0.28 – 2.35)	
5 – 15	10 / 24	0.46 (0.11 – 7.16)	
>15	50 / 79	1.06 (0.37 – 3.09)	
TGIF2 - RAB51F			0.4234
None	51 / 107	0.26 (0.13 – 0.52)	
0 – 5	54 / 120	1.86 (0.79 – 4.36)	
5 – 15	10 / 24	0.50 (0.14 – 6.51)	
>15	50 / 79	1.51 (0.62 – 3.72)	
HK2			0.8768
None	51 / 107	2.59 (1.53 – 4.38)	
1 – 5	54 / 120	3.74 (2.11 – 6.66)	
5 – 15	10 / 24	3.37 (1.78 – 6.91)	
> 15	50 / 79	3.76 (2.07 – 6.92)	

n = number of cases; N = total number of patients; ¶ Some categories were collapsed due to the small sample size.

We wanted to evaluate the interactions between gene expression and treatment doses by dividing gene expression into two classes (higher or lower than the mean expression in controls), but given the low number of events in the strata, analyses could not be carried out.

Table S6. Demographic and clinical characteristics of patients included in the matched-case control study for heart failure outcome (n=148)

	Cases (n=74)		Controls (n=74)	
Gender, n (%)*				
Female	32	(43.2)	32	(43.2)
Age at primary cancer diagnosis in years				
Median (IQR)	4	(1 – 7)	4	(2 - 7)
[0 – 5[44	(59.5)	45	(60.8)
[5 – 10[21	(28.4)	21	(28.4)
[10 – 15[7	(9.5)	6	(8.1)
[15 – 22[2	(2.7)	2	(2.7)
Type of malignancy, n (%)*				
Lymphoma	17	(23.0)	17	(23.0)
Neuroblastoma	14	(18.9)	14	(18.9)
Bone sarcoma	14	(18.9)	14	(18.9)
Nephroblastoma	10	(13.5)	10	(13.5)
Soft tissue sarcoma	11	(14.9)	11	(14.9)
Gonad	3	(4.1)	3	(4.1)
Others	5	(6.8)	5	(6.8)
Follow-up after diagnosis in years, n (%)				
Median (IQR)	18	(12 – 23)	35	(25 – 42)
[0 – 10]*	15	(20.3)	0	(0.0)
[10 – 20[26	(35.1)	5	(6.8)
[20 – 30[24	(32.4)	20	(27.0)
[30 – 40[9	(12.2)	26	(35.1)
[40 – 50[0	(0.0)	14	(18.9)
≥ 50	0	(0.0)	9	(12.2)
Year of primary cancer diagnosis, n (%)				
Before 1970	0	(0.0)	15	(20.3)
[1970, 1979[33	(44.6)	20	(27.0)
[1980, 1989]	33	(44.6)	24	(32.4)
[1990, 2000]	8	(10.8)	15	(20.3)
Anthracyclines, n (%)				
No	7	(9.5)	30	(40.5)
Yes	67	(90.5)	44	(59.5)
Cumulative anthracycline doses (mg/m²)				
Median (range)	362.3	(106.2-1207.5)	280.4	(84.4-546.1)
None	6	(8.1)	29	(39.2)
]0 – 250]	14	(18.9)	21	(28.4)
]250 – 360]	20	(27.0)	15	(20.3)
>360	34	(46.0)	9	(12.2)
Heart Radiations				
No	27	(36.5)	29	(39.2)
Yes	47	(63.5)	45	(60.8)
Mean heart radiation doses (Gy)				
Median (range)	2.5	(0.0 – 31.7)	1.2	(0.0 – 41.6)
None	27	(36.5)	29	(39.2)
]0 – 5]	28	(37.8)	32	(43.2)
]5 – 15]	3	(4.1)	5	(6.8)
>15	16	(21.6)	8	(10.8)

IQR: Interquartile range, Gy: Gray, * variables used for the matching between cases and controls, * In cases, the follow-up starts from the cancer diagnosis and ends at heart failure development, which can occur before the first five years of follow-up.

Table S7. Comparison of NFE2L2 gene expression levels before and after exposure to anthracyclines

	Mean (min-max)	Comparison	Mean difference (95% CI)	Adjusted p-value[¶]
Ctrl	1.00 (1.00 – 1.00)			
Ant 1 μM	1.09 (0.81 – 1.73)	Ctrl vs. Ant 1 μ M	-0.09 (-0.37 – 0.19)	0.7586
Ant 10 μM	0.35 (0.18 – 0.63)	Ctrl vs. Ant 10 μ M	0.65 (0.37 – 0.93)	0.0022
		Ant 1 μ M vs. Ant 10 μ M	0.74 (0.46 – 1.02)	0.0032
Kruskal-Wallis $\chi^2= 13.97$, p-value = 0.0009				

Ctrl: control, Ant: Anthracyclines, CI: Confidence interval, [¶]Dunn’s Test post-hoc for multiple comparisons

Table S8. Effects of cumulative anthracycline doses and mean heart radiation doses on cardiac diseases (main endpoint) and heart failure (secondary endpoint) estimated from a conditional logistic regression adjusted on year of cancer diagnosis

Variables	n/N	Odds Ratio (95% Confidence Interval)	p-value
Cardiac Diseases (primary outcome)			
Year of cancer diagnosis			0.8262
< 1970	17 / 46	1.00	
[1970-1979]	57 / 100	1.53 (0.60 – 3.87)	
[1980-1989]	67 / 132	1.42 (0.46 – 3.62)	
[1990-2000]	24 / 52	1.57 (0.51 – 4.85)	
Cumulative anthracycline doses (mg/m²)			0.0008
None	57 / 132	1.00	
]0 - 250]	32 / 77	1.03 (0.49 – 2.17)	
]250 - 360]	25 / 48	2.04 (0.88 – 4.76)	
>360	51 / 73	3.85 (1.74 – 8.52)	
Mean heart radiation doses (Gy)			0.0037
None	51 / 107	1.00	
]0 – 5]	54 / 120	0.86 (0.45 – 1.63)	
]5 – 15]	10 / 24	0.74 (0.26 – 2.12)	
>15	50 / 79	2.81 (1.33 – 5.92)	
Heart Failure (secondary outcome)			
Year of cancer diagnosis[¶]			0.6171
< 1980	33 / 68	1.00	
[1980-2000]	41 / 80	1.32 (0.44 – 4.00)	
Cumulative anthracycline doses (mg/m²)^{¶¶}			
None	7 / 37	1.00	
]0 - 250]	13 / 33	9.40 (1.10 – 80.58)	
>250	54 / 78	40.27 (4.70 – 344.77)	
Mean heart radiation doses (Gy)^{¶¶}			0.4086
None	27 / 56	1.00	
]0 – 15]	31 / 68	1.32 (0.44 – 3.94)	
>15	16 / 24	1.86 (0.63 – 12.96)	

n = number of cases; N = total number of patients;