HAL
open science

Délai de prise en charge des cancers du sein dans la Somme selon l'évolution de l'incidence, le niveau socio-économique, le type de découverte, et répercussion sur le stade au diagnostic

Bruno Sawczynski

To cite this version:

Bruno Sawczynski. Délai de prise en charge des cancers du sein dans la Somme selon l'évolution de l'incidence, le niveau socio-économique, le type de découverte, et répercussion sur le stade au diagnostic. Médecine humaine et pathologie. 2018. dumas-02139251

HAL Id: dumas-02139251

https://dumas.ccsd.cnrs.fr/dumas-02139251

Submitted on 24 May 2019

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.

UNIVERSITÉ DE PICARDIE JULES VERNE

FACULTÉ DE MÉDECINE D'AMIENS

Année 2018 Thèse n°2018-167

THÈSE

Pour l'obtention du Diplôme d'État de Docteur en Médecine

Spécialité Santé Publique et Médecine Sociale

DÉLAI DE PRISE EN CHARGE DES CANCERS DU SEIN DANS LA SOMME SELON L'ÉVOLUTION DE L'INCIDENCE, LE NIVEAU SOCIO-ÉCONOMIQUE, LE TYPE DE DÉCOUVERTE, ET RÉPERCUSSION SUR LE STADE AU DIAGNOSTIC

Présentée et soutenue publiquement le 16 octobre 2018

Par Bruno SAWCZYNSKI

Devant un jury composé de :

Monsieur le Professeur Olivier GANRY, président du jury et directeur de thèse

Monsieur le Professeur Denis CHATELAIN

Monsieur le Professeur Bruno CHAUFFERT

Monsieur le Professeur Fabrice SERGENT

Monsieur le Docteur Salah El Din ZERKLY

À Monsieur le Professeur Olivier GANRY, président du jury et directeur de thèse,
Professeur des Universités - Praticien Hospitalier
(Épidémiologie, économie de la santé et Prévention)
Responsable du service d'Épidémiologie, hygiène hospitalière et santé publique
Pôle "Biologie, pharmacie et santé des populations"

Merci d'avoir accepté de diriger ce travail et de présider ce jury. Je vous remercie pour votre humanité et votre soutien dans les moments les plus difficiles. Un grand merci pour les bons conseils que vous m'avez donnés au cours de mon cursus. Veuillez recevoir, Monsieur GANRY, toute ma gratitude.

À Monsieur le Professeur Denis CHATELAIN,

Professeur des Universités-Praticien Hospitalier

(Anatomie et cytologie pathologique)

Vous me faites l'honneur de faire partie du jury de cette thèse, je vous en suis sincèrement reconnaissant.

À Monsieur le Professeur Bruno CHAUFFERT,

Professeur des Universités-Praticien Hospitalier

Chef du service d'Oncologie médicale

CHU D'Amiens

Merci de me faire l'honneur d'avoir accepté de juger ce travail, je vous en suis sincèrement reconnaissant.

À Monsieur le Professeur Fabrice SERGENT,

Professeur des Universités-Praticien Hospitalier

(Gynécologie-Obstétrique)

Vous me faites l'honneur et le plaisir de faire partie du jury de cette thèse. Vous trouverez ici l'expression de mes sincères remerciements et de mon respect le plus profond.

À Monsieur le Docteur Salah El Din ZERKLY,

Assistant Spécialiste

Service d'Épidémiologie, hygiène hospitalière et santé publique

CHU Amiens Picardie

Merci de me faire l'honneur et le plaisir de faire partie de ce jury. Merci pour tous tes conseils, ton aide précieuse et ton soutien inconditionnel. Merci pour toutes ces discussions, qu'elles gravitent autour de la médecine ou non. Ne change pas.

Remerciements personnels

À **mes parents**, pour votre amour, votre soutien indéfectible, pour tout ce que vous avez fait pour nous (oui j'inclus la fratrie). Merci de nous avoir inculqué certaines valeurs et de nous avoir tout donné. Vous avez fait de moi ce que je suis aujourd'hui en grande partie. Je ne vous remercierai jamais assez pour tout cela. Je vous aime.

À **François et Claire**, mes frère et sœur. Merci pour votre soutien et tous ces bons moments passés ensemble, hier à se chamailler et aujourd'hui autour d'une bonne table.

À **mes grands-parents**, pour certains aujourd'hui disparus. Merci pour l'amour que vous me portez et pour votre soutien.

À mes pièces rapportées, **Leïla et Yann**. Comment pourrais-je vous oublier ? J'ai beaucoup de chance de vous avoir. Ne changez pas.

À **Orene**, ma nièce préférée.

À **Ghizlane**, Un immense merci pour tout. J'aurais pu faire 60 pages juste sur toi alors je ne vais pas tout énumérer. Ton tajine de veau aux pruneaux et ton poulet aux olives vont me manquer.

À **Nadir**, pour nos soirées foot, tes leçons de bricolages pour m'aider à m'en sortir et tes plans jardinage foireux, merci. En espérant que ton niveau tennistique ne baisse pas en mon absence…

À **Ines** et **Zyad**, vos deux petits monstres.

À **Mathilde**, mon ponuche préféré. Tellement de choses à dire, et si peu de place. Reste comme tu es. J'te kiffe.

À **Abdel Majax**, parce que t'es le plus gros d******* que la Terre ait jamais porté.

À **Jilali**, le musicien et kabyle du groupe. Tahia Djazair !

À **Amate**, tu resteras à jamais la meilleure interne que j'ai eue. Pour ton tire-bouchon thaïlandais et ta conception du lavage de la vaisselle. Pour ta bonté, ta gentillesse et ta générosité. Merci

À **Karim**, même si tu es loin je ne t'oublie pas.

À **Salih**, l'oignon fait la force mon ami !

À **Assim**, le seul mec que je connaisse qui travaille en robe.

À **Adile, Yasmine** et **Sarah**, même si je ne sais pas vraiment pourquoi je vous cite.

À mes amis de l'externat : **Jiss**, **Soufiane**, **Samy**, **Clément**, **J-Y**, **Muhanad** et ceux que j'oublie…

À mes amis de St Quentin : **Amel**, **Charlotte**, **Chloé**, **Clémence**, **Hind**, **Mouna** et **Adrien**. Nous avons passé d'agréables soirées ensemble, en espérant qu'il y en aura d'autres. Une mention particulière pour **Charlotte**, merci d'avoir relu ce travail et d'y avoir apporté quelques corrections pertinentes.

À celles et ceux que j'ai croisé de près ou de loin pendant mon internat et avec qui j'ai passé de bons moment : **Alexandra**, **Estelle**, **Gaëtan**, **Mumu**, **Guillaume**, **Sébastien**.

Au **Dr Aurore Prud'homme**, pour tout ce que vous m'avez appris.

Au **Dr Benoît Vaysse**, parce que toi aussi tu m'en as beaucoup appris.

Au **Dr Bagadule**, parce que tu as bien le droit à une deuxième dédicace, m****, quand même !

À **Mme Christine Bué et Mr Alain Marchand**, pour votre gentillesse et pour la confiance que vous m'avez accordée pendant toutes ces années.

À **Mr Eric Trécherel**, sans qui ce travail n'aurait pas été possible. Un immense merci pour le temps que tu as toujours pris pour répondre à mes questions ou sur mes requêtes et pour ton soutien.

Au **Dr Bénédicte Lapôtre-Ledoux**, merci pour vos conseils et votre soutien.

À **Mmes Astrid Causel**, **Christine Cotte**, **Claudie Cussey**, **Agnès Thuilliez** et **Nadia Muguet** pour avoir dû remettre le nez dans beaucoup de dossiers pour ce travail sans m'en avoir tenu rigueur.

À **Del Piero**, **Pirlo** et cie, pour leurs coups-francs et autres actions d'anthologie et toutes les émotions par lesquelles ils m'ont fait passer.

TABLE DES MATIÈRES

1. INTRODUCTION 19

1.1. Épidémiologie du cancer du sein 19

1.2. Facteurs de risque du cancer du sein 21

1.3. Dépistage organisé du cancer du sein 21

1.4. Stade au diagnostic 23

1.5. Objectifs de l'étude 23

2. MATÉRIEL ET MÉTHODE 24

2.1. Le registre du cancer 24

2.2. Les IRIS et score EDI 24

2.3. Schéma de l'étude 26

2.4. Critères d'inclusion 27

2.5. Critères d'exclusion 27

2.6. Données recueillies 28

2.7. Critère de jugement principal 30

2.8. Critères de jugement secondaires 30

2.9. Gestion des données manquantes 30

2.10. Analyses statistiques 31

2.11. Considérations éthiques et légales 33

3. RÉSULTATS 34

3.1. Analyse descriptive de tous les cas de cancers entre 2004 et 2014 34

3.2. Analyse descriptive des cas de cancers avec délai de prise en charge 38

3.3. Délai de prise en charge en fonction du taux d'incidence ajusté 51

4. DISCUSSION 56

4.1. Du recueil de données et des données manquantes 56

4.2. Du tri à plat 57

4.3. Des résultats du modèle de régression linéaire multiple 58

5. CONCLUSION 59

RÉFÉRENCES BIBLIOGRAPHIQUES 60

ANNEXES 64

TABLE DES ILLUSTRATIONS

Tableau 1 : Liste des variables recueillies

Tableau 2 : Liste des variables créées

Tableau 3 : Correspondance entre TNM et Stade UICC

Tableau 4 : Calcul du grade SBR

Tableau 5 : Effectifs et répartition des cas de cancer du sein dans la Somme entre 2004 et 2014

Tableau 6 : Paramètres de la variable âge

Tableau 7 : Effectifs et répartition des modes de découverte des cancers du sein

Tableau 8 : Effectifs et répartition des grades SBR

Tableau 9 : Paramètres de la variable âge

Tableau 10 : Effectifs et répartition des cas de cancer du sein entre 2004 et 2014

Tableau 11 : Effectifs et répartition des cas de cancer du sein entre 2004 et 2014 selon le mode de diagnostic

Tableau 12 : Paramètres de la variable Délai de prise en charge

Tableau 13 : Score EDI en fonction du mode de découverte

Tableau 14 : Effectifs et répartition de la classification SBR

Tableau 15 : Effectifs et répartition de la classification SBR en fonction du mode de diagnostic

Tableau 16 : Effectifs et répartition des stades UICC

Tableau 17 : Paramètres de la variable EDI

Tableau 18 : Score EDI en fonction du stade UICC

Tableau 19 : Variables et leurs p values

Tableau 20 : Résultats du modèle de régression linéaire multiple

Figure 1 : Évolution du taux d'incidence brut au cours du temps, pour 100 000 personnes-années, en France et dans la Somme

Figure 2 : Répartition de l'âge au diagnostic (histogramme)

Figure 3 : Âge moyen au diagnostic en fonction du grade SBR

Figure 4 : Histogramme de la répartition des cas de cancer du sein en fonction du score EDI

Figure 5 : Flow chart de l'étude

Figure 6 : Répartition des effectifs par année

Figure 7 : Effectifs des cas de cancer du sein par année et par type de découverte

Figure 8 : Répartition du délai de prise en charge, en jours

Figure 9 : Évolution du délai de prise en charge au cours du temps

Figure 10 : Évolution du délai de prise en charge au cours du temps selon le mode de découverte

Figure 11 : Délai de prise en charge en fonction du score EDI

Figure 12 : Délai de prise en charge en fonction du score EDI pour les cancers diagnostiqués par dépistage individuel

Figure 13 : Délai de prise en charge en fonction du score EDI pour les cancers diagnostiqués par dépistage organisé

Figure 14 : Délai de prise en charge en fonction du score EDI pour les cancers diagnostiqués par présence de signes cliniques

Figure 15 : Grade SBR et âge

Figure 16 : Score EDI en fonction du grade SBR

Figure 17 : Effectifs des stades UICC

Figure 18 : Répartition des scores EDI (quintiles)

Figure 19 : Répartition des scores EDI en fonction de l'âge

Figure 20 : Répartition des scores EDI en fonction du mode de découverte

Figure 21 : Effectifs des cancers diagnostiqués par signes cliniques par année, en fonction du stade UICC

Figure 22 : Effectifs des cancers diagnostiqués par dépistages individuel et organisé par année, en fonction du stade UICC

Figure 23 : QQplot de la normalité de la variable délai de prise en charge

Figure 24 : QQplot de la normalité de la variable délai de prise en charge après transformation logarithmique

Figure 25 : Diagrammes des conditions de validité du modèle

ABRÉVIATIONS

ANAES : Agence Nationale d'Accréditation et d'Evaluation en Santé

DI : Dépistage Individuel

DO : Dépistage Organisé

EDI : European Deprivation Index

ENCR : European Network of Cancer Registries

EPCI : Etablissements Publics de Coopération Intercommunale

HAS : Haute Autorité de Santé

INCa : Institut National Du Cancer

IRIS : Ilots Regroupés pour l'Information Statistique

RCP : Réunion de Concertation Pluridisciplinaire

SBR : Scarff Bloom et Richardson

SC : Signes Cliniques

TNM : Tumor Node Metastasis

UICC : Union for International Cancer Control

1. INTRODUCTION

1.1. Épidémiologie du cancer du sein

Selon GLOBOCAN, le cancer du sein est le second cancer le plus fréquent dans le monde et, de loin, le plus fréquent chez les femmes avec 1.67 million de nouveaux cas de cancer diagnostiqués chaque année [1] et en 2012, 6.3 millions de femmes vivaient avec un cancer du sein diagnostiqué au cours des cinq années précédentes [2]. C'est le cancer le plus courant chez les femmes dans les pays développés avec 794 000 nouveaux cas par an. Les taux d'incidence varient presque de quatre fois dans les différentes régions du monde, avec des taux allant de 27 à 92 pour 100 000.

Le cancer du sein est la cinquième cause de décès par cancer (522 000 décès) et le cancer le plus fréquemment diagnostiqué chez les femmes dans 140 des 184 pays couverts par GLOBOCAN dans le monde. Il représente maintenant un cancer sur quatre chez les femmes et la seconde cause de mortalité par cancer dans les pays les plus développés (198 000 décès, 15.4%) après le cancer du poumon. Le cancer du sein est aussi l'une des principales causes de décès par cancer dans les pays les moins développés (324 000 décès, 14.3% du total). Le rang des taux de mortalité entre les régions du monde est inférieur à celui de l'incidence en raison de la survie plus favorable du cancer du sein dans les pays développés (à forte incidence) [1-3].

En France, 48 763 nouveaux cas ont été estimés en 2012, ce qui situe le cancer du sein au 2ème rang le plus fréquent chez la femme. Le taux d'incidence standardisé est de 88.0 pour 100 000 personnes-années. Avec 11 886 décès, ce cancer se situe au 3ème rang des décès. Le taux de mortalité standardisé est de 15.7 pour 100 000 [4].

Dans le département de la Somme, le nombre de cas de cancer du sein chez les femmes a augmenté entre 2010 et 2014, passant de 547 en 2010 à 681 en 2014, le taux d'incidence standardisé sur la population mondiale passant de 93.9 à 113.2 pour 100 000 personnes-années sur la même période [5]. Ce taux avait augmenté progressivement augmenté entre 1982 et 2005, passant de 57.1 sur la période 1982-2002 à 107.2 sur la période 2002-2005, diminuant sur la période 2006-2009 (98.7 pour 100 000 personnes-années). Dans le même temps, le taux de mortalité standardisé est passé de 20.9 sur la période 1982-1985 à 21.2 sur la période 2006-2009 [6].

Les délais de prise en charge diagnostique et thérapeutique sont très importants pour l'efficacité du traitement et de la qualité de vie des personnes atteintes de cancer. En effet, Smith et al. [7] ont évoqué l'impact négatif sur la survie d'un délai de traitement supérieur à 6 semaines, tandis que Shin et al. [8] ne rapportaient un effet délétère sur la survie que lorsque le délai entre le diagnostic et la chirurgie excédait 12 semaines.

Le délai de prise en charge des cancers du sein a été étudié dans plusieurs régions de France (Alsace, Auvergne, Franche-Comté, Lorraine, Nord-Pas-de-Calais, Picardie, Poitou-Charentes, PACA) en 2011, étude menée par l'INCa (Institut National du Cancer), en collaboration avec les Réseaux Régionaux de Cancérologie ainsi que les Observatoires Régionaux de Santé. Ce délai était segmenté en plusieurs parties : le délai d'accès au diagnostic (entre la mammographie et la biopsie) et le délai d'accès à la chirurgie (entre la RCP pré-opératoire et la chirurgie). Ces délais étaient en moyenne de 18 et 23 jours, constituant un délai moyen entre la mammographie et la chirurgie de 41 jours [9], alors que l'ANAES (Agence Nationale d'Accréditation et d'Evaluation en Santé) préconise une prise en charge de première intention d'un mois entre la mammographie et le traitement [10].

Plus récemment, en 2013, Molinié et al. ont étudié le délai de prise en charge entre la première imagerie mettant en évidence le cancer du sein et le premier traitement, mettant en évidence une médiane de délai de 34 jours [11].

Enfin, en 2018, Mimouni et al. ont également étudié ce délai de prise en charge dans une étude unicentrique au Maroc (Institut National d'Oncologie de Rabat) qui montrait un délai moyen de 69 jours [12].

Le taux d'incidence du cancer du sein dans la Somme étant en augmentation ces dernières années, la question du délai de prise en charge s'est posée. En effet, nous nous sommes questionnés sur le fait que le système de santé, qui présente une stabilité depuis plusieurs années, aurait pu être impacté par cette augmentation de l'incidence des cancers du sein et ainsi être saturé. Ceci aurait pu avoir une répercussion sur le délai de prise en charge des patientes atteintes de cancer du sein, et on aurait ainsi pu observer une augmentation de ce délai.

1.2. Facteurs de risque du cancer du sein

Les facteurs reconnus du cancer du sein sont l'âge, les antécédents personnels et familiaux, les prédispositions génétiques, la consommation d'alcool, de tabac, le surpoids ou l'obésité.

D'autres facteurs, comme le diabète de type 2, sont fortement suspectés comme étant des facteurs aggravants [13-22].

De plus, l'environnement socio-économique est connu pour représenter un déterminant de santé important [23-25].

Concernant le cancer du sein, l'incidence et la mortalité sont plus fortes dans les classes socio-économiques plus élevées [26], mais le pronostic y est meilleur que dans les classes inférieures qui présentent un stade au diagnostic plus avancé et une mortalité relative plus élevée [27].

1.3. Dépistage organisé du cancer du sein

Le dépistage organisé (DO) du cancer du sein a été généralisé en France en 2004, après avoir été progressivement mis en place pendant plus d'une décennie dans plusieurs départements pilotes. Dans la Somme, un de ces départements, le programme existe depuis 1991.

Le DO s'effectue tous les 2 ans, est gratuit pour les patientes, et comprend un examen clinique, au moins deux clichés radiographiques par sein, la comparaison avec les clichés antérieurs et une deuxième lecture par deux radiologues différents [28].

Ne sont concernées que les femmes ayant un niveau de risque « moyen » de développer un cancer du sein, c'est à dire les femmes de 50 à 74 ans, sans antécédent personnel ni familial de cancer du sein.

Certaines femmes ont un niveau de risque « élevé » de développer un cancer du sein, il s'agit des femmes ayant des antécédents de cancer du sein ou de l'utérus et/ou de l'endomètre, certaines affections du sein (comme l'hyperplasie atypique ou une affection proliférative bénigne), ou les femmes ayant été exposées à une irradiation thoracique à haute dose avant l'âge de 30 ans (traitement d'un lymphome de Hodgkin par exemple).

Enfin, d'autres femmes ont un risque « très élevé » de développer un cancer du sein. Il s'agit des femmes ayant des prédispositions génétiques (mutations familiales BRCA1 et BRCA2) [29].

Malgré des résultats incontestables (bien que moindres que prévus) concernant la réduction de la mortalité par cancer du sein, tournant autour de 20% pour les femmes invitées [30-32], le DO est souvent sous le feu des controverses [33]. Le problème du surdiagnostic (défini par les cancers diagnostiqués et traités qui n'auraient jamais cliniquement évolué) est omniprésent et très difficile à évaluer avec nos connaissances actuelles.

En France, la cible de 75% de couverture par le DO n'est pas atteinte (nécessaire pour une diminution optimale de la mortalité et un rapport coût/efficacité favorable) avec une couverture nationale à 51.5% en 2015, soit plus de 2.5 millions de femmes dépistées, qui semble s'être stabilisée et peu évoluer ces dernières années [3]. À ce chiffre il convient d'ajouter environ 10% de femmes se faisant dépister en dehors du circuit du DO [34].

Dans le département de la Somme, la participation est un peu plus élevée que la moyenne nationale, avec 56.5% de participation au DO en 2011-2012, ainsi qu'un taux moyen de mammographies réalisées de manières individuelles un peu plus élevé qu'ailleurs dans la plupart des établissements publics de coopération intercommunale (EPCI) du département [35]. Près de 25 000 femmes sur les 42 000 femmes de 50 à 74 ans domiciliées dans la Somme ont participé au dépistage organisé en 2014 [28].

Duport et al. en 2007, ont montré que les femmes réalisant leur dépistage de façon individuelle correspondaient à des femmes ayant un niveau sociodémographique élevé, un meilleur suivi médical, que les femmes ne réalisant aucun dépistage correspondaient à des femmes ayant un niveau sociodémographique bas et un suivi médical irrégulier, et qu'enfin les femmes bénéficiant du dépistage organisé étaient les femmes ayant une position intermédiaire aux deux groupes précédents [36].

Par ailleurs, nous pouvons nous poser la question de l'impact du dépistage du cancer du sein, organisé ou non, sur le délai de prise en charge d'une certaine population. En effet, les femmes dont le cancer a été diagnostiqué du fait de la présence de signes cliniques, par crainte du cancer fortement médiatisé du fait des campagnes sur le dépistage, pourraient consulter aujourd'hui plus vite après l'apparition des signes cliniques qu'il y a 10 ans. Ainsi, les cancers

du sein diagnostiqués par la présence de signes cliniques seraient moins avancés aujourd'hui qu'il y a 10 ans.

1.4. Stade au diagnostic

Concernant le stade au diagnostic, directement lié au pronostic de la maladie, plusieurs études [37-38] montrent un recul net des stades chez les participantes au dépistage. Notamment un recul des stades T2 et plus, ainsi que des cancers N+ et M+.

1.5. Objectifs de l'étude

L'objectif principal de cette étude était de mettre en évidence une modification, si elle existait, du délai de prise en charge des cancers du sein chez les femmes de 50 à 74 ans dans la Somme entre 2004 et 2014, en fonction de l'évolution de l'incidence du cancer dans la population.

Les objectifs secondaires étaient :

- Mettre en évidence une évolution des stades au diagnostic chez les femmes ayant un cancer du sein diagnostiqué en dehors du dépistage, sur des signes cliniques, traduisant ainsi une « contamination » du discours ambiant sur le dépistage et entraînant une volonté de consultation plus rapide après l'apparition des symptômes
- Mettre en évidence une différence de délai de prise en charge selon l'EDI (European Deprivation Index), le type de découverte, le grade SBR (Scarff Bloom et Richardson), le stade UICC (Union for International Cancer Control).

2. MATÉRIEL ET MÉTHODE

2.1. Le registre du cancer

Le registre du cancer de la Somme a été créé en 1982. Il a pour mission principale de recenser tous les cas de cancer diagnostiqués dans la population du département. Son travail est le résultat de deux activités complémentaires :

- d'une part la recherche, la validation et le chaînage des informations médicales, démographiques et administratives,
- d'autre part la mise en forme de cette information pour la constitution d'une base de données informatisée directement analysable [39].

Les premiers registres français ont été créés à partir de 1975 sur des initiatives individuelles. En 1986, la création du Comité National des Registres a permis d'inscrire les registres dans une politique nationale. Les registres de cancers généraux se sont alors regroupés en association au sein du réseau français des registres de cancer, FRANCIM, qui a pour objectif d'harmoniser les pratiques d'enregistrement, de coordonner et de faciliter les travaux réalisés par les registres de cancer existants, de fournir à la communauté les indicateurs épidémiologiques utiles à la connaissance et à la prise en charge des cancers en lien avec différents partenaires institutionnels [40].

2.2. Les IRIS et score EDI

En 1999, l'INSEE avait développé un découpage du territoire en mailles de taille homogène appelées IRIS2000, pour Ilots Regroupés pour l'Information Statistique et qui faisait référence à la taille visée de 2000 habitants par maille élémentaire.

Depuis, l'IRIS constitue la brique de base en matière de diffusion de données infra-communales. Il doit respecter des critères géographiques et démographiques et avoir des contours identifiables sans ambigüité et stables dans le temps. Les communes d'au moins 10 000 habitants et une forte proportion des communes de 5 000 à 10 000 habitants sont découpés en IRIS. Ce découpage constitue une partition de leur territoire. La France compte environ 16 100 IRIS dont 650 dans les DOM. Par extension, afin de couvrir l'ensemble du territoire, on assimile à un IRIS chacune des communes non découpées en IRIS.

On distingue 3 types d'IRIS :

- IRIS d'habitat : population entre 1 800 et 5 000 habitants. Ils sont homogènes quant au type d'habitat et leurs limites s'appuient sur les grandes coupures du tissu urbain
- IRIS d'activité : regroupent plus de 1 000 salariés et comptent au moins deux fois plus d'emplois salariés que de population résidente
- IRIS divers : grandes zones spécifiques peu habitées et ayant une superficie importante.

Au 1er janvier 2008, 92% des IRIS étaient des IRIS d'habitat et 5% des IRIS d'activité. Le découpage en IRIS peut être affecté par les modifications de la géographie communale (fusions de communes, créations ou rétablissement de communes…) [41].

L'EDI a été construit en trois étapes :

- La construction d'un indicateur individuel de défavorisation en utilisant les besoins fondamentaux des personnes, lui-même créé à partir d'une enquête européenne EU-SILC (European Union – Statistics on Income and Living Conditions) menée en 2006 par Eurostat (the European Commission's Statistical Office) reposant sur un questionnaire standardisé impliquant des interviews annuelles d'un panel représentatif de ménages et d'individus et prenant en compte la pauvreté objective et subjective
- L'identification et la dichotomisation de variables disponibles à la fois au niveau individuel et dans le recensement
- La construction de l'indice écologique de défavorisation, l'EDI, en sélectionnant les variables associées à l'indicateur individuel

Dans l'enquête EU-SILC, les individus devaient indiquer si leur ménage possédait certains biens ou utilisaient certains services, et pour ceux qu'ils ne possédaient/n'utilisaient pas, ils devaient indiquer si c'était pour des raisons financières ou pour d'autres raisons. Les biens/services qu'une minorité (soit moins de 50% des ménages) ne possédaient/n'utilisaient pas parce qu'ils n'avaient pas les moyens financiers étaient considérés comme des besoins fondamentaux et étaient par conséquent retenus.

Par exemple, 95% des ménages français n'avaient pas les moyens financiers de se payer une piscine, donc « posséder une piscine » n'était par conséquent pas considéré comme des

besoins fondamentaux. Par contre, 7% des ménages français n'avaient pas les moyens de manger de la viande ou du poisson ou un équivalent végétarien une fois tous les deux jours, donc « manger de la viande ou du poisson ou un équivalent végétarien une fois tous les deux jours » était par conséquent considéré comme un besoin fondamental, sa non-réalisation reflétant la défavorisation.

Plusieurs variables ont été retenues et une formule a ainsi été établie pour le calcul du score EDI :

Score = 0.11 x « Surpeuplement » + 0.34 x « Pas d'accès au chauffage central ou électrique » + 0.55 x « Non propriétaires » + 0.47 x « Chômage » + 0.23 x « Nationalité étrangère » + 0.52 x « Pas d'accès à une voiture » + 0.37 x « Ouvriers non qualifié-ouvriers agricoles » + 0.45 x « Ménage avec au moins 6 personnes » + 0.19 x « Faible niveau d'étude » + 0.41 x « Familles monoparentales ».

Les scores EDI sur l'ensemble du territoire français allaient de -5.3 à 20.52, une moyenne à 0 et un écart-type à 1.8, les scores les plus élevés correspondant aux IRIS les plus défavorisés [42].

2.3. Schéma de l'étude

Il s'agissait d'une étude épidémiologique observationnelle, descriptive, rétrospective et monocentrique.

Dans un premier temps, l'ensemble de la population des femmes atteintes de cancers du sein diagnostiqués dans le département de la Somme entre 2004 et 2014 était décrit grâce aux données disponibles pour cette population : le taux d'incidence, comparativement au taux d'incidence de l'ensemble de la population française, les effectifs et la répartition des diagnostics par année, la répartition de l'âge au diagnostic, les modes de découverte du cancer, la répartition du grade SBR et enfin la répartition des scores EDI. Ceci a permis d'obtenir un aperçu de la population des femmes atteintes de cancer du sein dans la Somme sur cette période.

Dans un second temps, nous avons étudié le délai de prise en charge des femmes atteintes de cancer du sein dans la Somme en fonction du taux d'incidence brut de cette même pathologie,

mais également en fonction du score EDI, du mode de découverte, de l'âge au diagnostic, du stade TNM au diagnostic (via le stade UICC), du grade SBR. Ces facteurs ont été utilisés comme facteurs d'ajustement sur le taux d'incidence.

Le mode de découverte du cancer était défini de la façon suivante :

- Dépistage individuel, réalisé par la patiente, prescrite par le médecin traitant ou le gynécologue
- Dépistage organisé par ADEMA80
- Mammographie réalisée en dehors de tout dépistage, en présence de signes cliniques.

2.4. Critères d'inclusion

La base de données est constituée par l'ensemble des cas de cancers du sein collectés dans le département de la Somme par le registre des cancers entre le 1er janvier 2004 et le 31 décembre 2014.

2.5. Critères d'exclusion

Concernant la première partie de ce travail, aucun critère d'exclusion n'a été retenu, toutes les femmes étant incluses. En revanche, pour la seconde partie, les femmes pour lesquelles nous n'avions pas de date de mammographie positive étaient exclues, ainsi que les femmes pour lesquelles la date de premier traitement n'était pas connue.

De plus, les femmes dont le type de découverte du cancer du sein était fortuit, faisait suite à une autopsie, était inconnu ou non renseigné dans la base de données du registre étaient également exclues de l'étude dans la deuxième partie de l'étude.

Enfin, le mode de découverte étant un facteur d'ajustement, et le dépistage organisé étant un des modes de découverte du cancer, les femmes non concernées par le dépistage organisé étaient exclues de l'étude.

2.6. Données recueillies

Les données ont été obtenues via la base de données du registre du cancer de la Somme. Une première extraction a été réalisée avec le logiciel Movibase® (version du 04/10/2011) le 22/03/2018. Les dates de mammographies réalisées via le dépistage organisé, et donc collectées par ADEMA80, ont été obtenues via cette structure par le numéro de tumeur (numéro à 9 chiffres commençant par l'année d'incidence du cancer). Via la base d'ADEMA, nous avons également reçu les dates de chirurgie, les classifications cTNM, pTNM et yTNM concernant les cancers dépistés par ADEMA80.

Les IRIS utilisés étaient les IRIS datant de 2017. Ceux-ci sont régulièrement remis à jour par l'INSEE (Institut National de la Statistique et des Etudes Economiques), mais très peu de modifications sont apportées : regroupements de communes…

Les scores EDI utilisés dataient de l'étude EU-SILC (en particulier de sa version française, SRCV Statistique sur Revenus et Conditions de Vie) réalisée en 2006.

Au total, 20 variables composaient la base de données des variables recueillies. D'autres ont été construites avec certaines de ces variables.

Variables recueillies	Type / Modalité
Numéro de tumeur	Numérique
Âge au diagnostic	Numérique
Année de diagnostic	Date au format aaaa
Type de découverte	**Catégorielle multiple** 1 Dépistage Individuel 2 Dépistage Organisé 3 Signes Cliniques
cT	Alphanumerique
cN	Alphanumerique
cM	Alphanumerique
pT	Alphanumerique
pN	Alphanumerique
pM	Alphanumerique
yT	Alphanumerique
yN	Alphanumerique
yM	Alphanumerique
SBR	**Catégorielle multiple** 1 Grade I 2 Grade II 3 Grade III
Iris	Numérique

Tableau 1. Liste des variables recueillies

Adresse	Alphanumérique
Date de mammographie	Date au format jj/mm/aaaa
Date de 1er traitement	Date au format jj/mm/aaaa
Taux d'incidence brut	Numérique
Score EDI	Numérique

Tableau 1. Liste des variables recueillies

D'autres variables (3) ont été créées via les précédentes :

Variables créées	Type / Modalité
TNM	Alphanumérique
Délai de prise en charge	Numérique
Stade UICC	**Catégorielle multiple** 0 I IIA IIB IIIA IIIB IIIC IV

Tableau 2. Liste des variables créées.

Le stade UICC a été calculé via le stade TNM correspondant, suivant les recommandations de l'HAS (Haute Autorité de Santé) et de l'INCa [43].

Stade UICC	TNM
0	Tis N0 M0
I	T1 N0 M0
IIA	T0 N1 M0 ; T1 N1 M0 ; T2 N0 M0
IIB	T2 N1 M0 ; T3 N0 M0
IIIA	T0 N2 M0 ; T1 N2 M0 ; T2 N2 M0 ; T3 N1 M0 ; T3 N2 M0
IIIB	T4 N0 M0 ; T4 N1 M0 ; T4 N2 M0
IIIC	Tous T N3 M0
IV	Tous T Tous N M1

Tableau 3. Correspondance entre TNM et stade UICC

Le grade SBR correspond à l'agressivité de la tumeur. Il est défini par l'examen anatomopathologique d'une biopsie de celle-ci. Trois paramètres morphologiques sont alors évalués :

- L'architecture tumorale (différenciation cellulaire)
- La forme et la taille du noyau de la cellule
- L'activité mitotique.

Chacun de ces critères est évalué et une note allant de 1 à 3 lui est attribuée.

Critère	Note 1	Note 3
Architecture	La tumeur contient beaucoup de structures bien formées	La tumeur contient pas ou peu de structures bien formées
Noyau	Les noyaux de la tumeur sont petits et uniformes	Les noyaux de la tumeur sont gros et leur taille et leur forme varient
Activité mitotique	Les cellules de la tumeur se divisent lentement (faible nombre de mitoses)	Les cellules de la tumeur se divisent rapidement (grand nombre de mitoses)

Tableau 4. Calcul du grade SBR

Le grade SBR correspond à la somme des notes obtenues, on obtient ainsi un score global classé de I à III.

- Le grade I correspond aux tumeurs les moins agressives
- Le grade III correspond aux tumeurs les plus agressives
- Le grade II est un grade intermédiaire [44].

2.7. Critère de jugement principal

Le critère de jugement principal était le délai de prise en charge du cancer du sein durant la période 2004 – 2014. Ce délai était constitué du délai entre la date de la mammographie de diagnostic du cancer et la date du premier traitement.

2.8. Critères de jugement secondaires

Les critères de jugement secondaires étaient composés par le score EDI, le stade UICC, le grade SBR, le type de découverte du cancer.

2.9. Gestion des données manquantes

La classification TNM comportait de nombreuses données manquantes. La prise en charge de ces données manquantes nécessite des explications. L'utilisation de la pTNM renseignée par le registre du cancer a été privilégiée quand celle-ci était disponible. Parmi les 1690 cas de

cancers du sein et ayant une TNM renseignée, 15 cas avaient bénéficié d'un traitement néo-adjuvant. Pour ces cas, la classification yTNM a été utilisée.

Concernant tous les autres cas, la pTNM renseignée par le registre a été utilisée lorsqu'elle était présente. Lorsqu'elle était absente, la TNM extraite de la base d'ADEMA 80, et fournie par cette dernière, était utilisée.

Nous avons donc ainsi obtenu un stade TNM pour 897 (53%) cas de cancers du sein. Afin de récupérer d'autres données manquantes, nous avons utilisé l'algorithme pour déterminer le stade ENCR (European Network of Cancer Registries) pathologique du cancer du sein, utilisé par l'ensemble des registres du cancer du réseau FRANCIM. (cf. Annexe 1) [45].
Par exemple, 202 cancers du sein étaient classés pT1N0Mx ou pT1N8Mx (N8 correspondant pour les registres à un curage réalisé mais une information sur ce curage non retrouvée). Ces femmes n'avaient pas bénéficié de traitement néo-adjuvant. Parmi ces 202 cas, 164 (81.2%) cas avaient un reliquat chirurgical classé 0 ou 1 (sans reliquat ou reliquat microscopique) et ainsi grâce à l'algorithme, les Mx de ces cas ont pu être reclassés en M0, ce qui a permis par la suite de faire correspondre au stade UICC.

2.10. Analyses statistiques

Dans un premier temps, nous avons décrit la base de données de toutes les femmes atteintes d'un cancer du sein diagnostiqué dans la Somme entre 2004 et 2014 (taux d'incidence, diagnostics par année, âge au diagnostic, type de découverte, grade SBR et score EDI).

Puis, dans un second temps, nous avons utilisé les analyses descriptives habituelles afin de décrire la base de données des femmes atteintes de cancer du sein pour lesquelles un délai de prise en charge était disponible : moyenne, écart-type, médiane et étendue pour les variables quantitatives, puis effectifs et pourcentages pour les variables qualitatives.

Visuellement, histogrammes et diagrammes en barres ont été utilisés pour représenter les différentes variables.

Concernant les cancers du sein diagnostiqués du fait de la présence de signes cliniques (et en comparaison aux cancers diagnostiqués par dépistage), l'analyse du stade au diagnostic a été faite par regroupement de stade UICC, par manque d'effectifs de certains stades pour ce mode

de découverte. Ainsi, les stades UICC 0 et I ont été regroupés, tout comme les stades IIA et IIB, et les stades IIIA, IIIB, IIIC et IV également.

Un modèle de régression linéaire multiple a été utilisé pour étudier la relation entre le délai de prise en charge et le taux d'incidence brut, ajusté sur les autres variables : le type de découverte, l'âge au diagnostic, le grade SBR, le stade UICC, le score EDI.

Ce modèle de régression linéaire multiple a été utilisé car la variable dépendante, la variable à expliquer (appelée Y) était quantitative, et qu'il y avait plusieurs variables explicatives (appelées X_i). Le modèle de régression linéaire multiple exprime l'espérance ou la moyenne de Y.

Le modèle était exprimé ainsi :

$$E(Y/X_1,X_2,X_3) = \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \varepsilon$$

Ceci peut être lu comme la moyenne de Y (variable à expliquer, soit ici le délai de prise en charge) sachant X_1, X_2, X_3 (variables explicatives).

- α : ordonnée à l'origine de la droite de régression
- β_i : coefficients de chaque variable explicative
- ε : bruit, ou résidus

Effectuer la régression linéaire de Y sur X_1, X_2, X_3 consiste à déterminer α, β_1, β_2, β_3. C'est en testant si $\beta_i = 0$ que l'on teste l'association entre la covariable X_i et Y.

Le résultat du test n'est valide que sous certaines conditions essentielles:

- Les résidus c'est-à-dire les erreurs ou « écarts » entre les valeurs observées de Y et leur estimation dérivée du modèle suivent une distribution normale de moyenne nulle
- La variance du bruit ne doit dépendre ni des valeurs de la variable à expliquer, ni des variables explicatives (hypothèse d'homoscédasticité)
- Le bruit doit être un « vrai » bruit, c'est-à-dire qu'il ne doit pas avoir de structure de corrélation évidente.

Ces hypothèses ont pu être vérifiées par des tests ou de manière plus pratique à l'aide de graphiques : distribution des résidus et graphique des résidus en fonction des covariables (la dispersion doit être homogène autour de 0) et QQplot représentants les quantiles de la distribution normale.

La sélection des variables dans le modèle s'est faite par des analyses univariées à travers des tests statistiques (ANOVA pour les variables catégorielles à plus de 2 modalités, coefficients de corrélation pour les variables quantitatives) aux conditions de validité adaptées à chaque test. Les variables présentant une p-value globale inférieure à 0.20 et dont les conditions de validité étaient acceptables lors de cette étape étaient ensuite insérées dans le modèle multivarié. L'aspect de redondance des informations apportées par deux variables fortement « liées » entre-elles fut exploré à travers des tests de corrélation (entre deux variables quantitatives) ou par le V de Cramer (entre deux variables catégorielles) allant de 0 (indépendance totale) à 1 (dépendance totale). Si une forte corrélation était présente, la sélection d'une seule variable sur les deux était requise.

Toutes les analyses ont été réalisées grâce au logiciel de traitement des données numériques *RStudio* (version 1.1.383) et aux packages suivants : prettyR, gplots, plyr et psy.

2.11. Considérations éthiques et légales

Ce travail s'inscrit dans le cadre de l'utilisation des données des registres du cancer du réseau FRANCIM. Toutes les autorisations nécessaires ont été recueillies dans le cadre de l'activité de ces registres.

3. RÉSULTATS

3.1. Analyse descriptive de tous les cas de cancers entre 2004 et 2014

Entre le 1er janvier 2004 et le 31 décembre 2014, 6064 cancers du sein ont été diagnostiqués chez les femmes dans le département de la Somme. Le taux d'incidence brut, la répartition des diagnostics par année, l'âge au diagnostic, le type de découverte, le grade SBR ainsi que le score EDI de ces patientes ont été décrits afin d'obtenir une vision d'ensemble de cette population.

Taux d'incidence :

Le taux d'incidence brut a été utilisé pour observer l'évolution du taux d'incidence sur l'ensemble des cancers du sein entre 2004 et 2014. Cette évolution peut être représentée de la façon suivante :

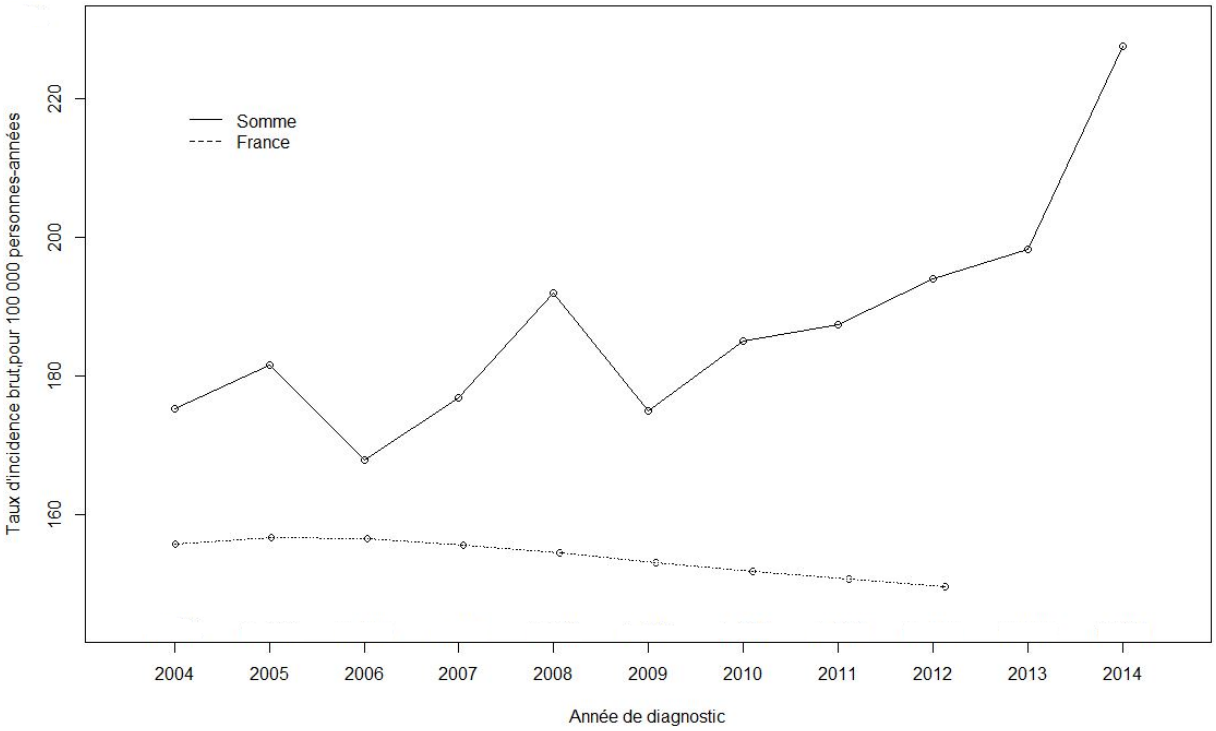


Figure 1. Evolution du taux d'incidence brut au cours du temps, pour 100 000 personnes-années, en France et dans la Somme

Le taux d'incidence est d'une part plus élevé dans le département de la Somme par rapport à celui de la France et d'autre part le taux d'incidence en France est en diminution depuis 2005 et ce jusqu'en 2012 alors que ce taux a présenté de multiples variations depuis 2004 dans la Somme avec une tendance générale à l'augmentation, particulièrement ces dernières années.

Diagnostics par année :

Dans cette même population, nous avons pu établir le nombre de diagnostics par année, qui se répartit de la manière suivante :

Années	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Effectifs	509	527	489	515	560	515	547	553	576	592	681
Pourcentages	8.4	8.7	8.1	8.5	9.2	8.5	9	9.1	9.5	9.8	11.2

Tableau 5. Effectifs et répartition des cas de cancer du sein dans la Somme entre 2004 et 2014

Concernant l'âge au diagnostic, ces cancers ont touché des femmes de 17 à 103 ans, pour une moyenne et une médiane de 61 ans. Aucune valeur n'était manquante concernant l'âge au diagnostic.

Âge au diagnostic :

Minimum	Maximum	1er Quartile	Médiane	Moyenne	3ème Quartile	Ecart-type	Valeurs manquantes
17	103.1	51.1	61	61.6	71.6	13.8	0

Tableau 6. Paramètres de la variable AGE

Graphiquement, nous avons pu représenter cette répartition sous la forme d'un histogramme.

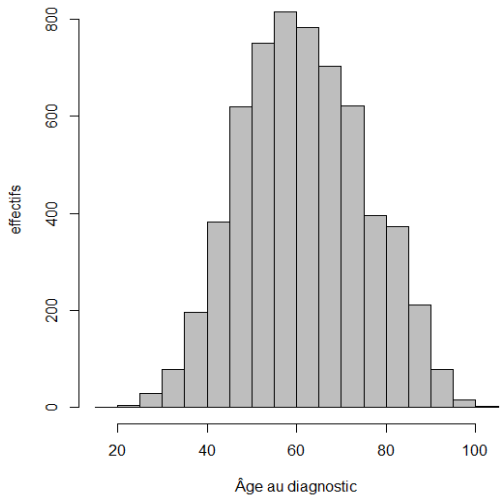


Figure 2. Répartition de l'âge au diagnostic (histogramme)

Type de découverte :

Le registre du cancer, dans son fonctionnement, inscrit le type de découverte du cancer pour chaque cas déclaré. Ces types de découverte sont classés ainsi :

- 1 : dépistage SAI (individuel) (DI)
- 2 : dépistage organisé (DO)
- 3 : signes cliniques (SC)
- 4 : découverte fortuite
- 5 : autopsie
- 8 : autre
- 9 : inconnu

Sur la période 2004-2014, le type de découverte est réparti de la façon suivante :

	Individuel	Organisé	Clinique	Fortuit	Autopsie	Autre	Inconnu
Effectifs	730	1074	1277	119	2	11	2851
Pourcentages	12	17.7	21.1	2	0.03	0.2	47

Tableau 7. Effectifs et répartition des modes de découverte des cancers du sein

Grade SBR :

Le grade SBR correspond à l'agressivité de la tumeur (cf. fin du paragraphe 2.6). Sur les 6064 cas de cancer, le grade SBR n'a pu être retrouvé pour 1338 cas (soit 22%) par le registre. Les 4726 grades connus étaient répartis de la façon suivante :

Grades SBR	I	II	III	Total
Effectifs	1038	2536	1152	4726
Pourcentages	22	53.7	24.3	100

Tableau 8. Effectifs et répartition des grades SBR

Le grade SBR a ainsi pu être comparé à l'âge au diagnostic.

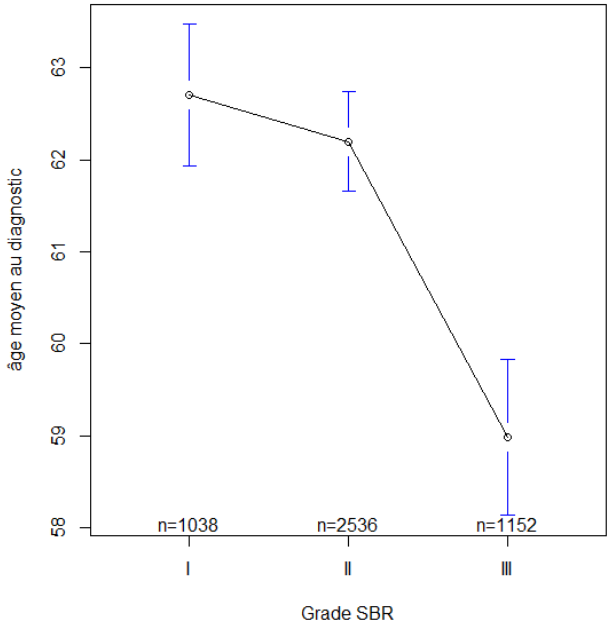


Figure 3. Âge moyen au diagnostic en fonction du grade SBR

On s'est aperçu que plus le grade SBR est élevé, plus l'âge au diagnostic est bas.

Score EDI :

Nous avons décrit les cas de cancers selon le score EDI (en fonction des quintiles de ce score). Il était à noter que la majeure partie des cancers du sein dans la Somme appartenaient aux 4ème et 5ème quantiles, donc aux scores EDI les plus élevés c'est-à-dire les niveaux socio-économiques les plus faibles.

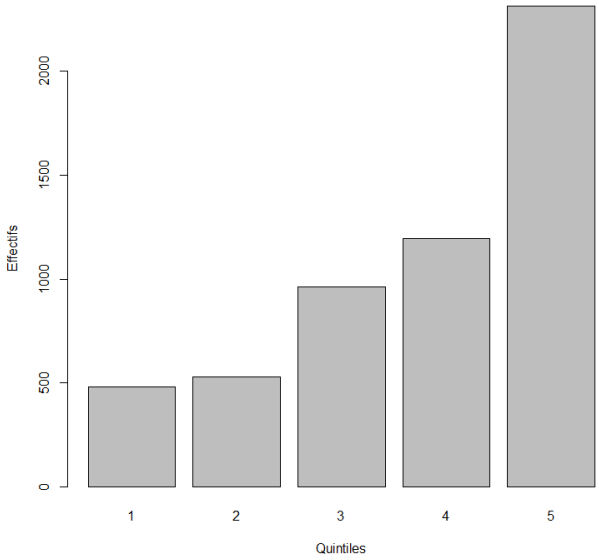


Figure 4. Histogramme de la répartition des cas de cancer du sein en fonction du score EDI (quintiles)

3.2. Analyse descriptive des cas de cancers avec délai de prise en charge

Le type de découverte étant un critère d'analyse pour cette seconde partie de notre étude, et le dépistage organisé étant un des types de découverte du cancer du sein, il fallait, pour pouvoir comparer les groupes, exclure les patientes de moins de 50 ans et celles de plus de 75 ans, puis exclure les patientes pour lesquels les modes de découverte du cancer étaient inconnus ou non renseignés. Nous avons également choisi d'exclure les patientes pour lesquelles le diagnostic de cancer faisait suite à une autopsie ou une découverte fortuite ou autre (132 patientes pour ces 3 modes de découvertes réunis), ainsi que les femmes présentant un second cancer du sein.

Enfin, puisque nous voulions étudier le délai de prise en charge, il était nécessaire d'avoir un début et une fin pour ce délai, ont donc été exclues les patientes pour lesquelles nous n'avions pas de date de mammographie, puis celles pour lesquelles nous n'avions pas de date de début de traitement.

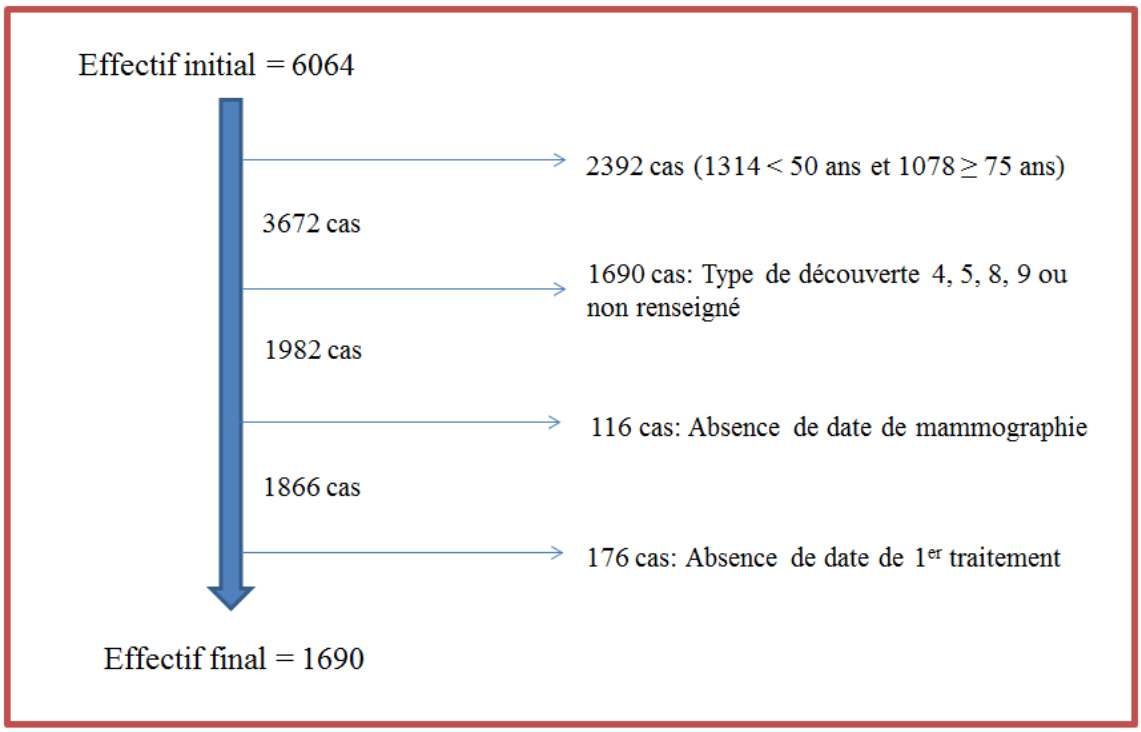


Figure 5. Flow chart de l'étude

La base de données était constituée de 23 variables et de 1690 lignes correspondant à autant de femmes (unités statistiques) atteintes d'un cancer du sein et respectant les critères d'inclusion de notre étude.

Âge au diagnostic :

L'âge au diagnostic était de 50 à 75 ans, avec une moyenne et une médiane à 61 ans. Aucune donnée n'était manquante pour la variable AGE.

Minimum	Maximum	Médiane	Moyenne	Ecart-type	Valeurs manquantes
50	74.9	61.6	61.9	7.1	0

Tableau 9. Paramètres de la variable AGE

Diagnostics par année :

Concernant les années de diagnostic, aucune valeur n'est manquante. La répartition des 1690 cas sur la période 2004-2014 cas se faisait comme suit :

Années	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Effectifs	101	108	123	140	158	139	161	162	198	182	218
Pourcentages	6	6.4	7.3	8.3	9.3	8.2	9.5	9.6	11.7	10.8	12.9

Tableau 10. Effectifs et répartition des cas de cancer du sein entre 2004 et 2014

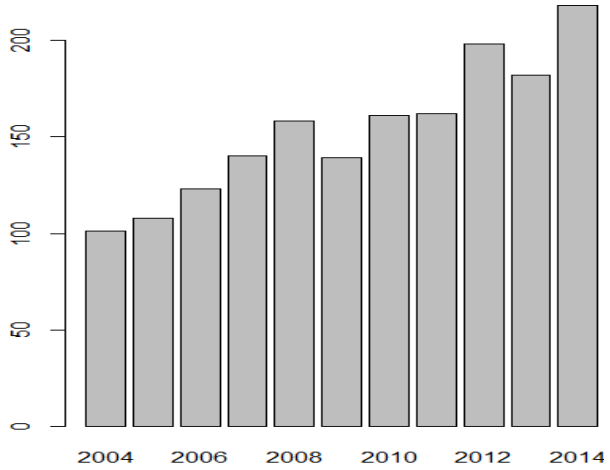


Figure 6. Répartition des effectifs par année

Type de découverte :

Les cancers du sein étaient diagnostiqués de 3 façons différentes : via le dépistage organisé (réalisé par l'association ADEMA 80), via le dépistage individuel ou du fait de la présence de signes cliniques. Il était réparti comme suit :

Années	Diagnostic	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Total
Eff.	Individuel	9	18	26	14	25	30	18	6	19	25	18	208
	ADEMA	90	79	89	121	120	91	130	155	162	126	180	1343
	Clinique	2	11	8	5	13	18	13	1	17	31	20	139
%	Individuel	8.9	16.7	21.1	10	15.8	21.6	11.2	3.7	9.6	13.7	8.3	
	ADEMA	89.1	73.1	72.4	86.4	75.9	65.5	80.7	95.7	81.8	69.2	82.6	
	Clinique	2	10.2	6.5	3.6	8.2	12.9	8.1	0.6	8.6	17.1	9.2	

Tableau 11. Effectifs et répartition des cas de cancer du sein entre 2004 et 2014 selon le mode de diagnostic

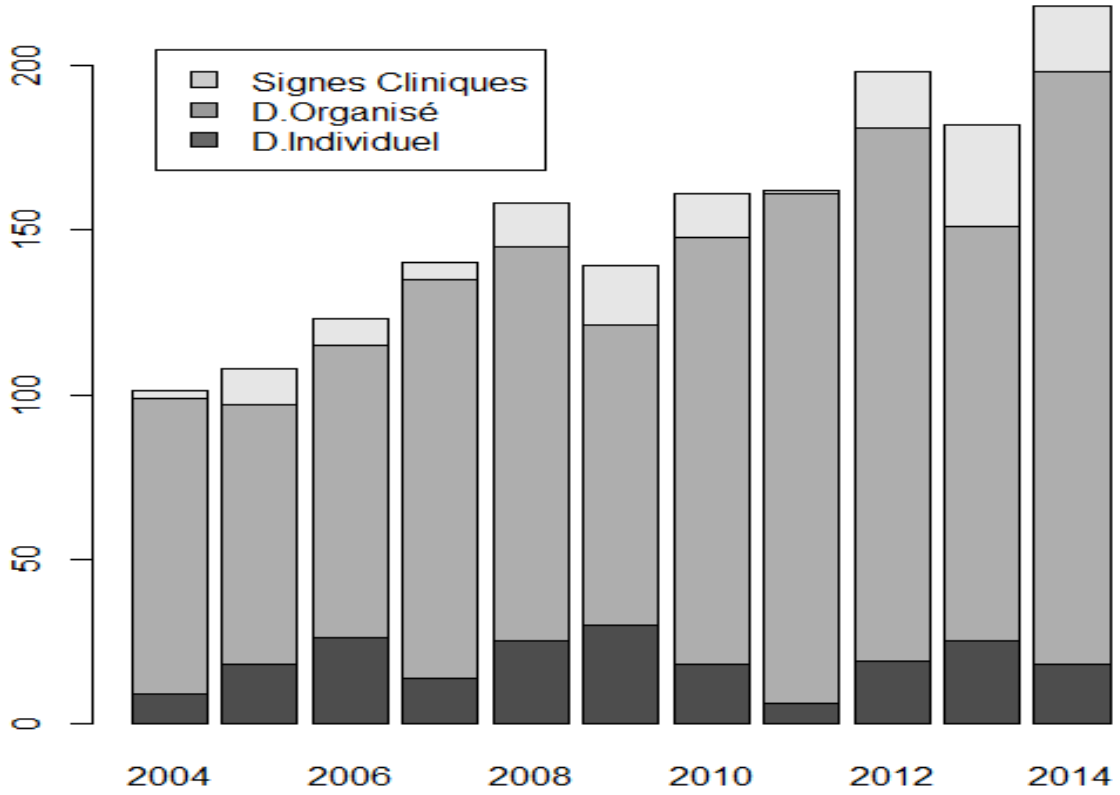


Figure 7. Effectifs des cas de cancer du sein par année et par type de découverte

Les 3 groupes à comparer étaient donc répartis de la façon suivante :

- Diagnostic par dépistage individuel : 208 cas de cancer (12.3%)
- Diagnostic par dépistage organisé : 1343 cas de cancer (79.5%)
- Diagnostic du fait de la présence de signes cliniques : 139 cas de cancer (8.2%)

Délai de prise en charge :

Le délai de prise en charge est défini par le délai entre la mammographie de dépistage et la date du 1er traitement, quel qu'il soit, en jours. La prise en charge minimum était de 7 jours, 295 au maximum, 50 jours en moyenne. 83 femmes atteintes de cancers du sein (soit 4.9%) avaient bénéficié d'une prise en charge supérieure ou égale à 100 jours.

Minimum	Maximum	1er Quartile	Médiane	Moyenne	3ème Quartile	Ecart-type	Valeurs manquantes
7	295	33	44	49.9	62	26.6	0

Tableau 12. Paramètres de la variable Délai de prise en charge

Il peut également être représenté de la manière suivante :

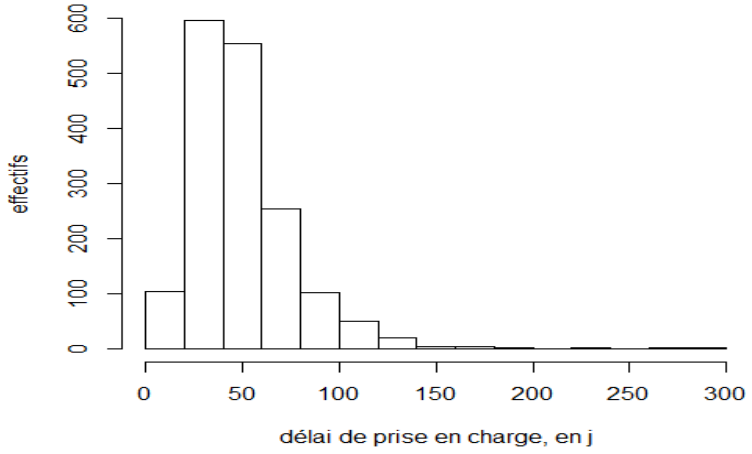


Figure 8. Répartition du délai de prise en charge, en jours

Le délai de prise en charge a également pu être étudié par année. Il est représenté par le graphique suivant :

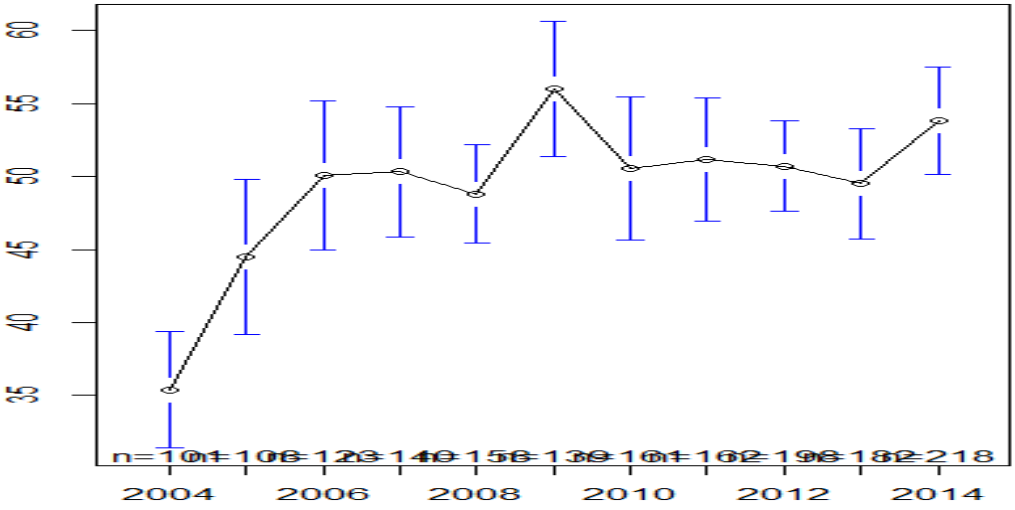


Figure 9. Evolution du délai de prise en charge (en jours), au cours du temps

Ce délai de prise en charge a considérablement augmenté entre 2004 et 2006, passant de 35 à 50 jours en moyenne, pour se stabiliser par la suite entre 2006 et 2014 (avec une augmentation du délai en 2009). En comparaison du taux d'incidence brut dans le département de la Somme sur la même période qui était plutôt stable entre 2004 et 2009, et qui a augmenté régulièrement jusqu'en 2013, puis a subi une forte augmentation. Il était difficile de mettre en évidence graphiquement une relation entre ce taux d'incidence et le délai de prise en charge. Le modèle de régression linéaire multiple présenté par la suite a permis de répondre statistiquement à cette question.

L'analyse du délai de prise en charge étant faite pour tous les types de diagnostics confondus, nous avons aussi pu décrire l'évolution de ce délai au cours du temps pour chaque mode de découverte. Ce délai de prise en charge apparaît plus important chez les femmes dont le cancer a été diagnostiqué par dépistage individuel que chez les femmes dont le cancer a été diagnostiqué par dépistage organisé ou par la présence de signes cliniques.

Ceci peut être représenté de la manière suivante :

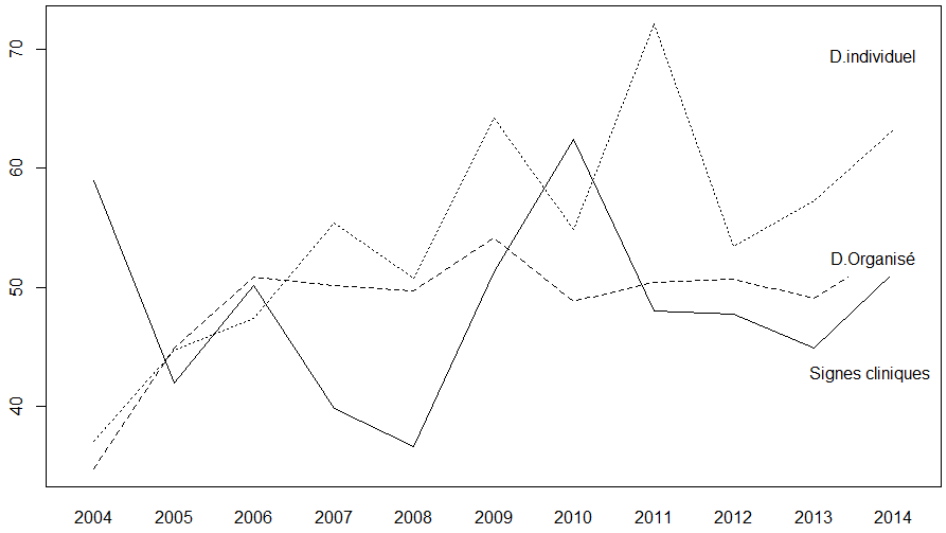


Figure 10. Evolution du délai de prise en charge (en jours) au cours du temps selon le mode de découverte

Délai de prise en charge selon le score EDI :

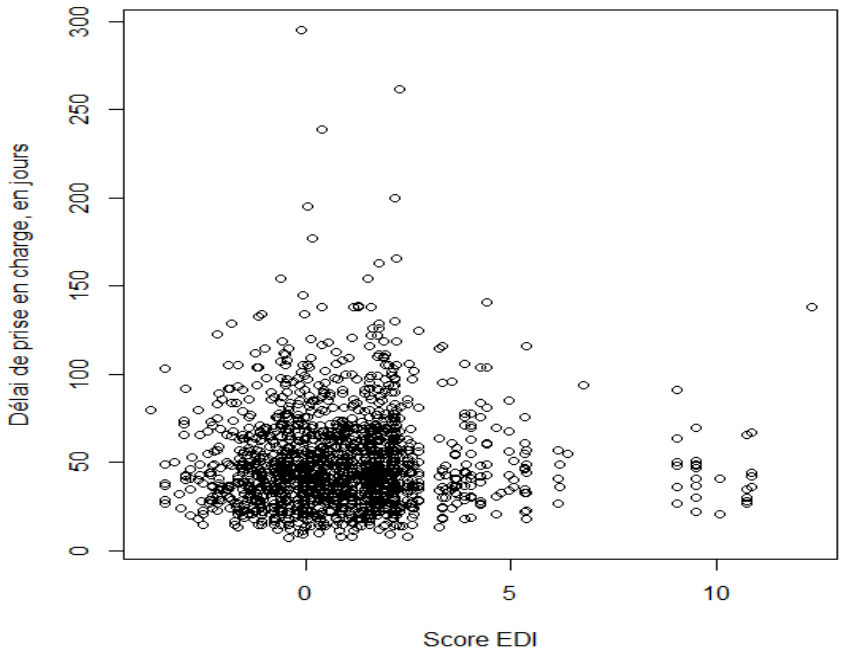


Figure 11. Délai de prise en charge en fonction du score EDI

Visuellement, le score EDI et le délai de prise en charge n'apparaissaient pas liés, le nuage de points ne présentant pas de direction globale évidente. Un test de corrélation a été réalisé par la suite pour déterminer ou non l'existence d'une corrélation. Ce test entre le délai de prise en charge (logarithme du délai, cf. ci-après) et le score EDI donnait un p = 0.29, les deux variables n'étaient donc pas corrélées. Le délai de prise en charge selon le score EDI et le type de découverte a également été décrit.

Délai de prise en charge selon le score EDI et le type de découverte :

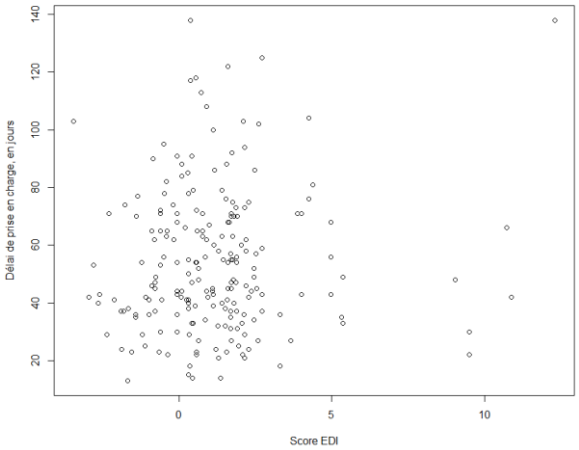


Figure 12. Délai de prise en charge en fonction du score EDI pour les cancers diagnostiqués par <u>dépistage individuel</u>

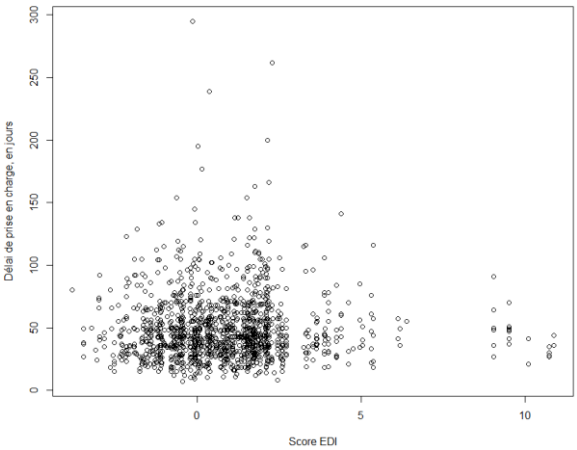


Figure 13. Délai de prise en charge en fonction du score EDI pour les cancers diagnostiqués par <u>dépistage organisé</u>

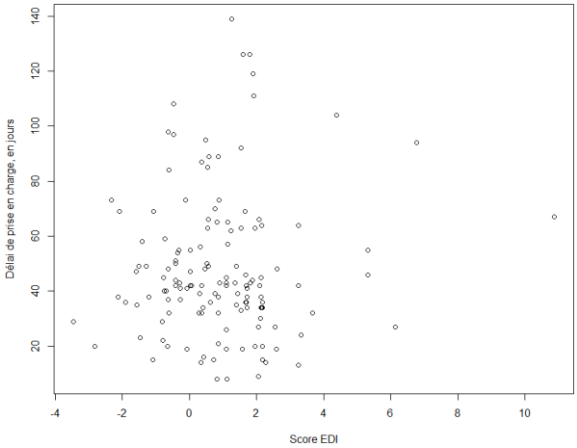


Figure 14. Délai de prise en charge en fonction du score EDI pour les cancers diagnostiqués par présence de <u>signes cliniques</u>

Aucune liaison n'apparaît entre le délai de prise en charge des cancers du sein et les scores EDI pour aucun mode de découverte. Il semblait tout de même intéressant de savoir si les scores EDI étaient différents selon les modes de découverte des cancers du sein.

Une régression linéaire simple a donc été réalisée. Le résultat était le suivant :

	Estimate	P value	IC 95%
Dépistage Organisé	-0.33	0.03	[-0.63 ; -0.03]
Signes Cliniques	-0.28	0.20	[-0.72 ; 0.15]

Tableau 13. Score EDI en fonction du mode de découverte

Les conditions de validité du modèle ont été vérifiées et respectées.

Ce modèle montrait que les femmes dont le cancer du sein était diagnostiqué par dépistage organisé présentaient un score EDI inférieur de 0.33 au score EDI des femmes dont le cancer du sein était diagnostiqué par le dépistage individuel, ceci de façon significative (p=0.03). Ce modèle montrait également une diminution du score EDI pour les femmes dont le cancer du sein était dépisté par la présence de signes cliniques, mais ceci de façon non significative. En d'autres termes, les femmes dont le cancer du sein a été diagnostiqué par dépistage organisé et par la présence de signes cliniques avaient un niveau socio-économique très légèrement supérieur aux femmes dont le cancer du sein a été diagnostiqué par dépistage individuel (résultat statistiquement significatif pour le DO mais non significatif pour les SC).

Classification SBR :

Parmi les 1690 cas de cancer, seuls 315 (18.6%) stades SBR étaient manquants. La classification SBR était répartie parmi les cas de cancer du sein comme suit :

Grades SBR	I	II	III	Total
Effectifs	369	731	275	1375
Pourcentages	26.8	53.2	20.0	100

Tableau 14. Effectifs et répartition de la classification SBR

Grades SBR / Modes de découvertes	I	II	III	Total
Dépistage individuel	54 (3.9%)	81 (5.9%)	23 (1.7%)	158 (11.5%)
Dépistage organisé	296 (21.5%)	573 (41.7%)	222 (16.1%)	1091 (79.3%)
Signes cliniques	19 (1.4%)	77 (5.6%)	30 (2.2%)	126 (9.2%)
Total	369 (26.8%)	731 (53.2%)	275 (20%)	1375 (100%)

Tableau 15. Effectifs et répartition de la classification SBR en fonction du mode de découverte

Le grade SBR a aussi pu être comparé en fonction de l'âge moyen au diagnostic. A priori, plus les patientes étaient jeunes, plus le grade SBR était élevé.

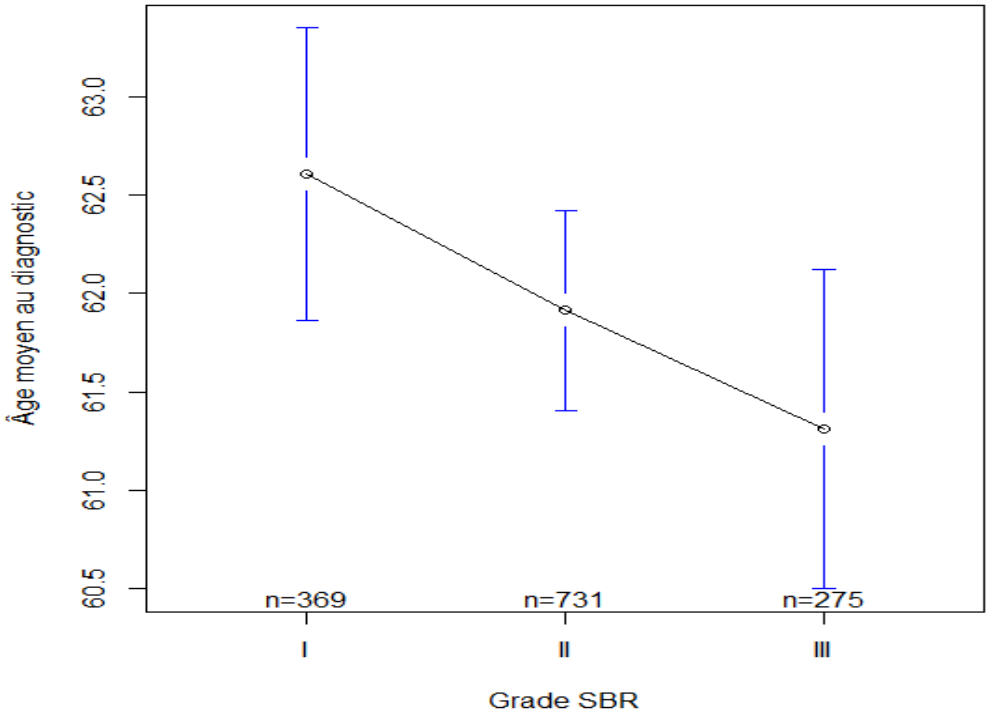


Figure 15. Grade SBR et âge

Nous avons aussi décrit le score EDI en fonction du grade SBR. Les patientes ayant un niveau socio-économique plus faible apparaissaient comme ayant un grade SBR plus élevé.

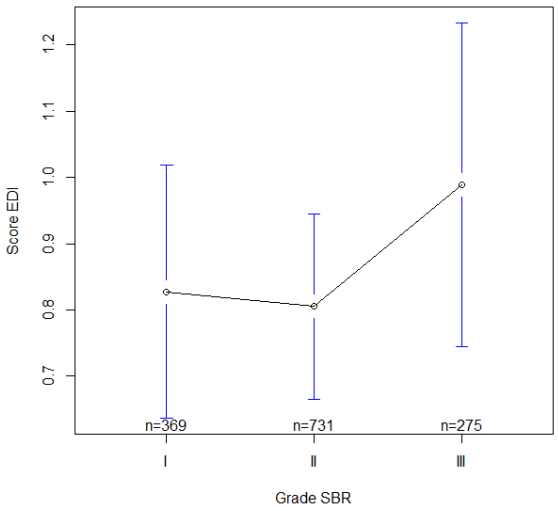


Figure 16. Score EDI en fonction du grade SBR

Stade UICC :

Concernant le stade UICC, 572 valeurs étaient manquantes (soit 33.8%), n'ayant pu être classées, un ou plusieurs items de la classification TNM étant manquants dans la base. Les cancers ayant une classification TNM, et donc un stade UICC étaient répartis de la façon suivante.

Stades UICC	0	I	IIA	IIB	IIIA	IIIB	IIIC	IV	Total
Effectifs	114	626	225	76	53	3	9	12	1117
Pourcentages	10.2	56	20.1	6.8	4.7	0.3	0.8	1.1	100

Tableau 16. Effectifs et répartition des stades UICC

Visuellement, les effectifs des stades UICC peuvent être représentés de la façon suivante.

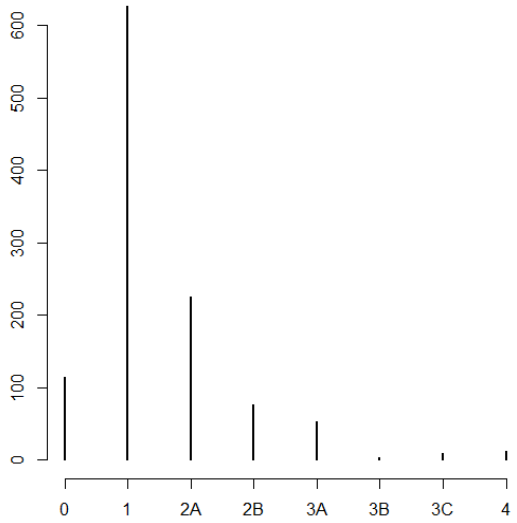


Figure 17. Effectifs des stades UICC

Score EDI :

Le score EDI s'échelonnait de -3.80 à 12.31, pour une moyenne à 0.87. Aucune valeur n'était manquante concernant ces scores.

Minimum	Maximum	1er Quartile	Médiane	Moyenne	3ème Quartile	Ecart-type	Valeurs manquantes
-3.80	12.31	-0.45	0.77	0.87	1.80	2.04	0

Tableau 17. Paramètres de la variable EDI

La répartition des scores EDI au sein de notre population peut être représentée de la façon suivante. Les quintiles les plus élevés sont les plus représentés.

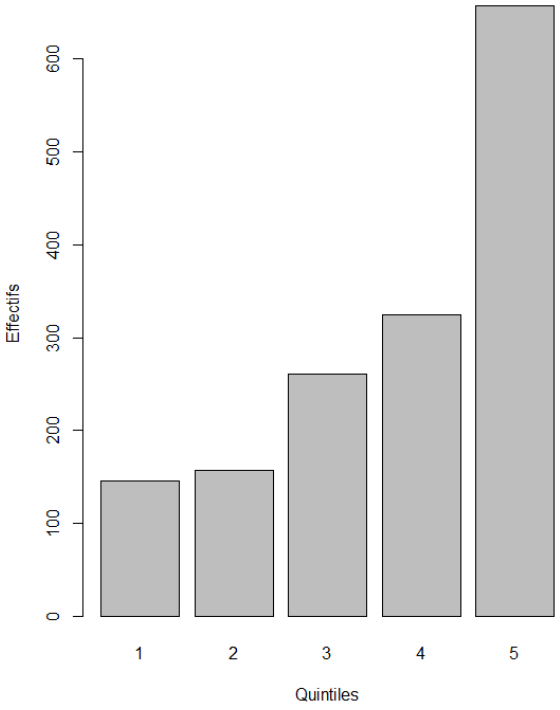


Figure 18. Répartition des scores EDI (quintiles)

Nous avons également représenté la répartition des scores EDI en fonction de l'âge au diagnostic. (cf. ci-après)

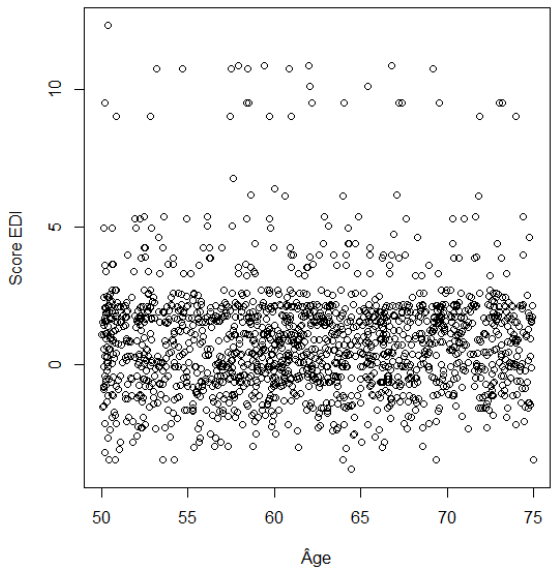


Figure 19. Répartition des scores EDI en fonction de l'âge

Ce diagramme permet de mettre en évidence plusieurs points. En effet, la grande majorité des cas était représentée par les scores EDI les plus faibles (nuage de points en bas du diagramme), et il n'apparaît pas, a priori, de différence entre l'âge au diagnostic et le score EDI (nuage de points horizontal).

Nous avons également représenté le score EDI en fonction du mode de découverte du cancer, qui peut être représenté visuellement. Les modes de découverte sont les suivants : 1 = dépistage individuel, 2 = dépistage organisé, 3 = signes cliniques.

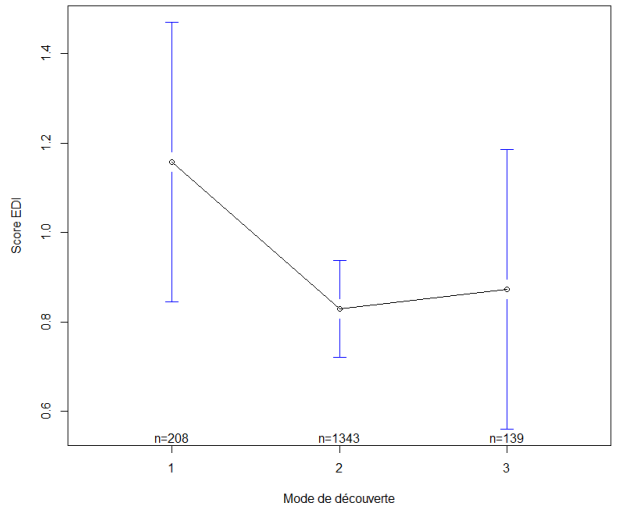


Figure 20. Répartition des scores EDI en fonction du mode de découverte

Stade UICC pour les cancers diagnostiqués du fait de la présence de signes cliniques :

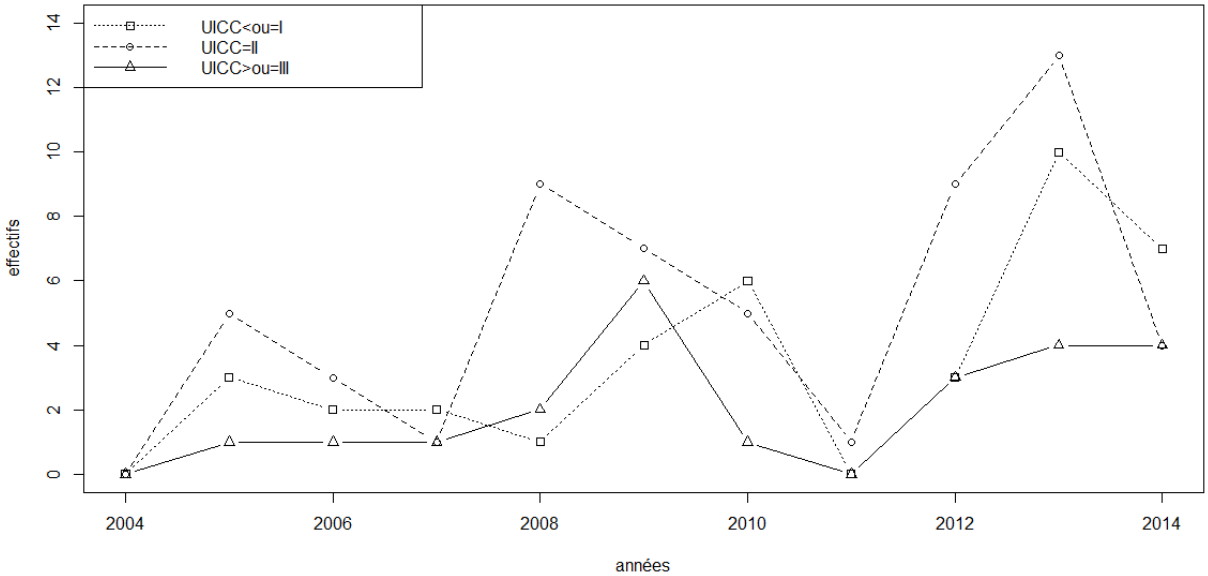


Figure 21. Effectifs des cancers diagnostiqués par signes cliniques par année, en fonction du stade UICC

Pour les cancers diagnostiqués du fait de la présence de signes cliniques, les stades UICC ont été regroupés du fait du manque d'effectifs (Stades 0 et I, stade IIA et IIB, stades IIIA IIIB IIIC IV). Les effectifs pour chaque catégorie présentaient les mêmes tendances, quand les effectifs d'une catégorie augmentaient ou diminuaient, les effectifs des autres faisaient de même. Notre hypothèse de « contamination » n'était pas vérifiée ici, il n'y avait pas d'amélioration des stades au diagnostic pour les cancers diagnostiqués par la présence de signes cliniques au cours du temps.

Nous avons pu comparer ces effectifs avec ceux des femmes dont les cancers ont été diagnostiqués par dépistage (organisé et individuel confondus). Nous avons donc pu voir que les stades les plus faibles (UICC 0 et I) deviennent plus fréquents au cours du temps, alors que les stades les plus élevés (IIIA, IIIB, IIIC et IV) sont constants.

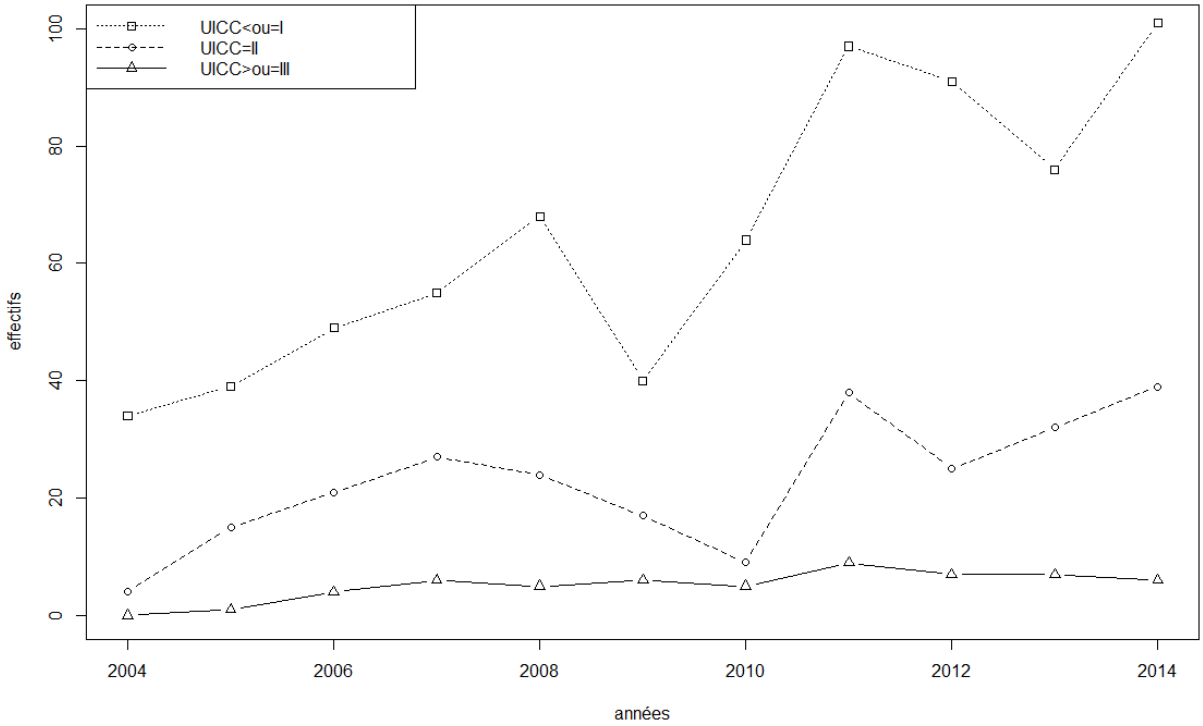


Figure 22. Effectifs des cancers diagnostiqués par DO et DI par année, en fonction du stade UICC

Stade UICC et EDI :

Le score EDI a été étudié selon le stade UICC par un modèle de régression linéaire simple. La variable à expliquer était le score EDI et la variable explicative le stade UICC (regroupé de la façon suivante : 0, I, II, III, IV). Ce modèle n'a montré aucune modification statistiquement significative du score EDI lorsque le stade UICC augmentait.

	Estimate	P value	IC 95%
UICC I	-0.29	0.16	[-0.70 ; 0.12]
UICC II	-0.13	0.55	[-0.57 ; 0.31]
UICC III	-0.12	0.69	[-0.75 ; 0.50]
UICC IV	-0.34	0.58	[-1.56 ; 0.88]

Tableau 18. Score EDI en fonction du stade UICC

3.3. Délai de prise en charge en fonction du taux d'incidence ajusté

Délai de prise en charge et EDI :

Les analyses suivantes ont été réalisées dans le but de mettre en évidence une différence du délai de prise en charge du cancer du sein en fonction du taux d'incidence, en ajustant sur les autres variables présentes dans la base de données afin d'éviter d'éventuels facteurs de confusion.

Nous avons commencé par analyser la normalité de la variable à expliquer Y, le délai de prise en charge.

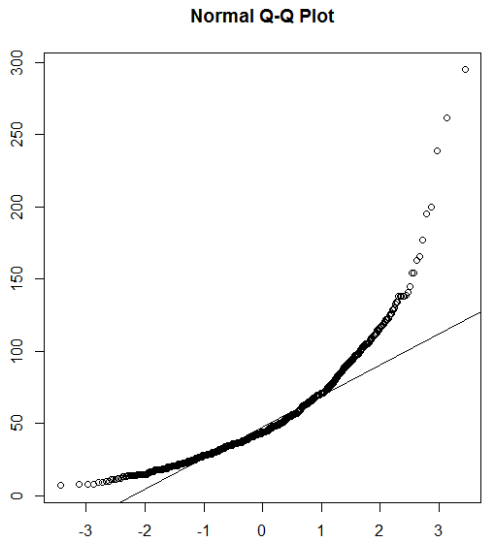


Figure 23. QQplot de la normalité de la variable délai de prise en charge

Les points déviaient fortement de la droite de normalité. La distribution des valeurs de la variable du délai de prise en charge ne suivait pas une loi normale. L'utilisation de la transformation logarithmique permettait une normalisation de la distribution de ces valeurs.

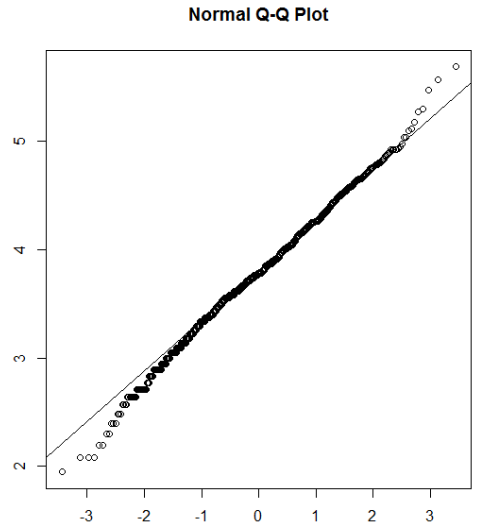


Figure 24. QQplot de la normalité de la variable délai de prise en charge après transformation logarithmique

La variable a donc été modifiée en « logarithme du délai », ce qui permettait d'avoir une distribution normale de la variable à expliquer Y.

Délai et Âge :

Le test de corrélation entre le log du délai et la variable âge retrouvait une p value à 0.1327, soit un $p \leq 20\%$, ce qui nous permettait de retenir la variable âge dans le modèle.

Délai et EDI :

Le test de corrélation entre le log du délai et la variable du score EDI retrouvait un $p = 0.29$, soit un $p > 20\%$, ce qui ne nous permettait pas de retenir la variable de l'EDI dans le modèle.

Délai et Incidence :

Le test de corrélation entre le log du délai et la variable du taux d'incidence brut retrouvait un $p = 0.0003911$, soit un $p \leq 20\%$, ce qui nous permettait de retenir la variable du taux d'incidence dans le modèle.

Délai et type de découverte :

La variable du type de découverte étant catégorielle à plus de 2 modalités, un test de Student n'était pas possible à utiliser, il a donc fallu faire une régression linéaire simple avec une seule variable explicative, le type de découverte.

Le résultat de ce test montrait une p value = 0.001726 soit $\leq 20\%$, la variable type de découverte était donc retenue pour le modèle final.

Délai et SBR :

De la même manière que pour le type de découverte, une régression linéaire a été utilisée pour le grade SBR.

Le résultat de ce test retrouvait une p value égale à 1.699×10^{-8}. Le p étant $\leq 20\%$, la variable du grade SBR était retenue pour le modèle final.

Délai et UICC :

Une régression linéaire simple a aussi été utilisée pour la variable du stade UICC. Le p valait alors 1.166×10^{-7}, donc $\leq 20\%$, ce qui nous a fait retenir la variable pour notre modèle final.

Les variables explicatives et leurs p values en fonction du délai de prise en charge pouvaient être synthétisées dans le tableau suivant.

Variables	p value
Score EDI	0.29 *
Âge au diagnostic	0.1327 **
Type de découverte	0.001726 **
Taux d'incidence brut	0.0003911 **
Stade UICC	1.166×10^{-7} **
SBR	1.699×10^{-8} **

Tableau 19. Variables et leurs p values

* : variable non retenue dans le modèle, car p value >20% en analyse univariée

** : variable retenue dans le modèle, car p value ≤ 20% en analyse univariée

Nous allons vérifier la redondance afin de savoir si certaines variables sont très corrélées entre elles, et, le cas échéant, choisir la variable la plus pertinente parmi elles dans le modèle.

Aucune corrélation n'a été retrouvée entre les variables. Par exemple, un V de Cramer a été réalisé entre les variables du grade SBR et du stade UICC. Ce test retrouvait un V de Cramer à 0.16, ce qui signifie que la corrélation était faible entre ces deux variables, donc elles pouvaient être conjointement introduites dans le modèle de régression.

Le modèle de régression linéaire multiple donnait alors le résultat suivant.

	Estimate	**P value**	**IC 95%**	**Variation du délai de prise en charge**
Taux d'incidence brut	0.003	0.002 **	[0.001 ; 0.005]	0.27 %
Age au diagnostic	0.007	0.001 **	[0.003 ; 0.011]	0.71 %
Type de découverte = 2 (DO)	-0.075	0.097 #	[-0.164 ; 0.014]	-7.23 %
Type de découverte = 3 (signes cliniques)	-0.110	0.099 #	[-0.241; 0.021]	-10.44 %
UICC I	-0.303	0.148	[-0.715 ; 0.108]	-26.17 %
UICC IIA	-0.345	0.101	[-0.758 ; 0.068]	-29.19 %
UICC IIB	-0.290	0.176	[-0.712 ; 0.131]	-25.22 %
UICC IIIA	-0.359	0.099 #	[-0.788 ; 0.068]	-30.23 %
UICC IIIB	-0.712	0.036 *	[-1.380 ; -0.045]	-50.96 %
UICC IIIC	-0.168	0.516	[-0.677 ; 0.341]	-15.49 %
UICC IV	-0.132	0.602	[-0.632 ; 0.366]	-12.44 %
SBR II	-0.122	0.001 **	[-0.195 ; -0.049]	-11.51 %
SBR III	-0.219	2.2e-06 ***	[-0.309 ; -0.129]	-19.68 %

Tableau 20. Résultats du modèle de régression linéaire multiple

*** : < 0.001 ** : < 0.01 * : < 0.05 # : 0.05 < p < 0.1

Il fallait ensuite vérifier les conditions de validité du modèle, soit une distribution normale des résidus, une variance du bruit indépendante (ne dépendant ni de la variable à expliquer ni des variables explicatives) et un bruit sans structure de corrélation évidente. Ceci peut être vérifié par les graphiques suivants :

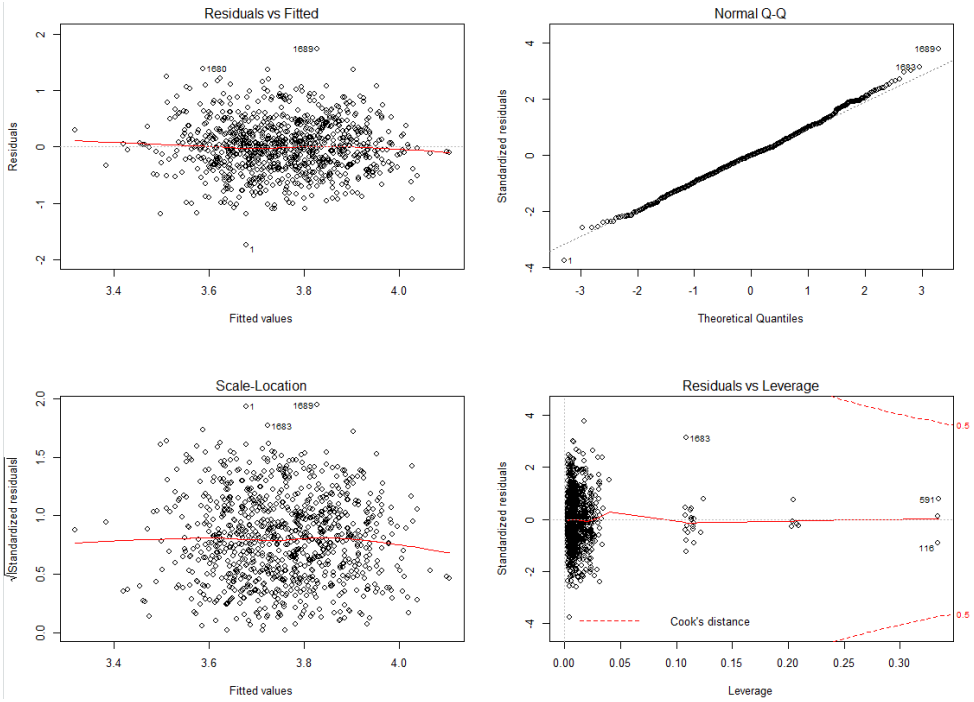


Figure 25. Diagrammes des conditions de validité du modèle

En haut à gauche : le nuage n'a aucune forme et les résidus semblent distribués aléatoirement. La courbe de lissage (ligne continue) suit dans l'ensemble la ligne pointillée horizontale, c'est-à-dire que les prédictions basées sur le modèle linéaire sont globalement justes et collent aux données.

En bas à gauche : l'homoscédasticité (condition d'égalité des variances) est représentée. La ligne doit être la plus horizontale possible sans cassure, ce qui était acceptable dans notre cas.

En haut à droite : les résidus avaient une distribution normale, les points s'écartaient peu de la diagonale (droite de Henry) qui traduit la relation linéaire positive entre les quantiles observés et les quantiles théoriques de la loi gaussienne.

En bas à droite : ce graphique permet de diagnostiquer quelles observations se distinguent trop des autres et pourraient être écartés de l'étude. Les observations 1683, 591 et 116

auraient pu mériter une attention particulière et auraient pu être écartées de l'analyse mais restaient cependant à l'intérieur de l'intervalle de confiance de Cook.

Ces graphiques nous ont donc permis d'affirmer que les conditions de validité du modèle étaient respectées.

L'analyse du modèle de régression linéaire multiple montrait finalement plusieurs résultats, significatifs ou non :

- Lorsque le taux d'incidence brut augmentait de 1/100 000, le délai de prise en charge augmentait de 0.27%, toute autre variable incluse dans le modèle égale par ailleurs, et ceci de façon statistiquement significative
- Lorsque le type de découverte était par DO, le délai de prise en charge était diminué de 7.23% par rapport au délai de prise en charge d'un cancer diagnostiqué par dépistage individuel, toute autre variable égale par ailleurs, mais ceci de façon statistiquement non significative
- Lorsque le type de découverte était par la présence de SC, le délai de prise en charge était diminué de 10.44 % par rapport au délai de prise en charge d'un cancer diagnostiqué par dépistage individuel, toute autre variable égale par ailleurs, mais ceci de façon statistiquement non significative
- Plus le grade SBR augmentait, plus le délai de prise en charge du cancer diminuait (de 11.5% pour le grade II et de près de 20% pour le grade III), toute autre variable égale par ailleurs, ceci de façon statistiquement significative
- L'augmentation du stade UICC ne montrait pas de réelle tendance à la diminution du délai de prise en charge, ces délais étant plus ou moins diminués selon les stades UICC sans continuité d'un stade à l'autre. Ces résultats n'étaient pas statistiquement significatifs.

4. DISCUSSION

4.1. Du recueil de données et des données manquantes

Le recueil de données a été réalisé grâce au Registre du Cancer de la Somme, ce qui a permis d'obtenir une exhaustivité des cas de cancer du sein dans le département ainsi que les variables nécessaires à notre étude.

Le travail du registre du cancer étant essentiellement basé sur la production des taux d'incidence ainsi que sur les données de survie, les variables concernant ce travail n'étaient donc pas toujours retrouvées dans la base du registre, notamment le stade TNM, le type de découverte du cancer, le grade SBR, la date de la mammographie correspondant au dépistage et la date du premier traitement. De nombreux retours sur dossiers papiers et/ou informatisés ont été réalisés au cours de ce travail, ce qui a permis de récupérer de nombreuses données, mais beaucoup restaient manquantes, dû au fait que ces mêmes informations n'étaient pas non plus présentes dans les documents médicaux disponibles (documents du registre et logiciel DxCare), impliquant alors de nombreuses exclusions de cas de cancers et donc une diminution de la puissance de l'étude. D'autres données ont été récupérées grâce aux différents croisements entre les bases du registre et les bases fournies par ADEMA80, ce qui a parfois été problématique, lorsque des données pour les mêmes patientes étaient présentes dans les deux bases mais différentes.

Certaines de ces données n'auraient de toute façon pas pu être retrouvées, n'étant pas toujours présentes dans les dossiers médicaux, comme par exemple la provenance de la mammographie de diagnostic, résultante d'un dépistage individuel, de signes cliniques ou du dépistage organisé ou de la classification TNM, pas toujours indiquée.

La gestion des données manquantes a concerné la classification TNM pour laquelle beaucoup de données étaient manquantes et donc pour laquelle une imputation était nécessaire. Celle-ci a été réalisée avec un algorithme défini par le réseau FRANCIM et donc utilisé nationalement par tous les registres. Cette imputation a pu entraîner des biais de classement potentiels, en classant les cancers dans des stades UICC qui n'étaient pas forcément les leurs au moment du diagnostic.

De manière plus générale, la non exhaustivité des variables de la base de données était retrouvée dans d'autres études, où, par exemple, entre 16 et 48% des données concernant la

classification TNM, 11.5% des modes de découverte du cancer, 25% des dates de mammographie étaient manquantes [9]. Ceci ne représente donc pas une difficulté du registre du cancer de la Somme, mais une difficulté générale concernant la capacité à récupérer les informations a posteriori.

4.2. Du tri à plat

Le délai de prise en charge retrouvé de 50 jours en moyenne était proche du délai de 41 jours mesuré par l'INCa en 2011 [9] mais assez éloigné des 34 jours retrouvés par Molinié et al [11] en 2013 ainsi que du mois préconisé par l'ANAES [10], qui semble être tout de même un objectif difficilement atteignable selon l'ensemble des études. Malgré ceci, le délai de prise en charge reste bien inférieur à celui d'un pays émergent comme le Maroc [12].

Enfin, le délai de prise en charge aurait pu être décomposé en délai au diagnostic, délai à la RCP, délai à la chirurgie, délai à la radiothérapie (avec ou sans hormonothérapie), pour être ensuite analysé et comparé aux délais dans d'autres études [9, 11, 12]. Celui-ci n'a pas été réalisé du fait du grand nombre de données manquantes déjà important et qui auraient pris une proportion encore plus importante après la segmentation de ce délai.

L'utilisation du score EDI peut être discutée. N'existant pas de gold standard concernant le niveau socio-économique, des études ont été menées en utilisant la catégorie socio-professionnelle [9] et ne montraient aucun résultat statistiquement significatif. L'EDI s'appuie sur des indicateurs individuels comme le statut de chômeur, l'accès à une voiture, le niveau d'étude…mais est calculé pour un IRIS, et donc est le même pour une ville ou un quartier, et ne reflète donc pas exactement le niveau socio-économique de chaque individu. La difficulté d'accès aux soins devrait pouvoir être mise en relation avec une pauvreté culturelle, un manque d'information et d'accès à l'information.

Le stade UICC, variable catégorielle ordonnée, a parfois été regroupé pour certaines analyses, <I, II ou >III pour cause de manque d'effectifs dans les unes ou les autres modalités selon l'analyse. Ce regroupement a respecté l'ordre des modalités de la variable mais ne suivait pas toujours une logique clinique, le stade IV étant le seul métastatique par exemple, était en commun avec le stade III, nous pouvons donc nous interroger sur la pertinence de ce regroupement.

4.3. Des résultats du modèle de régression linéaire multiple

Tout d'abord, la présence de données agrégées peut engendrer des erreurs dans l'analyse des résultats des variables concernées. En effet, deux variables : le score EDI (qui n'apparaissait pas dans le modèle de régression linéaire multiple) ainsi que le taux d'incidence brut sont deux variables de données agrégées. Le taux d'incidence brut était calculé et indiqué par année, et l'étude étant réalisée sur 11 années complètes, la variable du taux d'incidence brut ne pouvait prendre que 11 valeurs possibles, et ces valeurs ne dépendent pas des individus, mais seulement de leur année de diagnostic. Ceci a pu influencer, par la présence des facteurs de confusion, comme dans toute étude épidémiologique, les résultats de l'analyse du modèle.

D'autres études pourraient confirmer ou infirmer l'impact du taux d'incidence sur le délai de prise en charge, et d'autres modèles pourraient également être testés pour évaluer cet impact.

Le grade SBR semble impacter le délai de prise en charge. En effet, plus le grade SBR augmente, plus le délai de prise en charge diminue, ceci de façon significative. Nous pouvons donc penser que du fait de l'agressivité du cancer, la prise en charge est organisée de manière plus précoce.

Mais ceci n'est pas appuyé par le résultat de l'analyse du stade UICC sur le délai de prise en charge, car quand le stade UICC augmente, et donc quand le cancer est diagnostiqué à un stade plus avancé, nous n'avons pas montré de relation linéaire à la diminution du délai de prise en charge, et ceci de manière non significative. Seul le stade IIIC montrait une diminution significative de 50% du délai de prise en charge par rapport au stade 0, avec des effectifs très faibles pour ce stade. Une étude comprenant des effectifs plus importants pour chaque stade nous permettrait de définir une éventuelle relation entre le stade UICC et le délai de prise en charge.

Le type de découverte du cancer du sein diagnostiqué par dépistage organisé, ainsi que par la présence de signes cliniques semblent diminuer le délai de prise en charge par rapport aux cancers diagnostiqués par dépistage individuel, mais de façon non significative.

Enfin, les femmes ayant eu un cancer du sein diagnostiqué en raison de la présence de signes cliniques ne semblent pas consulter plus rapidement en 2014 que dix ans auparavant, les stades UICC au diagnostic étant toujours dans les mêmes proportions au fil des années. Ici encore, une étude avec plus de puissance serait nécessaire, les effectifs pour chaque stade UICC et chaque année étant trop faibles pour pouvoir conclure.

5. CONCLUSION

Notre étude est la seule étude française à avoir étudié le délai de prise en charge du cancer du sein en fonction du taux d'incidence. Lorsque le taux d'incidence augmente dans la population, le délai de prise en charge augmente de manière significative. Il faudra analyser l'effet du mode de découverte sur le délai de prise en charge, l'effet du grade SBR sur ce délai, et l'impact sur le stade au diagnostic du cancer, sur une population plus importante et contenant moins de données manquantes afin d'obtenir une meilleure puissance de l'étude.

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. GLOBOCAN 2012 v1.1, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2014. Available from: http://globocan.iarc.fr
2. Dernières statistiques mondiales sur le cancer. 12 décembre 2013. https://www.iarc.fr/fr/media-centre/pr/2013/pdfs/pr223_F.pdf
3. INCa, Les cancers en France, Les données, 2016. http://www.e-cancer.fr/ressources/cancers_en_france.
4. Binder-Foucard F, Belot A, Delafosse P, Remontet L, Woronoff AS, Bossard N. Estimation nationale de l'incidence et de la mortalité par cancer en France entre 1980 et 2012. Partie1 – Tumeurs solides. Saint-Maurice(Fra) : Institut de veille sanitaire, 2013. 122p.
5. Lapôtre-Ledoux B, Ganry O, Trecherel E, Incidence des cancers de la Somme pour la période 2010-2014. http://www.chu-amiens.fr/wp-content/uploads/2017/06/CHU-Amiens-Picardie-Registre-Cancer-Somme-Incidence-2010-2014.pdf
6. Incidence et mortalité 2006-2009, Evolution de 1982 à 2009, Registre du cancer de la Somme, Avril 2012. P52 http://www.chu-amiens.fr/wp-content/uploads/2016/12/CHU-Amiens-Picardie_Registre-Cancer-Somme_2012_-rapport_cancer_N2.pdf
7. Smith EC, Ziogas A, Anton-Culver H. Delay in surgical treatment and survival after breast cancer diagnosis in young woman by race/ethnicity. JAMA Surg 2013 ;148(6) :516-23.
8. Shin DW, Cho J, Kim SY, Guallar E, Hwang SS, Cho B, et al. Delay to curative surgery greater than 12 weeks is associated with increased mortality in patients with colorectal and breast cancer but not lung or thyroid cancer. Ann Surg Oncol 2013 ;20(8) :2468-76
9. INCa. Etude sur les délais de prise en charge des cancers du sein et du poumon dans plusieurs régions de France en 2011. www.ecancer.fr/content/download/63265/569429/file/ETUDELSEINPOU12.pdf
10. ANAES. Chirurgie des lésions mammaires : prise en charge de première intention.2002. https://www.has-sante.fr/portail/jcms/c_447405/fr/chirurgie-des-lesions-mammaires-prise-en-charge-de-premiere-intention

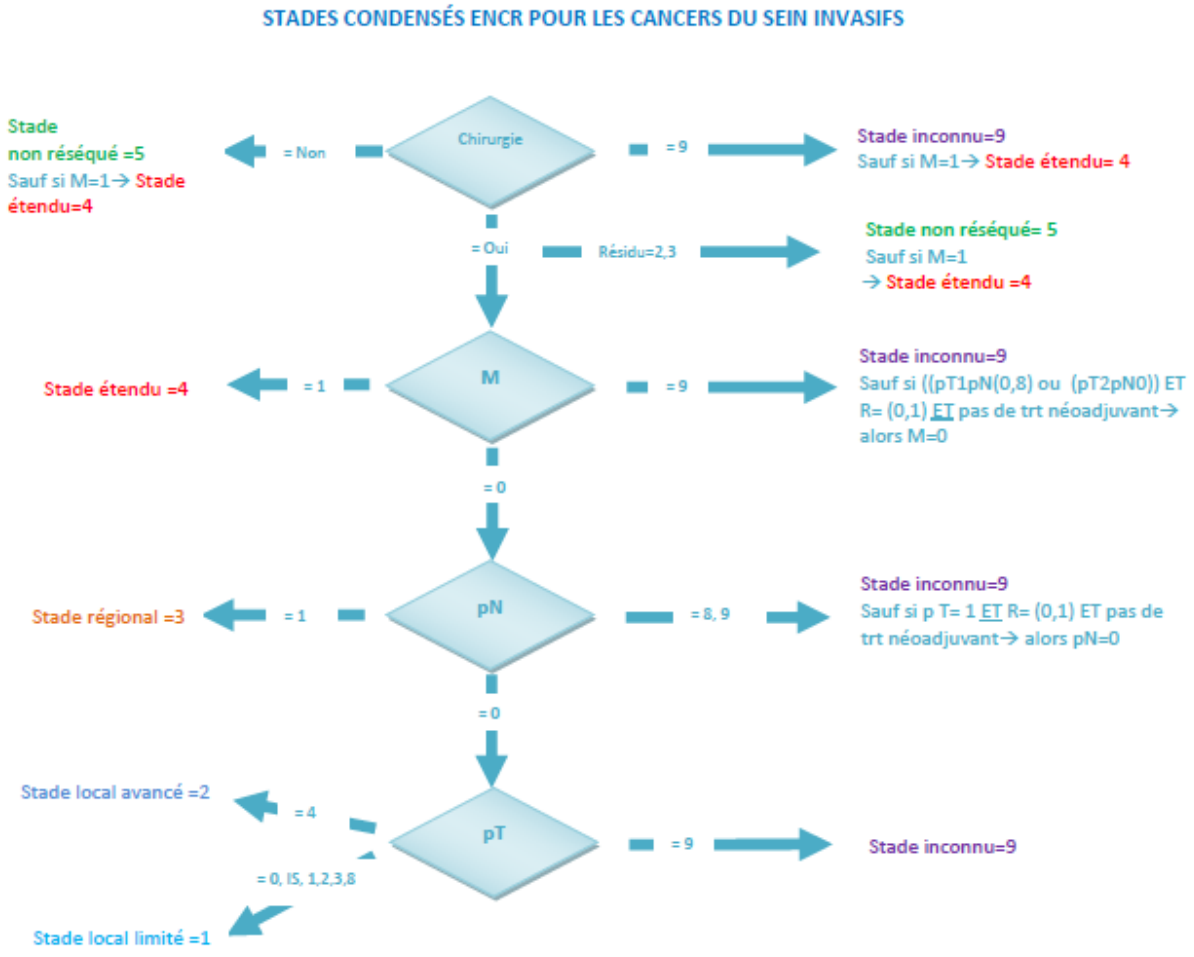
11. Molinié F et al.Waiting time disparities in breast cancer diagnosis and treatment : A population-based study in France. 2013
12. Mimouni M,Chaouki W, Errihani H, Benjaafar N. Analyse des délais de traitement du cancer du sein : expérience d'un centre de référence tertiaire au Maroc. Bull Cancer. 2018
13. Hanna R Brewer, Michael E.Jones, Minouk J.Schoemaker, Alan Ashworth, and Anthony J. Swerdlow. Family history and risk of breast cancer : an analisys accounting for family structure. 2017
14. Hamajima N, Hirose K, Tajima K, Rohan T, Calle EE, Heath CW, et al. Alcohol, tobacco and breast cancer-collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. British journal of cancer. 2002
15. Peto J, Easton D, Matthews F, Ford D, Swerdlow A. Cancer mortality in relatives of women with breast cancer : the OPCS study. Int J Cancer. 1996
16. Scheen AJ, Beck E, De Flines J, Rorive M. Obesity, insulin resistance and type 2 diabetes : risk factors for breast cancer. 2011
17. Renehan AG, Tyson M, Egger M, et al. Body-mass index and incidence of cancer : a systematic review and metaanalysis of prospective observational studies. Lancet, 2008
18. Rose DP, Vona-Davis L. Interaction between menopausal status and obesity in affecting breast cancer risk. Maturitas, 2010
19. Maruthur NM, Bolen S, Brancati FL, Clark JM. Obesity and mammography : a systematic review and metaanalysis. J Gen Intern Med, 2009
20. Smith U, Gale EAM. Does diabetes therapy influence the risk of cancer ? Diabetologia, 2009
21. Giovannucci E, Harlan DM, Archer MC, et al. Diabetes and cancer : A consensus report. Diabetes Care, 2010
22. Simon D, Balkau B. Diabetes mellitus, hyperglycaemia, 2010
23. Amarasinghe A, D'Souza G, Brown C, Oh H, Borisova T. The influence of Socioeconomic and Environmental Determinants on Health and Obesity : A West Virginia Case Study. Int J Environ Res Public Health. 2009 ; 6(8) : 2271-87
24. Merletti F, Galassi C, Spadea T. The socioeconomic determinants of cancer. Environ Health, 2011 ; 10 (Suppl 1) : S7
25. Braveman P. The Social Determinants of Health : It's Time to Consider the Causes of the Causes. Public Health Rep. 2014 ; 129(Suppl 2) : 19-31.

26. Strand, B.H, et al., 2007. The reversed social gradient : Higher breast cancer mortality in the higher educated comared to lower educated. A comparison of 11 European populations during the 1990s. European Journal of Cancer 43, 1200-1207.
27. F.Berger et al., 2011. Impact du statut socio-économique sur la gravité du diagnostic initial de cancer du sein. Revue d'Épidémiologie et de Santé Publique 3835,1.
28. Adema 80-Association pour le dépistage des maladies dans la Somme http://www.adema80.fr/
29. INCa. http://www.e-cancer.fr/Comprendre-prevenir-depister/Se-faire-depister/Depistage-du-cancer-du-sein/Les-niveaux-de-risque#toc-niveau-de-risque-lev-
30. The benefits and harms of breast cancer screening : an independant review, 2012. The Lancet 380, 1778-1786.
31. Uhry Z. et al., 2011. Modelling the effect of breast cancer screening on related mortality using French data. Cancer Epidemiology 35, 235-242
32. Broeders M. et al., 2012. The impact of Mammography Screening on Breast Cancer Mortality in Europe : A Review of Observational Studies. J Med Screen 19, 14-25
33. Hill C. 2014. Dépistage du cancer du sein. La Presse Médicale 43, 501-509
34. Haute Autorité de Santé – La participation au dépistage du cancer du sein des femmes de 50 à 74 ans en France https://www.has-sante.fr/portail/jcms/c_1194998/fr/la-participation-au-depistage-du-cancer-du-sein-des-femmes-de-50-a-74-ans-en-france
35. OR2S – Cancer du sein, données sanitaires et sociales, rapport 2015.
36. Duport N, Ancelle-Park R, Boussac-Zarebska M, Uhry Z, Bloch J. Facteurs d'adhésion au dépistage organisé du cancer du sein : étude FADO-sein.2007
37. Puliti D. et al., 2017. Advanced breast cancer rates in the epoch of service screening : The 400,000 women cohort study from Italy. European Journal of Cancer. 75, 109-116
38. Autier P. et al., 2009. Advanced Breast Cancer and Breast Cancer Mortality in Randomized Controlled Trials on Mammography Screening. JCO 27, 5919-5923.
39. http://www.chu-amiens.fr/chercheurs/le-registre-du-cancer-de-la-somme/
40. INCa. http://lesdonnees.e-cancer.fr/Informations/Sources/SOURCE-Reseau-FRANCIM
41. IRIS, définition, méthode et qualité. https://www.insee.fr/fr/metadonnees/definition/c1523
42. Pornet, C. et al., 2012. Construction of an adaptable European transnational ecological deprivation index: the French version. J Epidemiol Community Health.

https://doi.org/10.1136/jech-2011-200311
43. https://www.has-sante.fr/portail/upload/docs/application/pdf/2010-02/ald_30_gm_ksein_vd.pdf
44. http://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancer-du-sein/Les-grades-du-cancer
45. http://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Stade-au-diagnostic-des-cancers-du-sein-du-colon-et-du-rectum-Etude-realisee-a-partir-des-registres-des-cancers-du-reseau-FRANCIM

ANNEXES

Algorithme pour déterminer le stade ENCR pathologique du cancer du sein



L'algorithme est construit à partir des valeurs des variables décrites dans le document 'Variables_Stade2013'

Le stade ENCR est composé à partir des éléments du pTNM et ne concerne que les cas invasifs (coder ENCR=0 pour les CIS).

Si traitement néoadjuvant : le stade est constitué avec les mêmes variables et les cas sont repérés par la variable NEOADJ 1= Oui

Annexe 1 : Algorithme pour déterminer le stade ENCR pathologique du cancer du sein.