



HAL
open science

Les évaluations économiques en population dans les prises en charge des cancers : évaluation des coûts des lymphomes et de la leucémie myéloïde chronique

Michaël Mounié

► **To cite this version:**

Michaël Mounié. Les évaluations économiques en population dans les prises en charge des cancers : évaluation des coûts des lymphomes et de la leucémie myéloïde chronique. Cancer. Université Paul Sabatier - Toulouse III, 2019. Français. NNT : 2019TOU30057 . tel-02896471

HAL Id: tel-02896471

<https://theses.hal.science/tel-02896471>

Submitted on 10 Jul 2020

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.



THÈSE

En vue de l'obtention du DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par
Michael MOUNIE

Le 5 juillet 2019

Les évaluations économiques en population dans les prises en charge des cancers : évaluation des coûts des lymphomes et de la leucémie myéloïde chronique

Ecole doctorale : **EDMITT - Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse**

Spécialité : **Epidémiologie**

Unité de recherche :

LEASP - Laboratoire d'Epidémiologie et Analyses en santé publique

Thèse dirigée par
Laurent MOLINIER et Nicolas SAVY

Jury

M. Benoît DERVAUX, Rapporteur
Mme Emmanuelle LERAY, Rapporteur
M. Thomas RAPP, Examineur
M. Alain GRAND, Examineur
Mme Karine CHEVREUL, Examinatrice
Mme Nadège COSTA, Examinatrice
M. Laurent Molinier, Directeur de thèse
M. Nicolas SAVY, Co-directeur de thèse

Les évaluations économiques en population dans les prises en charge des cancers : évaluation des coûts des lymphomes et de la leucémie myéloïde chronique

Remerciements

La liste est longue et la peur d'oublier certaines personnes se fait ressentir. Je vais néanmoins essayer de minimiser mes oublis tout en maximisant ma reconnaissance.

Je remercie bien entendu M Benoit Dervaux et Mme Emmanuelle Leray d'avoir accepté de juger mon travail de thèse. Ce travail de relecture nécessite un temps important et je vous suis très reconnaissant d'avoir joué le jeu malgré des délais plus courts que la normale.

J'en profite pour remercier également l'ensemble du Jury, M Alain Grand, Mme Karine Chevreul et M Thomas Rapp, qui était également pressenti comme rapporteur.

Ensuite, sur le podium des remerciements c'est avec joie que je place mes directeurs de thèse, Laurent et Nicolas, ainsi que tout naturellement Nadège qui est également présente dans mon jury. Rien n'aurait été possible sans vous trois qui avez tout fait pour que l'aventure se réalise dans les meilleures conditions. Un grand merci de m'avoir fait découvrir le monde de la recherche, tous ses mécanismes, sa politique et ses retors. J'ai eu grand plaisir à travailler avec vous, j'espère que c'était réciproque et que ce plaisir se poursuivra. Je vous suis extrêmement reconnaissant d'avoir pris le temps de m'accompagner durant ces 3... 4... 5 années dans tous les domaines et notamment l'expression et la rédaction, mes bêtes noires.

J'en profite pour remercier toute l'Unité d'Evaluation Médico-Economique,

Romain avec qui je me suis particulièrement bien entendu. Nous avons pu avoir nos petits débats autant mathématiques, informatiques, statistiques que ludiques. J'ai été ravi de te chaperonner durant ton stage de M1 et de te savoir dans l'équipe pour la suite de tes études et le début de ta thèse.

Géraldine, tu es partie pour d'autres horizons, j'ai été ravi de collaborer avec toi pendant une majeure partie de mon doctorat. J'espère que tu trouveras chaussure à ton pied pour la suite et que M. R et M. Stata arrêteront de t'embêter.

Hélène, avec qui j'ai eu de très bons rapports. Merci pour ta bonne humeur et tes rires lointains qui remettent toujours d'aplomb. J'espère avoir l'occasion de collaborer avec toi un jour.

Je remercie également tout le Département d'Information Médicale de m'avoir accueilli à l'hôtel Dieu. J'ai passé de très agréables moments avec vous et j'espère pouvoir continuer à travailler à vos côtés. Un remerciement particulier aux DD du DIM : Dominique & Didier ainsi qu'au maître informatique de l'étage d'en dessous, Vincent, administrateur de nos postes et résolveur de tous problèmes numériques, à Laurent B médecin DIM avec qui j'ai eu l'occasion de passer d'agréables pauses cigarettes et à tous les autres (Emmanuelle, Marie-Christine, Thierry etc..).

Je remercie également Fanny, Bethy et Laure de m'avoir aidé dans les méandres des problèmes administratifs et avec qui j'ai vraiment apprécié communiquer.

Un immense merci à mes collègues de l'INSERM 1027,

Avec particulièrement l'équipe EQUITY qui m'a intégré dans ses rangs en vue de passer mon doctorat et avec qui j'ai tissé des liens forts. Je vous suis très reconnaissant et j'ai été ravi de travailler avec des gens aussi intéressants que généreux. Votre thématique de recherche a retenu toute mon attention, j'espère que vous poursuivrez dans cette direction et parviendrez à diminuer les inégalités toujours plus présentes dans la santé. J'ai particulièrement apprécié les présentations et débats animés des réunions d'équipe du mardi.

Je ne peux pas citer l'ensemble de l'équipe mais je remercie tout particulièrement Cyrille et Thierry, présent et ancien responsable de l'équipe qui m'ont intégré et convié à l'ensemble des réunions et activités de l'équipe.

Je remercie Sébastien avec qui j'ai été très heureux de travailler et de partir en congrès. J'aime me sentir dans une relation de confiance et d'amitié dans laquelle on peut communiquer librement. C'est bien plus agréable et efficace !

Une pensée également pour tous ceux avec qui j'ai eu la chance d'échanger tels que Charlie, Hélène, Raf, Pascale et tant d'autres...

Je continue mes remerciements en citant le projet CAPTOR sans quoi je n'aurais jamais commencé cette thèse. Malgré certaines modifications dans mon sujet originel, c'est dans le cadre de ce projet que j'ai pu rencontrer une majorité des personnes avec qui j'ai travaillé par la suite. Un grand merci donc à tous les participants du projet, et notamment au professeur Guy Laurent de m'avoir accordé sa confiance ainsi qu'à Fabien Despas qui m'a épaulé pendant ce début de doctorat. Malgré des réunions tardives le jeudi ou mercredi soir, je ne regrette en rien d'y avoir participé. Elles étaient très instructives et m'ont apporté une approche et un cadre différent de celui dans lequel je travaillais habituellement.

Je remercie bien évidemment mes collègues de pharmaco. Travailler ensemble s'est toujours déroulé dans la joie. Un grand merci notamment à Maryse pour ses conseils avisés et sa participation dans nombreux de mes travaux. Je remercie grandement Cécile avec qui j'ai eu la chance de collaborer sur une grande partie de mes recherches. Une pensée également pour Manuela, Aurore, Bérangère et l'ensemble des personnes avec qui j'ai eu l'occasion de travailler.

Je reconnais avoir loupé les dernières réunions mais j'étais en vacances tout de même ! Un grand merci à Grégory et Valérie pour la création de ce groupe d'échange entre maso de stats. C'était une riche idée ! Vous êtes tous les deux des personnes formidables et j'ai adoré échanger avec vous ! Que ce soit Valérie la bavarde ou Grégory le Breton ! A très vite pour plus de stats et plus de conneries ;) . Un petit coucou à tous les gens du groupe avec qui je m'entends très bien (Caroline, Mélanie, Virginie, Vanessa etc...)

J'adresse un énorme merci à Sandra ! Secrétaire INSERMienne au (très) grand cœur avec qui j'adore échanger. Merci beaucoup de m'avoir assisté dans à peu près toutes mes corvées administratives et de nous réserver de si beaux hôtels pour les congrès... Ne change en rien, t'es au top !

Un merci à Camille et Emilie que j'ai eu la joie de côtoyer dans le cadre professionnel et amical. J'espère que l'on gardera contact malgré ton départ Emilie. J'ai passé de très bons moments avec vous. J'ai été très heureux de vous rencontrer, d'échanger avec vous et de joindre ainsi l'utile à l'agréable.

Je tiens également à intégrer dans mes remerciements certaines personnes que je n'ai pas pu ranger dans mes cases « boulot » : Gauthier Chassang mon juriste préféré, Damien Mouly & Didier Raboisson, mes maitres de stage de master qui m'ont donné l'envie de rentrer dans le monde de la recherche, Agnès Requis & Martine Labryère de m'avoir aidé dans les étapes administratives associées à l'ED, Farida Ghib pour ses conseils d'avant soutenance, Martin Gauthier, hématologue avec qui j'ai apprécié collaborer et Grégory Pugnet, interniste avec qui j'ai eu la chance de m'associer dans les débuts de mon doctorat.

Je quitte le travail pour en venir à l'essentiel. Je ne remerciais jamais assez ma famille d'avoir fait de moi ce que je suis maintenant. Merci à mes parents de m'avoir inculqué vos notions/ préceptes de respect, de gentillesse, d'altruisme et tout simplement d'amour. J'espère que vous êtes fiers de moi comme je le suis de vous. Je remercie également mon frère avec que j'ai passé une enfance très fun. Malgré la distance, j'en garde de très bon souvenir et j'espère que l'on aura l'occasion de re-trainer ensemble dans le futur.

Je remercie également et très sincèrement tous mes ami(e)s. Merci d'être aussi cons qu'intelligents, d'être aussi débiles que sérieux, d'être aussi ouverts d'esprit, bordéliques que rigoureux, engagés que laxistes. J'ai passé, je passe encore et j'espère que ça continuera, d'excellents moments avec vous tous. J'espère n'oublier personne... Surtout pas les gens qui vont venir à la soutenance pour ne lire que les remerciements afin de s'y voir ^^ . Après avoir hésité à reprendre les remerciements de vos thèses pour plus de simplicité, je vais improviser...

Ainsi, une KassDéDi à (ne faites pas attention à l'ordre, ça n'a rien à voir avec mes préférences... quoique... !?) Ali l'ami maladroite des bouteilles, Matthieu le macroniste pétanqueur à toute heure, Tif le timide Blackbloc, Hélène l'amoureuse transi du soleil levant (c'était plus joli que « l'asiat' » :p), et Hadri mon précieux ex-colloc (KouliKouli ?), Fabien Statan l'AntiFa-Coreur, Sarah l'ornitho 🎵 non-chanteuse 🎵, Paulo l'ami des fleurs ♣, Rodolph le paumé squatcheur et sa FeuMeu Priscilla l'amie des vaches (mais pas forcément réciproque), PM le gros porc (ça t'apprendra de m'appeler 'végét' dans tes remerciements ;)), Laurent et Aicha nos nouveaux Toulousains, Alex Achier, euh Acquier pardon, le pro des néons, Idriss l'asperger du poker (et ça rime !), Bertille la scoliose et son homme de Pierre, Clémence la gentille des îles, Abdul le ravagé ravageur, Florian et Thomas mes camarades

informaticiens en jaune, La Benje mon petit ou grand frère selon les cas, Roch le slacker de l'extrême, Ben le nouveau roi des hippies, Thibaut et Adeline les plus jeunes mais plus avancés que le double H, Medhi le Rebeu-logiste accidenté, Samy le cachet de l'amour, Soufiane mon confrère stateu marocain et sa chérie du sud, Alexis L. pour son bon pain, Maxou mon dealeur de cartes Pokémon préféré et Cannelle sa chérie parfumée de nouvelle Calédonie.. Egalement une pensée pour Lionel et Touki.. et à tous ceux que j'ai oublié..

Je remercie bien entendu et pour finir, ma douce Céline, qui m'a accompagné durant la majorité de cette thèse. Je ne sais pas si j'ai été insupportable ou non mais merci d'avoir été là lorsque j'en ai eu besoin. J'espère que tu seras fière de môa lorsque tu me verras soutenir, transpirant sous ma chemise en train d'essayer de paraître intelligent... :). Comme promis je m'arrête là et je ne te fous pas trop la honte... ♥

Financements

Ce travail a reçu le soutien de l'Agence Nationale de la recherche (ANR) dans le cadre des investissements d'avenir (ANR-11-PHUC-001, Programme CAPTOR). Il a également reçu le soutien de l'Unité d'Evaluation Médico-Economique du CHU de Toulouse.

Sommaire

<u>LES EVALUATIONS ECONOMIQUES EN POPULATION DANS LES PRISES EN CHARGE DES CANCERS : EVALUATION DES COUTS DES LYMPHOMES ET DE LA LEUCEMIE MYELOÏDE CHRONIQUE</u>	<u>1</u>
<u>PARTIE 1. INTRODUCTION, OBJECTIFS ET STRUCTURE DE LA THESE</u>	<u>15</u>
<u>I. INTRODUCTION.....</u>	<u>16</u>
<u>II. OBJECTIFS ET STRUCTURE DE LA THESE.....</u>	<u>20</u>
II.1. OBJECTIFS DE LA THESE.....	20
II.2. STRUCTURE DE LA THESE.....	20
<u>PARTIE 2. ETAT DE L'ART</u>	<u>23</u>
<u>III. EVALUATIONS ECONOMIQUES DANS LE DOMAINE DE LA SANTE.....</u>	<u>24</u>
III.1. LES METHODES D'EVALUATION ECONOMIQUE	24
III.1.1. Analyse de coût ou « Cost of Illness »	27
III.1.2. Analyse de minimisation des coûts.....	29
III.1.3. Analyse coût efficacité.....	29
III.1.4. Analyse coût utilité	31
III.1.5. Analyse coût bénéfice.....	33
III.1.6. Analyse d'impact budgétaire.....	34
III.2. COUTS EN ECONOMIE DE LA SANTE	36
III.2.1. Coût d'opportunité	37
III.2.2. Typologie des coûts	37
III.2.3. Valorisation des coûts.....	40
III.2.4. Point de vue.....	47
III.2.5. Actualisation des coûts.....	48
III.3. ANALYSES DE SENSIBILITE	50
III.3.1. Analyses de sensibilité déterministes	52
III.3.2. Analyses de sensibilité probabilistes.....	54
III.3.3. Courbe d'acceptabilité.....	64
<u>IV. SOURCES DE DONNEES ECONOMIQUES</u>	<u>67</u>
IV.1. LES ETUDES EPIDEMIOLOGIQUES ET CLINIQUES	68
IV.1.1. Les différents types d'étude.....	69
IV.1.2. Le recueil des données dans le cadre des études épidémiologiques et cliniques	75
IV.2. LES REGISTRES	76
IV.3. LES BASES DE DONNEES MEDICO-ADMINISTRATIVES DU SYSTEME NATIONAL DES DONNEES DE SANTE (SNDS)	78
IV.3.1. Finalités du SNDS.....	79
IV.3.2. Structure du SNDS	80
IV.3.3. Accès au SNDS.....	89
IV.3.4. Intérêts, limites et usages des données du SNDS	91
IV.4. REGLEMENTATION ASSOCIEE A L'UTILISATION DES DONNEES DE SANTE.....	95
<u>V. MODELISATION DES COUTS.....</u>	<u>98</u>

V.1. OBJECTIFS ET HYPOTHESES GENERALES DES MODELES PREDICTIFS.....	101
V.1.1. Objectifs	101
V.1.2. Modèle linéaire simple et hypothèses fondamentales.....	102
V.2. PRISE EN COMPTE DE LA DISTRIBUTION DES COUTS.....	104
V.2.1. Coûts strictement positifs.....	104
V.2.2. Coûts positifs	108
V.3. PRISE EN COMPTE DES DONNEES CORRELEES.....	112
V.3.1. Modèles de population générale.....	113
V.3.2. Modèles mixtes	115
V.3.3. Comparaison entre les deux approches	117
V.4. SELECTION DES VARIABLES ET ADEQUATION DU MODELE	122
V.4.1. Sélection des variables	122
V.4.2. Vérification de l'adéquation du modèle	125
<u>VI. HEMOPATHIES MALIGNES.....</u>	129
VI.1. LYMPHOMES.....	130
VI.1.1. Carcinogénèse	130
VI.1.2. Anatomopathologie et symptomatologie	131
VI.1.3. Lymphomes non Hodgkiniens.....	132
VI.1.4. Lymphomes Hodgkiniens.....	143
VI.1.5. Coût des lymphomes en France.....	148
VI.2. LEUCEMIE MYELOÏDE CHRONIQUE	150
VI.2.1. Carcinogénèse	150
VI.2.2. Anatomopathologie et symptomatologie	151
VI.2.3. Epidémiologie et facteurs de risque	152
VI.2.4. Classification clinique	152
VI.2.5. Diagnostic.....	153
VI.2.6. Prises en charge thérapeutiques	154
VI.2.7. Coût de la Leucémie Myéloïde Chronique en France.....	157
<u>PARTIE 3. RESULTATS : EVALUATION DES COÛTS DES LYMPHOMES ET DE LA LEUCEMIE MYELOÏDE CHRONIQUE.....</u>	160
<u>VII. EVALUATION DES COUTS DES LYMPHOMES.....</u>	161
VII.1. RESUME DE L'ARTICLE "COSTS OF HODGKIN AND NON-HODGKIN LYMPHOMAS IN FRANCE: RESULTS BASED ON THE FRENCH NATIONAL HEALTH INSURANCE DATABASE"	161
VII.2. REAL WORLD COSTS OF HODGKIN AND THE MAIN B-CELL NON-HODGKIN LYMPHOMAS IN FRANCE ...	163
<u>2. METHOD.....</u>	166
2.1. Study design, setting and population.....	166
2.2. Identification of the study population.....	166
2.3. Costs estimates	167
2.4. Others data.....	167
2.5. Statistical analysis.....	168
<u>3. RESULTS.....</u>	168
3.1. Patient characteristics.....	168
3.2. Cost analysis	169
<u>4. DISCUSSION.....</u>	170

<u>VIII. EVALUATION DU COUT DE LA LEUCEMIE MYELOÏDE CHRONIQUE.....</u>	<u>182</u>
<u>IX. DISCUSSION GENERALE ET PERSPECTIVES.....</u>	<u>203</u>
IX.1. UN CONTEXTE EN PLEINE MUTATION	203
IX.1.1. Evolutions des prises en charge en cancérologie.....	203
IX.1.2. Place de l'évaluation économique dans le système de santé français	205
IX.1.3. Evolution des modes de financement : le financement au parcours de soins	207
IX.2. EVALUATION DES COUTS DES STRATEGIES DE PRISE EN CHARGE	209
IX.2.1. Fardeau économique des hémopathies malignes	209
IX.2.2. Estimation des coûts de prise en charge en population générale : un enjeu majeur pour le système de santé.....	211
IX.2.3. Modélisation des coûts et prise en compte de l'incertitude	213
<u>X. CONCLUSION ET PERSPECTIVES.....</u>	<u>216</u>
<u>BIBLIOGRAPHIE.....</u>	<u>218</u>
<u>ANNEXE.....</u>	<u>242</u>

Liste des abréviations

ACB : Analyses Coût-Bénéfice
ACE : Analyses Coût-Efficacité
ACS : Aide au paiement d'une Complémentaire Santé
ACU : Analyses Coût-Utilité
AIB : Analyse d'Impact Budgétaire
AIC : Akaike Informal Criterion
ALD : Affection de Longue Durée
AM : Assurance Maladie
AMC : Analyse de Minimisation des Coûts
AMM : Autorisation de Mise sur le Marché
ANSM : Agence Nationale de Sécurité du Médicament
ARS : Agence Régionale de Santé
ASD : Analyse de Sensibilité Déterministe
ASP : Analyse de Sensibilité Probabiliste
ATC : Anatomique, Thérapeutique et Chimique
ATIH : Agence Technique de l'Information sur l'Hospitalisation
BCA : Bias Corrected and Accelerated
BCMD : Base des Causes Médicales de Décès
CAPTOR : CANcer Pharmacology of Toulouse-Oncopole and Region
CCAM : Classification Commune des Actes Médicaux
CEAC : Cost-Effectiveness Acceptability Curve
CépiDc : Centre d'épidémiologie sur les causes médicales Décès
CEESP : Commission d'Evaluation Economique et de Santé Publique
CER : Comité d'Evaluation des Registres
CEREES : Comité d'Expertise pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé
CES : Collègue des Economistes de la Santé
CIM-10 : Classification Internationale des Maladie, 10ieme révision
CIP : Code Identifiant de Présentation
CMU-(C) : Couverture Maladie Universelle-(Complémentaire)
CNAMTS : Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés
CNIL : Commission Nationale de l'Informatique et des Libertés
CNR : Comité National des Registres
CNSA : Caisse Nationale de la Solidarité pour l'Autonomie
COI : Cost Of Illness
CPAM : Caisse Primaire d'Assurance Maladie
CPP : Comité de Protection des Personnes
CSBM : Consommation de Soins et de Biens Médicaux
CSM : Centres de Services Métiers
CSH : Cellules Souches Hématopoïétiques
DAP : Disposition A Payer
DCIR : Datamart de Consommation Inter Régime
DGOS : Direction Générale de l'Offre de Soins
DIM : Département d'Information Médicale
DM : Dispositif Médical
DMI : Dispositifs Médicaux Implantables
DRG : Diagnosis Related Groups
ECR : Essais Cliniques Randomisés
EGB : Echantillon Généraliste des Bénéficiaires
EGBS : Echantillon Généraliste des Bénéficiaires Simplifié

El : Effets Indésirables
EME : Evaluation Médico-Economique
ENC : Etude Nationale des Coûts
ENCC : Etude Nationale des Coûts à méthodologie Commune
FDep : French Deprivation index
GEE : Generalized Estimating Equation
GHM : Groupes Homogènes de Malades
GHS : Groupe Homogène de Séjour
GLM : Generalized Linear Model
HAD : Hospitalisation A Domicile
HAS : Haute Autorité de Santé
IC : Intervalle de Confiance
ICER : Incremental Cost-Effectiveness Ratio
ICUR : Incremental Cost-Utility Ratio
IID : Indépendamment et Identiquement Distribués
IJ : Indemnité Journalière
INDS : Institut National des Données de Santé
INSEE : Institut National de la Statistique et des Etudes Economiques
INSERM : Institut National de la Santé et de la Recherche Médicale
ISPOR : International Society for Pharmacoeconomics and Outcomes Research
ITK : Inhibiteurs de la Tyrosine Kinase
LBDGC : Lymphomes B Diffus à Grandes Cellules
LCM : Lymphomes à Cellules du Manteau
LF : Lymphomes Folliculaires
LFSS : Loi de Financement de la Sécurité Sociale
LH : Lymphomes Hodgkiniens
LL : Lymphomes Lymphocytiques
LLB-B : Lymphomes/leucémies LymphoBlastiques à précurseurs B
LLC : Leucémie Lymphoïde Chronique
LLP : Lymphomes LymphoPlasmocytaires
LMC : Leucémie Myéloïde Chronique
LNH : Lymphomes Non Hodgkiniens
LPPR : Liste des Produits et Prestations Remboursables
LT/NK : Lymphomes T/NK
LZM : Lymphomes de la Zone Marginale
MCF : Méthode des Coûts de Friction
MCH : Méthode du Capital Humain
MCO : Médecine Chirurgie Obstétrique et odontologie
MDPH : Maisons Départementales des Personnes Handicapées
MM : Myélome Multiple
MSA : Mutualité Sociale Agricole
MW : Macroglobulinémies de Waldenström
NABM : Nomenclature des Actes de Biologie Médicale
NGAP : Nomenclature Général des Actes Professionnels
NICE : National Institute for Health and Care Excellence
NIR : Numéro d'Identification au Répertoire
OMS : Organisation Mondiale de la Santé
ONDAM : Objectif National de Dépenses d'Assurance Maladie
PIB : Produit Intérieur Brut
PMSI : Programme de Médicalisations des Systèmes d'Information
PS : Professionnel de la Santé
QALY : Quality-Adjusted Life Year

RCP : Réunion de Concertation Pluridisciplinaire
RG : Régime Général
RNIAM : Répertoire National Inter Régime de l'Assurance Maladie
RNIPP : Répertoire National d'Identification des Personnes Physiques
RSI : Régime Social des Indépendants
RSS : Résumé de Sortie Standardisé
RUM : Résumés d'Unités Médicales
SA : Service Attendu
SLM : Sections Locales Mutualistes
SMR : Service Médical Rendu
SNDS : Système National des Données de Santé
SNIIRAM : Système National d'Information Inter Régimes de l'Assurance Maladie
SSR : Soins de Suite et de Réadaptation
T2A : Tarification A l'Activité
TEP : Tomographie par Emission de Positons
WTP : Willingness to pay
ZAGA : Zero Adjusted GAMMA
ZAIG : Zero Adjusted Inverse Gaussian

Liste des tableaux

Tableau 1. Différents types d'évaluation en santé.
Tableau 2. Avantage et limites des données du SNDS
Tableau 3. Différences entre modèles de populations générales et les modèles mixtes.
Tableau 4. Les Effets Indésirables dus aux traitements des LNH.
Tableau 5. Les Effets Indésirables dus aux traitements des LH.
Tableau 6. Effets indésirables rencontrés la 1ère année après l'initiation aux Inhibiteurs de Tyrosine Kinase de première ou seconde génération

Liste des figures

Figure 1. Analyse comparative de deux programmes de santé
Figure 2. Typologie des coûts en économie de la santé
Figure 3. Exemple de diagramme de Tornado
Figure 4. Ellipse de confiance au niveau $\alpha=5, 50, 95\%$ (i.e, $p=1- \alpha$)
Figure 5. Exemple d'ICER bootstrapé avec l'ellipse de confiance
Figure 6. Courbe d'acceptabilité
Figure 7. Alimentation et restitution du SNDS
Figure 8. Les différents accès au SNDS
Figure 9. Aspects réglementaires des études mobilisant des données individuelles
Figure 10. Représentation de Y en fonction de X avec la droite de régression
Figure 11. : Densité des lois gamma et log normales
Figure 12. : Densité de la loi inverse-gaussienne
Figure 13. Histogramme représentant une distribution Tweedie
Figure 14. Point influent et effet de levier
Figure 15. Répartition des principaux sous types de LNH (%) (France, 2012)
Figure 16. Les différents traitements des LH selon le stade

Liste des annexes

Annexe 1. Avantages et inconvénients des données issues des différents designs pour l'évaluation économique

Annexe 2 : Structure générale de la collecte des informations dans un registre

Annexe 3 : Evolution du SNIIRAM jusqu'au SNDS

Annexe 4. Principales informations contenues dans le PMSI

Annexe 5. Représentation de l'alimentation et des restitutions du SNDS

Annexe 6. Organisme ayant droit à un accès permanent au SNDS

Annexe 7. Synthèse du circuit d'accès ponctuel au SNDS

Annexe 8. Ecriture et résolution des modèles de population générale.

Annexe 9. Différents types de modèles mixtes.

Annexe 10. Etude des résidus

Annexe 11. Stades des lymphomes

Annexe 12. Evolution des traitements et de la survie des patients atteints de la LMC

PARTIE 1. INTRODUCTION, OBJECTIFS et STRUCTURE DE LA THESE

I. Introduction

Le système de santé français est confronté depuis de nombreuses années à de profondes mutations. Les difficultés croissantes de financement des dépenses de santé ont obligé les pouvoirs publics à engager de multiples réformes visant à limiter ces dernières à une part acceptable des ressources nationales. Ces réformes ont eu pour conséquences de soumettre régulièrement les acteurs du système de santé à de nouvelles contraintes budgétaires. Dans ce contexte de ressources limitées, des choix budgétaires s'imposent, rendant plus nécessaire encore la détermination de priorité et de choix stratégiques en termes de santé publique (1).

Les outils de l'évaluation économique sont connus des industriels depuis des années. Son application en médecine et en santé publique a débuté dans les années soixante et s'est développée à partir de 1970 en raison de contraintes économiques et d'une évolution technologique permettant l'amélioration des performances médicales, mais au prix de moyens accrus (2). Parmi les évaluations économiques dans le domaine de la santé, l'évaluation médico-économique (EME) des stratégies médicales s'est largement développée en Amérique du Nord et dans le Nord de l'Europe. Son utilisation en France est plus récente. Elle analyse de façon comparative les coûts et les conséquences médicales de plusieurs stratégies de prise en charge concurrentes. Ainsi, en cherchant à estimer le bilan économique net pour la société d'un programme de santé ou en permettant de choisir parmi plusieurs stratégies tendant au même but, les évaluations économiques des stratégies médicales peuvent aider à la prise de décisions dans le processus d'allocation des ressources au sein du secteur de la santé (3).

La cancérologie qui constitue un enjeu national, s'inscrit précisément dans cette problématique. Les prises en charge préventives, diagnostiques, thérapeutiques, et plus largement sociétales des patients atteints de pathologie cancéreuse, consomment des moyens de plus en plus importants (4). Elles gagnent en efficacité mais deviennent également plus onéreuses, et expliciter les choix thérapeutiques retenus devient capital. Les professionnels de santé sont ainsi confrontés à des

questions difficiles en matière de choix entre les stratégies de prise en charge. Ces questions portent notamment sur le développement de nouveaux créneaux de soins, sur les possibilités thérapeutiques offertes par des molécules onéreuses, voire très onéreuses, sur l'évaluation des stratégies de prise en charge diagnostiques et thérapeutiques, et plus généralement sur l'établissement des priorités. Le plan cancer 2014-2019 insiste sur « La recherche d'une efficacité accrue des organisations et d'une plus grande pertinence des soins qui permettra de conserver une ambition forte pour accélérer les progrès et toujours améliorer la qualité des soins » (5). Une absence d'évaluation des coûts des activités en fonction de leurs résultats peut être préjudiciable au rendement des ressources consommées par les Professionnels de Santé (PS). L'insuffisance des données économiques peut conduire à une distribution inappropriée de ces ressources, compte tenu de leur rareté. En apportant des informations complémentaires à l'évaluation clinique, l'EME peut aider l'ensemble des acteurs à identifier plus clairement les stratégies pertinentes en cancérologie.

Les EME des prises en charge des pathologies cancéreuses reposent majoritairement sur des données issues d'études ad hoc avec notamment les essais contrôlés (6). Centrées le plus souvent sur une séquence thérapeutique ou sur une prise en charge réalisée chez des patients sélectionnés, faisant l'objet d'une surveillance accrue et dans un nombre limité d'établissements spécialisés, ces études ne permettent pas de rendre compte véritablement de l'impact des recommandations ou des innovations thérapeutiques sur les pratiques (7). L'approche en population, fondée sur l'utilisation de données d'observation représentatives de la population générale, est intéressante car elle permet de produire des estimations construites sur l'observation de pratiques réelles (8). L'accès désormais facilité aux bases de données médico-administratives nationales, notamment au Système National des Données de Santé (SNDS), offre de réelles possibilités dans ce domaine (9). Ces bases permettent notamment de disposer pour chaque patient, du volume exhaustif des ressources consommées lors des prises en charge effectuées en secteur hospitalier comme en soins de ville. L'accès à ces bases de données facilite le partage des connaissances portant sur la qualité, la sécurité et la pertinence des soins, les pratiques professionnelles et le recours aux soins (10). En raison de leur exhaustivité, de

leur robustesse, de leur caractère représentatif des pratiques observées et de leur accès qui devrait être facilité, les données du SNDS vont être de plus en plus d'une utilité majeure dans la réalisation d'études médico-économiques et permettront de compléter les résultats provenant des études ad hoc (8) (11).

Toutefois, les bases populationnelles donnant accès à des données de vie réelle sont composées de populations hétérogènes. La mise en œuvre de comparaisons entre différentes populations peut être confrontée à de multiples biais (i.e. sélection, confusion) pouvant conduire à de fausses interprétations et directement affecter les choix des décideurs. Afin de pallier à ces biais, les évaluations peuvent être réalisées en faisant appel à des techniques de modélisation statistique. La modélisation permet de structurer les données en identifiant les différentes entités et dépendances entre elles, d'ajuster les estimations selon les caractéristiques de la population utilisée afin d'obtenir des résultats plus proches de la réalité et de guider les autorités de santé dans les décisions à prendre vis-à-vis de la collectivité (12). Différents modèles théoriques ont été construits et permettent de s'adapter à un grand nombre de situation. Ces modèles sont applicables à l'évaluation économique des stratégies de soins des pathologies cancéreuses. Le choix du modèle approprié dépendra de multiples facteurs tels que l'objectif poursuivi, la pathologie étudiée, les paramètres pris en compte, les données disponibles. Leur utilisation est essentielle à l'acquisition de connaissances jouant un rôle moteur lors de la conception et de l'interprétation des phénomènes. Dans l'EME des stratégies de prise en charge, les modèles peuvent aider à clarifier les conséquences de processus décisionnels et à simuler l'effet de programmes alternatifs. Sans se substituer aux données d'observation, ils utilisent et complètent ces dernières quand elles sont trop partielles pour pouvoir répondre aux préoccupations des acteurs du système de santé (13).

Les hémopathies malignes sont directement concernées par ces questions où des avancées thérapeutiques majeures ont été réalisées ces dernières années, notamment dans la Leucémie Myéloïde Chronique (LMC), la leucémie lymphoïde, les lymphomes et les myélomes (14). La prise en

charge thérapeutique des lymphomes et de la LMC a été profondément transformée par l'arrivée au cours des années 2000 de traitements innovants, particulièrement efficaces, mais très coûteux. L'introduction, respectivement, des Inhibiteurs de la Tyrosine Kinase (ITK) et de l'immunothérapie à base d'anticorps monoclonaux (rituximab) a permis d'améliorer significativement le pronostic de survie des patients atteints de LMC et de la majorité des lymphomes, mais au prix de moyens substantiels. Les conséquences économiques de ces cancers sont mal connues en France. A notre connaissance, aucune étude n'a évalué, à ce jour, le coût de la prise en charge thérapeutique de la LMC en France et aucune étude n'a évalué le coût de la phase active de traitement des lymphomes avec des données françaises.

Dans un système de santé qui cherche à gagner en efficience, les critères de choix doivent donc passer d'une étape relativement subjective reposant sur peu de données, à une meilleure connaissance des indicateurs de coûts et de résultats prenant en compte l'intégralité des prises en charge des cancers hématologiques. Dans le contexte actuel de transformation du système de santé marqué par un objectif d'amélioration de la qualité et de la pertinence des soins, par la nécessité de renforcer la coordination des soins et de faire évoluer les modes de financement tout en maîtrisant l'évolution des dépenses de santé, il nous a semblé intéressant d'analyser les coûts relatifs à ces deux pathologies. Cette thèse a été effectuée dans le cadre plus général du programme de recherche CAPTOR (Cancer Pharmacology of Toulouse-Oncopole and Region) financé par l'Agence Nationale de la Recherche sur les médicaments anti-cancéreux.

II. Objectifs et structure de la thèse

II.1. Objectifs de la thèse

Les objectifs de cette thèse sont :

- Etablir une synthèse des fondements théoriques des évaluations économiques en santé,
- Présenter les différentes sources de données permettant l'accès à des données utilisables pour l'évaluation économique,
- Présenter différentes approches de modélisations statistiques permettant la modélisation des données économiques,
- Décrire et modéliser les coûts des prises en charge des lymphomes et de la LMC en situation de vie réelle afin, d'une part, d'analyser le coût de la phase active des lymphomes, et d'autre part, d'analyser le coût de la LMC après l'initiation d'un traitement par ITK.

II.2. Structure de la thèse

La première partie de la thèse est constituée de deux chapitres, le chapitre I composé de l'introduction et du contexte puis le chapitre II qui présente les objectifs et la structure de la thèse.

La deuxième partie de la thèse décrit l'état des lieux relatif aux évaluations économiques en santé, aux sources de données permettant l'accès à des données économiques portant sur la consommation des soins, à la modélisation statistique des données économiques et aux lymphomes et à la LMC. Cette partie est structurée en quatre chapitres : les chapitres III, IV, V et VI.

Le chapitre III « Evaluations économiques dans le domaine de la santé » présente les méthodes et les outils des évaluations économiques dans le cadre de la santé. Il envisage notamment la mesure des coûts et la prise en compte de la variabilité des résultats.

Le chapitre IV « Sources de données économiques » décrit les sources des données permettant un accès à des informations économiques portant sur la consommation de soins des patients et l'évaluation économique des pathologies cancéreuses. Les données issues des études d'observations et d'interventions, les registres et les bases de données médico-administratives sont présentés. Nous détaillons dans cette partie la structure, les possibilités offertes, les intérêts et les limites de chacune des sources de données.

Le chapitre V « Modélisation dans le cadre d'une évaluation économique » présente diverses méthodologies permettant une modélisation correcte des données économiques. Les différentes approches considérant des données longitudinales sont notamment étudiées et mises en perspective dans ce chapitre. Nous explicitons également comment parvenir à un modèle final en adéquation avec l'objectif de l'étude et les données disponibles.

Le chapitre VI présente deux hémopathies malignes : Les lymphomes et la LMC. La carcinogénèse, l'épidémiologie, la classification clinique, les prises en charge et les effets indésirables des traitements associés à ces pathologies sont détaillés dans ce chapitre.

La troisième partie aborde l'évaluation économique des lymphomes et de la LMC.

Dans le chapitre VII, nous estimons et analysons les coûts des principaux types de lymphomes en France. Nous y décrivons la méthode, notamment fondée sur la construction d'un échantillon représentatif de patients issus des bases médico-administratives nationales et nous détaillons finement les coûts de la phase active de traitement.

Dans le chapitre VIII, nous estimons et analysons les coûts de la LMC à partir d'un échantillon régional de patients. Nous y décrivons la méthode fondée sur l'élaboration d'un modèle prenant en

compte l'aspect longitudinal des données provenant de la nature chronique de la pathologie, et les résultats particulièrement en matière d'impact du traitement sur les coûts de la prise en charge.

Dans le chapitre IX, nous mettons en perspective les résultats de nos travaux avec des éléments contextuels et méthodologiques. La discussion porte dans un premier temps sur les mutations du système actuel de soins, des évolutions des prises en charge en cancérologie, de la place des évaluations économiques dans ce contexte et de nouveaux modes de financement possibles. Dans un second temps, nous explicitons le fardeau économique grandissant apporté par les hémopathies malignes et la nécessité d'utiliser des données de vie réelle avec une modélisation adaptée utiles aux décideurs.

Enfin, nous concluons en décrivant nos perspectives de recherches dans le chapitre X.

PARTIE 2. ETAT DE L'ART

III. Evaluations économiques dans le domaine de la santé

Un ensemble d'outils et de méthodologies permettent de mettre en place une évaluation économique afin d'apporter aux acteurs du système de santé des informations économiques complémentaires aux données cliniques utiles à la prise de décision. De par le rôle de « porteurs de risques » des décideurs et de par le grand nombre de données d'entrée incertaines dans ces évaluations, il est essentiel d'avoir une mesure de la variabilité des résultats avant de prendre une décision finale. Nous expliciterons dans ce chapitre les différentes méthodologies associées aux évaluations économiques dans le domaine de la santé, comment valoriser les coûts, notamment du point de vue de l'Assurance Maladie (AM), et comment prendre en compte la variabilité des résultats.

III.1. Les méthodes d'évaluation économique

Selon la Haute Autorité de Santé (HAS), les EME mettent en regard les résultats attendus d'une intervention de santé avec les ressources consommées pour la produire (3). Leur objectif est de comparer les différentes options envisageables pour les décideurs selon leurs résultats et en fonction des ressources mobilisées, en vue d'une allocation optimale des ressources. Cet objectif équivaut à une recherche d'efficacité. L'EME fait ainsi face à deux logiques décisionnelles. Au niveau de la population entière, l'éthique impose de mener une réflexion sur la répartition optimale et équitable des ressources disponibles. Elle concerne un choix collectif. Au niveau médical, la déontologie suppose de pouvoir prendre en considération et d'appliquer à l'individu l'ensemble des données de la science disponible. Elle relève d'un choix individuel.

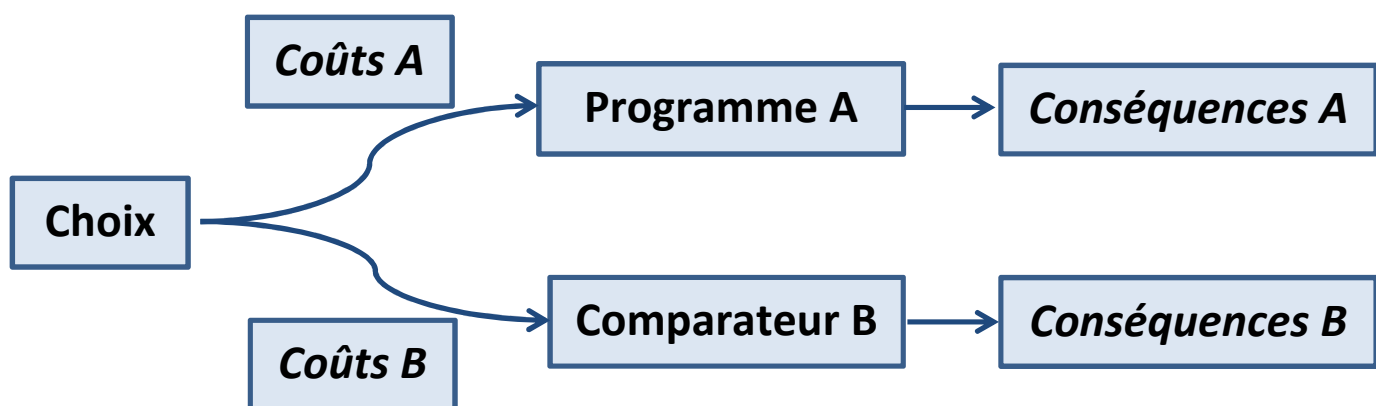
L'EME des programmes de santé peut être définie comme « l'analyse comparative d'options possibles, sur la base de leurs coûts comme de leurs conséquences » (15). Le premier travail associé à ces évaluations consiste donc à identifier, mesurer et comparer les coûts et les conséquences des

options envisagées. Ce travail intégrant la définition préalable des options, caractérise toutes les EME menées dans le secteur de la santé. Les évaluations économiques en santé apportent également des informations sur les coûts d'une pathologie, et les conséquences sur le budget d'un financeur de l'introduction et la diffusion sur le marché d'un nouveau programme de santé.

L'EME explore alors différentes facettes d'un tel programme. Selon Drummond, on peut les appréhender sous la forme d'input et d'output (15). Les inputs correspondent aux ressources consommées par le ou les programmes évalués. Les outputs représentent les conséquences du ou de ces programmes. Les conséquences peuvent être exprimées de différentes façons selon le type d'EME réalisée. Elles peuvent être représentées par des coûts évités ou supplémentaires, par une mesure de l'efficacité (morbidité, mortalité, survie), par une quantification de l'amélioration de l'état de santé (qualité de vie) ou par un bénéfice (prix de la vie humaine).

Soient A et B deux programmes mis en comparaison. On schématise le choix à faire entre les deux programmes selon la figure 1.

Figure 1. Analyse comparative de deux programmes de santé



La nature précise des coûts et des conséquences à prendre en compte et leurs évaluations seront discutées dans les paragraphes suivants. Cependant dès lors qu'on évalue deux programmes A et B, on compare en règle générale la différence entre les coûts à la différence entre les conséquences, dans le cadre d'une analyse différentielle.

Le tableau 1 permet de distinguer les différentes situations d'évaluation en santé communément rencontrées dans la littérature, à partir des réponses apportées aux deux questions suivantes :

- Etudie-t-on à la fois les coûts et les conséquences des options envisagées ?
- Y a-t-il une comparaison entre plusieurs options ?

Tableau 1. Les différents types d'évaluation en santé.

Etudie-t-on à la fois les coûts et les conséquences des options envisagées?

		Non		Oui
		Examen des conséquences seules	Examen des coûts seuls	
Y a-t-il une comparaison de 2 options ou plus?	Non	<i>Evaluation partielle</i>		<i>Evaluation partielle</i>
	Oui	Description des résultats	Description des coûts	Description coût-résultat
	Non	<i>Evaluation partielle</i>		<i>Evaluation économique globale</i>
	Oui	Evaluation de l'efficacité théorique ou pratique	Analyse des coûts	Analyse de minimisation des coûts Analyse coût-efficacité Analyse coût-utilité Analyse coût-bénéfice

Essai clinique randomisé

Coût de la pathologie

Réponse à la question de l'efficacité

Quand les deux réponses aux questions posées sont négatives (cadrant situé en haut et à gauche), on se retrouve dans le cas d'une absence de comparaison entre les options. L'évaluation du programme se limite à une description des résultats. Dans un cas, l'évaluation est restreinte à la description des conséquences cliniques. Dans l'autre cas, l'étude se limite à la description des coûts. On parle alors d'étude de type « Cost Of Illness » (COI).

Lorsqu'on décrit les résultats et les coûts d'un seul programme (cadrant situé en haut et à droite), l'évaluation porte le nom de description coût-résultat.

Le cadrant situé en bas et à gauche correspond aux situations dans lesquelles deux options ou plus sont comparées, sans que ni les coûts, ni les conséquences de chacun des programmes concurrents ne soient examinés simultanément. Il peut alors s'agir d'une évaluation de l'efficacité théorique ou de l'efficacité pratique, ou bien d'une analyse de coûts.

Dès lors que la réponse aux deux questions posées est positive, c'est-à-dire que les coûts et les conséquences de plusieurs programmes de santé sont à la fois évalués et comparés entre eux, on a alors recours aux évaluations économiques globales ou EME (cadran situé en bas et à droite) qui peuvent être déclinées en quatre types d'analyse : les Analyses de Minimisation des Coûts (AMC), les Analyses Coût-Efficacité (ACE), les Analyses Coût-Utilité (ACU) et les Analyses Coût-Bénéfice (ACB).

Nous présentons dans les paragraphes suivants les études de types COI ainsi que les différentes EME. Nous présentons également l'Analyse d'Impact Budgétaire (AIB) qui permet d'analyser l'accessibilité financière d'un programme de santé pour le payeur.

III.1.1. Analyse de coût ou « Cost of Illness »

Selon Jefferson et al, l'objectif des études COI est descriptif (16). Elles visent à détailler, évaluer et additionner les coûts d'une pathologie particulière dans le but d'étudier son fardeau économique. Une étude de type COI portant sur une maladie spécifique met en évidence les économies potentiellement réalisables si la maladie en question venait à disparaître (17). Ce type d'étude apporte des informations importantes sur les différentes composantes du coût et leur poids respectif dans le coût total de la pathologie étudiée. De telles informations peuvent par nature aider les décideurs à identifier des priorités de recherche et de financement en mettant en avant les domaines dans lesquels des économies potentielles peuvent être réalisées (18). Les estimations des coûts peuvent également aider à la réalisation d'EME. Les études COI apportent un grand nombre d'informations sur les coûts d'une pathologie qui peuvent ensuite être utilisées dans l'implémentation de modèles mobilisant des données agrégées issues de la littérature.

Afin de réaliser une étude COI, il est nécessaire d'identifier, de répertorier, de mesurer et d'évaluer les coûts d'une pathologie, et des conséquences qu'elle peut engendrer (complications, comorbidités...). Deux approches peuvent être développées: l'approche reposant sur la prévalence et

celle sur l'incidence. L'approche fondée sur la prévalence permet d'estimer le coût total d'une maladie contractée au cours d'une année donnée. Elle permet d'estimer le poids économique d'une pathologie, si possible pour la société. L'approche par incidence permet d'estimer les coûts induits par une maladie sur une période de temps, qui peut aller du diagnostic initial jusqu'à la guérison ou tout autre évènement, chez des patients nouvellement diagnostiqués. Elle est particulièrement adaptée à l'évaluation des conséquences économiques des pathologies car elle permet d'étudier finement les coûts des différentes phases de la prise en charge d'une pathologie (diagnostiques, thérapeutiques, complications, rémissions, fin de vie) (19). Cette approche offre également la possibilité de disposer d'une base de référence qui aidera à évaluer les innovations dans le domaine de la santé (17). L'approche par prévalence est préférablement utilisée dans les analyses d'impact budgétaire en raison du fait qu'elles considèrent, au sein d'une population, l'ensemble des individus malades (20).

Il existe néanmoins des contre-arguments à l'utilisation d'études de types COI. Certains auteurs critiquent le fait qu'identifier un domaine de dépense relié à une maladie ne suffit pas à justifier des changements dans les actions des pouvoirs publics (21). Compte tenu du fait que les maladies ne peuvent que très rarement être complètement éradiquées, les coûts totaux évalués ne peuvent être économisés en totalité. Une autre critique provient de la priorisation des pathologies étudiées dans le cadre des COI. Les pathologies présentant un lourd fardeau économique sont souvent étudiées même si nous ne savons pas s'il est possible de réaliser des économies en lien avec sa prise en charge. Les pathologies moins coûteuses pour la société sont moins étudiées alors que, dans certains cas, leur prise en charge pourrait être facilement optimisée, notamment par des interventions de prévention peu coûteuses mais apportant un gain clinique important (17). Les études COI contribuent néanmoins au fonctionnement d'un système de santé efficace en apportant des informations utiles et qui peuvent aider à la réalisation d'études comprenant un critère d'efficacité clinique.

III.1.2. Analyse de minimisation des coûts.

L'Analyse de Minimisation des Coûts (AMC) se présente comme un cas particulier de l'analyse coût-efficacité. Elle s'applique lorsque les conséquences de deux interventions sont comparables ou très proches. Elle repose sur une hypothèse très forte d'égalité de l'efficacité et de la tolérance (i.e. innocuité) des stratégies comparées. Elle revient à comparer simplement les coûts des deux options étudiées. Ainsi, l'intervention la moins coûteuse devrait être choisie. Dans ce cas, on peut évoquer la comparaison entre un médicament princeps et son générique par exemple.

Néanmoins, l'hypothèse d'égalité de l'efficacité est rarement totalement satisfaite. L'AMC est utilisée lorsque les niveaux d'efficacité sont jugés sensiblement équivalents. Il faut alors être en capacité de démontrer l'équivalence d'efficacité entre les options comparées. L'AMC est le plus souvent utilisée en pharmaco-économie.

III.1.3. Analyse coût efficacité

L'ACE est l'analyse la plus répandue parmi les EME. Elle se caractérise par la prise en compte simultanée des coûts et des conséquences des interventions étudiées. Ces analyses nécessitent l'utilisation d'un critère d'efficacité adapté. La HAS recommande l'utilisation de la durée de vie calculée à partir de la mortalité toutes causes confondues (3). Le plus souvent, un proxy de l'espérance de vie est utilisé. Pour les pathologies cancéreuses, la survie nette (i.e. relative) ou absolue (i.e. globale) ou le temps jusqu'à la progression de la maladie sont régulièrement retrouvés dans la littérature (22) (23). La survie nette est néanmoins à privilégier car elle permet de s'affranchir des éventuelles différences de mortalité dues à d'autres causes que celle de la maladie étudiée, et facilite ainsi la comparaison entre les programmes.

L'ACE évalue l'efficacité des interventions testées par rapport à une référence ou entre elles. Elle permet d'identifier l'intervention qui dégage le bénéfice médical maximal dans un cadre budgétaire fixé ou l'intervention qui permet d'atteindre un objectif médical donné au moindre coût (15). L'objectif est de classer les interventions par ordre de priorité en utilisant un critère d'efficacité opérationnelle, le Ratio Coût-Efficacité Différentiel ou « Incremental Cost-Effectiveness Ratio » (ICER) définit par:

$$ICER = \frac{C_B - C_A}{E_B - E_A} = \frac{\Delta C}{\Delta E}$$

C_A et C_B sont respectivement les coûts relatifs aux traitements A et B, et E_A et E_B respectivement les efficacités des deux traitements. L'ICER est étudié dès lors qu'une intervention est plus coûteuse ET plus efficace, ou, moins coûteuse ET moins efficace que celle à laquelle elle est comparée. La valeur de l'ICER est ensuite comparée à la Disposition A Payer (DAP) collective, « Willingness to pay » (WTP), du décideur, noté λ . Si l'ICER est inférieur au λ , la stratégie B sera préférée à la A et on dira qu'elle est coût-efficace, tandis que si l'ICER est supérieur à λ , on préférera la stratégie de référence A. L'ICER permet d'exprimer les efforts monétaires à consentir par la société pour gagner une unité d'efficacité supplémentaire.

La DAP est définie dans les systèmes de santé de chaque pays. En France, il n'existe aucun consensus sur la DAP. L'utilisation de la frontière d'efficacité relève de la décision publique. À l'issue de l'évaluation, les interventions sont donc qualifiées d'efficaces ou non sans préjuger de leur acceptabilité au regard de la DAP du décideur public (3). L'usage des résultats des ACE donne lieu à un cas particulier en France. En effet, il ne conduit pas à une décision finale portant sur l'acceptation ou le rejet d'un bien médical ou d'un programme. Ces études aident les pouvoirs publics à identifier les prises en charge les plus efficaces et à s'interroger sur les capacités dont dispose notre système de santé afin de les financer et à quel prix (24).

Néanmoins, l'ACE présente certaines limites. De par sa structure de ratio et la distribution non gaussienne des coûts, la distribution de l'ICER n'est en général pas gaussienne. Il n'est donc pas aisé de travailler sur cet indicateur que ce soit pour évaluer sa variabilité ou pour calculer une taille d'échantillon nécessaire pour disposer d'une puissance statistique acceptable (25) (26). L'ACE est également critiquée quant à l'utilisation du λ . En effet, le choix du λ n'est pas toujours réaliste car dépendant de plusieurs paramètres tels que la prévalence de la maladie étudiée ou le Produit Intérieur Brut (PIB) de la région ou du pays. Il n'est alors pas généralisable et n'a aucune valeur scientifique (27). De plus, il est parfois difficile de choisir un critère d'efficacité adapté, disponible et identique dans les deux options comparées. Enfin, ce type d'analyse ne prend pas en compte la qualité de vie du patient. L'Analyse Coût Utilité (ACU) pallie à ce problème en l'intégrant dans le calcul du ratio.

III.1.4. Analyse coût utilité

L'ACU présente de nombreuses analogies avec l'ACE quant à la construction des ratios, l'évaluation des coûts et son usage. Elle s'en différencie par la prise en compte de la qualité de vie des individus dans les effets de santé des programmes évalués. Elle est particulièrement indiquée dans l'étude des questions où la qualité de vie est un critère important pour juger des résultats d'une intervention.

L'ACU met en relation le coût différentiel des programmes avec l'amélioration différentielle de santé mesurée principalement en années de vie gagnées pondérées par la qualité de vie liée à la santé ou Quality-Adjusted Life Year (QALY). Elle se fonde sur la notion d'utilité. Dans l'approche QALY, l'ajustement par la qualité est fondé sur un ensemble de valeurs ou pondérations appelées utilités, une pour chaque état de santé possible, qui représentent la désirabilité relative de ces différents états (28). Ce concept permet de résumer en un seul indicateur une information quantitative (i.e. un nombre d'années de vie gagnées) et une information qualitative (i.e. une appréciation de la qualité

de vie) (29). Les années de vie sont ainsi pondérées par les utilités ou préférences respectivement accordées par les patients à chacun des états de santé.

Le Ratio Coût-Utilité Différentiel ou « Incremental Cost-Utility Ratio » (ICUR) se calcule de façon comparable à l'ICER :

$$ICUR = \frac{Coût(B) - Coût(A)}{QALY(B) - QALY(A)}$$

Ce ratio s'interprète de façon analogue à l'ICER quant à l'intervention la plus efficiente.

Les ACU font néanmoins l'objet de plusieurs critiques. Lorsque les interventions étudiées sont en lien avec des prises en charge de maladies chroniques, l'hypothèse selon laquelle l'utilité est indépendante du temps passé dans un état n'est pas toujours vérifiée. De plus, ces analyses jouent un rôle limité dans certaines populations, telles que celles recevant des soins palliatifs ou bien les populations âgées (30). Ces populations sont caractérisées par des espérances de vie faibles et sont susceptibles de ne pas profiter des bénéfices des programmes sur le long terme. L'utilité que la société attache à la santé est ainsi dépendante de l'individu qui vit cet état de santé. Deux individus pourraient juger différemment un même état de santé : la mesure de l'utilité est subjective (31).

Par construction, les QALY ne prennent pas en compte l'aversion face au risque. Dans cette approche, un individu est indifférent entre la possibilité de vivre pendant 10 ans avec une qualité de vie optimale ($u=1$) et la possibilité de vivre pendant 20 ans avec une qualité de vie moyenne ($u=0,5$), ce qui peut poser des problèmes tant dans la comparabilité que dans l'interprétation des résultats.

Les QALY sont des indicateurs génériques et sont moins sensibles que d'autres indicateurs spécifiques, directement conçus pour rendre compte des manifestations cliniques en lien avec les pathologies et de leur traitement. Les gains de qualité de vie peuvent ainsi être sous-estimés (30).

L'ACB peut pallier à certaines limites des ACU.

III.1.5. Analyse coût bénéfice

L'analyse coût-bénéfice est fondée sur l'estimation d'un bénéfice net, c'est à dire sur la différence entre les bénéfices et les coûts d'un programme. Elle compare les flux actualisés de bénéfices différentiels et de coûts différentiels du programme, la différence entre ces deux flux étant le bénéfice social net. Le but de ce type d'analyse est de savoir si les bénéfices du programme excèdent ses coûts, un bénéfice social net positif indiquant que le programme mérite d'être mis en œuvre (32). Elle ne s'applique pas seulement en économie de la santé, mais aussi à tous les secteurs de l'économie. L'ACB est la forme la plus aboutie d'EME, notamment en raison de la perspective la plus globale envisagée (33). Elle permet en effet d'évaluer différents choix stratégiques par rapport aux conséquences relatives à tous les acteurs affectés par ces décisions. C'est un outil analytique d'aide à la décision permettant d'allouer les ressources de manière socialement efficace (32).

Cependant, la valorisation des résultats de santé en termes monétaires, constitue la principale difficulté de ce type d'analyse. Tenter d'attribuer de façon explicite des valeurs monétaires aux résultats de santé est controversé. Toutefois, pour évaluer des résultats de santé en termes monétaires, trois approches ont été développées : l'approche dite du capital humain, l'approche par les préférences révélées et l'approche par les préférences déclarées de la disposition-à-payer. L'approche du capital humain pondère le temps en bonne santé par des poids monétaires en utilisant les taux de rémunération du marché. La valeur du programme est mesurée par la valeur actualisée des revenus futurs. La méthode permet ainsi d'attribuer une valeur aux états de santé (34). L'approche des préférences révélées consiste à mesurer le bénéfice des individus en observant leur comportement et l'importance qu'ils accordent au bien obtenu ou enlevé. L'objectif peut être par exemple d'étudier la relation entre certains risques sur la santé, liés à un travail dangereux, et les niveaux de salaires que les individus exigent pour accepter ce travail. La troisième approche se fonde sur le concept de disposition à payer. Elle mobilise des méthodes d'enquête destinées à présenter aux personnes interrogées des scénarios hypothétiques concernant le programme à évaluer. On

demande à ces personnes de réfléchir à l'existence d'un marché hypothétique, pour un programme ou un bénéfice de santé, et d'envisager combien elles seraient prêtes à payer au maximum pour ce programme ou ce bénéfice. Les consommateurs sont interrogés sur les biens qu'ils seraient prêts à sacrifier, pour obtenir les bénéfices du programme si ceux-ci étaient disponibles sur un marché.

L'ACB est, au moins en théorie, la plus achevée des techniques d'évaluation économique car elle permet de répondre directement à des questions d'efficacité allocative, contrairement aux autres techniques. Elle est pourtant difficile à mettre en œuvre et peu utilisée, en raison d'une part, des difficultés liées à l'évaluation monétaire des bénéfices d'un programme de santé, et d'autre part en raison de la quantité importante d'informations à recueillir.

Bien que très utiles, les EME (ACE, ACU et ACB) ne permettent pas de prendre en compte l'impact financier sur les dépenses globales de santé pour les financeurs du déploiement d'un programme de santé. L'analyse d'impact budgétaire apporte ces informations utiles aux décideurs.

III.1.6. Analyse d'impact budgétaire

L'AIB évalue les conséquences sur le budget d'un financeur, de l'introduction et de la diffusion sur le marché d'un programme de santé (35). Elle s'est développée dans les années 2000 avec les publications de Mauskopf (36) et Trueman (37) et gagne en importance en France à partir de 2008 avec la publication du « guide méthodologique pour la mise en place d'une analyse d'impact budgétaire » par le Collège des Economistes de la Santé (CES) (38). L'AIB est obligatoire en France pour les produits éligibles à l'EME et dont le chiffre d'affaires prévisionnel en 2ème année de commercialisation est supérieur ou égal à 50 millions d'euros. Outre cette obligation, elle est encouragée par la HAS lors de la soumission d'un dossier efficacité par une entreprise du médicament ou d'un fabricant de dispositif médical pour un avis dans le cadre de la

commercialisation d'un nouveau produit de santé. La HAS a publié son premier guide méthodologique en 2016 (35).

L'AIB développe une approche financière complémentaire aux analyses de l'efficience qui ne permettent pas de statuer sur les conséquences financières globales, pour un financeur spécifique, de l'adoption d'une intervention sur l'ensemble de la population d'un pays (39). Elle traite de l'accessibilité financière d'un programme de santé et mesure l'incidence financière sur le budget d'un agent payeur de sa mise en place ou de son retrait. Elle complète les évaluations économiques globales en déterminant les interventions les plus efficaces selon l'enveloppe budgétaire disponible (40). Quand les analyses d'efficience s'en remettent à une valeur seuil en deçà de laquelle on devrait accepter toute nouvelle technologie, l'AIB permet d'avoir des informations précises sur les conséquences budgétaires de chaque innovation et permet d'anticiper les moyens de financement nécessaires à son éventuel remboursement (15). Lorsqu'une intervention est validée comme efficace, l'AIB permet alors d'estimer ce que pourrait être l'impact financier de l'introduction et de la diffusion de cette intervention dans la population (41).

Les méthodes d'évaluation économique présentées dans cette partie nécessitent la prise en compte de données économiques. Les coûts associés aux ressources de santé consommées doivent être définis précisément et identifiés selon leur type. Nous allons alors définir plus en détail la notion de coût d'opportunité, la typologie des coûts et leur valorisation dans le contexte Français. Les différents points de vue possibles des analyses économiques ainsi que les préférences pour le présent seront détaillés en suivant.

III.2. Coûts en économie de la santé

L'économie de la santé se fonde sur le concept de coût d'opportunité. Différents auteurs se sont essayés à définir la notion de ce coût. Midy et Grignon en 2002 ont résumé ces définitions (42) :

- Williams (1970) : « Dans le sens commun, le coût d'un produit est ce que l'on sacrifie pour l'obtenir. (...) Ainsi, le coût d'une intervention médicale est représenté par la valeur (dans la meilleure alternative possible) de toutes les ressources consommées » (43).
- Posnett (1996) : Le coût est égal à la valeur des ressources utilisées par le programme de santé, « Cette valeur étant le reflet du coût d'opportunité pour avoir détourné ces ressources de leur meilleure utilisation possible » (44).
- Drummond (1998) : Le coût réel d'un programme « correspond (...) aux résultats de santé qui auraient été obtenus grâce à un autre programme si les ressources utilisées avaient été affectées à celui-ci plutôt qu'à celui-là » (15).
- Launois (1999) : «Le coût d'un traitement est une somme exprimée en unités monétaires de toutes les répercussions qu'une pathologie donnée peut avoir pour la société » (45).

Ces définitions font référence à la notion de choix et de sacrifice à faire pour un bien plutôt qu'un autre. On peut alors envisager le coût en économie de la santé comme la valorisation de l'ensemble des ressources détruites lors de leur consommation par un programme de santé alors qu'elles auraient pu être utilisées ailleurs, dans le cadre d'une solution plus avantageuse. Nous allons maintenant rentrer plus en détail dans la définition du coût d'opportunité, avant de définir la typologie des coûts.

III.2.1. Coût d'opportunité

Selon un grand nombre d'économistes, le véritable coût d'un bien est son coût d'opportunité, c'est-à-dire ce à quoi on renonce pour l'obtenir. Ce terme est un franglais provenant de sa traduction anglaise « opportunity cost » énoncé la première fois par David Green (46), repris par Davenport (47) et qui se traduirait plutôt par « coût de renoncement » en Français (48). Cette formulation met en avant le fait que le coût ne peut être évalué que par la comparaison entre l'utilisation de ressources engagées pour produire deux options et par les biens auxquels on renonce lorsqu'on procède à un choix entre les deux. C'est le principe de substitution (49) (50). En effet, le coût d'opportunité d'un bien est égal à la valeur générée par l'utilisation la plus rentable des ressources qui ont été engagées pour le produire. Le coût en l'état d'un bien n'existe pas. Le coût d'une décision ou d'un évènement peut être appréhendé par rapport à un scénario de référence. Le coût n'est pas une caractéristique objective d'un bien ou d'un service de santé. Il dépend du point de vue choisi, du contexte ou de la dimension temporelle. Selon Marshall, le raisonnement fondant le coût d'opportunité permet alors de répondre à la problématique de l'économie de la santé qui est de maximiser les résultats de santé dans le cadre de ressources contraintes (50).

Pour mesurer les coûts liés à une activité, il est important d'identifier correctement la combinaison des facteurs permettant d'en assurer sa production.

III.2.2. Typologie des coûts

Les coûts dans le champ de la santé se décomposent principalement en quatre catégories : les coûts directs, indirects, intangibles et informels (20).

III.2.2.1. Coûts directs

Les coûts directs correspondent aux ressources consommées par la prise en charge de la pathologie et de ses effets secondaires. On distingue alors sous l'appellation de coûts directs deux sous entités. Les coûts directs médicaux représentent les dépenses directement liées à la prise en charge de la pathologie étudiée telles que les dépenses associées aux séjours hospitaliers, aux soins ambulatoires, aux consultations et aux médicaments. Les coûts directs non médicaux représentent les dépenses non directement liées à la prise en charge évaluée, mais néanmoins nécessaire à sa mise en œuvre, telles que les frais liés aux transports, aux aides à domicile ou encore aux frais de garde des enfants (42). Les coûts directs non médicaux rassemblent également les coûts dits institutionnels liés à l'enseignement relatif à la santé, à la recherche et aux équipements (45).

III.2.2.2. Coûts indirects

Midy et Grignon ont synthétisé les définitions du collège des économistes de la santé et de Gold pour caractériser les coûts indirects comme la valorisation de l'ensemble des ressources non rémunérées au titre de la prise en charge mais qui sont pourtant rendues indisponibles pour d'autres activités (temps du patient ou de la famille, actes de soins et de ménage gratuit) (42) (51) (52). Les Indemnités Journalières (IJ) en tant que revenu de compensation, sont des coûts indirects comptabilisées lorsque l'on adopte le point de vue de l'AM. Nous reviendrons sur les différents points de vue dans un prochain paragraphe. Les coûts indirects peuvent alors être globalement définis comme étant les coûts en lien avec à la perte de productivité.

III.2.2.3. Coûts intangibles

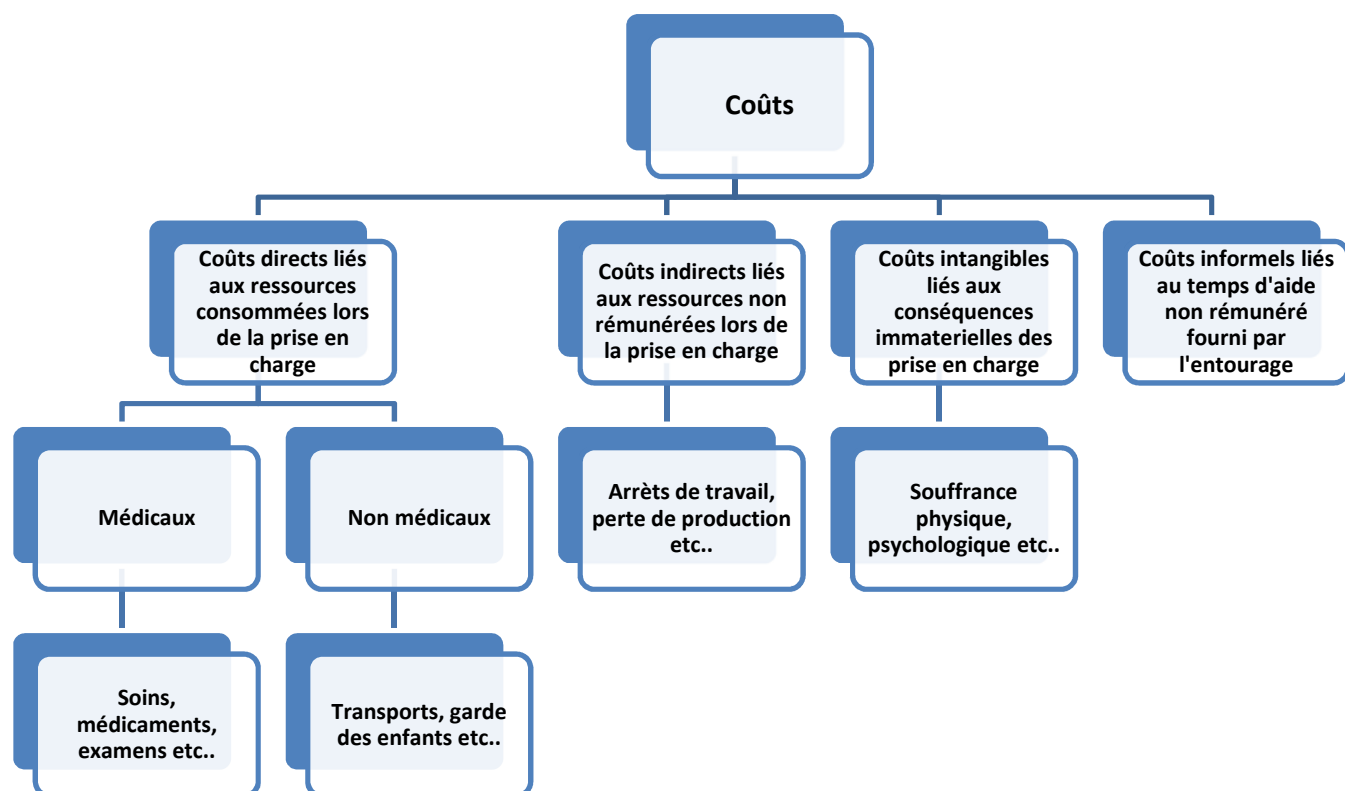
Les coûts intangibles représentent la valorisation des conséquences immatérielles de la maladie sur le patient et son entourage. Ils sont liés au stress, à l'anxiété, à la douleur, et de manière plus générale, à toute perte de bien-être et de qualité de vie vécue par le patient. Les coûts intangibles ne sont pas souvent directement intégrés dans les études économiques. Il est en effet admis de manière consensuelle que pour certaines pathologies, les conséquences immatérielles d'une prise en charge doivent être traitées comme une composante positive de cette prise en charge avec des mesures de qualité de vie plutôt que de les considérer directement dans les coûts.

III.2.2.4. Coûts informels

Les coûts informels représentent la valorisation du temps d'aide prodiguée par l'entourage d'un patient, mais non rémunéré pour cette activité (53). Il est de plus en plus important de les prendre en considération, notamment en raison du vieillissement de la population et de l'augmentation du nombre de personnes dépendantes atteintes de pathologies chroniques. La prise en compte de ces coûts permet de disposer d'informations complémentaires portant sur le fardeau économique supporté par la famille de patients atteints de ces pathologies chroniques (54).

La figure suivante inspirée la thèse de Costa N, résume les différents types de coûts rencontrés dans les évaluations économiques en santé (55).

Figure 2. Typologie des coûts en économie de la santé



La valorisation des coûts identifiés dans ce chapitre est présentée dans la partie suivante.

III.2.3. Valorisation des coûts

Deux approches permettant de valoriser les coûts sont habituellement envisagées. Une première approche est l'approche descendante ou « top-down » qui considère l'ensemble des dépenses de santé d'un pays et ventile ces dépenses selon le nombre de cas traités. La seconde est l'approche ascendante ou « bottom-up » qui estime les coûts associés à un patient traité et l'étend au groupe considéré (56). L'approche « bottom-up » est la méthode préférentiellement utilisée. Elle se fonde sur des données mesurées sur les patients, plus proches de la réalité. Elle permet de prendre en considération les caractéristiques médicales, sociales et démographiques des individus. Toutefois, lorsque les données disponibles sont limitées, l'approche « top-down » peut représenter la seule

méthode permettant une valorisation des coûts. L'approche « top-down » nécessite la connaissance des dépenses liées à la population ciblée totale ainsi que la possibilité de disposer de clefs de ventilation pour répartir ces dépenses au sein de la population étudiée.

La valorisation monétaire des coûts s'effectue en multipliant la quantité d'unités physiques de soins par le prix ou le tarif unitaire correspondant à chaque soin. Nous allons alors détailler dans les paragraphes suivants les prix ou tarifs unitaires les plus souvent utilisés pour valoriser les biens et services de santé en France.

III.2.3.1. Valorisation des coûts directs médicaux

III.2.3.1.1. Valorisation des actes médicaux et paramédicaux

En France, les actes médicaux sont valorisés à partir des tarifs de la Classification Commune des Actes Médicaux (CCAM) de l'AM. Les actes paramédicaux et certains actes médicaux sont valorisés à partir des tarifs de la Nomenclature Générale des Actes Professionnels (NGAP) de l'AM. Ces classifications rassemblent les actes médicaux techniques communs aux secteurs privé et public. La NGAP a été mise en place en 1972 afin de décrire l'activité médicale. Elle a été complétée par la CCAM en 2005. La décision de l'Union Nationale des Caisses d'Assurance Maladie (UNCAM) au moment de la mise en œuvre de la CCAM a été de conserver la NGAP pour les actes cliniques médicaux, des sages-femmes et des auxiliaires médicaux. La CCAM a parallèlement remplacé le Catalogue des Actes Médicaux (CdAM), qui était le versant hospitalier de la NGAP. L'objectif de la création de la CCAM était de permettre un rééquilibrage équitable des honoraires entre spécialités et d'améliorer la répartition des ressources allouées (57).

Les consultations de médecins généralistes ou de spécialistes, les actes de biologies ainsi que les actes paramédicaux sont généralement valorisés au moyen des tarifs appliqués par l'AM. Ils sont disponibles et régulièrement mis à jours sur le site de l'AM (www.ameli.fr). Les actes de biologies

sont de plus associés à un coefficient (i.e. « coefficient B ») issu de la Nomenclature des Actes de Biologie Médicale (NABM) qui permet de fixer le niveau de cotation.

III.2.3.1.2. Valorisation des médicaments et des dispositifs médicaux

En France, la valorisation des médicaments et des Dispositifs Médicaux (DM) est effectuée à partir des tarifs de remboursement fixés par l'AM lorsqu'ils sont admis au remboursement ou des prix unitaires déterminés par le laboratoire qui les commercialise.

Les médicaments admis au remboursement sont disponibles sur les bases des médicaments remboursables par l'AM (58). Dans le cadre d'un séjour hospitalier, un médicament peut être financé dans le tarif du séjour ou être considéré en supplément du séjour, « en sus », et être financé en complément. C'est généralement le cas lorsque le médicament est innovant et/ou très onéreux.

Les DM remboursés par l'AM sont inscrits sur la Liste des Produits et Prestations Remboursables (LPPR) (59). Le financement d'un DM peut également être intégré à un celui d'un acte médical, à un séjour hospitalier ou être financé en supplément de ce séjour (60).

III.2.3.1.3. Valorisation des séjours hospitaliers

III.2.3.1.3.1. Différents types de coûts

La valorisation du coût des séjours hospitaliers demande de faire des distinctions méthodologiques entre les différents types de coût à prendre en compte (61).

La première distinction concerne le coût moyen et le coût marginal. Le coût moyen représente le coût par unité produite (i.e. coût total divisé par le nombre d'unités de production). Le coût marginal constitue le coût supplémentaire induit par la dernière unité produite (62). La distinction entre ces

deux types de coûts est fondée sur le principe que la production d'un bien met en jeu des charges variables et des charges fixes, et que l'évolution de la proportion entre ces charges est importante pour l'interprétation des coûts. Ainsi, les coûts engagés dans le cadre d'un séjour hospitalier ne sont pas constants au cours du temps, et les coûts moyens et les coûts marginaux d'un séjour auront de fortes chances d'être différents. Ces notions de coût moyen et marginal peuvent grandement impacter les résultats des études, notamment celles concernant l'évaluation d'interventions visant à réduire la durée d'un séjour hospitalier.

La deuxième distinction concerne les coûts complets et les coûts partiels. Cette distinction nécessite de définir les notions de charges directes et indirectes. Les charges directes sont celles directement liées à l'activité étudiée. Elles sont imputables à une prise en charge clairement identifiée (consultations, services médicotechniques...). Les charges indirectes sont les charges partagées par plusieurs activités telles que les dépenses d'eau et d'électricité, de ménage ou de maintenance des établissements. Les coûts complets prennent en considération l'ensemble des charges (directes et indirectes) supportées par un établissement de soins tandis que les coûts partiels prennent en compte seulement une partie des coûts complets, qui regroupent le plus souvent des charges directes (62).

Différentes approches permettent d'évaluer les coûts à l'hôpital. Elles utilisent diverses sources de données (comptabilité analytique hospitalière, Etude Nationale des Coûts à méthodologie Commune (ENCC), Programme de Médicalisation des Systèmes d'Information (PMSI) et T2A ...) afin d'évaluer un coût réel ou un coût standard associé à un séjour (63).

III.2.3.1.3.1. Les méthodes d'évaluation des coûts à l'hôpital

Les coûts réels représentent les coûts réellement engagés par une structure pour produire un séjour hospitalier. Ils sont évalués principalement à partir des approches dites de « micro-costing » ou de

« gross-costing » (64). L'approche par « micro-costing » consiste à observer de façon détaillée en temps réel les ressources consommées à chaque étape de la prise en charge d'un patient. Les consommations de ressources sont ensuite valorisées monétairement à partir de diverses sources de données locales (comptabilité analytique hospitalière, système d'information hospitalier, données financières issus des directions et services hospitaliers). L'approche par « micro-costing » se justifie dans tous les cas où il n'existe pas de références de coût, notamment dans le cadre des innovations technologiques ou d'évolution de certaines composantes d'une prise en charge. Le coût réel d'un séjour hospitalier peut également être déterminé par une approche de « gross-costing ». Cette approche consiste à allouer le montant total des coûts supportés par un établissement de santé à un service particulier en utilisant des clés de répartitions prédéterminées. Cette approche est moins chronophage et moins onéreuse. Elle est utilisée par défaut en l'absence de données précises (64).

Les coûts standards sont évalués à partir des informations issues du PMSI, dans le cadre de la Tarification A l'Activité (T2A). Nous reviendrons plus en détail sur le PMSI dans la partie relative aux sources de données économiques (IV.3.2.3.1). La T2A fonde l'allocation des ressources aux établissements de soins publics et privés sur la nature et le volume de leur activité réalisée, mesurée pour l'essentiel à partir des données issues du PMSI. Les tarifs qui rémunèrent les séjours sont calculés sur la base des coûts de référence des Groupes Homogènes de Malades (GHM). Cette classification en GHM fut inspirée de la classification Américaine des « *Diagnosis Related Groups* » (DRG) (65). Chaque GHM est rattaché à un Groupe Homogène de Séjour (GHS) qui est le versant économique du GHM. Les GHS introduisent une notion tarifaire en rapport avec l'utilisation du PMSI pour le financement des établissements de santé dans le cadre de la T2A (66). Le tarif du GHM peut être complété par des suppléments tels que des médicaments onéreux, des dispositifs médicaux, des suppléments de réanimation, de surveillance continue ou de soins intensifs. Les tarifs des GHM sont fixés par le ministère de la Santé à partir de l'ENCC. L'étude Nationale des Coûts (ENC) a été créée en 1993 avec la participation des établissements de soins du secteur public. Elle a ensuite intégré, à partir de 2005, les établissements du secteur privé pour devenir l'ENCC. Les ENCC sont mises en

œuvre par l'Agence Technique de l'Information sur l'Hospitalisation (ATIH). Elles intègrent des établissements de santé publics et privés qui ont une activité dans les champs du MCO, des Soins de Suite et de Réadaptation (SSR) et de l'Hospitalisation A Domicile (HAD). L'ENCC a trois objectifs (66). Le premier est la construction d'une échelle de coûts par GHM, révisée annuellement, et servant de fondement pour les tarifs utilisés afin d'assurer le financement des établissements. L'ENCC publie ensuite les coûts moyens par GHM discrétisés par postes de charge permettant aux établissements de se comparer afin d'optimiser leur gestion interne. Enfin, l'ENCC constitue une base de données pour assurer l'évolution de la classification des GHM.

III.2.3.2. Valorisation des transports

Les transports sont pris en charge par l'AM sur prescription médicale pour des soins ou examens dans les cas décrits sur le site de l'AM (67).

La valorisation des dépenses de transports sanitaires dépend principalement des paramètres suivant : la distance entre le lieu de résidence du patient et son lieu de prise en charge et le type de véhicule utilisé. On retrouve les tarifs conventionnels des transports sanitaires sur le site dédié à l'administration française (68). L'AM applique un taux de remboursement de 65% du tarif de base minoré de la franchise médicale qui est établi à 2€ par trajet et dont le plafond journalier est de 4€. Lorsque le malade est en Affection Longue Durée (ALD), les transports sont remboursés en totalité si le transport est effectué en lien avec la pathologie pour laquelle le patient est en ALD et si le patient présente des incapacités ou des déficiences définies par le référentiel de prescription des transports (69).

Les transports en commun ou avec un véhicule personnel ne sont pris en charge que seulement si le patient remplit les conditions de prise en charge pour une ALD.

III.2.3.3. Valorisation des coûts indirects : liés à la sphère productive

Les approches développées pour valoriser les coûts associés à la perte de productivité sont multiples. Les méthodes les plus utilisées sont présentées ci-dessous.

La Méthode du Capital Humain (MCH) et la Méthode des Coûts de Friction (MCF) cherchent à placer une valeur monétaire relative au temps où la production de l'individu est impactée par la maladie. La MCH évalue les pertes de production en multipliant la quantité de travail perdue en raison de la maladie par le salaire brut de l'individu concerné. La MCF est supposée plus réaliste car elle envisage la possibilité de remplacement de l'individu malade par des collègues ou des personnes au chômage. Ainsi les pertes de production évaluées dans le cadre de de la MCF sont plus faibles car cette approche prend en compte la possible compensation de ces pertes (70). Dans le cadre de la MCF, les coûts indirects sont valorisés à partir de la capacité de l'entreprise à revenir au niveau de production initial et des pertes associées (71).

En France, les IJ sont versées par l'AM au titre des absences sur le lieu de travail en raison de la maladie (72). Un patient perçoit les IJ après un délai de carence de trois jours dans le secteur privé et d'un jour dans le secteur public.

Un salarié peut prétendre à une pension d'invalidité lorsqu'il est en incapacité totale ou partielle de travailler suite à un accident ou d'une maladie d'origine non professionnelle. Les montants des remboursements de ces pensions sont calculés sur la base du salaire annuel moyen des 10 meilleures années d'activité du salarié avec des minorations et majorations selon le salaire de l'individu (73).

III.2.3.4. Valorisation des coûts informels

Plusieurs méthodes de valorisation du temps d'aide apporté par l'entourage ont été développées. La méthode du coût d'opportunité valorise le temps passé de l'aidant par son revenu. Si l'aidant n'exerce pas d'activité rémunérée, le temps valorisé correspondra à une perte du temps de loisir valorisée entre 25 et 33 % du coût d'opportunité d'une personne exerçant une activité rémunérée

(74). La méthode du coût de remplacement ou « proxy good », valorise le temps d'aide informel au prix du marché du travail pour un aidant rémunéré tel qu'une aide-ménagère ou une auxiliaire de vie. Ainsi, l'aide informelle est valorisée selon les tâches effectuées par l'aidant (75).

Ces deux méthodes sont les plus communément utilisées (76).

L'analyse des coûts dépend par construction du point de vue envisagé dans l'étude. Nous envisageons ce point dans le paragraphe suivant.

III.2.4. Point de vue

Le point de vue considéré dans le cadre d'une évaluation des coûts constitue un élément particulièrement important car il définit le périmètre des coûts à prendre en compte dans l'étude. Une étude peut être réalisée du point de vue du patient, des centres de soins (hôpitaux, cliniques, médecins), des financeurs, notamment de l'AM obligatoire, des mutuelles, des assurances privées ou encore de la société (77).

Lorsqu'on se place du point de vue du patient, les coûts pris en compte seront les coûts médicaux restant à leur charge, c'est-à-dire les frais non remboursés par l'AM et les mutuelles. Le coût en lien avec l'aide informelle peut être inclus dans ce point de vue.

Lorsqu'on se place du point de vue des établissements de santé, les coûts considérés représentent les coûts qu'entraîne leur fonctionnement. Ainsi, les coûts considérés sont les coûts directement rattachés à un malade et à sa prise en charge dans le cadre de son séjour à l'hôpital ou les coûts directement liés au fonctionnement d'un service clinique.

Si on adopte le point de vue d'une assurance de santé (i.e. AM, mutuelles...), seules les dépenses remboursées par ces assurances ou mutuelles sont prises en considération.

Le point de vue le plus général est alors celui de la société. Il englobe toutes les charges telles que les dépenses en lien avec les secteurs sanitaires et médicosocial, les secteurs administratifs et de productions des soins ainsi que les coûts non médicaux supportés par les patients et les familles.

La HAS recommande d'adopter une perspective collective dans les EME afin de tenir compte de l'ensemble des parties prenantes concernées par les interventions étudiées dans le contexte du système français de santé (3). Il est néanmoins difficile voire impossible de prendre en compte les dépenses de santé incombant à tous les acteurs. Grignon et Midy parlent même « d'hypocrisie de la perspective sociétale », qu'il est possible d'approcher seulement à l'aide d'analyses macroéconomiques (45).

Les différentes méthodes d'EME décrites dans la partie précédente s'effectuent sur un horizon temporel plus ou moins important. Nous allons voir dans le prochain paragraphe comment prendre en compte cette dimension temporelle.

III.2.5. Actualisation des coûts

Les EME peuvent couvrir une durée importante pouvant aller jusqu'à la vie entière. La comparaison entre les interventions évaluées s'effectue par rapport à un moment précis dans le temps. Les coûts et les conséquences en matière de santé entre les interventions étudiées peuvent survenir à des temps différents (78). Il est également possible de chercher à comparer une intervention considérée comme étant la référence, anciennement évaluée, à une nouvelle intervention évaluée plus récemment. Il devient alors nécessaire de considérer toutes les possibilités relatives à la temporalité des résultats dans les comparaisons entre interventions étudiées.

L'actualisation des coûts rend possible la comparaison des résultats des programmes évalués à des moments différents, en ramenant leurs coûts et résultats futurs à leurs valeurs présentes (3). L'actualisation des résultats permet la prise en compte des préférences des individus pour le présent plutôt que pour le futur. En règle générale, il paraît plus intéressant de bénéficier des avantages d'un programme dans l'immédiat comme il paraît plus intéressant de supporter les coûts le plus tard possible. Pour un même programme, les coûts et conséquences peuvent apparaître à des moments différents. C'est notamment le cas pour un programme de prévention ou de dépistage d'une pathologie, dont les coûts, au moins pour partie, sont consentis tout de suite. Or, les coûts et les avantages futurs ne pèsent pas aussi lourd dans la décision que les coûts et les avantages présents. C'est ce qu'on appelle la préférence pour le présent (79). Il est plus facile de visualiser les coûts et effets de santé d'un programme dans le présent que dans le futur où cela nécessite un effort de projection. Cette notion s'applique aux coûts et aux conséquences en termes de santé.

Le calcul de l'actualisation d'un coût s'effectue à partir de la formule suivante :

$$C = \sum_{i=1}^n \frac{C_i}{(1+r)^i}$$

Avec :

C : Le coût total du produit actualisé au présent

C_i : Le coût du produit pour l'année i

r : Le taux d'actualisation considéré

Le facteur $(1+r)^i$ s'appelle le facteur d'actualisation pour l'année i . L'actualisation est communément appliquée sur une base annuelle. Ainsi, selon le commissariat général du plan en 2005 (80), le taux d'actualisation pour les EME est fixé à 4% en France. Ce taux dépendra néanmoins de l'horizon temporel considéré. Les résultats doivent être actualisés dès lors que l'horizon temporel

de l'étude est supérieur à 12 mois. Le taux d'actualisation décroît à partir de 30 ans, avec une décroissance continue et limitée à 2 %. Les taux d'actualisation varient également selon les pays. Etant données les différences des taux choisis selon les pays qu'ils l'utilisent, il est conseillé de réaliser des analyses de sensibilité relatives à la valeur du taux d'actualisation pris en compte. Il est également utile de présenter les coûts bruts afin de rendre les résultats présentés plus facilement utilisables dans d'autres analyses. Le taux d'actualisation prendra généralement les valeurs de 0, 3 et 5% dans les analyses de sensibilité (13). L'utilisation du même taux d'actualisation pour les coûts et les effets de santé est controversée car pouvant, dans certain cas, modifier la prise de décision (81). Il est néanmoins préconisé par la HAS d'actualiser les coûts et les effets de santé sur la base d'un même taux.

Actualiser les coûts et les résultats d'un programme de santé permet donc de disposer d'un point de référence normalisé et facilite la comparaison entre les études, notamment au niveau international.

Les résultats des évaluations économiques des programmes de santé sont soumis à une grande incertitude en raison de la quantité d'information à exploiter (coûts, conséquences des programmes) et des différents paramètres à prendre en considération (point de vue, actualisation, horizon temporel...). Il est essentiel de prendre en compte cette variabilité et d'en estimer l'impact sur les résultats en mettant en œuvre des analyses de sensibilité.

III.3. Analyses de sensibilité

Les résultats provenant des EME peuvent être soumis à une grande incertitude. Les effets de santé ainsi que les coûts pris en compte sont pour un grand nombre de raisons, incertains. Cette incertitude peut provenir (82):

- Des différentes méthodologies liées aux estimations des données d'entrée, ou à la construction des modèles : les méthodes pour estimer l'efficacité, la qualité de vie, les coûts et les paramètres d'entrée des modèles utilisés,
- Des erreurs d'échantillonnage ou de mesure relatives aux erreurs statistiques,
- Des différences dans les choix méthodologiques structurants : perspective, horizon temporel, taux d'actualisation,
- D'un manque de connaissance : hypothèses ou simplifications dans les données.
- D'une méconnaissance de la forme fonctionnelle d'un modèle, des interactions entre paramètres et résultats et de la validation de son adéquation aux données (i.e. incertitude structurelle)

Selon Briggs, la valeur d'une analyse médico-économique ne réside pas simplement dans sa capacité à générer une estimation ponctuelle¹ pour un résultat spécifique mais se situe aussi dans l'examen systématique de l'incertitude entourant ce résultat afin de prendre une décision finale (83). Les analystes doivent alors identifier les hypothèses critiquables afin d'effectuer des analyses complémentaires en modifiant les valeurs des paramètres découlant de ces hypothèses. Ces analyses complémentaires s'appellent alors des analyses de sensibilité.

Les décideurs ont un rôle de « porteurs de risques » lorsqu'il s'agit de décisions d'investissement public (84). Il est nécessaire d'avoir une mesure de la variabilité des résultats afin de pouvoir transférer ces décisions du niveau populationnel au niveau individuel. L'analyse de sensibilité permet de vérifier que les résultats et conclusions d'une étude sont suffisamment robustes et ne sont pas affectés par des choix discutables pris lors du processus d'analyse. Elle vérifie que les conclusions restent les mêmes lorsqu'on répète les analyses en modifiant les hypothèses posées à priori (85). Lorsque les résultats sont très variables selon les modifications des paramètres principaux, il est

¹ Estimation ponctuelle : lorsqu'on estime un paramètre par une seule mesure ≠ Estimation par intervalle : lorsqu'on estime un paramètre par un intervalle intégrant ainsi la variabilité due notamment à l'échantillonnage. Une estimation ponctuelle n'apporte pas d'information sur la précision de l'estimateur.

nécessaire de réduire cette incertitude en modifiant le champ de l'étude (i.e. résultats qui diffèrent selon les sous populations considérées) ou en améliorant les connaissances des différentes variables stratégiques impactant la variabilité des résultats.

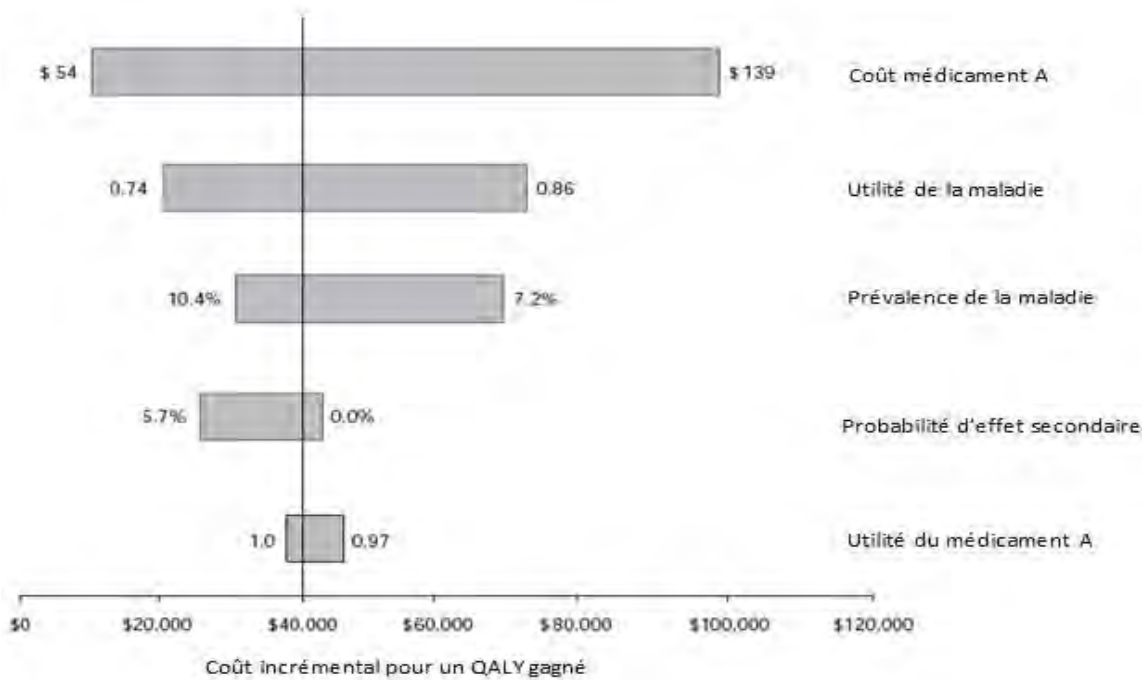
Il existe deux types d'analyse de sensibilité : Les Analyses de Sensibilité Déterministes (ASD) et les Analyses de Sensibilité Probabilistes (ASP).

III.3.1. Analyses de sensibilité déterministes

La mise en œuvre d'une ASD nécessite l'identification des paramètres incertains relatifs aux différentes hypothèses, leurs variations maximales, et dans le cas multivarié, les scénarios les plus extrêmes en combinant les différentes valeurs des paramètres. Ces informations sont donc des valeurs discrètes généralement issues de la littérature, des avis d'experts, d'observations ou de choix arbitraires.

L'ASD univariée consiste à sélectionner et à faire varier une à une les valeurs des paramètres afin d'évaluer leurs impacts sur le résultat final. Les diagrammes de « Tornado » facilitent cette sélection. Ce sont des diagrammes en bâton inversés listant les paramètres de haut en bas en fonction de l'impact de chacun d'eux sur le résultat final. La figure suivante est un exemple de diagramme de Tornado appliqué aux variations de l'ICUR (86).

Figure 3. Exemple de diagramme de Tornado



Le diagramme schématise la façon dont les variations des paramètres affectent le résultat (ICUR). Par exemple, plus le coût du médicament testé est important, plus l'ICUR sera élevé. De même, plus le médicament présente une utilité faible, plus le ratio est grand.

L'ASD multivariée consiste à faire varier les valeurs des paramètres simultanément afin d'en évaluer l'impact global sur les résultats. En combinant les valeurs extrêmes des paramètres sélectionnés, différents scénarios plus ou moins éloignés de l'estimation ponctuelle du résultat sont créés. Un panel de scénarios composé du plus vraisemblable, du plus pessimiste et du plus optimiste est généralement utilisé afin de visualiser l'étendue de la variation des résultats finaux. Il existe également l'analyse de seuil consistant à sélectionner les valeurs critiques des paramètres et déterminer quelles combinaisons de ces paramètres entraînerait un dépassement du seuil.

L'ASD est limitée (87) (88) (89):

- L'analyste choisit lui-même les paramètres rentrant dans l'analyse de sensibilité engendrant un biais de sélection,

- L'ASD ne permet de faire varier que peu de variables simultanément car les analyses deviennent rapidement trop lourdes à mettre en œuvre,
- Il est difficile d'interpréter les choix effectués par l'analyste quant aux différentes variations des paramètres pris en compte sachant qu'il n'existe pas de recommandations particulières, ni de standards,
- En faisant varier les paramètres un après l'autre ou simultanément de manière arbitraire, l'ASD ne prend pas en compte les possibles interactions entre les paramètres,
- Les ASD se basent sur des scénarios mais ne révèlent pas la probabilité d'occurrence de chaque scénario, c'est-à-dire sa vraisemblance.

C'est pour pallier à ces limites qu'il est également nécessaire de réaliser des Analyses de Sensibilité Probabilistes (ASP).

III.3.2. Analyses de sensibilité probabilistes

Les ASP permettent d'explorer la variabilité des conclusions d'une étude selon celle des paramètres d'entrée. En considérant les variations de l'ensemble des paramètres d'entrée de manière concomitante, l'ASP complète l'ASD (12). Elles constituent de nos jours la référence pour étudier la variabilité d'un résultat. Les ASP sont recommandées en Angleterre par le « National Institute for Health and Care Excellence » (NICE) comme en France par la HAS (3) (90).

Les ASP se basent principalement sur le choix des distributions statistiques relatives aux paramètres d'entrée impactant le résultat final. Ces distributions dépendent de la nature des paramètres. Les lois log normales et gamma sont notamment utilisées pour représenter les coûts car ils sont positifs et asymétriques vers la droite. Les utilités des états de santé sont souvent modélisées par une loi bêta, comprise entre 0 et 1 de même que les probabilités de transition entre différents états. Les temps passés dans un état peuvent être représentés par une loi exponentielle ou une loi de Weibull.

Les distributions attribuées aux variables d'entrée du modèle seront ensuite utilisées dans le cadre de simulations paramétriques (i.e. qui se réfèrent à des lois de probabilité définies par leur(s) paramètre(s) (moyenne, variance etc...)) ou non paramétriques. Elles permettent d'obtenir un résultat ponctuel à partir de la moyenne des résultats ainsi qu'une mesure de l'incertitude autour de ce résultat à partir des IC. Ces méthodes utilisant des procédés aléatoires afin d'estimer le résultat de l'étude et ses variations sont les méthodes de Monte-Carlo (87). Elles consistent simplement à réaliser un nombre important de simulations du résultat à partir des distributions des paramètres d'entrée afin de permettre de disposer d'une distribution de résultats en sortie (88).

Il existe cependant des problèmes relatifs aux distributions des résultats évalués. L'ICER par exemple, est la moyenne des quotients de la différence des coûts sur la différence d'efficacité. Les coûts ne sont pas distribués normalement, mais il peut être admis que les moyennes des différences suivent des lois normales si les tailles de population sont assez importantes (91). Néanmoins, sachant que le dénominateur peut être nul ou se rapprocher de 0, il est difficile de justifier par le théorème de la limite centrale que l'ICER suive une loi normale et qu'il est alors aisé d'en estimer son IC à 95% et donc sa variabilité (92).

Un ensemble de méthodes dites paramétriques ou non-paramétriques permettent de passer outre la non normalité des distributions des résultats telles que les coûts ou les ICER. Parmi les méthodes paramétriques, la méthode des ellipses de confiance et la méthode de Fieller posent des hypothèses sur les distributions des coûts et des effets, et permettent de calculer les IC autour des ratios (ICER, ICUR...). Les méthodes non paramétriques découlent de la méthode du bootstrap non paramétrique et sont applicables sur les coûts, les ICER ou tout autre paramètre d'intérêt. Une courbe d'acceptabilité est également régulièrement utilisée comme support graphique pour l'étude de l'incertitude autour de l'ICER en fonction du seuil d'efficacité fixé.

III.3.2.2. Méthode des ellipses de confiance

Van Hout et al. posent l'hypothèse que les coûts et l'efficacité sont corrélés et qu'il faut considérer le couple $(\Delta C, \Delta E)$ afin de prendre en compte cette corrélation (93). Soient respectivement $m_{C1}, m_{C2}, \sigma_{C1}^2, \sigma_{C2}^2, m_{E1}, m_{E2}, \sigma_{E1}^2, \sigma_{E2}^2$, les moyennes et les variances des coûts et de l'efficacité dans les deux interventions comparées. Van Hout et al. supposent que les différentiels de coûts et d'efficacité suivent des lois normales telles que :

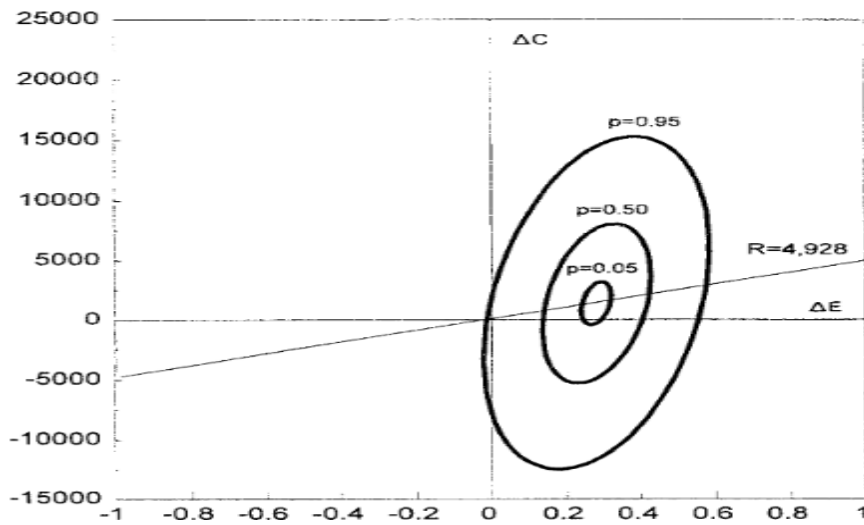
$$\Delta C \sim N(m_{C2} - m_{C1}, \sqrt{\sigma_{C2}^2 + \sigma_{C1}^2})$$

$$\Delta E \sim N(m_{E2} - m_{E1}, \sqrt{\sigma_{E2}^2 + \sigma_{E1}^2})$$

Et on peut considérer que le couple $(\Delta C, \Delta E)$ suit une loi normale bivariée.

La densité jointe des différentiels de coûts et d'efficacité peut être représentée graphiquement. Au même titre que la forme connue de cloche de la densité de la loi normale à une dimension, qui permet d'obtenir, à un niveau de risque donné, un IC en une dimension, sa forme en 2 dimensions permet alors d'obtenir, à un niveau de risque donné, une zone de confiance sous forme d'une ellipse appelée ellipse de confiance. En l'intégrant au plan coût efficacité, on obtient une représentation graphique facile d'interprétation. Ci-dessous, on peut trouver un exemple tiré de l'article de Van Hout et al. (93).

Figure 4. Ellipse de confiance au niveau $\alpha=5, 50, 95\%$ (i.e, $p=1- \alpha$)



R représente la valeur de l'ICER estimé. Il est au centre des ellipses. Ainsi, si on simule 100 valeurs de l'ICER, 95% des valeurs se trouveront dans l'ellipse telle que $p = 1 - \alpha = 0.95$. Une approximation de l'IC à 95% de l'ICER est alors donnée par les droites partant de l'origine et tangentes aux extrémités de l'ellipse.

Cette méthode ne pose pas d'hypothèse sur la distribution du ratio. Elle permet d'intégrer les corrélations entre les différentiels de coût et d'efficacité. Elle donne une bonne idée de l'incertitude autour de l'ICER et est souvent retrouvée dans la littérature.

III.3.2.3. Méthode de Fieller

Cette méthode découle du théorème de Fieller, permettant de calculer des IC pour des ratios (94). C'est une méthode paramétrique permettant de prendre en compte la corrélation entre les coûts et l'efficacité. Elle a été adaptée au calcul de l'IC de l'ICER par William et O'Brien et par Chaudhary et Stearns (95) (96).

Soit R , l'ICER construit à partir des ΔC et ΔE , les différences de coûts et d'efficacité.

On a alors :

$$R = \frac{\Delta C}{\Delta E}, \text{ ou } R \times \Delta E - \Delta C = 0$$

Avec une variance connue telle que :

$$\text{var}(R \times \Delta E - \Delta C) = R^2 \text{var}(\Delta E) + \text{var}(\Delta C) - 2R \text{cov}(\Delta E, \Delta C)$$

La méthode de Fieller suppose que le numérateur et le dénominateur du ratio estimé suivent une loi normale bivariée. Il est alors possible de générer une loi normale de moyenne 0 et de variance 1 :

$$\frac{R \times \Delta E - \Delta C}{\sqrt{R^2 \times \text{var}(\Delta E) + \text{var}(\Delta C) - 2R \times \text{cov}(\Delta E, \Delta C)}} \sim N(0,1)$$

Il est ensuite possible de calculer l'IC à 95% autour du R estimé. Cette méthode nécessite de connaître la valeur du R, représentant le maximum qu'un financeur est prêt à payer pour gagner une unité d'efficacité.

III.3.2.4. Méthodes de bootstrap non paramétriques

La méthode dite de « bootstrap » est une méthode de ré-échantillonnage développée par Efron (97) et utilisée par Mullahy et O'Brien pour calculer les bornes de l'IC autour de l'ICER (98) (99).

La méthode de bootstrap consiste à ré-échantillonner la population d'étude un grand nombre de fois et à calculer pour chaque réplique la statistique d'intérêt afin d'en obtenir une distribution. Cette méthode est utilisée lorsque la distribution de la statistique évaluée dans la population étudiée est inconnue. Elle permet d'approcher par simulation, la distribution réelle de la statistique mesurée. Elle substitue la distribution inconnue de la statistique étudiée par une distribution empirique issue d'un échantillon de la population étudiée (i.e. échantillon d'apprentissage). On réalise n tirages de sous échantillons de la population initiale avec remise afin de calculer n fois la statistique d'intérêt. La distribution empirique donne alors un poids $1/n$ à chacune des réalisations. La loi simulée de la statistique évaluée est une approximation asymptotiquement convergente de la loi initiale de la

statistique. Elle apporte des mesures d'estimation du biais, de la variance et de l'IC de l'estimateur sans avoir besoin de faire des hypothèses sur sa distribution.

Pour résumer, soit θ la statistique étudiée (par exemple, l'ICER) dans la population générale. Soit $\hat{\theta}$ l'estimation empirique de la statistique θ dans la population d'étude. $\hat{\theta}$ et θ sont dans notre cas toutes les deux inconnues. Si la distribution de $\hat{\theta}$ était connue ou approximable par une loi normale, $\hat{\theta}$ suffirait à caractériser θ . La méthode du bootstrap permet alors, à partir d'un ré-échantillonnage de la population d'étude et du calcul à chaque itération de la statistique étudiée, d'obtenir $\hat{\theta}^*$, la statistique estimée par bootstrap de $\hat{\theta}$. $\hat{\theta}^*$ se comporte alors comme $\hat{\theta}$ qui se comporte comme θ . La connaissance de $\hat{\theta}^*$ (i.e. sa distribution, sa moyenne, sa variance..) renseigne alors sur $\hat{\theta}$ et sur θ . Ce principe s'appelle le principe du « plug-in ».

Comme explicitée précédemment, la distribution des coûts et la distribution du différentiel de coût ne sont pas normalement distribuées et la distribution des ratios tels que l'ICER ou l'ICUR n'appartient à aucune loi connue. Nous considérons ici la méthode de bootstrap non paramétrique n'impliquant pas d'hypothèses sur les distributions des différentiels de coûts et d'efficacité ou du ratio entre les deux. Il existe également le bootstrap paramétrique qui implique des différentiels de coûts et d'efficacité suivant une loi normale bivariée et des simulations suivant ces lois pour approcher la statistique d'intérêt. La méthodologie est donc quasiment identique.

Soit l'ICER notre variable d'intérêt. Les méthodes présentées ci-dessous seraient identiques si l'on changeait de variable d'intérêt et que l'on considérait seulement des coûts ou un critère d'efficacité. L'ICER, de par sa construction, est constitué de 4 paramètres provenant de deux échantillons. Lors des itérations du bootstrap, c'est-à-dire lors du ré-échantillonnage, il est nécessaire de considérer le calcul de ces paramètres deux à deux avant de calculer les ICER bootstrapés.

Ainsi, soient C_1, E_1 le coût et l'efficacité moyens du traitement 1 calculés sur l'échantillon 1 de taille n_1 , et C_2, E_2 le coût et l'efficacité moyens du traitement 2 calculés sur l'échantillon 2 de taille n_2 . On a alors : $R = \frac{C_2 - C_1}{E_2 - E_1}$.

La méthodologie est la suivante :

- 1) Tirer aléatoirement avec remise un sous échantillon de taille n_1 sur l'échantillon 1 et calculer C_1^*, E_1^* , le coût et l'efficacité moyens de l'échantillon obtenu,
- 2) Tirer aléatoirement avec remise un sous échantillon de taille n_2 de l'échantillon 2 et calculer C_2^*, E_2^* , le coût et l'efficacité moyens de l'échantillon obtenu,
- 3) Construire l'ICER bootstrapé $R_i^* = \frac{C_2^* - C_1^*}{E_2^* - E_1^*}$ de l'i-ième itération,
- 4) Conserver la valeur de R_i^* et recommencer à l'étape 1 en tirant des échantillons avec remise. Par convention, il faut réitérer les étapes un minimum de 1000 fois (97).

Les R_i^* représentent ainsi la distribution de l'ICER bootstrapé. On notera R^* le vecteur de l'ensemble des R_i^* . Différentes méthodes permettent ainsi de construire des IC (100).

- La méthode d'approximation par loi normale consiste à considérer que la distribution de l'ICER bootstrapé suit une loi normale de moyenne égale à la moyenne de R^* (i.e. $\overline{R^*}$) et de variance égale à la variance de R^* . L'IC est ensuite calculé à partir de la moyenne des R^* , son erreur standard et les quantiles de la loi normale selon la formule suivante :

$$IC(\overline{R^*}) = \left(\overline{R^*} - z_{\frac{\alpha}{2}} \times se(R^*); \overline{R^*} + z_{\frac{\alpha}{2}} \times se(R^*) \right)$$

Cette méthode peut induire des erreurs en cas de non normalité de la distribution. Il n'est pas conseillé de l'utiliser à moins d'avoir réalisé un test de normalité au préalable².

- La méthode des percentiles est une méthode passant outre la non normalité de la distribution de R^* . Elle consiste simplement à prendre le $100 * \left(\frac{\alpha}{2}\right)$ ième % et le $100 * (1 -$

² Test de normalité: test de Shapiro-Wilk (loi normale) ou de Kolmogorov-Smirnov (toutes lois)

$\frac{\alpha}{2}$) ieme % des percentiles de la distribution des ICER bootstrapés et ordonnées de façon croissantes, pour respectivement la borne inférieure et supérieure de l'IC. Par exemple, pour 1000 itérations, on classe les ICER bootstrapés du plus petit au plus grand, et on sélectionne la 25ieme valeur et la 975ieme pour construire les bornes de l'IC de \overline{R}^* à un niveau de risque α de 5%. Cette méthode prend en compte la possible asymétrie des estimations bootstrapées mais si la distribution des R_i^* n'est pas symétrique l'erreur de couverture³ peut être tout de même importante.

- La méthode des biais corrigés ou « bias-correction » a été créée afin de prendre en compte l'erreur de couverture évoquée ci-dessus (101). Cette méthode corrige cette erreur à partir d'un facteur calculé selon la proportion d'estimations bootstrapées (i.e. proportion des R_i^*) inférieure au R observé à l'origine (échantillon initial). Ensuite, la méthode des percentiles est implémentée mais les percentiles utilisés seront modifiés en fonction du facteur trouvé. Ainsi, cette méthode s'appelle aussi « centring correction » car les percentiles utilisés seront centrés sur l'estimateur corrigé. Cette méthode apporte une meilleure appréciation de l'IC autour de la mesure d'intérêt.
- La méthode des biais corrigés et accélérés, « Bias Corrected and Accelerated » (BCA), considère l'erreur de couverture de la même façon que ci-dessus mais tient également compte de l'asymétrie de la distribution du R^* (101). Pour cela, cette méthode considère un facteur d'accélération qui ajuste l'asymétrie de la distribution avant d'utiliser les percentiles tels les méthodes précédentes. Elle permet une meilleure gestion de l'erreur de couverture dans les échantillons bootstrapés mais nécessite des calculs plus complexes.

Dans le logiciel de statistiques R, le package « boot » permet de calculer les IC à l'aide de la fonction « *boot.ci* » selon les méthodes explicitées ci-dessus (102).

³ Erreur de couverture : Erreur survenant lorsqu'il y a omission ou répétition d'une valeur aberrante dans l'échantillon de départ.

III.3.2.5. Comparaison des méthodes

Nous avons ainsi vu tout un arsenal de méthodes afin d'estimer l'incertitude autour d'une statistique d'intérêt de manière probabiliste. Des méthodes qui posent des hypothèses sur les distributions des coûts et des effets moyens et les méthodes de bootstrap non paramétriques.

Selon Polsky et al., la méthode de Fieller et la méthode du bootstrap non paramétrique des percentiles sont les plus précises pour estimer l'IC à 95% autour de l'ICER et apportent la meilleure approximation de ses variations (103). Briggs et al. comparent également les IC obtenus à partir de ces deux méthodes et concluent à l'équivalence de leurs résultats (83). Ils préconisent néanmoins d'utiliser la méthode tirée du théorème de Fieller pour deux raisons. Ils argumentent sur la supériorité des méthodes paramétriques comparées aux méthodes non paramétriques et sur le fait que l'IC provenant de la méthode du bootstrap, étant fondé sur des simulations, ne donne pas toujours le même résultat. Siani et al. mettent en avant l'instabilité des méthodes bootstrap lorsque la variable d'intérêt est un ratio et que le dénominateur est proche de 0 (104). Il peut en effet arriver que des dénominateurs très faibles fassent « exploser » la valeur de l'ICER conduisant à des valeurs aberrantes.

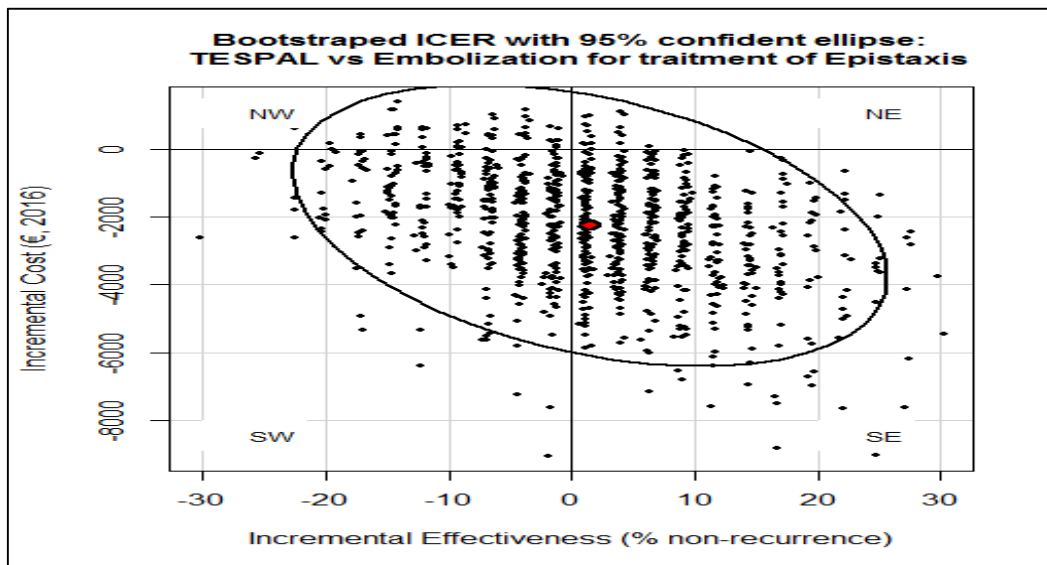
Dans leurs comparaisons, Polsky et al. et Briggs et al. prennent seulement en compte la méthode de bootstrap par percentile pour estimer l'IC à 95% de l'ICER alors que nous avons vu des méthodes plus poussées permettant de prendre en compte les biais dus à l'erreur de couverture et la forme de l'asymétrie de la distribution de la statistique ré-échantillonnée. Les deux raisons évoquées par Briggs et al. préconisant l'usage de la méthode de Fieller, sont de plus secondaires. Ils énoncent tout d'abord que les méthodes paramétriques sont supérieures aux méthodes non paramétriques alors que le choix entre les méthodes dépend plutôt des hypothèses, des objectifs, des risques et des paradigmes en lien avec l'étude et l'analyste (105). Les méthodes paramétriques imposent ainsi un grand nombre d'hypothèses selon les tailles d'échantillons et leur distribution alors que les méthodes non paramétriques permettent d'avoir une approche plus générale en passant outre les fréquentes

erreurs de spécifications des paramètres des modèles (i.e. coûts non gaussiens). De plus, les deux méthodes permettent d'obtenir une estimation de l'incertitude autour de la statistique d'intérêt à partir de l'IC. Le principe même d'une estimation est alors d'approcher la réalité mais pas de l'égaliser car inconnue. Le fait que l'approche bootstrap par simulation ne donne pas exactement les mêmes bornes estimées pour chaque réplification de la méthode n'est pas un réel problème mais plutôt une preuve de l'incertitude autour de la variable d'intérêt. Il est de plus possible de fixer les « graines » (i.e. la base) du générateur aléatoire associé au logiciel utilisé afin d'obtenir le même résultat lorsque l'on ré-implémente un code. Nous soulignons toutefois l'importance du travail de Siani et al. qui insistent sur la dangerosité d'utiliser du bootstrap lorsque l'efficacité différentielle des traitements comparés est proche de 0 (104).

Dans le cadre des analyses menées dans les études intégrées à notre travail de thèse, nous avons choisi de mettre en œuvre la méthode du BCA-bootstrap afin d'estimer les IC à 95% des ratios d'efficacité et des coûts. Cette méthode est facile à implémenter et permet de visualiser la distribution des mesures simulées. De plus, dans le cadre de l'ICER ou de l'ICUR une ellipse de confiance à 95% peut alors être utilisée en complément à la méthode de bootstrap afin de procurer une interface graphique facile d'interprétation afin que le lecteur/décideur ait une vision claire des variations du résultat final.

Un exemple de sortie graphique souvent retrouvée dans la littérature est représenté dans la figure suivante (106).

Figure 5 Exemple d'ICER bootstrapés avec l'ellipse de confiance.



L'objectif de l'étude associé à la figure ci-dessus était d'évaluer, chez des patients atteints d'épistaxis, l'efficacité de la ligature par rapport à l'embolisation. Afin d'évaluer l'ICER associé, le critère d'efficacité utilisé était la proportion de patient récidivant et les coûts pris en compte étaient les coûts directs médicaux limités aux séjours hospitaliers. On retrouve ainsi en rouge, l'ICER observé et en noir, l'ensemble des ICER simulés par bootstrap. Il y a ici 1000 répliquions bootstrap et donc 1000 points noirs. L'ellipse de confiance à 95% dessinée entoure 95% des ICER bootstrapés et permet d'avoir une vision graphique des possibles variations de l'ICER. Lorsque les différentiels d'efficacité sont faibles (i.e. $\sim 1/100$), la méthode de Fieller est néanmoins à utiliser.

Nous allons maintenant présenter un outil applicable à l'ICER : la courbe d'acceptabilité.

III.3.3. Courbe d'acceptabilité

La courbe d'acceptabilité ou « Cost-Effectiveness Acceptability Curve » (CEAC), est une courbe représentant l'évolution de la probabilité qu'un programme évalué demeure efficace par rapport à un programme alternatif en fonction du seuil d'efficacité choisi. Comme son nom l'indique cette

courbe s'utilise au cours d'une analyse d'efficience. Elle illustre graphiquement la problématique du choix d'une intervention par rapport à une autre en fonction du seuil d'acceptabilité retenu. Ainsi, afin d'obtenir les probabilités que le programme évalué soit coût-efficace par rapport aux différents seuils d'acceptabilités choisis, il est nécessaire d'utiliser plusieurs valeurs calculées de l'ICER. C'est un autre avantage de la méthode du bootstrap qui permet de calculer facilement ces probabilités. Ainsi pour construire la courbe d'acceptabilité, la méthodologie est la suivante :

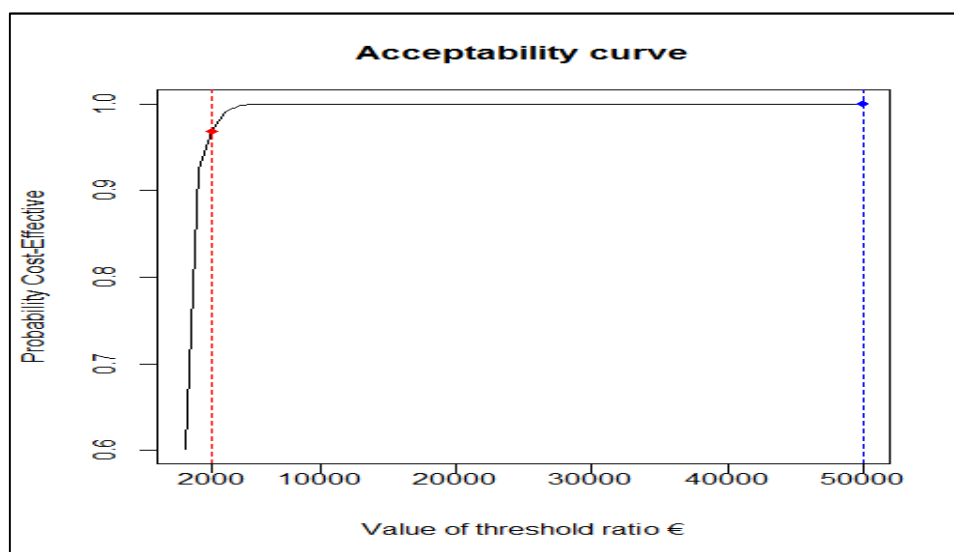
- 1) Réaliser les répliques bootstrap de l'ICER comme décrit précédemment,
- 2) Identifier les différentes valeurs du seuil d'acceptabilité à considérer,

Soient R_i^* , les ICER bootstrapés, et λ_j les différentes valeurs du seuil d'acceptabilité.

- 3) Pour chaque λ_j , évaluer le pourcentage de R_i^* tel que $R_i^* < \lambda_j$ (i.e. la probabilité que l'ICER soit coût-efficace),
- 4) Placer les différents pourcentages en ordonnée, et les valeurs des λ_j en abscisse.

En reprenant l'article de Costa et al., avec les données utilisées pour la représentation des ICER bootstrapés, on obtient (106):

Figure 6. Courbe d'acceptabilité



Les probabilités, comprises entre 0 et 1 (en ordonnée), évoluent en fonction de la valeur du seuil d'acceptabilité (en abscisse). Ainsi plus le seuil est grand, plus les ICER ont une probabilité d'y être inférieurs et la probabilité que la stratégie soit coût-efficace augmente. On a en rouge la valeur minimale du seuil choisis et en bleu sa valeur maximale. Le choix des variations du seuil d'acceptabilité est arbitraire et dépend de la valeur de l'ICER estimé et des normes utilisées pour qu'un décideur accepte de financer un programme. Ainsi, dans cette étude, le programme évalué a plus de 95% de chance d'être coût efficace par rapport à une disposition à payer du financeur de 2000€.

Nous avons vu l'importance de réaliser des analyses de sensibilité dans les évaluations économiques afin d'obtenir des mesures sur la variabilité des résultats. Cette étude de l'incertitude est indispensable dans les analyses descriptives ainsi que dans le processus conduisant à une prise de décision portant sur différents programmes évalués. Les ASD n'apportent qu'une information incomplète sur les variations des résultats. Dans ce cas, il est alors primordial de réaliser en complément des ASP prenant en compte les corrélations entre les paramètres incertains considérés simultanément. Dans ce type d'analyse, la méthode de Fieller, notamment si le dénominateur de l'ICER est proche de 0, et la méthode du bootstrap non paramétrique BCA, « bias corrected, accelerated », sont les plus performantes. La méthode du bootstrap permet de plus d'obtenir des interfaces graphiques faciles d'interprétations, à partir des répliques, notamment pour l'ICER (i.e. ellipse de confiance et courbe d'acceptabilité). Cette méthode fournit aux décideurs une assurance supplémentaire sur la robustesse des conclusions.

Les évaluations des coûts des programmes de santé présentent comme caractéristique commune la nécessité d'accéder à des données portant sur les ressources consommées lors des prises en charge. Dans ce cadre, différentes sources de données peuvent être mobilisées. Les données utilisées

dépendent principalement des objectifs de l'étude, de la méthode utilisée pour le calcul des coûts (top-down ou bottom-up) et des possibilités d'accès aux données.

IV. Sources de données économiques

Les objectifs définis dans le cadre d'une étude économique permettent de délimiter les catégories de coûts à prendre en considération et les données à recueillir. Les données peuvent être recueillies selon deux approches, l'approche bottom-up ou l'approche top-down.

L'approche ascendante (« bottom-up »), se distingue de l'approche descendante (« top-down ») notamment par le niveau d'agrégation des données nécessaires pour répondre à l'objectif de l'étude. L'approche ascendante combine des données économiques pour chaque patient inclus afin d'estimer le coût de la pathologie analysée. Elle nécessite l'accès à des données individuelles sur la population étudiée. L'approche descendante considère les dépenses de santé d'une entité (pathologie, centre de soins, pays...) et alloue ces dépenses selon le nombre de cas traités. Elle mobilise alors des données agrégées (56).

Il est également possible de dissocier les approches par prévalence et celles par incidence décrites dans la partie III.1.1. Dans le cadre de l'évolution économique d'une pathologie, l'approche par incidence utilise préférentiellement une approche ascendante nécessitant des données individuelles (107).

Ainsi, deux types de données peuvent être distingués: les données identifiées par individu (i.e. données individuelles) et les données agrégées. Chacune d'entre elles apporte des informations complémentaires et s'utilisent dans des contextes différents. Les données agrégées sont par nature des données correspondant à une population ciblée qui ont fait l'objet d'un traitement statistique. Elles sont généralement constituées à partir d'un ensemble de données individuelles et résultent de la combinaison de différentes mesures. Elles peuvent provenir d'études publiées dans la littérature

ou d'organismes, comme les organismes d'AM, qui produisent des données de consommations de soins de manière agrégée qui sont rendues accessibles au grand public dans le cadre des « open data » (108). Elles permettent de disposer d'informations portant sur des groupes d'individus ayant une ou plusieurs caractéristiques communes, mais ne permettent pas de réaliser une description détaillée du coût d'une pathologie et de ses variations en fonction de différentes caractéristiques individuelles. Néanmoins, elles permettent d'étudier l'offre et la demande de soins en fonction des secteurs d'activité, par période et/ou par zones géographiques. Elles peuvent également permettre d'alimenter des modèles statistiques dont l'objectif est d'extrapoler les conclusions d'une analyse sur une plus grande période et sur l'ensemble d'une population.

Les données individuelles sont issues d'observations qui n'ont pas été statistiquement traitées. Elles peuvent provenir d'enquêtes, de recherches, de sondages, de recensements, etc... Elles permettent une description fine des coûts et des variations associées aux caractéristiques des individus. En effet, la disponibilité d'informations individuelles portant sur les états de santé, les comportements, les consommations de soins, les facteurs de risque et sur les caractéristiques sociodémographiques constituent un atout majeur pour évaluer les actions dans les champs de la qualité et la sécurité des soins, l'organisation et l'efficience du système de santé, et plus globalement les politiques de santé publique (109). Dans ce contexte, ce chapitre présente les sources de données permettant l'accès à des informations de nature économique portant sur les consommations de soins individuelles des patients. Elles peuvent être issues d'études épidémiologiques, des registres ou des bases de données médico-administratives.

IV.1. Les études épidémiologiques et cliniques

Dans le domaine de la santé, différents schémas d'études existent. Nous allons dans un premier temps les décrire et présenter leurs forces et limites, notamment dans le cadre des évaluations

économiques. Dans un second temps, nous allons présenter les méthodes de recueil d'informations médicales et économiques généralement utilisés dans ces études.

IV.1.1. Les différents types d'étude

Les études épidémiologiques et cliniques sont des études mises en œuvre de façon ponctuelle afin de répondre à une question conjoncturelle (110). Elles se divisent principalement en deux types d'études: les études observationnelles et les études interventionnelles (111). Les études observationnelles regroupent les études descriptives et les études analytiques, et les études interventionnelles correspondent aux études quasi-expérimentales et aux études expérimentales.

IV.1.1.1. Les études observationnelles

L'objet des études observationnelles ne porte pas sur des expérimentations et l'investigateur n'a pas de contrôle sur l'attribution des expositions ou des facteurs étudiés (112). Les données issues de ces études sont ainsi le reflet des pratiques courantes.

IV.1.1.1.1. Les études descriptives

Les études descriptives sont mises en œuvre afin de recueillir des informations sur les fréquences d'apparition de phénomènes de santé et de mesurer leurs variations (112). Il n'y a pas de groupe de comparaison dans ces études. On distingue principalement parmi ces études, les études de série de cas et les études transversales.

Les études de série de cas recensent les caractéristiques des patients et les éléments cliniques de la maladie dont ils sont atteints. Elles permettent d'obtenir des informations afin de formaliser une hypothèse sur un phénomène de santé, l'effet d'une exposition ou d'un facteur de risque,

particulièrement lorsque peu de connaissances sont disponibles (i.e. maladie rare) ou lorsque un phénomène inhabituel se produit (i.e. nouvelle maladie) (113). Elles sont ainsi indispensables en pharmacovigilance (114). Ces études sont généralement réalisées sur un faible nombre de patients.

Les études transversales sélectionnent des individus représentatifs de la population d'intérêt à un moment donné afin de recueillir des informations sur la fréquence des cas présents, c'est-à-dire sur la prévalence de la maladie. Elles peuvent être considérées comme une photo instantanée de la situation épidémiologique relative à un phénomène en santé (110). De par la nature figée du recueil de données, ces études ne permettent pas de connaître l'incidence de la pathologie, d'étudier les aspects dynamiques des phénomènes observés ou de déduire des relations de causalité entre les événements observés (115). Ce type d'étude peut être analytique lorsque différents groupes d'exposition sont identifiés et comparés, mais de par la difficulté d'identifier des relations causales, elles sont rarement utilisées dans ce contexte.

Les études descriptives sont rapides et peu coûteuses à mettre en place (113) (115). Ce type d'enquête est peu utilisé dans le cadre des évaluations économiques car elles rencontrent différents biais. La population incluse n'est pas toujours représentative de la population cible engendrant des biais de sélection important. Il est de plus difficile d'identifier des liens causaux entre les variables, entravant les possibilités d'étude des variations des coûts. Dans ce cadre, les études analytiques apportent un cadre d'analyse plus poussé. Elles permettent d'analyser les relations entre les expositions et la maladie et ainsi d'étudier les variations de coûts associées.

IV.1.1.1.2. Les études analytiques

Les études analytiques cherchent à établir une relation spécifique entre la survenue d'un événement et une exposition (110). Ces études reposent sur la comparaison entre des groupes de patients suivis longitudinalement. Elles mettent en relation soit l'incidence de la maladie chez des sujets exposés et

non exposés, soit la fréquence de l'exposition chez des patients malades et sains. Elles sont principalement représentées par les études de cohorte et les études cas témoins.

Les études de cohortes mobilisent des sujets suivis dans le temps et regroupés en fonction de leur statut en lien avec l'exposition à des facteurs de risque (111). Elles permettent d'étudier les causes des maladies et l'évolution de cette maladie selon les différentes expositions. On distingue alors les cohortes en population qui visent à identifier les causes d'une maladie à partir d'une population saine, et les cohortes de malades qui s'attachent à étudier l'évolution d'une population malade selon les prises en charge effectuées (116). Les données issues d'études de cohortes, notamment de malades, peuvent être utilisées dans le cadre des évaluations économiques si les données ont été recueillies à priori (117).

Les études cas témoins comparent la fréquence d'une exposition chez des patients malades et chez des sujets indemnes de la maladie étudiée afin de mettre en évidence une association entre l'exposition et la maladie (118). Dans le cadre des évaluations économiques, la présence d'individus témoins permet notamment le calcul d'un coût différentiel entre les sujets sains et les sujets malades. Ce résultat représente le coût directement imputable à la maladie qui serait économisé dans l'éventualité où elle viendrait à disparaître. Ce type d'étude peut ainsi mettre à disposition des données utiles dans le cadre des études de type COI (119).

Dans le cadre des évaluations économiques, ces études ont pour avantage d'être basées sur l'observation de pratiques réelles qui permet l'obtention de résultats relevant de la vraie vie (120). Ces résultats peuvent alors guider les choix des décideurs dans le cadre de leurs décisions publiques (7) (8). Etant donné leur nature comparative, les données issues de ces études sont appropriées à la réalisation des évaluations économiques. Néanmoins, de par l'hétérogénéité des populations sélectionnées, des biais importants peuvent être présents (110). Une attention particulière aux caractéristiques des individus est nécessaire avant de généraliser les conclusions. Dans ce contexte,

les études interventionnelles apportent un cadre d'analyse comparatif rigoureux et une meilleure validité interne.

IV.1.1.2. Les études interventionnelles

L'objectif des études interventionnelles est d'évaluer l'effet d'une intervention sur un groupe de patient (112).

IV.1.1.2.1. Les études quasi-expérimentales

Les études quasi-expérimentales sont généralement utilisées lorsque la randomisation entre les groupes comparés est difficile ou impossible (121). Elles sont représentées par les études avant-après, ici-ailleurs et les séries chronologiques.

Les études avant-après mobilisent des données recueillies avant et après une intervention pour un groupe ayant reçu l'intervention et pour un groupe contrôle ne l'ayant pas reçu. Les études ici-ailleurs utilisent des données recueillies sur deux groupes de patients identiques à l'exception faite de l'intervention étudiée et des zones géographiques. La principale difficulté dans le cadre de ces études est d'obtenir une population témoin comparable en tout point à la population d'intervention (122). De plus, les informations recueillies au cours des deux périodes ou entre les différentes zones géographiques peuvent ne pas être comparables en termes de qualité et d'exhaustivité. Il est également possible que les pratiques autres que celle étudiée évoluent durant les périodes comparées (123).

Les études fondées sur l'approche des séries chronologiques se basent sur des observations répétées réalisées avant et après l'intervention étudiée. Il n'y a pas de groupe contrôle dans ces études et seule la tendance évolutive des indicateurs mesurés est comparée entre les périodes pré

intervention, intervention, et post intervention (122). Ces études permettent de s'affranchir des tendances ou des variations temporelles mais elles n'apportent aucune preuve de causalité.

Ces études sont rapides et peu coûteuses à mettre en œuvre. Néanmoins, sans randomisation, les résultats de ces études comportent des biais de sélection ou de confusion qui peuvent être importants (122).

IV.1.1.2.2. Les études expérimentales

Les études expérimentales sont regroupées sous le terme d'essais cliniques définis par l'Agence National de Sécurité du Médicaments (ANSM) comme des « recherches biomédicales organisées et pratiquées sur l'homme en vue du développement des connaissances biologiques ou médicales » (124). Les essais cliniques sont principalement représentés par les ECR, où les individus ont été répartis aléatoirement dans les différents bras de traitements. La randomisation permet d'obtenir des groupes comparables et de limiter au maximum les éléments pouvant introduire un biais afin d'évaluer exclusivement l'efficacité de la stratégie de santé étudiée. De plus, les patients sont sélectionnés sur des critères d'inclusion stricts et font l'objet d'une surveillance accrue. Les ECR évaluent ainsi l'impact d'une stratégie de santé dans les conditions optimales de comparaison et fournissent des résultats essentiels à son introduction sur le marché. Ils apportent un des meilleurs niveaux de preuve en médecine en privilégiant la validité interne à leur validité externe (125). La validité interne est l'indicateur permettant l'évaluation de la fiabilité des conclusions relatives aux données utilisées tandis que la validité externe est l'indicateur permettant l'évaluation de la généralisation des conclusions à l'ensemble de la population (126).

Les ECR présentent certaines limites, notamment dans le cadre des évaluations économiques (125) (127) (128). Ces études se focalisent généralement sur une prise en charge particulière et le périmètre du recueil d'informations, notamment économiques, peut être limité. Les patients inclus

dans les ECR sont sélectionnés selon des critères d'inclusion stricts, écartant généralement les individus aux caractéristiques particulières (i.e. femmes enceintes, individus âgés...). Cette méthodologie fait la force des ECR pour étudier l'effet d'un traitement mais constitue leur principale faiblesse dans le cadre de décision publique : la rigueur de leur schéma expérimental devient leur principale limite pour établir des conclusions dans des conditions de vie réelle (8). Etant donnée la surveillance accrue des patients, les effets secondaires sont détectés plus rapidement et/ou pris en charge de façon très précoce et les patients présentent une meilleure observance, pouvant fausser la généralisation des résultats. Cette difficulté à généraliser de telles conclusions à la population générale entrave les possibilités de choix pour les décideurs dans l'allocation des ressources. De plus, les ECR sont généralement limités dans le temps et les décideurs ne pourront pas toujours appréhender l'efficacité testée et les coûts mesurés au-delà de la durée de l'essai (129).

Selon Drummond, « le choix entre l'utilisation de données expérimentales ou d'observations est représenté par une série de compromis, entre validité interne et validité externe, entre précision spécifique ou pertinence publique, et entre solutions immédiates et études à long terme. S'il y a un principe directeur dans ce choix, c'est le suivant : nous ne devrions généralement pas utiliser les données d'observation pour établir ou attribuer une différence entre les traitements, mais elles peuvent être utilisées pour estimer les conséquences économiques d'une telle différence » (120). Les données d'observations sont ainsi mieux adaptées à l'évaluation économique d'une pathologie. Elles permettent de tenir compte du contexte dans lequel les décideurs se situent et des contraintes auxquels ils doivent faire face.

IV.1.2. Le recueil des données dans le cadre des études épidémiologiques et cliniques

Le recueil de données individuelles en recherche épidémiologique et clinique, peut être effectué de différentes façons. Dans le cadre des études d'observation, il est possible d'avoir recours à des questionnaires ou des interviews. Les questionnaires consistent en une série de questions visant à recueillir les variables définies dans le cadre du protocole et utiles aux analyses. Il existe un large éventail de possibilité d'administration des questionnaires (auto ou hétéro questionnaires), de type de questions (ouvertes, fermées, à choix unique ou multiple...) et de mode de passation (passation en présentielle au cours d'une visite de suivi, envoi postal, questionnaires en ligne) (110). Il est également possible de réaliser des interviews afin de collecter les données nécessaires aux analyses. Dans ce cas, les entretiens sont réalisés par une personne formée au recueil de données, lors d'entretiens face à face ou téléphoniques. Ils peuvent être individuels ou en groupes et réalisés nécessairement sur un ensemble d'individus ou de groupes d'individus. Ils permettent d'obtenir des informations sur des faits ou des représentations (130). Il existe différents types d'interviews telles que les interviews libres, semi dirigées, ou dirigées, selon le rôle de l'interviewer dans l'entretien. Dans le cadre des questionnaires ou des interviews, les populations doivent être sélectionnées soigneusement et l'étude des caractéristiques des individus ne souhaitant pas répondre aux questionnaires ou être entretenus (i.e. non répondants) est primordiale afin d'éviter certains biais de sélection.

Dans le cadre des ECR, les données sont généralement obtenues à partir d'un formulaire de recensement de données ou « Case Report Form » (CRF). Le CRF est un document papier ou numérique conçu en accord avec le protocole de l'essai et qui permet de recenser les variables qui y sont indiquées. Le CRF doit permettre de recenser des données suffisamment détaillées sans provoquer d'ambigüité et dans un même temps, éviter la redondance et le recueil de données inutiles (131).

Mettre en place les études décrites dans cette partie et réaliser des entretiens, utiliser des questionnaires ou des CRF est relativement long et coûteux. Il est alors possible de faire appel à d'autres sources de données déjà existantes et alimentées en routine afin d'obtenir des données économiques individuelles. Les registres et les bases de données médico-administratives peuvent ainsi être mobilisés à cette fin.

IV.2. Les registres

Un registre est une structure qui réalise « un recueil continu et exhaustif de données nominatives intéressant un ou plusieurs événements de santé dans une population géographiquement définie, à des fins de recherche et de santé publique, par une équipe ayant les compétences appropriées » (132).

Les premiers registres français datent de 1975 sur des initiatives individuelles dans un double objectif de surveillance et de recherche. Ils s'inscrivent dans une politique nationale de santé publique à partir de 1986 avec la création du Comité National des Registres (CNR), devenu le Comité d'Évaluation des Registres (CER) en 2014. Leur objectif principal est la réalisation d'études épidémiologiques permettant de comprendre et mieux contrôler l'apparition des pathologies en mesurant et comparant les incidences au cours du temps, selon différentes zones géographiques ou caractéristiques socio démographiques des patients (133). Les aspects cliniques (traitement, guérison, devenir des patients) n'étaient pas les préoccupations initiales. Les registres recensent de manière exhaustive et sans double comptage les patients atteints d'une pathologie d'intérêt (cancers, maladies cardio ou neuro vasculaires, malformations congénitales, maladies rares) afin de décrire sans biais l'incidence de leur maladie (134). Cela s'effectue en deux temps. Dans un premier temps, il y a la recherche, le chaînage et la validation des données médicales démographiques et

administratives. Dans un second temps, les informations obtenues sont mises en forme, constituant une base de données codées et analysables pour la production d'indicateurs statistiques. Ils permettent également de réaliser des études plus détaillées en recherchant des sources d'informations complémentaires (dossiers médicaux, sources démographiques..) pour répondre à des questions plus spécifiques (133). L'annexe 1 décrit la structure générale de la collecte des informations dans un registre.

Les registres élaborent et mettent en œuvre des bases exhaustives à partir desquelles il est possible de tirer des échantillons représentatifs des cas dans la population cible. Ils permettent d'obtenir des informations agrégées fiables, standardisées et représentatives de la population générale. Ils procurent un ensemble de données déclinées au niveau national et régional portant sur l'incidence, la prévalence et la mortalité des pathologies selon un ensemble de caractéristiques, notamment géographiques. Il est également possible de retourner aux données individuelles sur les cas identifiés et d'obtenir des informations sur la consommation de soins à partir des dossiers médicaux. Ces données sont alors fondées sur l'observation des pratiques réelles qui apportent des estimations de « vraie vie » et représentatives de la zone géographique ciblée. Les informations relatives aux consommations de soins seront toutefois très limitées. Ce sont généralement celles contenues dans les dossiers médicaux, c'est-à-dire les protocoles et les prescriptions mis en œuvre pour le patient (135). Toutefois, dans le cadre des registres du cancer, des enquêtes dites de « hautes résolutions », permettent d'obtenir des informations complémentaires en population générale sur l'efficacité médicales des traitements réalisés. Ces études s'effectuent sur des cohortes de patients présentant une même localisation, sur une année et une zone géographique donnée. Néanmoins, elles restent peu adaptées aux évaluations économiques des pathologies ou des prises en charge. L'exhaustivité des registres permet également l'étude des capacités d'algorithmes d'identification de cas au niveau du pays sur des bases médico-administratives nationales lorsque les codages discriminants d'intérêts sont identiques (136).

Le choix entre l'usage des données issues d'un registre ou d'une étude épidémiologique dépend de la balance des avantages et inconvénients de ces solutions (133). Les données provenant d'une étude permettront de répondre de façon précise à davantage de questions car l'étude sera construite pour répondre à l'objectif visé. Néanmoins, la réponse sera plus lente et le coût de la recherche élevé. Les exploitations d'un registre peuvent être rapides et peu coûteuses mais limitées par les données (données non collectées ou codées sous une forme inadaptée). Les bases médico-administratives du SNDS offrent, dans ce cadre, d'autres possibilités avec des informations exhaustives sur la consommation de soins des patients.

IV.3. Les bases de données médico-administratives du Système National des Données de Santé (SNDS)

En France, la Consommation des Soins et des Biens Médicaux (CSBM) est financée en grande partie par l'AM. Parmi les différentes branches de la sécurité sociale, la branche maladie qui finance les risques maladie, maternité, invalidité et décès de l'ensemble de la population française est gérée par l'AM. Elle garantit l'accès aux soins, accompagne les individus dans la préservation de leur santé (i.e. prévention, éducation thérapeutique) et participe à l'amélioration de l'efficacité du système de santé en mettant à la disposition des acteurs de la santé, des données portant sur la consommation des soins. L'AM prend en charge l'ensemble des individus affiliés à divers régimes de sécurité sociale tels que le Régime Général (RG), qui a absorbé le Régime Social des Indépendant (RSI) début 2018, la Mutualité Sociale Agricole (MSA) et des régimes spéciaux. L'AM dispose d'un accès à un ensemble de données médico-administratives caractérisant ses assurés et leurs consommations de soins. Depuis 1999, la Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés (CNAMTS) travaille à l'ouverture de l'accès à ces données et est chargée de la gestion du SNIIR-AM, appelé SNDS depuis la loi de modernisation du système de santé (9). Le SNDS est un système d'information qui exploite un ensemble de bases de données (i.e. bases de l'AM, bases du PMSI, Bases des Causes Médicales de

Décès (BCMD)...) et permet un accès à une quantité importante d'informations sur les assurés et leur consommation de soins remboursés. C'est un entrepôt de données ou « datawarehouse ». L'enjeu actuel est de permettre à un grand nombre d'utilisateur d'accéder de façon sécurisée à ces données afin d'en favoriser une exploitation et une analyse pluridisciplinaire (108).

IV.3.1. Finalités du SNDS

Selon l'article L1461-1 III. du Code de la santé publique, le SNDS a pour finalité la mise à disposition des données portant sur les assurés de l'AM afin de favoriser (9) (10) :

- les études, les recherches, les évaluations relatives à la santé présentant un caractère d'intérêt public et apportant de l'information sur la santé, l'offre de soins, la prise en charge médico-sociale et sa qualité. La comparaison des pratiques aux référentiels, l'évaluation des comportements de consommation de soins et des déterminants de la qualité des soins permet l'amélioration du système de santé
- la définition, la mise en œuvre et l'évaluation des politiques de santé et de protection sociale. Le flux de données accessible permet d'identifier et d'analyser les parcours de soins des patients ainsi que d'évaluer et surveiller leur état de santé et leur consommation de soins associée
- la connaissance des dépenses de santé, des dépenses de l'AM et des dépenses médico-sociales. Connaître la répartition des dépenses de santé selon les pathologies, les zones géographiques, la nature des dépenses, par catégorie de PS ou d'établissement apporte une meilleure gestion de l'AM et de la distribution des ressources
- l'information des professionnels, des structures et des établissements de santé ou médico-sociaux sur leur activité. La transmission d'informations pertinentes aux PS et aux centres de

soins sur leurs activités, leurs recettes, leurs prescriptions permet l'adaptation et l'amélioration des pratiques

- la surveillance, la veille et la sécurité sanitaire et l'évaluation de l'innovation dans les domaines de la santé ainsi que la prise en charge médico-sociale. Des données constamment disponibles et régulièrement actualisées permettent d'analyser la santé de la population dans le temps et dans l'espace, de détecter les événements de santé inhabituels et d'endiguer les possibles menaces pour la santé publique. La récurrence des informations contenues dans le SNDS donne la possibilité d'évaluer les innovations présentes sur le marché.

On retrouve également deux finalités prohibées. Il est interdit d'utiliser les bases de données pour :

- la promotion des produits de santé auprès des PS et des établissements de santé et pour l'exclusion de garanties des contrats d'assurance
- la modification de cotisation ou de primes d'assurance d'un individu ou d'un groupe d'individus présentant un même risque.

Le SNDS est doté d'une architecture complexe. Une bonne connaissance de son système d'alimentation, de son évolution dans le temps et selon les différents régimes, et des données qu'il contient est nécessaire avant toute utilisation (137).

IV.3.2. Structure du SNDS

Le SNDS est le prolongement du SNIIRAM qui constituait l'ancien système d'information de l'AM. Le SNIIRAM regroupait les informations issues des remboursements de soins effectués par l'ensemble des régimes d'AM.

IV.3.2.1. Du SNIIRAM au SNDS

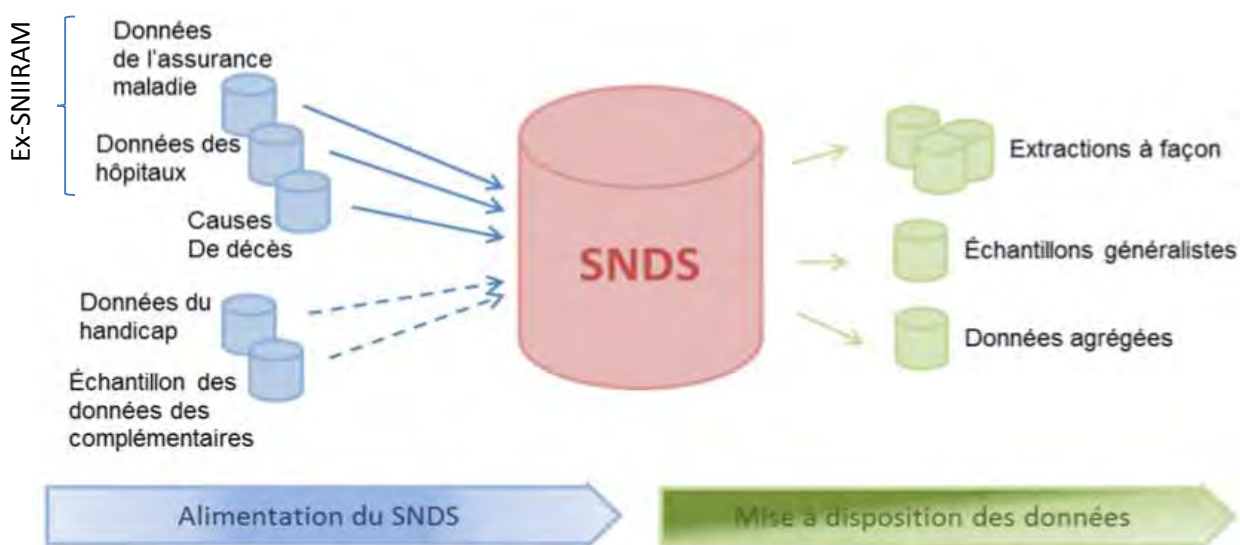
Le SNIIRAM a connu de profondes transformations depuis sa naissance jusqu'à la création du SNDS. Le document présenté en annexe 2 résume ses évolutions. Il a été fondé en 1998 et a été mis en œuvre à partir de 2003. Initialement, il regroupait les données portant sur la consommation de soins de ville relatives au RG, soit 76% de la population Française (138). Un échantillon du SNIIRAM à l'échelle 1/97 appelé « Echantillon Généraliste des Bénéficiaires » (EGB) et composé des données du SNIIRAM a été créé en 2005. La même année, le SNIIRAM a été chaîné avec les données hospitalières. En 2007, le SNIIRAM s'enrichit des données des bénéficiaires d'autres régimes de sécurité sociale tels que la MSA, le RSI, la sécurité sociale des militaires, des cultes et des clercs de notaire. De plus la même année, les données du PMSI sont accessibles dans l'EGB avec les dates des séjours hospitaliers au format mois/année. Dès 2009 dans le SNIIRAM, les dates de décès deviennent accessibles et les données hospitalières intègrent désormais les molécules onéreuses et les Dispositifs Médicaux Implantables (DMI) hors GHS, renforçant son potentiel pour les travaux de recherche. En 2010, les dates exactes de séjours sont renseignées dans l'EGB et en 2011, il devient inter-régime avec l'ajout des assurés du RSI et du MSA. En 2012, le SNIIRAM s'enrichit de 9 Sections Locales Mutualistes (SLM) lui permettant de couvrir 98.8% de la population Française (139). En 2013, l'EGB Simplifié (EGBS) est créé afin de faciliter l'utilisation de l'EGB. Finalement, en 2015 se rajoutent les SLM contenues dans le SNIIRAM, permettant à l'EGB de couvrir une population similaire. Le SNDS constitue le prolongement du SNIIRAM depuis le décret n° 2016-1871 du 26 décembre 2016 (140).

IV.3.2.2. Structure générale du SNDS

Le SNDS est un entrepôt de données rassemblant les données des assurés de l'AM chaînées avec les actes des PS et l'ensemble des prestations de soins en ville, les données d'activité hospitalière du PMSI, les causes médicales de décès du Centre d'Épidémiologie sur les causes médicales de Décès

(CépiDc) de l'Institut National de la Santé et de la Recherche Médicale (INSERM), les données médico-sociales des Maisons Départementales des Personnes Handicapées (MDPH), les données de la Caisse Nationale de la Solidarité pour l'Autonomie (CNSA) et un échantillon de données de remboursements des organismes d'AM Complémentaires. La figure 7 issue du site de l'INDS synthétise les sources de données composant le SNDS ainsi que les différentes restitutions possibles de ces données (141).

Figure 7. Alimentation et restitution du SNDS



Le SNDS est actuellement en cour de finalisation. Les données relatives au handicap et l'échantillon des données des organismes d'assurance complémentaire ne sont pas disponibles à ce jour (141).

Le SNDS offre trois ensembles de restitution possibles pour les utilisateurs. Des ensembles de données agrégées selon différentes thématiques sont accessibles en open data. L'EGB est également mis à disposition et permet l'accès à des données individuelles échantillonnées du SNDS. Enfin, il est possible de demander des extractions de données individuelles dans le cadre de projets spécifiques.

IV.3.2.3. Les bases de données du SNDS

Le SNDS est constitué actuellement de trois sources de données: les données d'activité hospitalière du PMSI, les données de consommation de soins inter-régimes de l'AM et la BCMD.

IV.3.2.3.1. PMSI

Le PMSI est un outil de description et de mesure médico-économique de l'activité hospitalière produite par les établissements de santé. Il est fondé sur le recueil systématique et le traitement automatisé d'informations médico-administratives structurées dans quatre champs spécifiques : le champ du MCO, le champ du SSR, le champ de l'HAD et le champ de la psychiatrie. Le principe général est de classer et regrouper les séjours hospitaliers dans des groupes (GHM en MCO) considérés comme étant homogènes en termes de prise en charge médicale mais aussi de ressources mobilisées pour assurer ces prises en charge. Il permet ainsi de disposer d'informations médicales standardisées et quantifiées caractérisant l'activité des établissements de santé. L'annexe 3 décrit les principales informations contenues dans le PMSI.

Sa mise en œuvre dans les établissements de santé est obligatoire depuis la loi du 31 juillet 1991 qui a imposé à ces mêmes établissements de procéder à l'évaluation et à l'analyse de leur activité. Il a pour objectif principal de mesurer l'activité hospitalière des établissements de santé et de déterminer leurs ressources pour le MCO et le SSR dans le cadre de la T2A. Dans le contexte de la procédure d'allocation budgétaire, il contribue ainsi au calcul du budget des établissements de santé. A ce titre, il est devenu un outil indispensable à la gestion des établissements de santé (142).

Le PMSI peut également être utile dans le cadre de la planification. Il peut aider les pouvoirs publics à améliorer la répartition de l'activité médicale et des plateaux techniques sur le territoire en fonction des besoins de la population. L'exploitation et l'analyse des bases PMSI présentent également un

intérêt majeur dans le cadre d'études épidémiologiques, de la surveillance sanitaire et d'évaluations économiques des stratégies de prise en charge.

Le Département d'Information Médicale (DIM) est chargé, au sein d'un établissement, de collecter l'information en lien avec les séjours hospitaliers et de gérer la base PMSI de l'établissement hospitalier, d'en contrôler l'exhaustivité et la qualité, d'en assurer son exploitation et sa diffusion aux différents acteurs, dans le cadre du respect des exigences réglementaires. A la demande de la communauté hospitalière, il réalise des analyses et contribue à différents travaux visant à confronter les données d'activité, de recettes et de coûts afin d'aider l'établissement dans sa gestion. Il a également vocation à aider la communauté hospitalière dans l'élaboration de sa stratégie médicale. A ce titre, il aide la communauté hospitalière à appréhender l'hôpital dans son environnement et sa place dans le dispositif de soins en réalisant des travaux fondés sur l'exploitation et l'analyse des bases du PMSI (143).

IV.3.2.3.2. Bases de l'AM

L'AM met à disposition un ensemble d'informations sur la consommation des soins remboursés. Ces informations sont notamment recensées à partir des feuilles de soins au niveau régional par les Caisses Primaires d'Assurance Maladie (CPAM). Elles sont ensuite regroupées et transmises quotidiennement aux bases nationales de l'AM. Pour les bénéficiaires du RG, les CPAM assurent l'alimentation des bases ERASME, bases régionales qui alimentent les bases nationales de l'AM. Les données recensées sont ensuite pseudonymisées (i.e. procédure FOIN (Fonction d'Occultation d'Informations Nominatives)) avant d'être regroupées. Les référentiels des individus sont regroupés dans le Répertoire National Inter régime des bénéficiaires de l'Assurance Maladie (RNIAM) qui permet de valider l'appartenance des individus à un régime d'AM. Cette validation s'effectue suite à la comparaison avec le Répertoire National d'Identification des Personnes Physiques (RNIPP), géré

par l'Institut National de la Statistique et des Etudes Economiques (INSEE). Les informations sur le statut civil des individus ainsi que le Numéro d'Identification au Répertoire (NIR) sont contenus dans les deux répertoires et permettent cette comparaison. Le RNIAM permet ainsi d'avoir une référence sur tous les individus affiliés à l'AM, consommateurs ou non des soins. Néanmoins, le NIR n'est pas directement contenu dans les bases de l'AM. Les autres référentiels couvrent les PS, les établissements sanitaires et médico-sociaux, les nomenclatures des procédures médicales (CCAM, NGAP), les régimes et allocations sanitaires, et le système de classification Anatomique, Thérapeutique et Chimique (ATC) des médicaments prescrits remboursés par les régimes nationaux d'AM.

Les données contenues dans les bases de l'AM sont multiples. On retrouve les informations suivantes portant sur :

- les individus :
 - o l'âge, le sexe, la complémentaire santé, la commune et le département de résidence, la date de décès, si bénéficiaire de la Couverture Maladie Universelle- (Complémentaire) (CMU-(C)) ou de l'Aide au paiement d'une Complémentaire Santé (ACS) et un indice de défavorisation sociale (le « French Deprivation index » (FDep) (144)).
- la consommation de soins de ville :
 - o toutes les prestations remboursées avec le codage détaillé de chaque prestation (actes médicaux, biologie, dispositifs médicaux, médicaments).
 - o l'ensemble des informations relatives à la quantité, au prix et au remboursement de la prestation.
 - o les dates de soins et de remboursements.
- l'offre de soin :
 - o spécialité du prescripteur, lieu d'exécution, statut.

- les pathologies traitées :
 - o diagnostics codés selon la Classification Internationale des Maladie, 10ieme révision (CIM-10), pour les personnes ayant fait une demande d'ALD et les personnes hospitalisées.
 - o des informations médicales (GHM, médicaments traceurs, actes techniques, examens biologiques ou dispositifs médicaux).
- les prestations en espèces :
 - o les IJ avec les codes CIM-10 associés pour les arrêts de plus de 6 mois, les accidents de travail, les maladies professionnelles, les pensions d'invalidité.

Il sera néanmoins impossible d'avoir accès simultanément à certaines informations dites « sensibles » car pouvant permettre la ré-identification des individus. Ces informations sont :

- la date de naissance (Année + Mois),
- la date des soins (Année + Mois + jour),
- la date de Décès (Année + Mois + jour),
- la commune de résidence.

IV.3.2.2.3. Base des causes médicales de décès

Le CépiDc est un laboratoire de l'INSERM qui produit des statistiques sur les causes médicales de décès en France depuis 1968 (145). Il participe à la diffusion de ses données et aux recherches sur les causes de décès. Ce laboratoire travaille notamment avec l'INSEE afin d'établir les certificats de décès.

Le CépiDC réalise des analyses épidémiologiques descriptives sur des pathologies spécifiques. Il évalue la qualité des données, ses variations et les biais existants (certification électronique du médecin, mesure de qualité, comparaison à d'autres bases de données...), évalue les facteurs de variations de la mortalité (chaleur, alcool, niveau socio-économique, contexte de résidence...), et la

qualité du système de soins (indicateur de mortalité évitable liée au système de soins, calcul d'indicateurs de mortalité ajustés représentatif de la qualité des soins...) (145).

Le CépiDC assure l'exploitation de la BCMD en France (146). La base couvre l'ensemble des décès survenus sur le territoire français depuis 1968. Les informations contenues dans la BCMD sont recueillies à partir du certificat et du bulletin de décès. Cette base anonymisée classe les décès selon l'affection responsable du décès ainsi que les autres facteurs y ayant contribué. Les causes médicales de décès sont codées selon la Classification Internationale des Maladies (i.e. de 1968 à 1978 : CIM-8 ; de 1979 à 1999 : CIM-9 ; depuis 2000 : CIM-10). La base contient également des informations non médicales liées au décès et à l'individu comme le lieu et la date du décès, l'âge et le sexe du défunt, son statut marital, sa catégorie socio-professionnelle pour les actifs.

IV.3.2.3. Domaines de restitution du SNDS

Comme nous l'avons précédemment exposé (figure 7), le SNDS est un entrepôt de données qui possède trois domaines de restitution : les données agrégées, des données échantillonnées et des données individuelles sur l'ensemble de la population couverte par l'AM. Les données agrégées mises à disposition sont strictement anonymes, utilisables par tous et décrites sur le site du SNDS (147). Nous présentons ici les restitutions permettant d'avoir l'accès à des informations de nature économique portant sur les consommations de soins individuelles.

IV.3.2.3.1. Données individuelles échantillonnées

L'Echantillon Généraliste des Bénéficiaires (Simplifié) (EGB(S)) permet l'accès à des données individuelles sur les bénéficiaires, sur les PS et sur les établissements. C'est un échantillon à l'échelle 1/97^{ième} de la population couverte par l'AM, représentatif de la population générale en termes d'âge,

de sexe et de couverture. Il est composé d'environ 660 000 individus et permet l'accès aux non consommateurs. Il couvre les trois principaux régimes de sécurité sociale (RG, RSI, MSA) depuis 2011. Il est alimenté de façon dynamique tous les mois à partir des données individuelles des bénéficiaires, des PS ou établissements (DCIR, AMOS) et tous les trois mois sont mis à jours les individus inclus dans l'échantillon (naissance, émigration, immigration, décès). Un temps de latence d'environ 2 ans est néanmoins à prévoir avant d'avoir accès aux données du PMSI de l'année en cours. De plus, les données de mortalités ne sont pas toujours correctement remontées pour les régimes autres que le RG et les informations sur les ALD pour la MSA ne sont pas exhaustives avant 2014. L'EGB contient les données individuelles des bénéficiaires et des PS ainsi que tous les détails et codages affinés sur les prestations et sur les données hospitalières, notamment pour les données MCO. Les données sont conservées sur un horizon temporel de 20 ans. L'EGB est particulièrement adapté à des travaux portant sur le long terme et sur des maladies avec une prévalence/incidence importante (148). Les données relatives aux individus ne rentrant plus dans les critères d'inclusion de l'EGB sont tout de mêmes conservées dans l'échantillon, facilitant ainsi la réalisation d'études longitudinales (11). L'EGB Simplifié (EGBS) est une version simplifiée de l'EGB. On y retrouve les mêmes informations mais regroupées afin d'obtenir une seule ligne par prestation associée à un bénéficiaire, prescrite et exécutée par un PS ou un établissement.

IV.3.2.3.2. Données individuelles non échantillonnées

Le SNDS donne accès à des données individuelles portant sur la totalité des bénéficiaires de l'AM et sur des PS et des établissements de soins.

L'univers AMOS, contient des données portant sur les PS et des établissements en exécution et en prescription et des données agrégées sur les bénéficiaires. Il permet de connaître le profil, la démographie, la géographie et l'activité en volume et les montants correspondant aux honoraires des PS libéraux exerçant en cabinet de ville ou en établissement privé. Cet axe permet de décrire

leurs patientèles. Il apporte des informations relatives aux remboursements selon les différents organismes.

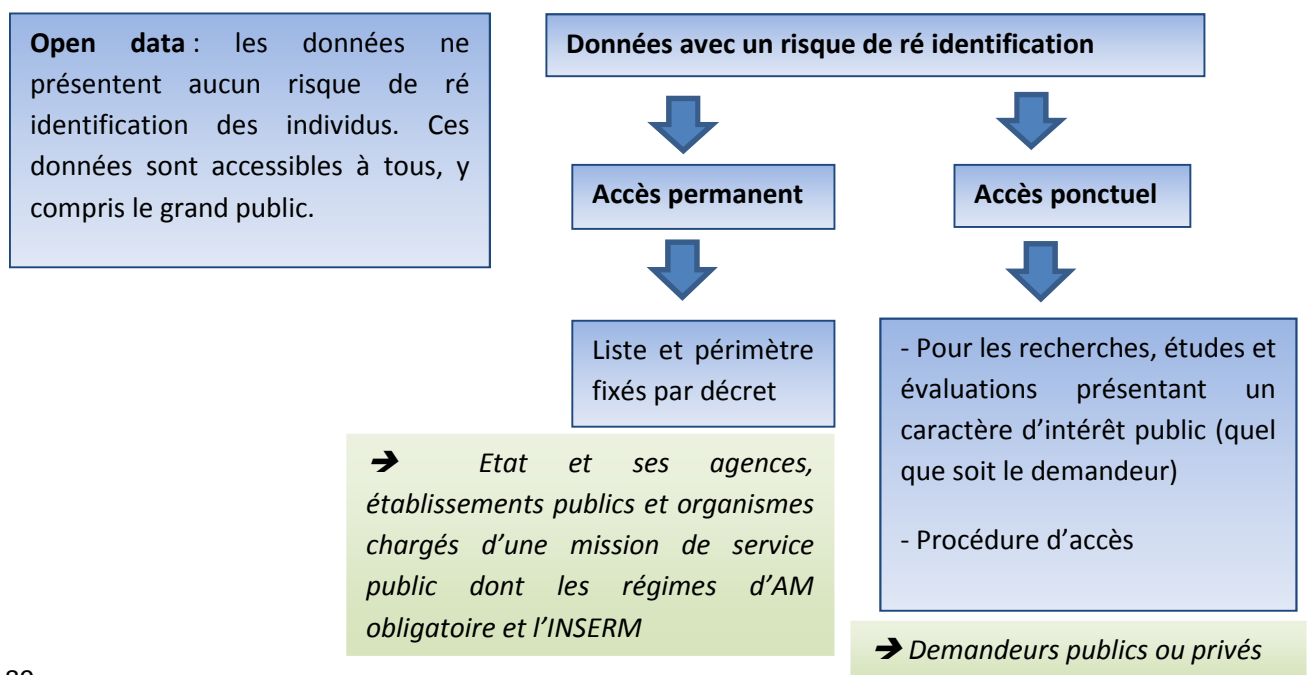
Le Datamart de Consommation Inter Régime (DCIR) permet l'accès à des données individuelles portant sur les bénéficiaires, les PS et sur les établissements. Ce datamart contient tous les codages affinés des prestations, les données hospitalières du PMSI, les causes médicales de décès, les numéros des PS pour tous les régimes relevant de l'AM à l'exception des régimes de l'Assemblée Nationale et du Sénat. Il permet de réaliser des études sur la consommation des soins de chaque bénéficiaire.

L'annexe 4 schématise l'alimentation et les domaines de restitutions du SNDS (140).

IV.3.3. Accès au SNDS

Le SNDS offre ainsi trois domaines de restitution possibles. Les accès à ces trois ensembles vont dépendre naturellement des données demandées. Les accès aux SNDS sont alors de trois types : l'accès en open data, les accès permanents et les accès ponctuels. La figure suivante extraite d'une présentation dans le cadre d'un séminaire de l'INSERM résume les différents accès au SNDS (149).

Figure 8. Les différents accès au SNDS



Selon le site officiel du SNDS : « Toute personne ou structure, publique ou privée, à but lucratif ou non, pourra dès avril 2017 accéder aux données du SNDS sur autorisation de la Commission Nationale de l'Informatique et des Libertés (CNIL), en vue de réaliser une étude, une recherche ou une évaluation présentant un intérêt public » (150). Ainsi des accès ponctuels sont accordés à tous, privé comme public, sous réserve que la recherche effectuée poursuive un but non lucratif et qu'elle soit validée par les autorités compétentes. Les demandes d'autorisation d'accès aux données du SNDS sont à déposer à l'Institut National des Données de Santé (INDS). Cet institut a, entre autres, pour missions de veiller à la qualité des données de santé et aux conditions générales de leur mise à disposition dans le respect de la loi Informatique et libertés. Il traitera toute demande d'autorisation à l'exception des recherches biomédicales (151). Nous reviendrons sur les procédures d'accès au SNDS dans la partie IV.4. sur la réglementation associée aux données de santé.

Des accès permanents au SNDS sont automatiquement accordés à certains organismes gouvernementaux, institutions publiques et autorités de service public. Ces accès sont fixés par un décret (décret n° 2016-1871) (152). L'annexe 5 rassemble les différents ayant-droits permanents au SNDS. Les accès aux bases seront limités aux seuls besoins du projet concerné et ne pourront être utilisés qu'à cette fin.

Le principe de l'accès au SNDS semble donc élargi par rapport au SNIIRAM. Tout individu ou entité peut maintenant y accéder dès lors que ses recherches sont réalisées pour l'intérêt public. Néanmoins, la notion « d'intérêt public » n'étant pas clairement définie, pourra donner lieu à de multiples interprétations. Les nouvelles conditions d'accès aux bases de données médico-administratives donnent ainsi un rôle et une responsabilité prépondérante aux acteurs participant au processus d'autorisation. Ils vont devoir identifier chez les multiples demandeurs, quels projets présentent un caractère d'intérêt public et ne dépassent pas sur les finalités proscrites. Il est encore

difficile de démontrer l'efficacité de ces nouvelles procédures d'accès étant donnée leur récence (151).

IV.3.4. Intérêts, limites et usages des données du SNDS

IV.3.4.1. Intérêts, et limites du SNDS

Le SNDS présente un intérêt majeur pour la recherche en santé en France. C'est un des systèmes d'information le plus complet au monde, traitant un volume considérable d'informations, et apportant un matériel très utile pouvant aider à l'amélioration du système de santé (11) (137). Il faut pour cela se l'approprier et comprendre ses possibilités comme ses limites. Le tableau suivant, adapté de Mouly et al, résume les principaux avantages et inconvénients des données du SNDS (11).

Tableau 2. Avantage et limites des données du SNDS

	Intérêts	Inconvénients
Architecture	Vision globale du parcours de soins des patients. Informations économiques détaillées. Bases de données alimentées de manière prospective. Profondeur importante	Architecture complexe, en évolution permanente.
Population incluse	Large part de la population française (98.8% depuis 2015). Base de données en « population générale » fondée sur des données de vraie vie. Causes médicales de décès renseignées.	Peu de données sociales (CMU-C, ACS, Fdep). Ajout des SLM tardif dans l'EGB. Identifiant unique seulement à partir de 2010 (RG), 2011 (RSI, MSA) et 2015 (SLM).
Soins de ville	Données individuelles et détaillées sur la consommation de soins en ville. Dates exactes de prescription et d'exécution des soins. Détails sur le dispensateur des soins. Détails sur les délivrances des médicaments : date, code CIP, code ATC, quantité, montant.	Pas d'information sur les indications des soins ou sur les résultats des examens cliniques. Pas d'information sur les médicaments en vente libre. Disponibilité d'informations sur les médicaments prescrits mais pas sur ceux consommés.
Soins hospitaliers	Données individuelles et détaillées sur les soins hospitaliers dans tous les centres privés et publics. Détails sur les codes diagnostics, procédures, molécules onéreuses et dispositifs médicaux implantables (en sus de GHS), soins palliatifs, néonatal...	Codes diagnostics parfois peu fiables. Pas de données cliniques détaillées. Pas de détails sur l'exposition médicamenteuse autre que les molécules onéreuses. Pas de données sur les consultations si aucune procédure n'est réalisée.
Accès	Informations agrégées en open data. Accès permanent facilité. Accès ponctuel ouvert selon le respect des finalités d'usage.	EGB : Usage de SAS guide imposé et temps de compilation des codes long (passages par de multiples serveurs).

SNIIRAM : Système National d'Information Inter Régime de l'Assurance Maladie ; EGB : Echantillon Généraliste des Bénéficiaires ; ACS : Aide au paiement d'une Complémentaire Santé ; CMU-(C) : Couverture Maladie Universelle-(Complémentaire) ; FDep : French Deprivation index ; CIP : Code Identifiant de Présentation ; ATC : Anatomique, Thérapeutique et Chimique ; GHS : Groupe Homogène de Santé ; PSY : Psychiatrie, HAD ; Hospitalisation A Distance ; SSR : Soins de Suite et Réanimation

Les informations contenues dans le SNDS sont exhaustives et représentatives des prises en charge effectuées sur l'ensemble de la population française. Elles permettent ainsi de disposer de résultats sur les pratiques observées dans des conditions de « vie réelle ». Toutefois, les données du SNDS n'ont pas initialement été collectées à des fins de recherche et peuvent faire l'objet d'erreurs de mesure ponctuelles ou systématiques. Ces erreurs peuvent impacter l'identification de la population d'étude, des expositions, des événements ou des covariables (137). Le SNDS s'étant enrichi continuellement, une attention particulière doit être portée sur la période d'étude afin de maîtriser les biais dus aux caractéristiques des bénéficiaires non inclus. De plus, certains individus ne

possèdent pas d'identifiant unique (i.e. un enfant) en raison de leur régime d'appartenance et/ou de la période d'extraction, pouvant perturber l'identification d'une population incidente, notamment pédiatrique.

Afin de réaliser une étude, les populations ciblées sont définies selon des algorithmes appliqués sur les diverses tables de consommation de soins du SNDS. Ces algorithmes peuvent apporter des biais de classement importants s'ils ne sont pas étudiés à priori, notamment si leur sensibilité et spécificité ne sont pas suffisamment importantes. La CNAMTS et diverses équipes travaillent notamment sur la création d'algorithmes d'identification de cas incidents ou prévalents combinant les tables ALD et les codes diagnostiques de la CIM-10 disponibles dans le PMSI, les procédures médicales et l'exposition médicamenteuse (136) (153). Les travaux réalisés par la CNAMTS sur le sujet sont disponibles sur le site de l'AM où 56 groupes non exclusifs de pathologies sont identifiées à partir d'algorithmes sur les bases du SNDS (154). Il est également possible d'extraire des populations témoins, indemnes des pathologies étudiées, dans l'EGB, car donnant accès aux individus non consommateurs (148).

Le SNDS est fondé sur des données de remboursement et ne recense pas les médicaments non remboursés ou administrés dans le cadre d'un séjour hospitalier (hors médicaments rétrocedés, en sus de GHS ou DMI). Toute la consommation médicamenteuse et les DM ne seront donc pas disponibles. De plus, il se peut qu'un médicament ou un DM prescrit ne soit pas dispensé et qu'un médicament dispensé ne soit pas consommé, faussant encore certaines mesures sur l'observance, la consommation ou la durée de prise en charge.

Le SNDS ne contient pas de données médicales (résultats cliniques et résultats d'examen paracliniques) ou relatives aux comportements à risques des individus (alcool, tabac...). Il contient peu de données de nature socio-économiques exception faite des populations à faible revenu (CMU-C, ACS), et peu de données géographiques (commune de résidence). Néanmoins les données en lien avec la commune de résidence permettent le calcul d'un indice de défavorisation (i.e. « Deprivation ») sociale (144). Le futur ajout des données médico-sociales des maisons

départementales des personnes handicapées apportera un complément d'information aux données socio-économiques déjà disponibles sur les individus. Le SNDS permet toutefois la mesure de covariables utiles aux études. A titre d'exemple, Bannay et al. ont adapté et validé un algorithme permettant d'estimer le score de comorbidité de Charlson à partir de diverses sources de données du SNDS, notamment des bases du PMSI (155).

IV.3.4.2. Usages des bases de données du SNDS

Un des objectifs de la création du SNDS était d'ouvrir les données de consommation de soins aux différents acteurs de la santé afin d'en favoriser un usage pluridisciplinaire (108).

En épidémiologie, de par la grande proportion de la population française incluse et son caractère prospectif, le SNDS peut permettre d'évaluer des incidences et des prévalences sur tout le territoire et selon différentes entités géographiques (11). Il est possible de travailler sur la mortalité de façon précise avec le récent ajout des causes médicales de décès. Néanmoins, il est trop tôt pour appréhender correctement cette possibilité.

Les données du SNDS sont des données d'observations, fondées sur des pratiques réelles qui permettent la mise en œuvre de recherches pharmacologiques. Ces données permettent d'évaluer l'utilisation, l'efficacité et la sécurité des médicaments ainsi que de comparer les pratiques aux référentiels de traitement. L'efficacité des médicaments peut être étudiée par le biais d'évènements graves et de la mortalité (i.e. codes de diagnostics hospitaliers, procédures médicales particulières ou exposition médicamenteuse). Elles permettent l'évaluation de l'innocuité des médicaments après commercialisation, en se focalisant sur l'occurrence des événements indésirables (139).

En économie de la santé, les bases de données du SNDS sont un atout majeur pour la réalisation d'études économiques (11) (137) (156). Ces bases permettent l'accès aux données de

remboursements extrêmement détaillées d'une large part de la population française. On y retrouve des informations sur les remboursements relatifs aux coûts directs médicaux (actes médicaux, actes paramédicaux, équipements médicaux, séjours hospitaliers, médicaments), aux coûts directs non médicaux (transports) et aux coûts en lien avec la perte de productivité (IJ, pensions). Le SNDS permet un suivi des dépenses remboursées par l'AM de façon longitudinale et récente. Il apporte une connaissance précise des ressources effectivement mobilisées et remboursées dans la prise en charge d'une pathologie. Il peut être également possible d'évaluer le coût net des pathologies étudiées en constituant des cohortes de patients témoins à partir de l'EGB. Le SNDS est de fait particulièrement adapté à l'évaluation du fardeau économique des pathologies, dès lors qu'elles sont identifiables dans les bases (119). Le personnel de la CNAMTS a par exemple utilisé les algorithmes d'identification cités dans la section précédente afin d'évaluer les poids économiques relatifs aux cancers en France à partir du SNDS (156). Couvrant actuellement plus de 98% de la population française, le SNDS permet d'obtenir des résultats dotés d'une très bonne validité externe. Le SNDS permet ainsi de guider les décideurs dans leur choix (137). Les résultats obtenus pourront de plus être très utiles pour alimenter des modèles statistiques.

Les multiples champs d'application du SNDS en font un outil d'une importance majeure dans l'exploitation et l'analyse des données en santé publique. Il s'inscrit parfaitement dans la réflexion actuelle développée dans le cadre du « Big Data » fondée sur un chaînage de diverses bases de données de tailles importantes qui apporte des résultats complémentaires à ceux des données issues des études décrites dans la section IV.1. avec des niveaux de preuves élevés (7).

IV.4. Règlementation associée à l'utilisation des données de santé

Les résultats issus d'études sur la santé, des travaux menés dans les registres et à partir des bases médico-administratives du SNDS sont fondées sur l'exploitation de données individuelles et

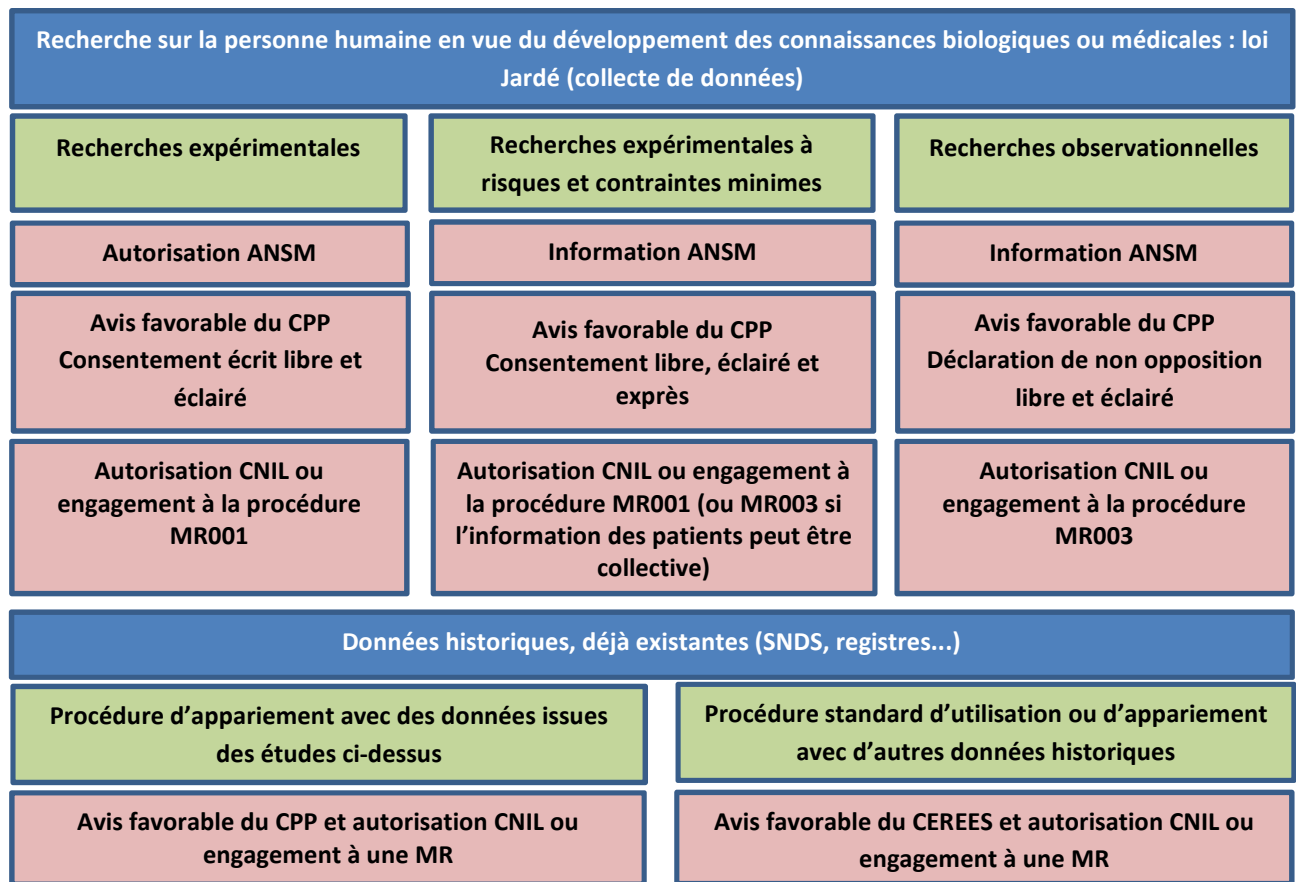
personnelles qui nécessitent d'être en conformité avec la loi sur les recherches impliquant la personne humaine, la protection de ces données et sur l'accord des individus inclus. Depuis 2016, la loi Jardé relative aux recherches impliquant la personne humaine est entrée en application et toutes recherches sur la personne humaine en vue du développement des connaissances biologiques et médicales doivent obtenir l'avis favorable d'un Comité de Protection des Personnes (CPP) ou du Comité d'Expertise pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé (CEREES) dans le cas où les données utilisées sont des données historiques (i.e. déjà existantes). Ces études doivent également obtenir une autorisation de la part de la CNIL ou s'engager à suivre une Méthodologie de Référence (MR) (157) (158). Dans le cadre de recherche sur la personne humaine, on distingue trois cas:

- Les recherches expérimentales nécessitent un consentement écrit libre et éclairé des individus participant à l'étude, un avis favorable du CPP et une autorisation de l'Agence National de Sécurité du Médicament (ANSM).
- Les recherches expérimentales à risques et contraintes minimales⁴ nécessitent un consentement libre, éclairé et exprès des individus participant à l'étude, un avis favorable du CPP et une simple information à l'ANSM (envoi du résumé et de l'avis favorable du CPP).
- Les recherches observationnelles nécessitent une déclaration de non opposition de la part du patient, un avis favorable du CPP et une simple information à l'ANSM (envoi du résumé et avis favorable du CPP).

La figure ci-dessous résume les autorisations nécessaires aux différents types d'études menées (157).

⁴ « Le caractère minime des risques et contraintes liés à la réalisation de la ou des interventions ... s'apprécie au regard de l'âge, de la condition physique et de la pathologie éventuelle de la personne se prêtant à la recherche, ainsi que de la fréquence, de la durée et des éventuelles combinaisons de ces interventions ». Arrêté du 3 mai 2017 fixant la liste des recherches mentionnées au 2° de l'article L. 1121-1 du code de la santé publique. Disponible : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000034598769&dateTexte=20180417>

Figure 9. Aspects réglementaires des études mobilisant des données individuelles



ANSM : Agence nationale de sécurité du médicament et des produits de santé ; CPP : Comité de Protection des Personnes ; CNIL : Commission nationale de l'informatique et des libertés ; CEREES : Comité d'Expertise pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé ; MR : Méthodologie de Référence.

Les recherches portant sur des données déjà existantes, telles qu'issues de registres ou des bases du SNDS rentrent dans la deuxième partie du schéma. Il n'est pas nécessaire d'informer l'ANSM pour ces études. Dans le cas d'un appariement des données déjà existantes au sein du SNDS ou d'un registre avec une étude observationnelle ou expérimentale, il est nécessaire de soumettre la recherche au CPP avant d'obtenir une autorisation de la CNIL. Dans le cas où des données historiques sont utilisées, ou lorsqu'elles sont appariées avec d'autres données historiques, il est nécessaire d'obtenir une autorisation de la CNIL, qui repose sur avis préalable de CEREES. Dans le cadre du SNDS, cette autorisation est obtenue à partir d'une demande à l'INDS qui les centralise pour faciliter et réduire le temps de réponse de la CNIL. Il existe également trois procédures simplifiées qui ne nécessitent pas d'accord CNIL. Une de ces procédures simplifiées utilise une des MR qui prévaut comme un accord

CNIL et demande seulement une inscription au répertoire public tenu par l'INDS. Les MR ont été modifiées dans le cadre de la Règlement général sur la protection des données et publié le 13 juillet 2018 où trois MR ont été rajoutées (159). Il n'est pas nécessaire d'informer l'ANSM pour ces études. L'annexe 6 schématise les procédures d'accès aux données du SNDS (160).

Les données extraites du SNDS sont conformes aux pratiques courantes mais leur utilisation nécessite certaines précautions. Les individus sélectionnés dans les études et leurs caractéristiques ne sont pas contrôlées, engendrant différents biais (i.e. sélection, confusion). Dans ce contexte, la modélisation statistique, permettant d'ajuster les résultats aux caractéristiques des individus, est d'autant plus essentielle. La HAS recommande l'utilisation de modèles adaptés afin de s'affranchir de l'hétérogénéité des populations et des pratiques (3). De plus, selon le rapport de l'« International Society for Pharmacoeconomics and Outcomes Research » (ISPOR), « les conclusions d'une étude sont conditionnelles aux hypothèses et aux données sur lesquelles le modèle s'est construit » (161). Les données permettant le développement d'une étude économique sont particulières en raison de leur distribution et leur nature (i.e. coûts non gaussien, données longitudinales) et nécessitent l'usage de méthodologies appropriées. Il est primordial de connaître la nature, les possibilités et les hypothèses sous-jacentes des modèles appliqués aux problématiques rencontrées afin d'utiliser le ou les plus pertinent(s).

V. Modélisation des coûts

La modélisation statistique pour l'évaluation économique dans le domaine de la santé est primordiale. Elle permet d'une part de réaliser des prédictions sur une variable d'intérêt selon les caractéristiques individuelles de la population et d'estimer leurs effets. D'autre part, elle permet de

simuler des processus⁵ et d'étudier les répercussions de différentes actions et paramétrisations sur les résultats finaux. On distingue ainsi deux types de modélisation permettant d'évaluer le coût d'une pathologie : les modèles de décision et les modèles de prédiction.

Les modèles de décision simulent des enchaînements d'évènements, généralement au cours du temps, afin d'étudier les conséquences associées aux décisions prises. La structure mathématique appliquée (arbres de décision, modèles multi-états, modèles individus-centrés...) représente le parcours d'un patient à travers divers états ou actions et évalue les résultats (i.e. efficacité, coûts, autres) selon chaque scénario créé. Ils sont régulièrement utilisés dans le domaine de la santé car ils permettent de faire des choix rationnels dans l'incertitude apportée par des arbitrages importants (162). Ils peuvent être employés pour estimer le coût d'une pathologie mais intègrent difficilement toute la complexité des prises en charge et des coûts associés. Dans ce contexte, ils interviennent pour extrapoler les résultats d'études, notamment d'interventions, qui ne prennent pas en compte toutes les options de traitement possibles, des populations représentatives et/ou une période suffisamment importante pour saisir les différences d'effets ou de coûts (163). Ces modèles mobilisent des données agrégées provenant de diverses sources telles que les probabilités de changer d'état ou de recevoir un traitement, les distributions de temps passé dans un état, les coûts associés...

Les modèles de prédiction synthétisent les informations contenues dans les données et identifient les dépendances entre les différentes entités (i.e. individus, caractéristiques...) (3). Ils établissent un lien entre un ensemble de caractéristiques individuelles connues (i.e. covariables) et une variable d'intérêt, permettant de prédire sa valeur ainsi que les effets associés à chaque caractéristique. Dans le cadre d'analyses mobilisant des données de vie réelle telle que décrites dans le chapitre précédent, ces modèles vont ajuster les estimations selon les paramètres de l'étude. Les biais

⁵ Processus (Larousse) : Enchaînement ordonné de faits ou de phénomènes, répondant à un certain schéma et aboutissant à quelque chose.

présents dans les études (i.e. sélection, confusion) seront considérés afin d'obtenir des résultats plus proches de la réalité.

Ces deux types de modélisation sont complémentaires. La véracité des conclusions d'un modèle de décision dépend des informations qui l'alimentent. Les modèles prédictifs sont alors particulièrement adaptés pour fournir une base de données agrégées conditionnellement aux caractéristiques des patients. Ils permettent l'évaluation des probabilités d'occurrence des événements étudiés ou des transitions entre états, représentant le parcours de soin du patient, et les intègrent dans les modèles de décision. Les modèles de prédiction décrivent finement les coûts de la pathologie étudiée et les variables l'impactant en fonction des données individuelles accessibles. Les modèles de décision sont ainsi généralement utilisés lorsque les données individuelles ne sont pas disponibles et/ou pour extrapoler les résultats à des populations plus larges ou sur un horizon temporel plus long.

Nous présentons dans ce chapitre les différents modèles permettant d'évaluer le coût d'une pathologie à partir de données individuelles. Selon l'objectif de l'étude et la nature des données, différents modèles prédictifs peuvent être utilisés. Il est essentiel de connaître les hypothèses sous-jacentes à ces modèles afin d'utiliser les plus pertinents et d'obtenir des résultats approchant au mieux la réalité.

V.1. Objectifs et hypothèses générales des modèles prédictifs

V.1.1. Objectifs

Les modèles prédictifs étudient les faits (i.e. données) présents et passés afin de faire des hypothèses sur des événements futurs. L'étude des données permet d'obtenir des estimations relatives aux comportements des individus et de leurs caractéristiques. L'objectif de ce type de modélisation est alors double. On distingue la visée explicative et la visée prédictive.

Ces modèles permettent d'une part d'étudier et de quantifier l'effet des variables du modèle sur la variable d'intérêt : c'est une étude à visée explicative. Les valeurs des autres variables sont prises en compte dans les estimations et les effets estimés sont dits « ajustés » par les autres variables. D'autre part, ces modèles permettent de prédire la probabilité de survenue d'un événement d'intérêt en fonction des variables explicatives : c'est la visée prédictive. Selon la visée de l'étude, la méthodologie de sélection des variables à prendre en considération dans le modèle ne sera pas la même. Nous reviendrons sur la sélection des variables afin de parvenir au modèle final dans la partie V.4.1.

Ces modèles s'appliquent dans de nombreux domaines. En médecine, ils peuvent s'utiliser pour évaluer les probabilités d'apparition d'une maladie chez un patient selon un ensemble de caractéristiques (i.e. visée prédictive) ou pour étudier le risque apporté par chacune (i.e. visée explicative). Dans le cadre de l'évaluation économique, ces modèles s'implémentent principalement pour étudier les variations de coûts relatives aux caractéristiques des individus (i.e. visée explicative).

Les modèles statistiques prédictifs utilisés sont des modèles de régression qui impliquent des variables explicatives et une variable à expliquer (ou « dépendante », ou « réponse »). Ainsi, les variables explicatives vont permettre de décrire les variations et de prédire la variable à expliquer. Les modèles de régression sont régis par un ensemble d'hypothèses fondamentales que l'utilisateur se doit de respecter et de vérifier pour s'assurer une modélisation valide menant à des estimations

pertinentes. L'étude des coûts brise certaines de ces hypothèses et des modèles particuliers doivent alors être mis en œuvre.

V.1.2. Modèle linéaire simple et hypothèses fondamentales

Soient Y , la variable à expliquer, et X le vecteur associé à la variable explicative. Nous nous plaçons ici dans le cas de la régression linéaire simple (i.e. univariée), mais les hypothèses sont les mêmes pour la régression linéaire multiple (i.e. multivarié).

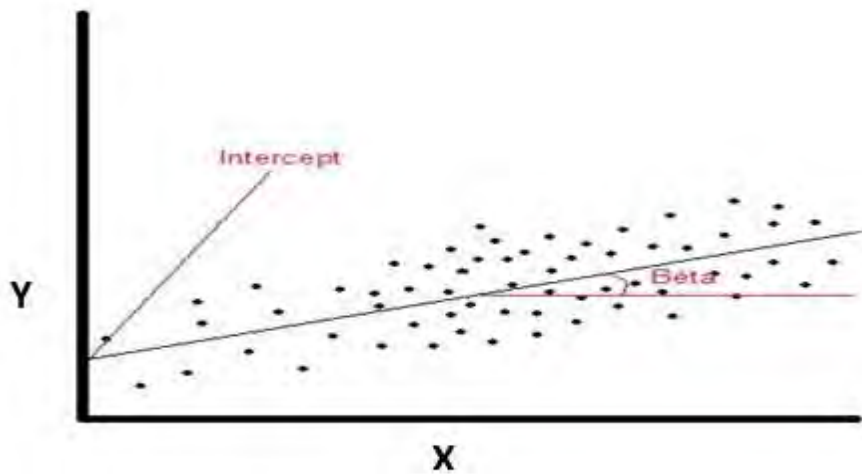
Equation relative au modèle linéaire simple :

$$Y_i = \beta X_i + \varepsilon_i = \beta_0 + \beta_1 X_{i,1} + \varepsilon_i$$

avec β les coefficients de régression tel que $\beta = (\beta_0, \beta_1)$, $X_i = (1, X_{i,1})^t$ et Y_i sont respectivement les valeurs de la variable explicative et à expliquer, ε_i les résidus du modèle (i.e. erreur aléatoire) et distribués selon une loi normale telle que $\varepsilon_i \sim N(0, \sigma^2)$, \forall ⁶ individus i . Le β se divise, dans le cas de la régression univariée, en un intercept (i.e. β_0) et en un coefficient de régression représentant l'effet de la variable $X_{i,1}$ sur la variable Y_i (i.e. β_1). L'estimation des paramètres du modèle permet de tracer la droite de régression. Elle représente le lien linéaire entre l'espérance de la variable à expliquer et les variables explicatives de l'équation ci-dessus. L'intercept constitue l'ordonnée à l'origine de la droite de régression et le coefficient β_1 représente la pente de la droite de régression (figure 10). Les βX_i sont également appelés un/des prédictor(s) linéaire(s). On peut représenter l'équation ci-dessus par la figure suivante.

⁶ \forall : « pour tous »

Figure 10. Représentation de Y en fonction de X avec la droite de régression



Le nuage de point représente les valeurs de la variable Y en fonction des valeurs de X et la droite de régression l'estimation du modèle.

Les hypothèses devant être vérifiées pour implémenter un tel modèle sont les suivantes :

- Hypothèse d'indépendance : sachant X_i , les Y_i sont indépendamment distribués. $Y_i \perp Y_j | X_i, X_j, \forall i \neq j$. Il n'y a pas de corrélation entre les Y_i
- Hypothèse de linéarité : Il existe une relation linéaire entre Y et X telle que $E(Y_i | X_i) = \beta X_i$
- Hypothèse de la variance constante : La variance de chacun des Y_i est la même, quelles que soient les valeurs des X_i . On a : $Var(Y_i | X_i) = \sigma^2, \forall i$ (i.e. homoscedasticité).
- Hypothèse de normalité : Les Y_i sachant X_i sont distribués selon une loi normale.

Il est possible de combiner ces hypothèses en posant : « sachant X_i , les Y_i sont Indépendamment et Identiquement Distribués (IID) selon une loi normale de moyenne $\mu_{Y_i | X_i} = E(Y_i | X_i) = \beta X_i$ et de variance σ^2 » ou « $Y_i | X_i \sim N(\mu_{Y_i | X_i}, \sigma^2) iid$ ». Il est également possible de considérer les hypothèses associées aux résidus : les résidus ε_i doivent être *iid* selon une loi normale de moyenne nulle ($E(\varepsilon_i) = 0$) et de variance constante ($Var(\varepsilon_i) = \sigma^2$).

Les coefficients β estimés sont également appelés « effets fixes » car étant constants pour la population d'étude.

V.2. Prise en compte de la distribution des coûts

Comme vu précédemment, les coûts ne sont pas distribués normalement (i.e. non gaussien) (91). Dans le domaine de la santé notamment, les coûts relatifs aux consommations de soins sont généralement positifs ou strictement positifs et étirés vers des valeurs élevées. Par exemple, les individus en bonne santé peuvent n'encourir que peu de coûts tandis que les patients atteints d'une maladie particulière peuvent encourir des coûts médicaux extrêmement élevés (164). De plus, la variance des coûts est positivement corrélée avec sa moyenne : plus les coûts sont importants en moyenne, plus la variance associée est grande.

V.2.1. Coûts strictement positifs

La distribution particulière des coûts ne respecte pas certaines des hypothèses fondamentales des modèles de régression linéaire. La distribution représentant les données n'est plus adaptée : la loi normale est symétrique, avec une variance constante et peut prendre des valeurs négatives. De plus, l'homoscédasticité n'est plus respectée (i.e. hétéroscédasticité) et le lien entre les variables explicatives et la variable d'intérêt n'est plus linéaire.

Dans ce cadre, il convient de faire intervenir les modèles linéaires dits « généralisés » ou « Generalized Linear Model » (GLM). Ces modèles remplacent la distribution normale représentant la variable d'intérêt par une distribution adaptée (i.e. avec une meilleure adéquation aux données) et

d'utilisent une fonction de « lien » reliant la variable réponse (Y) aux prédicteurs linéaires (βX) construits à partir des variables explicatives.

Les lois de probabilité utilisées dans le cadre d'un GLM appartiennent généralement aux lois de la famille exponentielle définies par leur densité qui peut s'écrire sous la forme (165):

$$f(y; \theta, \varphi) = \exp\left[\frac{y\theta - b(\theta)}{a(\varphi)} + c(y, \varphi)\right]$$

$f(y; \theta, \varphi)$ représente la densité de probabilité pour une variable continue. $a()$, $b()$, $c()$ sont des fonctions connues, y la variable aléatoire, θ le paramètre naturel et φ le paramètre de dispersion.

Dans certains cas, φ est connu et il reste à trouver θ ; sinon $a(\varphi) = \frac{\varphi}{\omega}$, où ω est un poids fixé connu.

On a notamment $E(Y) = b'(\theta)$ et $Var(Y) = a(\varphi)b''(\theta)$ ⁷.

Soient μ l'espérance de Y et $\eta = X\beta$ le prédicteur linéaire. Le lien entre l'espérance de Y et les prédicteurs linéaires s'effectue à partir de la fonction g tel que : $\eta = g(\mu)$.

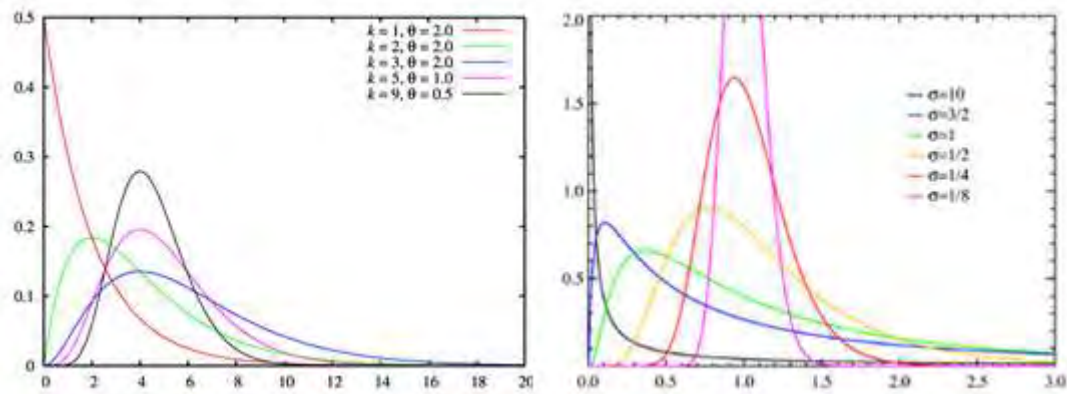
Une fonction de lien pour laquelle $\eta = \theta$, est appelée une fonction de lien canonique. On se retrouve alors avec :

$$E(Y) = b'(\theta) = b'(\eta) = b'(X\beta)$$

Les deux distributions relatives à l'étude des coûts les plus communément retrouvées dans la littérature internationale sont la loi gamma et la loi log-normal. Ces deux lois prennent en considération des valeurs strictement positives, sont étirées vers la droite, et ont une variance proportionnelle à la moyenne. La figure 11 représente les distributions de ces lois selon différentes paramétrisations (166) (167).

⁷ b' et b'' représentent les dérivés première et seconde de b

Figure 11 : Densité des lois gamma et log normales.



Selon les paramètres considérés, les courbes sont plus ou moins étirées vers la droite, représentant la faible proportion d'individus engendrant les coûts les plus importants. En pratique, les paramètres des lois sont estimés par le modèle à partir des données. L'utilisateur n'a pas besoin de les spécifier.

Nixon et Thompson en 2004 (168) comparent 4 modélisations distinctes sur 4 jeux de données représentant des coûts en matière de santé. Les 4 modélisations sont réalisées à partir des lois normales, log-normales, gamma et log-logistiques. Les auteurs concluent sur la nécessité de ne pas utiliser une simple loi normale afin de représenter les données. Les performances des autres modélisations dépendent ensuite des données utilisées. Le choix de la distribution représentant les données doit ainsi se faire avec précaution. Néanmoins, la loi gamma convient d'avantage aux différents jeux de données.

Fu et Moncher en 2004 (169) comparent les distributions log normales et gamma afin de modéliser des données économiques. D'après les simulations réalisées, les auteurs concluent en la supériorité générale de la distribution gamma qui permet d'obtenir des estimations moins biaisées, malgré des données en mauvaise adéquation avec la distribution sélectionnée.

Une fonction de lien est régulièrement utilisée afin de d'établir une relation linéaire entre l'espérance de la variable réponse et les prédicteurs linéaires (i.e. variables explicatives). Cette fonction de lien permet également de s'assurer que les valeurs prédites par le modèle respecteront

la nature d'origine (i.e. positivité pour les coûts) de la variable réponse. La fonction de lien logarithmique est donc souvent utilisée dans l'étude des coûts car cette fonction est définie sur l'ensemble des réels strictement positifs.

Une distribution gamma couplée a une fonction de lien logarithmique est ainsi régulièrement retrouvée dans la littérature afin d'évaluer les variations des coûts relatives aux différentes covariables. Cette loi s'adapte généralement bien à toutes les données économiques. Le modèle s'écrit alors :

$$g(Y_i) = \beta X_i + \varepsilon_i$$

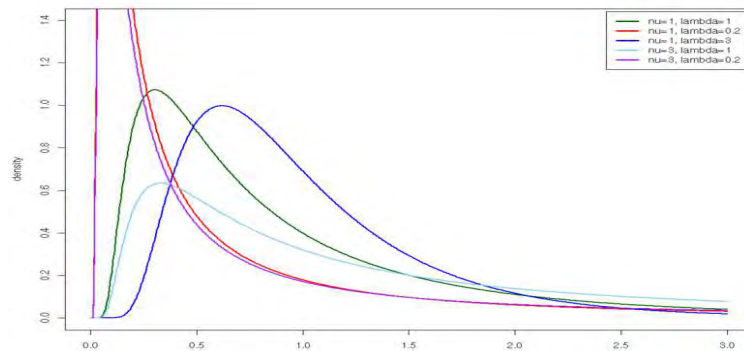
Ou :

$$g(E(Y_i)) = \beta X_i$$

Avec g , la fonction de lien, $g() = \log()$ et Y suivant une loi gamma.

Nous soulignons tout de même qu'un modèle s'attache à représenter les données. Si les modèles de régression gamma permettent généralement une bonne adéquation avec les données économiques, il se peut que d'autres distributions soient plus adaptées dans certains cas particuliers. Il en va de même pour l'utilisation de la fonction de lien logarithmique qui peut ne pas être adaptée si certaines covariables ne sont pas liées de façon log-linéaire avec la variable réponse (170). Comme fréquemment en statistiques, le choix final reviendra toujours à l'analyste et dépendra ainsi des données et des objectifs mais sera rarement généralisable. La distribution inverse-gaussienne par exemple peut permettre une meilleure représentation des observations lorsque les données de coûts sont très volatiles (171). La figure 12 représente la densité de la loi inverse-gaussienne (172).

Figure 12 : Densité de la loi inverse-gaussienne



Ces distributions permettent d'étudier des données strictement positives mais leurs usages rencontrent des difficultés lorsque la variable étudiée prend des valeurs égales à 0. Afin de prendre en compte ces valeurs nulles, deux méthodologies peuvent être employées : les « tweedie models » et les « Zero Adjusted GAMMA (ZAGA) models » ou « Zero Adjusted Inverse Gaussian (ZAIG) models » selon la distribution utilisée.

V.2.2. Coûts positifs

Certains patients sont dits « non consommateurs » et le coût associé à leur prise en charge peut ainsi être nul. Les GLM classiques avec une distribution gamma (i.e. log normale, inverse gaussienne, etc) ne sont alors plus utilisables car prenant en compte seulement des données strictement positives. Dans ce contexte, les coûts positifs et nuls peuvent être représentés respectivement comme une distribution combinant une partie discrète et une partie continue. La partie discrète représente les valeurs nulles (i.e. probabilité qu'une valeur soit nulle ou non) et la partie continue les valeurs positives.

V.2.2.1. Modèles Tweedies

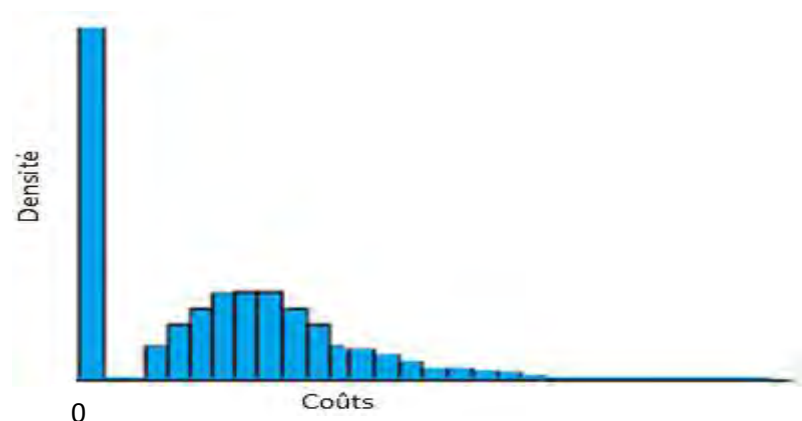
La modélisation de cette distribution à l'aspect discret et continu peut être réalisée par l'usage des modèles Tweedies. Ces modèles tirent leurs noms de la distribution de la variable réponse nommée « Tweedie distribution » et décrite pour la première fois par Tweedie en 1984 (173). Cette distribution particulière appartient à la famille des « Modèles de dispersion exponentielle » ou « Exponential dispersion models ». Ces distributions sont une sous-catégorie des lois de la famille exponentielle avec une condition sur le rapport entre la moyenne et la variance des données.

Soit Y la variable aléatoire représentant les coûts étudiés. Y s'écrit comme :

$$Y_i = \begin{cases} 0, & \text{si } M = 0 \\ X_1 + \dots + X_M, & \text{si } M > 0 \end{cases}$$

Avec M suivant une loi de Poisson⁸, les X_i suivant des lois gamma avec M et X_i indépendants. La somme de lois gamma étant une loi gamma, Y suit alors une loi de poisson-gamma. Si $M = 0$, $Y = 0$, ce qui représente la probabilité que les coûts soient nuls (174). Si $M > 0$, Y , conditionnellement à M , suit une loi gamma représentant les coûts étudiés.

Figure 13. Histogramme représentant une distribution Tweedie



⁸ La loi de Poisson est une loi de probabilité discrète qui décrit le comportement du nombre d'événements se produisant dans un intervalle de temps fixé.

La principale limite de ces modèles est de ne pas permettre l'estimation de façon concomitante de la probabilité que les coûts soient nuls et des valeurs des coûts non nuls (175). La probabilité au niveau des valeurs nulles ne peut pas être caractérisée par une fonction des variables explicatives. Ce qui devient problématique dans les cas où les valeurs nulles s'expliquent par une caractéristique des individus qui est comptabilisée dans le modèle de régression (i.e. patient non exposé à une comorbidité coûteuse, patient jeune, etc..).

V.2.2.2. Modèles ZAGA (i.e. ZAIG)

Pour plus de clarté, nous présentons seulement les modèles ZAGA dans cette partie car uniquement la distribution diffère dans le cadre des modèles ZAIG. Les modèles ZAGA modélisent les valeurs nulles et positives de la variable à expliquer simultanément et permettent ainsi d'ajuster la probabilité qu'une valeur soit nulle par les covariables du modèle. La distribution discrète des valeurs nulles est représentée par une loi de Bernoulli et les valeurs continues par une distribution gamma.

Soit C_i une variable aléatoire suivant une loi de Bernoulli, égale à 0 avec une probabilité $1 - \pi_i$ et supérieure à 0 avec une probabilité π_i .

$$C_i \begin{cases} = 0 & \text{avec une probabilité } 1 - \pi_i \\ > 0 & \text{avec une probabilité } \pi_i \end{cases}$$

Soit $g(C_i)$, la densité de la distribution des données strictement positives et étirées vers la droite (i.e. gamma ou inverse-gaussienne) (176). La variable réponse Y_i suit une loi de probabilité discrète et continue:

$$Y_i = f(C_i) = \begin{cases} 1 - \pi_i, & \text{si } C_i = 0 \\ \pi_i g(C_i), & \text{si } C_i > 0 \end{cases}$$

La variable réponse est ensuite mise en lien avec les prédicteurs linéaire (i.e. $\beta_i X_i$) et contrairement aux modèles Tweedie, la probabilité qu'une valeur soit nulle, est évaluée en fonction des covariables.

Soit $g(C_i)$ la densité de la loi gamma :

$$g(C_i) = \frac{C_i^{\left(\frac{1}{\sigma^2}-1\right)} \exp\left(-\frac{C_i}{\sigma^2\mu_i}\right)}{(\sigma^2\mu_i)^{\frac{1}{\sigma^2}} \Gamma\left(\frac{1}{\sigma^2}\right)}$$

Avec σ un scalaire et μ_i l'espérance de la loi gamma.

L'espérance et la variance de la distribution ZAGA sont définies comme:

$$E(Y_i) = \pi_i \cdot \mu_i$$

$$Var(Y_i) = \pi_i \cdot \mu_i^2 (\pi_i + \sigma^2)$$

Et les covariables sont intégrées via l'expression logit⁹ suivante:

$$\pi_i = \frac{\exp(x_i^T \beta_\pi)}{1 + \exp(x_i^T \beta_\pi)}$$

Et la relation :

$$\mu_i = \exp(x_i^T \beta_\mu)$$

Où les β_π et les β_μ sont les coefficients de regression pour les π_i et les μ_i , et x_i le vecteur des variables explicatives.

Ces modèles nécessitent un temps de calcul relativement important dû au nombre de paramètres à estimer. Ils permettent d'ajuster la probabilité d'avoir un coût nul en fonction des covariables et ainsi d'obtenir des mesures de risque pour chacune des modalités des caractéristiques des individus (176).

⁹ $Logit(p) = \ln\left(\frac{p}{1-p}\right)$

V.3. Prise en compte des données corrélées

Une des hypothèses fondamentales des modèles de régression est l'indépendance des Y_i , c'est-à-dire que les différentes valeurs de la variables réponse Y ne sont pas corrélées entre elles. Il est pourtant fréquent d'étudier des données mesurées de façon répétée dans le temps, ou des données provenant de différents groupes échantillonnés. Ainsi, il y a de grandes chances qu'au sein d'un même individu durant la période totale de suivi ou bien pour les différents groupes, les valeurs de la variable d'intérêt soient corrélées.

Par exemple, il est commun d'analyser les coûts associés à une pathologie de manière longitudinale, notamment dans le cas d'une maladie chronique, afin de visualiser l'effet du temps sur les coûts. Ainsi, chaque individu aura une mesure de coût à chaque période de temps définie. Les mesures des coûts associés à un même individu sont ainsi corrélées entre elles.

De même, il est régulièrement nécessaire d'étudier les variations de coûts relatives à une maladie chez des patients recrutés dans différents pays, régions ou centres de soins. Sachant que les systèmes de santé diffèrent d'un pays à l'autre, les références de traitement et les habitudes peuvent varier d'une région ou d'un centre de soins à l'autre, il est rationnel de penser que les coûts relatifs à la maladie pour les individus dans un même pays, région ou centre de soins soient corrélés entre eux.

Soient i un individu et j une période de temps considérée, ou bien, i un sous-groupe de la population, et j les individus de ce sous-groupe avec $i = 1 \dots K$ et $j = 1 \dots n_i$. K représente le nombre de groupe dans les données et n_i le nombre d'observation dans chaque groupe.

On a:

$$Y = Y_{i,j} = Y_{1,j} \dots Y_{K,j} = Y_{1,1}, \dots, Y_{1,n_1}, Y_{2,1}, \dots, Y_{2,n_2}, \dots, Y_{K,1}, \dots, Y_{K,n_K}$$

Les valeurs $Y_{i,j}$ sont alors corrélées entre elles et l'hypothèse d'indépendance des Y n'est plus respectée.

Deux types de modèle permettent de prendre en compte cette corrélation : les modèles marginaux (i.e. modèles en population générale, « population average models ») et les modèles mixtes (i.e. modèles à effets aléatoires ou multiniveaux). Ces deux types de modèle ont des méthodologies différentes. Les modèles marginaux modifient la procédure d'estimation des coefficients de régression tandis que les modèles mixtes modifient directement l'écriture du modèle. Ces modèles s'appliquent également dans le cadre des GLM, avec une fonction de lien et une distribution non normale.

V.3.1. Modèles de population générale

Les modèles de population générale ont été développés par Liang et Zeger en 1986 (177). L'écriture de ces modèles est identique aux modèles linéaires classiques ou généralisés mais la procédure d'estimation diffère afin de tenir en compte des données corrélées. Les auteurs introduisent le concept des « Equations d'Estimation Généralisées » (« Generalized Estimating Equation » (GEE)) qui remplace les équations dites de « scores » (i.e. équation à résoudre afin d'obtenir les estimateurs du maximum de vraisemblance des coefficients β) sous l'indépendance. Les GEE font intervenir une matrice de corrélation de travail qui corrige les inférences selon la corrélation au sein de la variable à expliquer. De plus, Liang et Zeger proposent d'estimer la variance des paramètres de régression estimés à l'aide d'un estimateur « sandwich » robuste. Ces estimateurs sont convergents malgré une mauvaise spécification de la matrice de corrélation de travail (178).

L'écriture du modèle et la procédure d'estimation (GEE) appliquées aux GLM sont détaillées en annexe 7 (179).

La matrice de corrélation doit être spécifiée par l'utilisateur qui devra la sélectionner selon sa perception des corrélations dans les données. Il existe différentes matrices de corrélation. Les plus couramment utilisées sont :

- La matrice de corrélation échangeable : $corr(Y_{i,j}, Y_{i,j'}) = \alpha, \forall j \neq j'$.

Les données sont identiquement corrélées pour tous les groupes. C'est la matrice de corrélation la plus simple qui s'utilise lorsque qu'il n'y a pas de logique remarquable dans les corrélations. On l'utilise également lorsque les données sont composées de peu de groupes mais de beaucoup d'individus dans chaque groupe.

- La matrice de corrélation non structurée : $corr(Y_{i,j}, Y_{i,j'}) = \alpha_{jj'}, \forall j \neq j'$.

Cette matrice estime toutes les corrélations entre les individus et les inclut dans les estimations des variances. Ainsi, chaque paire d'observations aura sa propre corrélation. Il se peut que la matrice estimée ne soit pas inversible ce qui peut causer des problèmes lors des calculs numériques.

- La matrice de corrélation autorégressive : $corr(Y_{i,j}, Y_{i,j'}) = \alpha^{|j-j'|}, \forall j \neq j'$.

Cette matrice associe une corrélation plus importante aux mesures proches par rapport aux mesures éloignées. Elle s'utilise lorsqu'on peut ordonner les valeurs de la variable à expliquer dans chaque groupe. C'est le cas avec des données longitudinales où plusieurs mesures sont recensées au cours du temps sur un même individu.

- Pas de matrice de corrélation : $corr(Y_{i,j}, Y_{i,j'}) = 0, \forall j \neq j'$.

On se retrouve exactement dans le cadre d'un modèle linéaire (généralisé ou non) traditionnel. La matrice de corrélation utilisée est la matrice identité.

Ces modèles sont faciles à implémenter avec le logiciel R. La fonction « *geeglm* » du package « *geepack* », s'utilise de manière identique à un modèle linéaire traditionnel avec en complément, la spécification d'une matrice de travail et de la variable créant la corrélation des données à expliquer (180).

Les modèles marginaux permettent de traiter des données corrélées en corrigeant les inférences d'un modèle à partir d'une structure de corrélation à spécifier. Cette approche ne modélise pas la corrélation entre les sujets des différents groupes et donc n'effectuent pas des prédictions pour

chaque groupe. Ils permettent alors d'évaluer l'effet des variables considérées globalement dans toute la population d'étude (i.e. « population average »).

Une autre méthodologie permet de prendre en considération des données corrélées : les modèles mixtes. Nous reviendrons sur les avantages et inconvénients des modèles marginaux lors de leur comparaison avec les modèles mixtes.

V.3.2. Modèles mixtes

L'étude de la variabilité des données est essentielle dans un modèle statistique. Les modèles mixtes constituent une méthodologie plus élaborée afin de l'étudier. Ils introduisent le concept d'effets aléatoires.

L'exemple suivant a pour objectif d'appréhender plus facilement la différence entre les effets fixes et les effets aléatoires. Soit un échantillon de 90 patients sur lesquels on souhaite tester, l'effet de 3 médicaments sur la maladie. On sépare les 90 individus en trois groupes de 30 auxquels on associe un médicament différent. Trois mesures dans le temps sont ensuite réalisées pour chaque individu afin d'étudier l'évolution de l'effet des médicaments. Deux facteurs peuvent alors affecter les résultats observés: l'effet du type de médicament administré ainsi que l'individu concerné. L'effet du médicament sur la maladie est un effet fixe car les médicaments étudiés et utilisés dans l'étude sont contrôlés. Néanmoins, les 30 individus sont un échantillon de la population générale et il est nécessaire de prendre en compte dans les inférences la variabilité de l'effet des médicaments due aux patients. Cette variabilité représente une des composantes inconnues de la variabilité totale. Le facteur « patient » engendrant cette variabilité est alors considéré comme un facteur à effet aléatoire.

Les modèles mixtes sont appelés ainsi car ils combinent les effets fixes et les effets aléatoires. Les effets aléatoires permettent de prendre en compte les composantes non observées dans les données et communes à plusieurs observations (i.e. variabilité intra groupe). Selon l'exemple précédent, les composantes non observées seraient les caractéristiques inconnues des patients pouvant affecter les effets estimés des médicaments. La variabilité totale de la variable d'intérêt est alors séparée en deux: celle due aux effets aléatoires (différente pour chaque patient) et celle que l'on affecte aux erreurs (identique pour chaque patient). Contrairement aux modèles à effets fixes, les modèles mixtes modélisent explicitement la variabilité inobservée dans les groupes, relative aux caractéristiques communes de chaque groupe. Ils fournissent ainsi des interprétations plus riches des résultats (181). L'hétérogénéité dans les données au sein des groupes est alors représentée par une loi de probabilité.

La formulation générale d'un modèle linéaire mixte est (182):

$$Y_{i,j} = \beta X_{i,j} + b_i Z_{i,j} + \varepsilon_{i,j}$$

Avec $\varepsilon_{i,j} \sim N(0, \sigma_e^2)$ et $b_i = (b_{0,i}, b_{1,i})^T \sim N(0, B)$ et B la matrice de variance covariance des b_i . Ou $X_{i,j}$ est le vecteur des variables explicatives, β est le vecteur des effets fixes, $Z_{i,j}$ est un sous vecteur de $X_{i,j}$. Pour rappel, $i = 1 \dots K$ et $j = 1 \dots n_i$. K étant le nombre de groupe.

Les b_i représentent les effets aléatoires et modélisent l'hétérogénéité des valeurs de la variable à expliquer dans chacun des groupes. Ne pas prendre en compte cette hétérogénéité apporterait des estimations biaisées. De plus, les effets aléatoires permettent de mesurer les effets relatifs aux groupes et de considérer les paramètres de « nuisance » (i.e. composantes non observées apportant l'hétérogénéité) comme un objet d'intérêt en tant que tel. Les modèles mixtes posent l'hypothèse que les effets aléatoires sont indépendants des variables explicatives (181).

L'annexe 8 présente graphiquement les explications ci-dessus et les deux principaux types de modèles mixtes : les modèles mixtes à intercepts aléatoires et les modèles mixtes à intercepts et pentes aléatoires (183).

Les modèles mixtes nécessitent la spécification de la distribution conjointe de la variable réponse, c'est-à-dire à la fois la spécification de la distribution conditionnelle de la variable à expliquer par rapport aux effets aléatoires (i.e. $f(Y_{i,j}|b_i)$) et la distribution à priori des effets aléatoires (i.e. b_i). Ainsi, les inférences dans ces modèles sont complexes car le nombre de paramètre à estimer est important. Les effets aléatoires modélisent directement la corrélation dans les groupes considérés et permettent de faire des inférences conditionnellement aux groupes. On appelle également ces modèles des modèles conditionnels ou « subject specific models » en opposition aux modèles marginaux ou « population average model » vus précédemment.

Deux approches permettent de considérer la non indépendance des valeurs de la variable à expliquer. Nous allons les mettre en comparaison afin de permettre à l'utilisateur de choisir la méthode la plus adaptée à sa problématique.

V.3.3. Comparaison entre les deux approches

Les deux approches proposées pour étudier des données corrélées amènent à des résultats sensiblement identiques mais qui peuvent avoir des interprétations différentes. Nous nous plaçons ici dans le cadre des GLM (i.e. fonction de lien et distribution non normale).

Concernant les modèles mixtes, l'espérance de la variable à expliquer est définie conditionnellement aux groupes (i.e. effets aléatoires). Dans le cadre des modèles mixtes à intercept aléatoire, on a (184):

$$E(Y_{i,j}|b_i) = g'(X_{i,j}^T\beta^* + b_i)$$

Avec g' la dérivée de la fonction de lien canonique g associée à la distribution des $Y_{i,j}$.

Dans le cadre des modèles marginaux, l'espérance de la variable à expliquer est évaluée sur toute la population (i.e. marginale). On a :

$$E(Y_{i,j}) = g'(X_{i,j}^T \beta)$$

Si l'hypothèse d'indépendance est respectée, la variance des b_i est nulle et les paramètres β^* et β sont égaux. Les GLM sont un sous ensemble des modèles mixtes où les effets aléatoires ont une variance nulle.

Dans les autres cas (i.e. indépendance non respectée) et si g' n'est pas une fonction linéaire des b_j , on a (184):

$$E(Y_{i,j}) = E\left(E(Y_{i,j} | b_j)\right) = E(g'(X_{i,j}^T \beta + b_j)) \neq g'(X_{i,j}^T \beta)$$

Et les paramètres β^* et β sont différents.

Dans le cadre des modèles mixtes, l'estimation d'un coefficient de régression β^* mesure l'effet moyen d'une variable explicative, les autres variables explicatives et la valeur de l'effet aléatoire étant maintenues constantes. L'interprétation des β^* est conditionnelle aux valeurs des variables explicatives et des effets aléatoires. Les effets sont alors mesurés conditionnellement au sous-groupe considéré.

Dans le cadre des modèles marginaux, l'estimation d'un des coefficients de régression β mesure l'effet moyen d'une variable explicative, les autres variables explicatives étant maintenues constantes. Les effets sont alors mesurés sur l'ensemble de la population. La valeur de β dépend des intensités de liaison entre les variables explicatives et la variable réponse et du degré d'hétérogénéité dans la population (184).

Le tableau suivant, inspiré de Hubbard et al, résume les principales différences entre les utilisations des modèles de populations générales et les modèles mixtes (185).

Tableau 3. Différences entres modèles de populations générales et les modèles mixtes.

Population average model	Mixed model
Applicable aux données longitudinales et corrélées	Applicable aux données longitudinales et corrélées
Se focalise sur les coefficients de régression	Se focalise sur les coefficients de régression et sur les composantes de la variance.
Evalue la modification de la variable à expliquer selon la modification d'une unité de la variable explicative, dans toute la population	Evalue la modification de la variable à expliquer selon la modification d'une unité de la variable explicative, selon le groupe considéré
Besoin de spécifier une matrice de corrélation	Besoin de spécifier la distribution des effets aléatoires (car non observés) et la distribution conditionnelle de la variable à expliquer
Inférences simples et convergentes malgré une spécification de la matrice de corrélation de travail erronée (variance des paramètres estimés à l'aide d'un estimateur de variance sandwich robuste)	Inférences complexes
	Permet de faire des prévisions sur un groupe donné
Nécessite un nombre de groupe important	Nécessite un nombre de groupe moindre
Mauvaise prise en compte des données manquantes aléatoires	Bonne prise en compte des données manquantes aléatoires
Variables explicatives ne dépendants pas du temps	Variables explicatives dépendantes du temps

Les modèles mixtes impliquent des hypothèses invérifiables sur la distribution des effets aléatoires et leur indépendance avec les variables explicatives. Dans ce cadre, ces modèles peuvent amener à des estimations trompeuses avec des inférences biaisées (181). Ces biais peuvent également provenir de la spécification de la distribution de la variable à expliquer par rapport aux effets aléatoires (i.e. $Y_{i,j}|b_i$) et des effets aléatoires (i.e. b_i) qui sont à priori non observés (185). Les modèles marginaux nécessitent uniquement la spécification d'une matrice de corrélation de travail. De plus, malgré une spécification erronée de cette matrice, les inférences sur les paramètres seront tout de même valides. Ainsi les modèles marginaux donneraient des meilleures approximations de la réalité (184) (185).

Ces propos sont néanmoins à nuancer car les estimateurs des paramètres de régression des modèles mixtes ne sont que peu biaisés lorsque la structure de covariance ou la distribution des effets aléatoires est mal spécifiée (184) (186) (187).

De plus, certains auteurs préconisent l'usage des modèles conditionnels plutôt que des modèles marginaux notamment pour éviter trois types de problèmes (188).

Le premier provient des variables omises ayant un effet spécifique pour certains groupes. Dans le cadre des modèles marginaux, les inférences sont réalisées sur l'ensemble de la population. L'hétérogénéité des effets des variables non incluses associée aux groupes n'est ainsi pas prise en compte lors des estimations. Dans le cadre des modèles mixtes, les inférences sont réalisées pour chaque groupe et les effets associés aux variables omises sont pris en compte dans les estimations des effets aléatoires (i.e. b_i), relativement à chaque groupe. Ainsi, les modèles marginaux ne capturent pas la variabilité des effets spécifiques aux groupes. Ce problème s'observe notamment lorsque les variables explicatives incluses dans le modèle sont dépendantes du temps. Les effets des variables omises sur ces variables à travers les différents groupes ne sont pas contrôlés dans les modèles marginaux (189) (190).

Le deuxième problème provient du paradoxe de Simpson. Ce paradoxe représente l'inversion de l'effet d'une variable lorsque l'analyse est effectuée dans chaque groupe ou si elle est réalisée sur l'ensemble de la population d'étude. Les modèles marginaux estiment l'effet des paramètres du modèle globalement et ne sont pas aptes à visualiser les effets dans chacun des groupes. Le troisième problème provient des erreurs de mesure de Berkson. Ce sont des erreurs aléatoires dites de sélections. Elles sont notamment présentes en épidémiologie lorsque les différents groupes comparés sont exposés différemment à un facteur d'exposition. De façon identique au paradoxe de Simpson, les modèles marginaux ne prendront pas en compte les possibles différences d'exposition entre les groupes.

Les modèles mixtes nécessitent des calculs complexes car ils estiment un grand nombre de paramètres (i.e. effets aléatoires et paramètres associés). Ils permettent cependant une meilleure

gestion des données manquantes. En effet, les estimateurs à partir des GEE convergent seulement si les données manquantes sont complètement aléatoires¹⁰ tandis que les modèles mixtes prennent en compte les données manquantes aléatoires¹¹ (184). Les modèles marginaux sont également plus sensibles au nombre de groupes et d'individus dans chaque groupe dans l'étude (191).

Les modèles marginaux (« population average ») et conditionnels (« subject specific ») constituent deux approches distinctes afin de travailler sur des données corrélées. Ces modèles ont tous les deux des avantages et des inconvénients et nous retiendrons principalement leurs différences d'interprétation et donc d'application. Il est nécessaire de sélectionner le plus adapté aux données et aux objectifs visés. Les modèles mixtes seront ainsi plutôt conseillés dans une optique explicative visant à évaluer l'effet de chaque variable conditionnellement aux effets aléatoires (i.e. aux groupes). En effet, ces modèles réalisent des inférences conditionnelles en modélisant la corrélation dans les groupes considérés. Les modèles marginaux seront plutôt conseillés dans une optique pragmatique visant à évaluer l'effet moyen des variables explicatives sur la variable à expliquer globalement dans la population considérée et non pas les effets sur les groupes. Ces modèles permettent de réaliser des inférences marginales (184).

Dans le cas d'un modèle log-linéaire, les effets estimés sont sensiblement différents et la distinction entre les interprétations est d'autant plus importante (192). C'est notamment le cas lorsque l'on travaille avec des coûts et une fonction de lien logarithmique.

Par exemple, si on veut étudier l'effet d'une complication sur les coûts d'une maladie au cours du temps, les modèles marginaux permettent d'évaluer l'effet moyen de la complication sur le coût de la maladie globalement au cours du temps.

Par exemple, si on veut étudier l'effet de covariables sur les coûts d'une pathologie dont les patients ont été recrutés dans plusieurs pays, les modèles mixtes seront mieux adaptés car ils permettent d'évaluer l'effet de chaque covariable sur la variable d'intérêt dans chacun des pays.

¹⁰ MCAR: Missing Completely at Random.

¹¹ MAR: Missing At Random

Les modèles permettant de prendre en compte des données corrélées peuvent être combinés avec les modèles Tweedies (voir V.2.2.1.) ou des modèles ZAGA (i.e. ZAIG) (voir V.2.2.2). Dans le cadre des modèles Tweedies, Swan T. combine ces modèles avec les GEE des modèles marginaux et Yang & al les utilisent avec un modèle mixte (193) (194). Dans le cadre des modèles ZAGA (i.e. ZAIG), Lei & al ont développé un modèle combinant les modèles mixtes à ces distributions (195).

Sélectionner une modélisation adéquate aux données et aux objectifs de l'étude ne suffit pas pour obtenir des résultats pertinents. Il est essentiel de parvenir à un modèle final répondant à la visée de l'étude et de sélectionner correctement les variables à inclure. Une étape de vérification des hypothèses des modèles est également nécessaire.

V.4. Sélection des variables et adéquation du modèle

La problématique associée à la sélection des variables est commune à tous les modèles de régression et s'applique donc aux modèles présentés ci-dessus. Cette sélection est particulièrement importante car elle permet d'obtenir un modèle final répondant correctement aux objectifs de l'étude. Parallèlement à cette sélection des variables, il est primordial de vérifier que les hypothèses des modèles sont respectées pour s'assurer une modélisation valide menant à des estimations cohérentes. Autrement, le modèle peut mener à des tests et estimations incorrectes et donc à des conclusions erronées.

V.4.1. Sélection des variables

L'ensemble des variables disponibles ne seront pas utilisées dans l'analyse des données. Il est nécessaire de faire des choix parmi les variables explicatives afin de parvenir à un modèle final permettant de tirer des conclusions correctes sur les données et répondant à l'objectif visé. Ce

modèle final doit conserver un nombre cohérent de variables explicatives significatives afin d'expliquer la variable réponse tout en prenant compte au mieux les potentiels facteurs de confusion.

Trois procédures automatiques sont notamment retrouvées dans la littérature : la procédure d'élimination progressive (i.e. « backward selection »), d'introduction progressive (i.e. « forward selection ») et de sélection pas à pas (i.e. « stepwise selection »). Elles se basent soit sur le modèle complet (i.e. avec toutes les variables explicatives) et en retire à chaque étape selon la significativité¹² de chacune (backward selection), soit sur un modèle univarié en ajoutant (forward selection) ou en ajoutant et remplaçant (stepwise selection) à chaque étape des variables. La sélection ou le retrait d'une variable s'effectue selon son seuil de significativité et selon la qualité du modèle mesuré généralement à partir du critère d'information d'Akaike¹³, « Akaike Informal Criterion »(AIC).

La sélection ou l'élimination de variables dans ces méthodes se basent sur des critères statistiques, sans tenir compte du contexte social, épidémiologique et de l'objectif de l'étude. Les variables sélectionnées ne sont pas toujours les plus aptes à décrire et expliquer les données mobilisées en vue de répondre à l'objectif fixé. Les différentes procédures automatiques des variables sont alors à bannir, notamment dans le cadre de la santé publique, faisant intervenir des paramètres socioculturels, psychosociaux ou physiques (184). Comme indiqué dans la section V.1.1., la visée de l'étude (i.e. explicative ou prédictive) détermine comment sélectionner les variables afin d'atteindre le modèle final.

Si l'objectif de l'étude est d'étudier l'effet d'une variable explicative sur la variable réponse (i.e. visée explicative), il est essentiel d'examiner les relations entre la variable à expliquer et les covariables ainsi que les relations entre covariables. L'étude de ces liens permet de prendre en compte les biais de confusion identifiables. Il est également crucial de détecter les colinéarités entre les variables

¹² Significativité : dépend de la valeur de la p-value et du seuil de probabilité choisi

¹³ AIC : Mesure la qualité d'un modèle à partir de la valeur de la vraisemblance utilisée dans les estimations.

explicatives, c'est-à-dire les variables ayant des significations proches. La colinéarité peut fausser les résultats des tests, les estimations des coefficients et les conclusions de l'étude (196).

Si l'objectif de l'étude est de prédire la survenue d'un évènement d'intérêt en fonction d'un ensemble de variables explicatives (i.e. visée prédictive), il est nécessaire de sélectionner les variables les plus pertinentes et discriminantes tout en trouvant l'équilibre entre complexité et reproductibilité des prédictions (184). A titre d'exemple, cette modélisation permet d'obtenir des probabilités de transition d'un état à un autre qui s'utilisent généralement dans les modèles de décision. L'objectif est de produire une prédiction proche de la réalité. Les préoccupations dues aux effets des variables explicatives colinéaires, non significatives ou des biais de confusion seront ainsi moindres. L'analyste doit sélectionner des variables pertinentes suivant ses hypothèses et transposables pour d'autres populations et d'autres analyses.

Il est recommandé en pratique d'étudier de façon univariée l'impact de chacune des variables explicatives sur la variable réponse ainsi que les variables explicatives sur la variable explicative d'intérêt (celle dont on veut connaître l'effet). Un seuil moins restrictif de p-value à 20% est souvent utilisé pour conserver une variable dans l'analyse multivariée. Néanmoins, certaines variables considérées comme déterminantes doivent être conservées dans les modèles malgré leur non significativité. Elles peuvent être cliniquement importantes, ou apporter de la confusion sur les autres variables observées ou non (i.e. âge, sexe). La matrice de corrélations des variables explicatives permet de visualiser quelles variables sont liées entre elles. L'étude des modifications de l'effet des variables (coefficients β et p-value) lors de l'ajout ou de l'élimination d'une autre variable permet d'identifier leurs impacts sur les autres paramètres. Si les coefficients et p-value sont grandement modifiés, l'analyste doit se questionner sur la pertinence de l'ajout ou de l'élimination de cette variable.

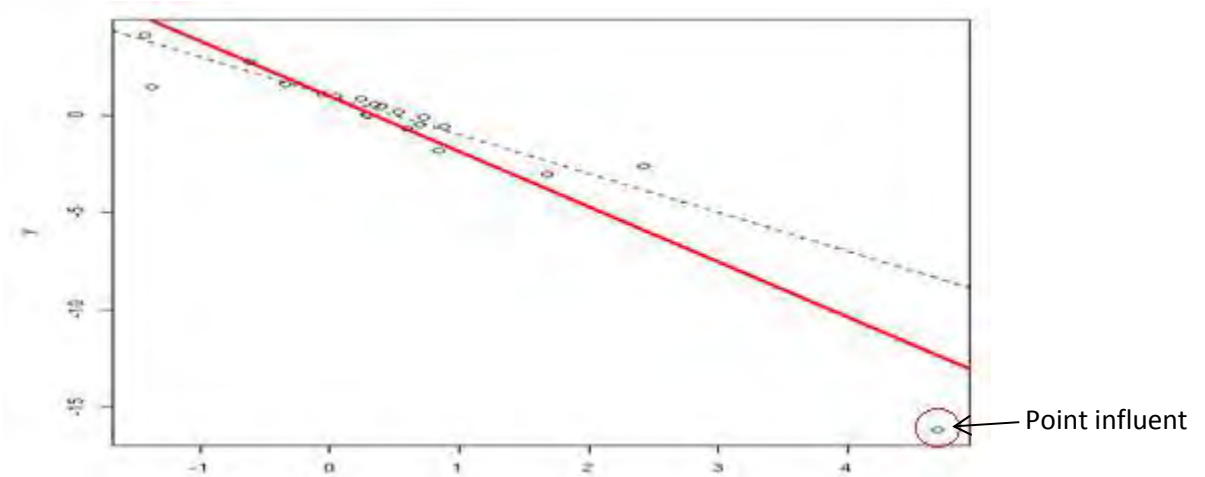
V.4.2. Vérification de l'adéquation du modèle

Les modèles de régression sont régis par un ensemble d'hypothèses fondamentales que l'utilisateur se doit de respecter.

A priori le processus de sélection des variables, l'étude graphique de la variable à expliquer (histogramme, QQ-plot...) ou à partir de tests statistiques (Shapiro-Wilk, Kolmogorov-Smirnov, ...) permet de choisir une loi en adéquation aux données. L'hypothèse de linéarité entre une variable explicative quantitative et la variable à expliquer peut se vérifier à partir du nuage de point entre elles. Ce graphique permettra d'identifier le comportement de la variable explicative par rapport à la variable étudiée. Il sera parfois nécessaire d'utiliser des transformations sur la variable explicative afin d'obtenir une meilleure linéarité. De plus, il est possible de transformer les variables quantitatives en variables qualitatives (ordonnées ou non) afin de passer outre la non linéarité des relations entre variables.

Certaines valeurs extrêmes ou « points influents » peuvent fausser les estimations dans le modèle. Elles peuvent d'une part prendre des valeurs importantes mais être valides ou d'autre part être aberrantes et provenir d'erreurs de mesure. Ces valeurs extrêmes peuvent alors exercer un effet de levier comme explicité sur la figure 14. Si ces valeurs sont très peu nombreuses, il est conseillé de les enlever de l'étude, dans l'autre cas, il sera nécessaire de transformer la variable.

Figure 14. Point influent et effet de levier



La figure ci-dessus représente en rouge la droite de régression estimées avec un point influent et en pointillée celle en le supprimant. La présence du point influent fausse les estimations dans le modèle.

La plupart des hypothèses des modèles linéaires simples se vérifient graphiquement notamment par l'étude des résidus.

Les résidus sont définis comme :

$$\varepsilon_i = Y_i - \beta X_i = Y_i - E(Y_i|X_i)$$

Ils ne sont pas observés mais sont estimés dans le modèle :

$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta} X_i$$

Si le modèle est en bonne adéquation avec les données alors les résidus estimés devraient satisfaire les hypothèses fondamentales des modèles de régressions. L'étude des résidus va permettre de vérifier si la normalité et l'homoscédasticité des Y est respectée et la provenance de l'hétéroscédasticité dans les variables explicatives. L'annexe 9 présente divers outils graphiques pour diagnostiquer un modèle linéaire simple.

Dans le cadre des GLM, les résidus ne sont pas systématiquement Gaussiens. Leur étude n'apporte pas d'information et il est nécessaire d'étudier les résidus « standardisés » définis par:

$$\hat{\varepsilon}_{pearson,i} = \frac{Y_i - \hat{\beta}X_i}{\sqrt{V(\hat{\beta}X_i)}}$$

Où $V(\hat{\beta}X_i) = \sigma^2 \times (1 - h_{ii})$, avec h_{ii} le levier associé à l'observation i . Le levier représente l'influence de l'observation i sur la i -ième valeur prédite. On a $\hat{Y}_i = \hat{\beta}X_i = \sum_{j=1}^n h_{ij} Y_j$.

Soit la statistique de Pearson :

$$X_p^2 = \sum (\hat{\varepsilon}_{pearson,i})^2$$

Et la déviance définie comme la variation de la log-vraisemblance entre le modèle saturé¹⁴ et le modèle final.

$$D = 2 * [L_{MS} - L_{MF}] = 2 * [L(Y, Y) - L(Y, \mu)]$$

Avec L_{MS} la log-vraisemblance du modèle saturé et L_{MF} celle du modèle final.

L'étude de la déviance et de la statistique de Pearson permet de vérifier l'adéquation à un modèle. Il faut alors étudier le rapport de D ou X_p^2 sur le nombre de degrés de liberté du modèle final à la valeur 1. Si le rapport entre les deux est plus petit que 1, le modèle est jugé satisfaisant sinon il y a un problème de sur dispersion.

Les modèles présentés dans ce chapitre permettent de modéliser le coût d'une pathologie et son lien avec les variables explicatives de façon appropriée selon sa distribution, sa nature et l'objectif de l'étude. Ces méthodes prédisent le coût de la pathologie étudiée et évaluent ses variations selon les caractéristiques des individus. Couplé à des données populationnelles d'observations, le recours à une modélisation adéquate fournit des résultats en vie réelle non biaisés et des conclusions

¹⁴ Modèle saturé : modèle qui aurait le meilleur ajustement aux données possible. $L(Y, Y)$ car les prédictions égalent les valeurs observées dans un tel modèle.

Degrés de liberté : $n-p-1$ avec n le nombre de valeur, p le nombre de paramètre (i.e. variable).

pertinentes aux décideurs afin de les guider objectivement dans leurs choix. Dans ce contexte, les prises en charge des pathologies cancéreuses constituent un fardeau économique pour la collectivité. Ce fardeau nécessite d'être évalué afin de permettre aux autorités de santé de parvenir à une allocation efficiente des ressources dans le domaine. Les hémopathies malignes sont notamment concernées par cette problématique en raison des fréquentes innovations thérapeutiques rencontrées et de leur caractère onéreux (14). Ces pathologies ont des prises en charge très disparates qui nécessitent d'utiliser différents modèles statistiques afin d'évaluer correctement leurs coûts.

VI. Hémopathies malignes

Les cancers sont des pathologies caractérisées par une prolifération cellulaire anormale à l'intérieur d'un tissu sain de l'organisme. Ils constituent une des principales causes de morbidité et de mortalité dans le monde, comptabilisant en 2012, 14 millions de nouveaux cas et représentant la 2^{ème} cause de décès avec 8,2 millions de morts (197). En Europe, 2,6 millions de personnes ont été diagnostiquées avec un cancer en 2012 et 1,3 millions sont décédées des suites d'un cancer. Les coûts (directs et indirects) associés à ces pathologies représentaient 83.2 à 86 milliards d'euros en 2014 (198) (199). En France, le nombre de nouveaux cas de cancer a été estimé à 399 500 en 2017 (200). Le coût direct et indirect des patients atteints d'un cancer en 2014 et bénéficiaire du RG a été estimé à 13.5 milliard d'euros (156).

Les hémopathies malignes (i.e. cancers du sang ou hématologiques) sont des cancers qui se développent à partir des cellules sanguines : globules blancs, globules rouges et plaquettes. Ces pathologies se développent secondairement à des altérations des cellules souches hématopoïétiques à différents stades de leurs maturations. En 2012, l'incidence des hémopathies malignes a été estimée à 35 000 nouveaux cas, représentant environ un dixième des nouveaux cas de cancer en France (201). Elles sont principalement représentées par les leucémies aiguës (lymphoblastiques, myéloblastiques), les Lymphomes Hodgkiniens (LH), les Lymphomes Non Hodgkiniens (LNH), la Leucémie Myéloïde Chronique (LMC), le Myélome Multiple (MM), la maladie de Waldenström et la Leucémie Lymphoïde Chronique (LLC) (202). Les hémopathies ont des caractéristiques génétiques, morphologiques et cliniques distinctes nécessitant des prises en charge très variées. Les avancées thérapeutiques majeures, les modes de prises en charge pluridisciplinaires et la qualité grandissante des moyens diagnostics ont permis une amélioration significative de la survie tout en modifiant grandement les parcours de soins (202). Dans ce contexte, les lymphomes et la LMC ont notamment vu leur prise en charge révolutionnée avec l'arrivée de nouveaux traitements dans les années 2000. Les hémopathies malignes sont prises en charge à 100% par l'AM dans le cadre de l'ALD 30 relative

au cancer. Dans ce contexte de transformation des prises en charge à partir de nouvelles thérapeutiques très onéreuses, et du peu d'informations économiques disponibles en France sur le sujet, nous nous intéressons à ces deux entités dans ce chapitre.

VI.1. Lymphomes

Les lymphomes sont des cancers du système lymphatique, le principal élément du système immunitaire de l'organisme. Ils se développent à partir des cellules de la famille des globules blancs, les lymphocytes. Ce sont les cancers les plus fréquents parmi les hémopathies, représentant 3 à 4% des cancers dans le monde (201). Ils se divisent en deux grandes catégories : les LH et les LNH. Le LH est la première forme de la maladie à avoir été identifiés par Thomas Hodgkin en 1832. Les LNH représentent la majorité des lymphomes (203).

VI.1.1. Carcinogénèse

Les lymphomes prennent naissance dans le système lymphatique, partie centrale du système immunitaire. Il est composé de vaisseaux lymphatiques qui permettent le transport de la lymphe, le liquide contenant les globules blancs, et de ganglions lymphatiques qui ont pour rôle de filtrer la lymphe afin d'en extraire les bactéries, virus et toutes autres substances étrangères. Ces ganglions sont répartis à travers tous le corps mais régulièrement retrouvés le long des veines et des artères. On peut les percevoir au niveau de l'aîne, du cou et des aisselles. D'autres organes, appelés organes du système lymphatique (rate, moelle osseuse, thymus, ...), participent à cette tâche. Lorsqu'un nombre important de bactéries est filtré, les ganglions ou les organes du système lymphatique peuvent gonfler et devenir douloureux.

Les globules blancs sont de trois types : les granulocytes, les lymphocytes et les monocytes. Les lymphomes sont ainsi des cancers du système lymphatique issus d'erreurs dans la fabrication des lymphocytes. Les lymphocytes cancéreux se divisent alors plus rapidement et/ou ont des durées de vie plus importantes que les cellules normales. Ils sont produits dans le système lymphatique et ils s'observent dans les organes du système lymphatique, notamment au niveau des ganglions. Ils peuvent néanmoins s'étendre à tous les organes du corps. Les lymphocytes cancéreux une fois regroupés forment des tumeurs, appelées lymphomes (204).

La différence fondamentale entre les LH et les LNH provient des lymphocytes cancéreux. Le LH est caractérisé par la présence de lymphocytes aux caractéristiques morphologiques précises (taille, noyaux..) appelés les cellules de Reed-Sternberg (204). Les LH sont des lymphomes définis par des cellules cancéreuses de type B tandis que l'on distingue chez les LNH des cellules cancéreuses de type B, T ou NK.

VI.1.2. Anatomopathologie et symptomatologie

Les lymphomes peuvent être localisés dans tout le corps et se manifester par des symptômes cliniques et biologiques variés (205). Cependant leur survenu se traduit généralement par l'augmentation du volume d'un ou de plusieurs ganglions (i.e. adénopathie). Ils se situent le plus souvent au niveau du cou ou des aisselles. Ils entraînent parfois des douleurs et des paralysies de par la pression exercée sur les nerfs due au gonflement. Selon la localisation, les ganglions auront différents effets. Lorsqu'ils sont présents au niveau abdominal, le patient peut ressentir des ballonnements ou des maux de dos tandis qu'ils peuvent être responsables d'un gonflement des jambes lorsqu'ils se situent au niveau de l'aîne. Ainsi, les symptômes vont dépendre de la partie du corps où se développe la tumeur. Toutefois dans de rares cas, les ganglions ne sont pas perceptibles. Les organes internes peuvent également être enflés, notamment la rate et le foie. Un patient atteint

d'un lymphome peut présenter un syndrome inflammatoire biologique et parfois une anomalie de l'hémogramme (cytopénie, hyperlymphocytose, thrombocytose). Une fièvre inexplicée durant une période supérieure à trois semaines et la présence de démangeaison (i.e. prurit) sont également des signes associés au lymphome. Les signes cliniques d'un lymphome ne sont pas homogènes pour tous les malades. De plus, les symptômes ne sont pas spécifiques au lymphome et peuvent être déclenchés par d'autres maladies. Ainsi, la présence d'un lymphome est parfois difficile à établir.

En cas de suspicion de lymphome, l'examen anatomopathologique au microscope de la biopsie d'un ganglion suspect permet de confirmer son diagnostic. Il permet de définir si le lymphome est Hodgkinien ou non, s'il est nodulaire ou diffus, s'il touche des petites ou des grandes cellules ainsi que le type des cellules affectées. Tous ces éléments vont permettre de définir le traitement approprié.

VI.1.3. Lymphomes non Hodgkiniens

VI.1.3.1. Epidémiologie et facteurs de risque

En France, les LNH représentent sensiblement 90% de tous les lymphomes avec une incidence estimée en 2011 à 11 600 nouveaux cas (201). Ils se placent au 6^{ème} rang des cancers les plus fréquents (204). L'âge médian au diagnostic est de 64 ans chez l'homme et de 70 ans chez la femme. La survie à 5 ans des patients est très variable selon le LNH considéré et est décrite ci-dessous.

Les LNH sont classés en tant que néoplasmes à cellules B (85% des cas) et T ou NK (15% des cas) selon le type de lymphocytes touchés par les mutations. Néanmoins, il existe plus de 40 sous-types avec des caractéristiques génétiques, morphologiques et cliniques distinctes (203). La classification des LNH est ainsi complexe. Elle a été établie afin de permettre aux médecins des différentes spécialités d'avoir une approche du diagnostic similaire et un langage international commun (204). Nous allons présenter les principales formes de LNH et les données épidémiologiques associées. Ces

informations sont extraites des rapports de l'institut national de veille sanitaire, fondées sur les données des registres des cancers du réseau FRANCIM de 2012 (201) (206).

Les formes les plus fréquentes de LNH sont :

- Les Lymphomes Folliculaires (LF)

L'incidence des LF a été estimée à 2 530 nouveaux cas en 2012 soit 7.9 cas pour 100 000 personnes années. Les FL sont des lymphomes de type B et les plus communs des LNH indolents. Ils touchent les adultes d'âge moyen ou avancé avec une médiane à 64 ans chez l'homme et 66 ans chez la femme. Les LF ont une évolution lente et sont le plus souvent diagnostiqués à un stade avancé (stade III et IV). Ils sont de bon pronostic avec une survie nette à 5 ans estimée à 87%. Ils peuvent évoluer en LBDGC. Le code CIM-10 associé est le C82.

- Les Lymphomes de la Zone Marginale (LZM)

L'incidence des LZM a été estimée à 1 772 nouveaux cas en 2012 soit 2.7 cas pour 100 000 personnes années, dont 49% chez l'homme. Les patients atteints d'un LZM ont un âge médian de 70 ans pour les hommes et de 74 ans chez les femmes. Ce sont des LNH de type B indolents pouvant évoluer en LNH agressifs. Ils sont de bons pronostics avec une survie nette à 5 ans de 86%. On distingue parmi les LZM, les lymphomes des tissus lymphoïdes associés aux muqueuses, ou lymphomes MALT, reliés au code C88.4 et en lymphomes à cellules B de la zone marginale ganglionnaire reliés au code C83.0.

- Les Lymphomes LymphoPlasmocytaires (LLP) / Macroglobulinémies de Waldenström (MW)

Les LLP sont des LNH de type B caractérisés par la présence de cellules lymphoplasmocytaires dans la moelle osseuse. La présence d'une protéine particulière supplémentaire (i.e. immunoglobuline monoclonale) dans la moelle définit la maladie de Waldenström. Ces maladies sont alors souvent regroupées bien que la MW soit une forme particulière de lymphome. L'incidence de cette entité a été estimée à 1 247 nouveau cas en 2012, soit 1,9 cas pour 100 000 personnes années, dont 64%

chez l'homme. Les patients atteints d'un LPP ou d'un MW, ont un âge médian de 73 ans pour les hommes et de 74 ans chez les femmes. Ces lymphomes sont considérés comme indolent et ont une survie nette 5 ans après le diagnostic estimé à 76%. Les LLP sont associés au code C83.0 et la MW au code C88.0.

- Les Lymphomes Lymphocytiques (LL) et les Leucémies Lymphoïdes Chroniques (LLC)

Le LL et le LLC sont considérés comme proches et diffèrent de par la localisation initiale des cellules tumorales. On retrouve généralement 75% de LLC et 25% de LL. L'incidence de cette entité a été estimée à 4 664 nouveau cas soit 6.6 pour 100 000 personnes années dont 60% chez les hommes. Ils touchent des personnes âgées avec une médiane d'âge de 71 ans chez les hommes et 74 ans chez les femmes. La survie nette des patients atteints de LLC ou de LL a été estimée à 83%. Ces pathologies ont une évolution et une prise en charge particulière. La LLC est surnommée la « maladie des trois tiers » car environ un tiers des patients ne nécessitent pas de traitement et n'en n'auront jamais besoin, environ un tiers ne nécessitent pas de traitement mais la maladie évoluera et pourra être traitée et environ un tiers qui nécessitent un traitement dès le diagnostic car leur LLC est évolutive. Les LL sont regroupés avec les C83.0 tandis que la LLC est classée par le code C91.1. La LLC est parfois considérée comme une leucémie de par la localisation des cellules cancéreuses dans la moelle osseuse et est souvent mise de côté comme une hémopathie maligne particulière (202).

- Les Lymphomes B Diffus à Grandes Cellules (LBDGC)

Ce sont les LNH les plus fréquemment diagnostiqués. Les LBDGC ont une incidence estimée à 4 096 nouveaux cas soit 6,7 nouveaux cas pour 100 000 personnes années dont 60% survenant chez l'homme. Ils représentent 60% des formes agressives de LNH. La médiane d'âge de survenue est de 69 ans chez l'homme et de 74 ans chez la femme. La survie nette 5 ans après le diagnostic a été estimée à 57%. Les LBDGC peuvent provenir de l'évolution d'un autre type de LNH tel que les FL, les LZM ou les LLC. Les LBDGC sont reliés au code CIM-10 C83.3.

- Les Lymphomes T/NK à cellules matures (LT/NK)

Ils regroupent un ensemble d'entités différentes. L'incidence de ces lymphomes a été estimée à 1 419 nouveaux cas soit 2.7 cas pour 100 000 personnes années, dont 61% retrouvés chez l'homme. C'est une maladie de l'adulte avec des patients d'âge médian 67 ans chez les hommes et 66 ans chez les femmes. Cette entité regroupe l'ensemble des LNH de type T et plus de 60% de ces lymphomes sont des lymphomes T périphériques. La survie nette 5 ans après le diagnostic a été estimée à 62%. Les codes CIM-10 associés sont les codes C84 et C86.

- Les Lymphomes à Cellules du Manteau (LCM)

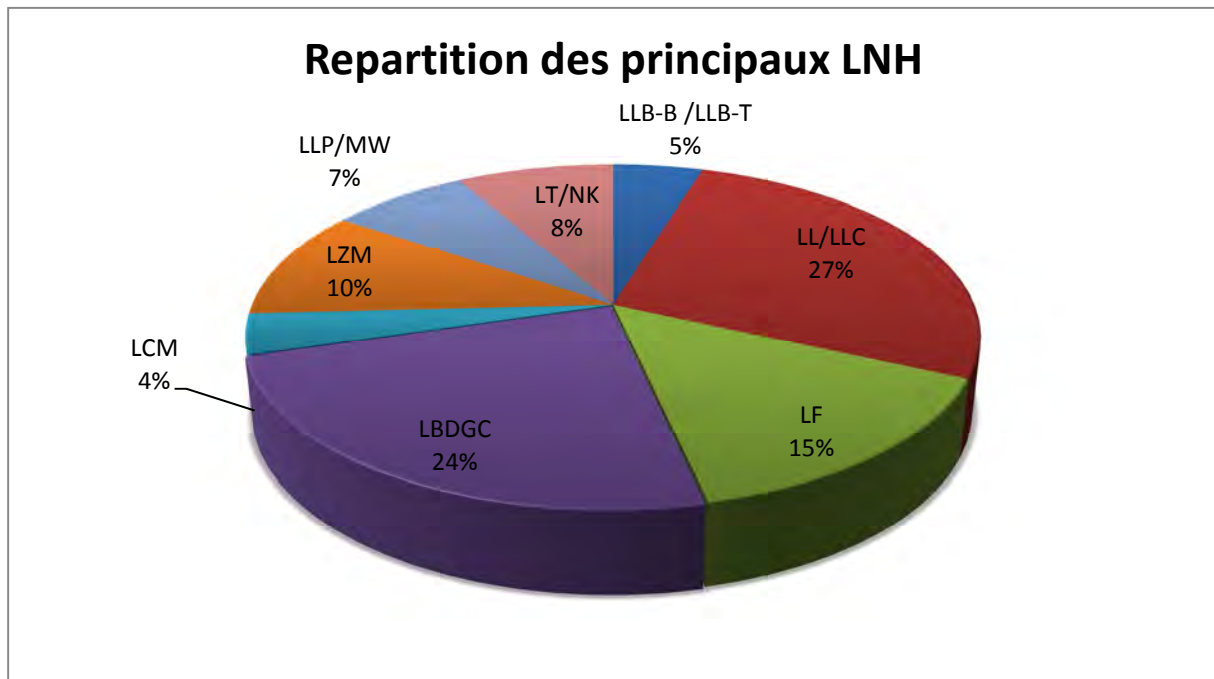
L'incidence de ces lymphomes a été estimée à 659 nouveaux cas, soit 1 cas pour 100 000 personnes années dont 75% chez l'homme. L'âge médian des patients est de 74 ans. C'est un lymphome agressif avec une survie nette 5 ans après le diagnostic qui a été estimée à 48%. Le code CIM-10 associé est le C83.1.

- Les Lymphomes/leucémies LymphoBlastiques à précurseurs B ou T (LLB-B ou LLB-T)

Ce sont des LNH de types B ou T selon les cellules atteintes. Ces LNH sont rares et touchent de jeunes adultes. Ils apparaissent lorsque les néoplasmes se créent au cours des phases très précoces de la différenciation de la cellule. Selon le nombre de région touchée dans la moelle osseuse, cette maladie est soit un lymphome soit une leucémie. Il est alors difficile de la classer avec le reste des LNH. Son incidence a été estimée à 810 nouveaux cas en 2012 dont 60% chez l'homme. C'est un cancer de mauvais pronostic et la survie à 5 ans après le diagnostic a été estimée à 44%. Les LLB-B se classent avec les C83.5 et les leucémies lymphoblastiques aiguës avec les C91.0.

La figure suivante représente la répartition des principaux LNH selon les données d'incidence ci-dessus.

Figure 15. Répartition des principaux sous types de LNH (%) (France, 2012)



FL : Lymphomes Folliculaires ; LZM : Lymphomes de la Zone Marginale ; LLP : Lymphomes LymphoPlasmocytaires ; MW : Macroglobulinémies de Waldenström ; LBDGC : lymphomes diffus à grandes cellules B ; LT/NK : lymphomes T/NK à cellule mature ; LLB-B ou LLB-T : Lymphomes/leucémies LymphoBlastiques à précurseurs B ou T ; LL : Lymphomes Lymphocytiques ; LLC : Leucémies Lymphoïdes Chroniques ; LCM : Lymphomes à Cellules du Manteau.

Plusieurs facteurs de risque ont été mis en évidence tels que les infections chroniques virales (VIH, hépatite C) et bactériennes, les immunodépresseurs chroniques (greffes d'organes, maladies auto immunes), l'exposition à certaines substances toxiques (dioxines, pesticides agricoles) ou encore des antécédents de chimiothérapies (207).

VI.1.3.2. Classification clinique

La classification d'un LNH s'effectue selon son type, son grade et son stade. Ces informations permettent aux PS de connaître les éventuelles évolutions de la maladie ainsi que le traitement le plus adapté.

Le type de LNH correspond à la lignée de ces cellules cancéreuses. Il est défini à partir de l'examen anatomopathologique des tissus prélevés lors de la biopsie. Les deux principaux types de LNH sont les LNH à cellules B, ou à cellules T selon le type de lymphocytes présentant une mutation. On ne retrouve que rarement des lymphomes touchant les cellules NK qui sont généralement rattachés aux LNH T. Un réseau national anatomopathologie appelé le réseau « Lymphopath » a été créé en 2010 dans le cadre du plan cancer 2009 – 2013. Il regroupe une trentaine de centres experts. Ce réseau permet de réaliser une double lecture systématique et gratuite pour tout nouveau cas de lymphomes afin d'en certifier le diagnostic (204). Selon Laurent et al., les diagnostics de lymphomes après relecture par des pathologistes experts sont plus pertinents dans 20 % des cas (208).

Le grade du lymphome correspond à son degré d'agressivité. Il dépend des caractéristiques morphologiques et immunologiques des cellules cancéreuses (204). Le grade est défini selon l'apparence et le comportement des cellules cancéreuses. Il indique la vitesse d'évolution de la tumeur et permet de choisir un traitement adapté. On distingue les lymphomes indolents (i.e. de bas grade ou de faible malignité) et les lymphomes agressifs (i.e. de haut grade ou de malignité intermédiaire ou élevée). Les lymphomes indolents représentent 40 à 50% des LNH. Ils ont un développement lent, allant de quelques mois à plusieurs années. Le traitement peut conduire à des rémissions complètes mais les rechutes sont fréquentes. Ces lymphomes ont toutefois un bon pronostic de survie qui se rapproche de celui de la population générale. Néanmoins, 15 à 40% des LNH indolents se transforment en un LNH agressif. Les lymphomes agressifs représentent 50 à 60% des cas de LNH.

Le stade du lymphome permet de décrire le degré d'extension de la maladie dans l'organisme. La classification d'Ann Arbor créée en 1971 par des experts américains, allemands et français, à Ann Arbor dans le Michigan, et modifiée en 1998 à Cotswolds en Angleterre, est principalement utilisée (209). Il existe cependant d'autres classifications pour les LNH extra-ganglionnaires, plus rares. L'évolution des LNH se divise alors en 4 stades. Les stades I et II sont dits localisés alors que les stades

III et IV sont considérés comme étendus ou disséminés. Les stades de la maladie sont détaillés en annexe 10.

VI.1.3.3. Diagnostic

Le diagnostic d'un patient atteint d'un LNH est sensiblement identique à celui d'un patient atteint d'un LH, étant donné qu'ils ne diffèrent que par la présence des cellules caractéristiques de Reed-Sternberg.

Tout d'abord, un examen clinique permet de suspecter ou non la présence d'un lymphome chez un individu. Il se base principalement sur la présence de ganglions, d'organes internes gonflés (rate et foie), de la présence de liquide dans la poitrine ou dans l'abdomen, de fatigue, de fièvre et de paralysies localisées. Après cet examen et s'il y a suspicion de la maladie, des examens complémentaires sont réalisés. Des examens sanguins et une biopsie d'un ou plusieurs ganglions sont effectués afin de détecter la présence de leucocytes cancéreux et de confirmer le diagnostic.

Un bilan d'extension est ensuite réalisé pour situer toutes les localisations du lymphome dans le corps. L'imagerie médicale, comprenant de la radiographie, l'utilisation d'un scanner et de la Tomographie par Émission de Positons (TEP), permet d'évaluer le degré d'extension de la maladie. Des examens de la moelle osseuse ou du liquide céphalo-rachidien permettent de vérifier si le lymphome ne s'est pas étendu dans la moelle ou dans le système nerveux.

Un bilan général du patient est également réalisé avec pour objectif de détecter les autres pathologies éventuellement présentes, en particuliers cardiaques (204). L'ensemble des examens réalisés apporte des informations essentielles afin de définir le type, le grade et le stade du lymphome. Ces éléments ainsi que l'état général du malade indiquent comment définir le traitement le plus approprié pour le patient.

VI.1.3.4. Prises en charge thérapeutiques

VI.1.3.4.1. Traitements

Une fois, le diagnostic de LNH confirmé, sa prise en charge est définie dans le cadre de Réunion Concertation Pluridisciplinaire (RCP) où se réunissent les PS en charge du patient. L'ensemble des décisions est ensuite communiqué et débattu avec le patient. Il est ainsi informé de toutes les options thérapeutiques disponibles dans sa situation, des bénéfices attendus, des risques et des Effets Indésirables (EI) potentiels, ainsi que de la possibilité et des modalités de participation à un essai clinique (207). Différentes prises en charge selon la pathologie et les autres paramètres associés (grade, stade, type...) existent tel que la chimiothérapie, la radiothérapie, l'immunothérapie et la greffe de cellules souches.

La référence pour la première phase active de traitement de la majorité des LNH est une immunochemiothérapie, c'est-à-dire l'association d'une chimiothérapie et d'une immunothérapie par anticorps monoclonal. Il est préconisé pour tous les LNH B agressifs (i.e. LBDGC) et pour certains LNH B indolents (i.e. FL) selon l'avancement de la maladie, le retentissement clinique et biologique du lymphome et l'état général du patient. Les LNH B indolents peu avancés (stade 1 et 2) peuvent simplement être surveillés et traités lorsque la maladie se dégrade. Pour le LF, en cas de réponse positive à la phase active de soins, un traitement d'entretien à base d'anticorps monoclonal est proposé sur une période maximale de 2 ans. Les LNH de type T agressifs sont généralement traités par chimiothérapie (207).

La chimiothérapie ou l'immunochemiothérapie est administrée selon des protocoles prédéfinis (nombre de cures, temps entre les cures, combinaison de médicaments utilisés...). Il est important de suivre ces protocoles afin d'optimiser les résultats du traitement (210). En pratique, il est peu commun d'atteindre une observance maximale étant donnée la toxicité des traitements. De plus, un

nombre important de caractéristiques au niveau du patient et des centres de soins peuvent expliquer les variations entre les référentiels de traitement et la pratique (210) (211). Les protocoles de chimiothérapie ou d'immunochimiothérapie sont très variés. Le référentiel Oncomip, réseau de cancérologie de Midi-Pyrénées, synthétise la littérature internationale et décrit les différents protocoles (212). Une unique dose de chimiothérapie ou d'immunochimiothérapie ne détruit qu'une partie des cellules cancéreuses et il est nécessaire d'administrer des doses répétées afin d'en supprimer le maximum. Les traitements s'administrent alors en cures, suivies d'une phase de repos permettant au patient de récupérer. L'ensemble de la cure couplé à une phase de repos représente un cycle. Le protocole de référence pour l'immunochimiothérapie associe généralement 6 à 8 cycles d'une chimiothérapie de type CHOP (Cyclophosphamide, Hydroxy, Doxorubicine, Vincristine, Prednisone) à un anti corps monoclonal anti-CD20 (i.e. rituximab) et communément appelé R-CHOP. Le rituximab a été introduit dans les années 2000 et intervient dans la plupart des protocoles. Il a profondément modifié la prise en charge des LNH en améliorant la survie des patients. Le temps entre les cures est généralement de 14 ou 21 jours pour le traitement des LNH agressifs (i.e. LBDGC) et de 21 ou 28 jours et plus pour le traitement des LNH indolents (LF) (212). Le temps entre les cures dépend notamment de l'état du patient. Ces traitements s'administrent par voie orale ou s'injectent par voie intraveineuse, intramusculaire ou sous cutanée (204). Ils sont dispensés dans le cadre d'une hospitalisation classique lorsqu'il est nécessaire de surveiller et d'ajuster le traitement ou lorsque la chimiothérapie est intensive, en hospitalisation de jour lorsque le traitement le permet ou plus rarement à domicile par voie orale ou intraveineuse.

La chimiothérapie ou l'immunochimiothérapie peut s'accompagner d'une greffe de cellules souches lorsque les doses administrées sont importantes et que les Cellules Souches Hématopoïétiques (CSH) sont touchées (212). Les greffes sont très éprouvantes et ne sont pas recommandées pour tous les malades.

Les LNH peuvent également se traiter par radiothérapie. Cette prise en charge consiste à exposer le patient à des rayons ou des particules de hautes énergies pour détruire les tumeurs. La radiothérapie est le plus souvent administrée pour les LNH indolents et précoces (stade I ou II) mais elle peut également être utilisée pendant ou après la chimiothérapie ou l'immunochimiothérapie pour les LNH qui évoluent rapidement (204).

Suite à la phase active de traitement, un suivi est organisé afin de détecter de possibles récurrences, des complications tardives (i.e. toxicité cardiaque, hépatite etc..) ou un second cancer. Cela permet également de veiller à la qualité de vie du patient et de faciliter sa réinsertion socioprofessionnelle. Le suivi est réalisé par l'équipe référente du traitement et adapté au lymphome considéré. Il consiste généralement en une surveillance pendant 5 ans avec des consultations tous les 3 mois pendant 2 ans puis tous les 6 mois pendant 3 ans (207).

Un patient peut récidiver en cas de résistance au traitement et selon le type de lymphome. Les lymphomes indolents récidivent généralement plus souvent que les agressifs (207). Les lymphomes indolents peuvent alors rester indolents et être simplement surveillés où recevoir un traitement identique à la première phase. S'ils se transforment en lymphomes agressifs, ou si les lymphomes sont agressifs à l'origine, le traitement initial est alors renouvelé à fortes doses. C'est le plus souvent suite à ce traitement intensif qu'il sera nécessaire de réaliser une autogreffe de cellules souches. Si un traitement à fortes doses n'est pas envisageable, un traitement palliatif afin de réduire l'intensité des symptômes est à prévoir.

VI.1.3.4.2. Effet indésirables des traitements

Les traitements dispensés ont une toxicité élevée et engendrent un nombre important d'Effets Indésirables (EI). Il est alors essentiel de les prendre en charge car ils affectent directement l'état de santé du patient, sa qualité de vie et peuvent impacter la bonne observance du traitement dispensé

et donc la survie du patient. Le tableau suivant résume les principaux EI retrouvés chez les patients selon le type de traitement utilisé (207):

Tableau 4. Les Effets Indésirables dus aux traitements des LNH.

Immunochimiothérapie	Radiothérapie
Toxicité digestive (nausée, vomissement, douleur abdominale, constipation, diarrhée)	Mucite et/ou xérostomie en cas d'irradiation des cavités ORL ;
Toxicité hématologique (anémie, neutropénie fébrile, thrombopénie)	Diarrhée en cas d'irradiation de la cavité abdominale
Toxicité muqueuse (Mucite)	Erythrodermie en fin de traitement
Toxicité cutanée (prurit, rash) Alopécie	Pancytopenie transitoire (si une partie du territoire hématopoïétique est irradiée).
Toxicité neurologique (neuropathie périphérique à la fois sensorielle et motrice)	Nausées et/ou vomissements (en cas de volume d'irradiation important)
Accès fébriles	Asthénie
Infections bactériennes, virales	
Réactions liées à la perfusion, angio-oedème	
Toxicité cardiaque (troubles du rythme, décompensation cardiaque)	
Asthénie	

Les EI surviennent généralement entre les séances d'immunochimiothérapies ou de radiothérapie. Ils sont détectés au retour du patient dans le centre de soins lors de la cure suivante. Ils sont ainsi pris en charge tardivement et peuvent entraver le bon déroulement des cures de traitements au détriment de la qualité de vie du patient et de l'efficacité du traitement. Ils peuvent également augmenter les coûts de prise en charge pour l'AM (hospitalisations plus longues, nombre de cures et consommation médicamenteuse plus importantes, pertes de productivité) et pour la société dans sa globalité. Il est essentiel de prendre en charge les EI de façon précoce et d'améliorer le suivi des patients (213).

Les LNH constituent un ensemble de pathologies hétérogènes avec des pronostics de survie et des stratégies thérapeutiques variables. Ils reçoivent généralement des traitements lourds caractérisés par une forte toxicité qui engendre un nombre important d'EI. Ces EI nécessitent une prise en charge précoce afin d'affecter au minimum le bon déroulement des soins et son efficacité. Les LH se différencient des LNH par la présence des cellules caractéristiques de Reed-Sternberg, par les populations atteintes et par le fait que le LH est plus souvent localisé et de meilleur pronostic (214).

VI.1.4. Lymphomes Hodgkiniens

VI.1.4.1. Epidémiologie et facteurs de risque

En France, les LH représentent environ 10% de tous les lymphomes avec une incidence estimée en 2012 à 1 880 nouveaux cas (201). C'est un cancer touchant une population jeune avec une médiane de survenue de 32 ans chez la femme et 42 ans chez l'homme. La courbe d'incidence du LH dessine un premier pic important vers l'âge de 25 ans et un second plus faible entre 75 et 80 ans. Le LH a un bon pronostic avec une survie à 5 ans de 83% chez l'homme et 88% chez la femme (206).

L'entité la plus commune est le LH classique qui représente 95% de tous les LH. Il se divise en 4 sous types histologiques : le LH scléro-nodulaire, à cellularité mixte, riche en lymphocytes et à déplétion lymphocytaire. Ils sont codés respectivement avec les codes CIM-10 : C81.1, C81.2, C81.4, C81.3. Ces différents sous types de LH ont néanmoins la même prise en charge. Il existe également le LH à prédominance lymphocytaire, forme nodulaire (C81.0) et d'autres formes spécifiques de LH (enfant, femme enceinte, patient très âgé) (C81.7) mais sont plus rarement retrouvés (215). Certains facteurs de risque liés au LH sont connus. Ainsi, les personnes atteintes du VIH, du virus de d'Epstein-Barr ou des membres de la même famille qu'un individu atteint d'un LH ont plus de risque de développer la maladie (214) (216).

VI.1.4.2. Classification clinique

La classification par stade du LNH est également utilisée pour le LH. C'est l'élément principal déterminant le traitement. Ainsi, 4 regroupements sont généralement utilisés :

- Les stades localisés (I et II) sus diaphragmatiques (si la tumeur se situe au-dessus du diaphragme)
- Les stades localisés (I et II) sous diaphragmatiques
- Les stades disséminés III-A
- Les stades disséminés III-B et IV

Les LH sont des lymphomes de type B indolents.

VI.1.4.3. Diagnostic

Comme indiqué en VI.1.3.3, le diagnostic du LH est identique à celui du LNH. Les examens anatomopathologiques permettent ainsi d'identifier la présence de cellules de Reed-Sternberg et suffisent dans la majorité des cas pour confirmer le diagnostic du LH. La structure particulière de ces cellules est en effet reconnaissable au microscope. Cet examen permet également de définir les caractéristiques de ces cellules et le sous type de LH (217).

VI.1.4.4. Prises en charge thérapeutiques

VI.1.4.4.1. Traitements

Les LH sont généralement traités par chimiothérapie associée ou non à de la radiothérapie. Le nombre de cycles de chimiothérapie dépendra du stade de la maladie, des facteurs pronostiques et de la localisation de la tumeur lorsque le LH est de stade localisé (214).

Le traitement des LH de stade I ou II et sus-diaphragmatiques associe une phase de chimiothérapie de 2 à 4 cycles à des séances de radiothérapie un mois après sur les ganglions originairement

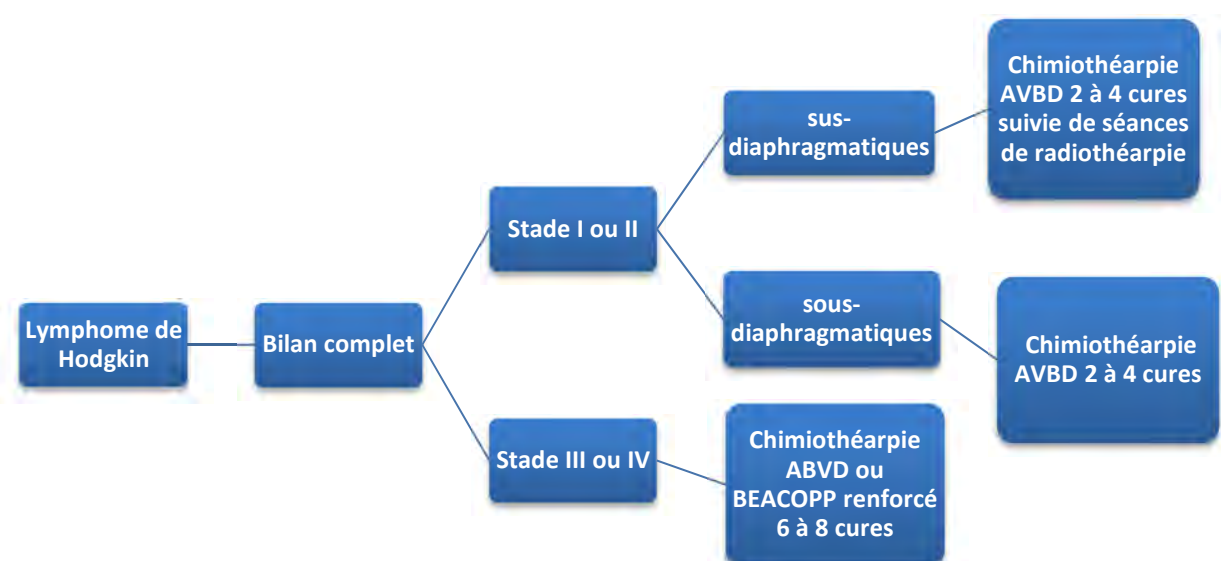
atteints. La radiothérapie devient optionnelle pour les LH de stade I ou II et sous diaphragmatiques. Les patients atteints de LH de stades disséminés (III ou IV) recevront de 6 à 8 cures de chimiothérapie et exceptionnellement de la radiothérapie en cas de réponse partielle de la chimiothérapie.

Le protocole de chimiothérapie de référence, tous stades confondus, est le protocole ABVD (adriamycine, bléomycine, vinblastine, dacarbazine). Les LH disséminés pourront également se voir administrer du BEACOPP (Bléomycine, etoposide, adriamycine, cyclophosphamide, vincristine, procarbazine, prednisone) renforcé. La plupart des traitements du LH sont réalisés en hospitalisation de jour.

Une évaluation pendant le parcours et en fin de traitement est réalisée afin d'identifier la réponse thérapeutique du traitement. Le patient peut répondre complètement ou partiellement au traitement si les signes de la maladie ont totalement ou en partie disparus. La maladie peut également être stable lorsque le traitement n'a pas d'impact sur les anomalies tumorales ou progresser lorsque la maladie s'intensifie. Les formes réfractaires de LH (maladies stables ou progressives) nécessiteront une deuxième ligne de traitement (216).

La figure suivante résume les différentes options thérapeutiques des patients atteints d'un LH.

Figure 16. Les différents traitements des LH selon le stade



Un suivi post cancer est également organisé et comprend une visite de suivi tous les 3 mois pendant 2 ans, tous les 6 mois pendant 3 ans, annuellement jusqu'à la 10ième année et enfin tous les 2 ans. Lors de ce suivi, l'état général du patient sera vérifié afin de confirmer la rémission ou de détecter une récurrence et d'identifier les éventuelles complications thyroïdiennes, cardiaques et l'apparition d'un cancer du sein ou du poumon.

De la même manière que pour les LNH, un patient atteint d'un LH peut rechuter. Dans ce cas, un traitement identique avec des doses plus intenses est généralement proposé en association avec de l'auto greffe de CSH.

VI.1.4.4.1. Effets indésirables des traitements

De façon identique aux LNH, le traitement des LH est caractérisé par une forte toxicité et engendre un nombre important d'EI. Ils proviennent ainsi de la chimiothérapie et/ou de la radiothérapie et seront sensiblement identiques à ceux développés dans le cadre d'un LNH.

Le tableau suivant inspiré du guide méthodologique produit par la HAS résume les principaux EI dus aux traitements des LH (214) (216).

Tableau 5 : Les effets indésirables dus aux traitements des LH

Chimiothérapie	Radiothérapie
Toxicité digestive (nausées, vomissements, douleurs abdominales, constipation)	Mucite ou xérostomie en cas d'irradiation des cavités ORL pouvant entraîner une dysphagie
Toxicité hématologique (anémie, neutropénie fébrile, thrombopénie)	Diarrhée en cas d'irradiation de la cavité abdominale
Douleurs osseuses sous facteur de croissance	Erythème
Alopécie	Asthénie
Anorexie	nausées ou vomissements en cas de volume d'irradiation important
Pigmentation cutanée	
Asthénie	
Mucite (sous BEACOPP)	

De la même façon que pour les EI provoqués par les traitements du LNH, une prise en charge inter-cure permet d'optimiser l'efficacité des traitements et la qualité de vie des patients.

Le LH est une entité moins complexe que le LNH et comptabilise seulement 10% de tous les lymphomes. Cette pathologie nécessite un traitement lourd qui affecte d'autant plus la vie professionnelle des patients qu'ils sont jeunes et actifs. Il est ainsi important d'en étudier tous les retentissements.

VI.1.5. Coût des lymphomes en France

Etant donnée leur incidence importante, des traitements onéreux et toxiques et un grand nombre d'EI en découlant, les lymphomes entraînent un coût de prise en charge non négligeable. En France, seulement deux études ont évalué le coût de ces pathologies. Ces études se sont focalisées sur les LNH.

La première étude de Best et al. a évalué le coût de la prise en charge des patients atteints de LBDGC sur un horizon temporel de 15 ans, selon s'ils étaient traités par CHOP ou par R-CHOP (218). Le point de vue utilisé était celui de l'AM. Les coûts considérés étaient les coûts directs médicaux limités aux coûts des séjours hospitaliers relatifs à l'administration des traitements, aux coûts des traitements, à la prise en charge des effets indésirables associés, à la surveillance, aux récurrences, aux transplantations et aux soins palliatifs. Les auteurs ont développé un modèle alimenté par des données provenant de différentes sources. Les données issues la cohorte NHL 98-5 ont permis d'estimer les distributions du nombre d'hospitalisation associés aux cures de chimiothérapie et aux effets indésirables. Ces séjours hospitaliers ont ensuite été valorisés à partir des GHS. Les coûts des autres postes de dépenses pris en compte ont été valorisés à partir de la littérature. Le coût moyen associé à la prise en charge d'un patient atteint de LBDGC et traité par R-CHOP a ainsi été évalué à 41 952 € sur 15 ans. Cette étude rencontre certaines limites. La consommation de soins de ville ainsi que les coûts directs non médicaux et les coûts indirects n'ont pas été pris en compte. De plus, la cohorte NHL 98-5 est composée de patients recrutés dans 86 centres répartis en France, Suisse et en Belgique (219). Les données de consommations estimées à partir de cette cohorte ne représentent donc pas intégralement le contexte Français.

La seconde étude de Deconick & al. a évalué le coût de la prise en charge des LF en rémission complète ou partielle suite au traitement d'induction, selon s'ils recevaient un traitement d'entretien par rituximab ou s'ils étaient surveillés (220). La perspective utilisée était celle de l'AM et les coûts

considérés étaient limités aux coûts directs médicaux. Un modèle de Markov a été implémenté sur l'horizon temporel de la vie entière. Les données sur la consommation de soins, les probabilités de transition entre les états et les temps passés dans chacun des états ont été recensées à partir de l'essai multicentrique EORTC 20981, de la littérature et d'une étude observationnelle sur les habitudes de prescription en France menée par « TNS Healthcare ». La consommation des ressources a été valorisée à partir des tarifs de remboursement de l'AM et des GHS. Le coût de la prise en charge des LF à partir du traitement d'entretien par rituximab sur l'horizon de la vie entière a été estimé à 71 314 €. Les coûts associés au traitement d'induction n'ont pas été considérés, ainsi que les coûts directs non médicaux et les coûts indirects. De plus, l'essai EORTC 20981 est une étude qui a recruté des patients dans 261 centres dont seulement deux en France. Les données issues de cet essai ne sont pas représentatives du contexte Français (221).

Ces études sont des études d'efficience dans lesquelles le coût a été évalué afin de l'associer à un critère d'efficacité. Elles n'évaluent pas le coût réel de la pathologie comme c'est le cas dans le cadre des études COI décrites dans la partie « III.1.1. » et à partir de données individuelles permettant de décrire finement les différents postes de dépenses. De plus, elles ne mobilisent pas de données d'observation permettant d'obtenir des résultats concordant aux pratiques réelles.

A notre connaissance, aucune étude française n'a évalué le coût de la prise en charge du LH, que ce soit dans le cadre d'une EME ou d'une étude COI.

Le lymphome, comme la majorité des cancers, impacte grandement la vie des malades et celle de leur entourage. Le diagnostic, les traitements et les EI associés peuvent provoquer des vulnérabilités physiques et psychologiques importantes et ainsi induire de l'anxiété, des dépressions ou d'autres morbidités psychiatriques (222). Ces pathologies peuvent également entraîner des modifications négatives dans la vie professionnelle des patients, telles que des démissions ou des retraites anticipées. Tous ces phénomènes participent à l'isolement des patients et impactent directement leur qualité de vie (223). L'entourage peut ainsi jouer un rôle fort dans la vie du patient et fournir un

temps d'aide non négligeable, tant pendant la période active de traitement que durant l'après cancer. Dans ce cadre, le lymphome induit un fardeau économique d'autant plus important si l'on se place du point de vue de la société. A titre d'exemple, une étude américaine a évalué le coût de l'aide informelle pour les patients survivants atteints de LNH, valorisée selon le salaire médian national, pendant les deux ans suivant le diagnostic à 59 613 \$ (224).

VI.2. Leucémie myéloïde chronique

Le terme « leucémie » fait référence à une maladie du sang caractérisée par une prolifération anormale des globules blancs ou leucocytes. Le terme « myéloïde » indique que l'origine de la maladie se situe dans la moelle osseuse au niveau des cellules myéloïdes, cellules immatures progressant en globules rouges, blancs ou en plaquettes. Le terme « chronique » signifie que la maladie s'installe lentement et progressivement (225).

VI.2.1. Carcinogénèse

Les différentes cellules sanguines (globules blancs, globules rouges et plaquettes) sont issues de la différenciation des CSH. Les CSH sont des cellules indifférenciées dites « immatures » en comparaison aux cellules sanguines différenciées dites « matures ». Elles sont créées dans la moelle osseuse grâce à leur capacité d'auto-renouvellement. Une CSH se divise ainsi en une CSH immature et une CSH en voie de différenciation. Une fois différenciée, la cellule sanguine mature passe dans le sang (226). Un patient atteint de LMC est ainsi caractérisé par une production anormale de cellules souches myéloïdes dans la moelle osseuse qui sont en cours de différenciation en granulocytes, un des types de globule blanc. Les granulocytes cancéreux passent ensuite dans le sang mais ne tiennent plus leur

rôle de défenseur immunitaire. De plus, cette surproduction endigue la bonne production des globules rouges et de plaquettes.

La production anormale de ces cellules souches myéloïdes immatures est liée à la fusion de deux chromosomes (9 et 22) des CSH. Ces deux chromosomes échangent mutuellement leurs gènes, c'est le processus de translocation. L'échange des gènes ABL (chromosome 9) et BCR (chromosome 22) entraînera l'apparition du chromosome de « Philadelphie ». Ce chromosome produit un gène BCR-ABL présent uniquement dans les cellules de la maladie et retrouvé chez 95% des LMC (227). Ce gène fabrique ensuite une enzyme, la tyrosine kinase, qui est responsable de la surproduction de ces globules blancs qui ont un renouvellement et une prolifération accélérée.

VI.2.2. Anatomopathologie et symptomatologie

La survenue d'une LMC se traduit généralement par trois signes cliniques visibles : une augmentation de la taille de la rate (splénomégalie), une fatigue intense accompagnée de fièvre et d'un amaigrissement et des saignements accentués au niveau des plaies, des gencives ou du nez. La fatigue ressentie provient de la diminution du nombre de globules rouges dans le sang (i.e. anémie) et les saignements sont dus à la diminution du nombre de plaquettes. Ces phénomènes sont expliqués par le remplacement des cellules normales de la moelle osseuse par des cellules leucémiques. Néanmoins les signes cliniques ne sont pas toujours présents. La LMC peut alors être détectée tardivement, une fois la maladie avancée et les signes cliniques visibles.

La numération de la formule sanguine (i.e. hémogramme) consiste à comptabiliser tous les éléments présents du sang et permet généralement d'identifier la présence de la LMC. Il est également possible d'observer la vitesse de prolifération des globules blancs à partir de cet examen.

Le chromosome de Philadelphie est ensuite recherché dans les cellules leucémiques à partir de la biopsie de la moelle osseuse. Différentes méthodes cytogénétiques ou moléculaires permettent de détecter ce chromosome dans l'échantillon ainsi prélevé (225).

VI.2.3. Epidémiologie et facteurs de risque

La LMC est le syndrome myéloprolifératif le plus fréquent, représentant 15 à 20% des cas de leucémie (228). C'est une maladie rare. Son incidence a été estimée en France en 2012 à 807 nouveaux cas (201). Elle survient plus souvent chez l'homme (59%) que chez la femme, et l'âge médian de son apparition est de 54 ans. De nos jours, l'espérance de vie des patients atteints de LMC se rapproche de l'espérance de vie de la population générale du même âge. La survie observée et la survie nette 5 ans après le diagnostic ont été estimées respectivement à 80% et 85% en France en 2012 (206). La LMC fait partie des hémopathies malignes ayant les meilleurs pronostics de survie, au même niveau que le LH, les LF ou d'autres hémopathies lymphoïdes chroniques.

Peu de facteurs de risques de la LMC sont connus aujourd'hui. Des incidences plus élevées ont été constatées suite aux bombardements d'Hiroshima et Nagasaki. L'exposition à de fortes doses de radiation est ainsi suspectée d'augmenter le risque de développer une LMC. Ce risque augmente également avec l'âge (225).

VI.2.4. Classification clinique

Contrairement aux lymphomes décrits dans les parties précédentes, classés selon la localité des lymphocytes cancéreux, le degré et la vitesse d'extension de la maladie, les cellules malignes leucémiques circulent dans le sang et sont donc réparties dans tout le corps. Le pronostic ne dépend pas de l'extension de la maladie. On parle alors de phase pour la LMC et on en distingue trois (229).

- La phase chronique. C'est la première phase de la maladie où la majorité des patients sont diagnostiqués. Elle dure en moyenne 4 à 5 ans en l'absence de traitement. Il y a encore peu de globules blancs anormaux dans la moelle osseuse et le sang et peu de symptômes ou bien ils sont insidieux. Les principaux symptômes visibles sont l'altération de l'état général (fatigue, amaigrissement, fièvre), l'augmentation de la taille de la rate et des signes de leucostase (i.e. accumulation anormale de globules blancs dans les vaisseaux capillaires sanguins créant des caillots) avec exceptionnellement du priapisme.
- La phase accélérée : c'est la transition entre la phase chronique et blastique. Cette phase dure généralement entre 12 et 18 mois mais peut être quasi inexistante lorsque la phase blastique est dite « explosive ». Elle correspond à une augmentation de la proportion de globules blancs anormaux et de l'accentuation des signes cliniques.
- La phase blastique : c'est la dernière phase de la maladie et la leucémie est considérée comme aigüe. La proportion de globules blancs anormaux est très importante et la moelle osseuse ne fonctionne plus correctement. Cette phase s'accompagne d'une accentuation des signes cliniques d'accélération de la maladie. Environ 10% des patients atteignent cette phase. Ces patients sont généralement les patients diagnostiqués tardivement ou qui n'ont pas répondu aux traitements.

VI.2.5. Diagnostic

Une LMC peut être suspectée chez des patients sans symptômes visibles, simplement suite à des résultats anormaux lors d'examens sanguins. Dans le cas contraire, un examen clinique permet d'identifier une éventuelle LMC si le patient présente : une forte fatigue, des saignements et des ecchymoses marquées et une augmentation du volume de la rate provoquant des gênes et douleurs abdominales et intestinales, ainsi qu'une satiété précoce (225). Suite à l'examen clinique du patient

et à la suspicion de LMC, des examens cytologiques sont réalisés dans le sang à partir d'un hémogramme et dans la moelle osseuse à partir d'une biopsie. Des taux anormaux de globules blancs (i.e. hyperleucocytose) et de plaquettes (i.e. thrombocytose) sont recherchés dans le sang ainsi qu'une augmentation du nombre de cellules dans la moelle osseuse (i.e. l'hyperplasie myéloïde). Des examens cytogénétiques sont effectués afin de mettre en évidence le chromosome de Philadelphie et des examens moléculaires sont réalisés pour identifier le transcrite BCR-ABL (230).

VI.2.6. Prises en charge thérapeutiques

VI.2.6.1 Traitements

La prise en charge de la LMC a grandement évolué depuis le premier traitement proposé à base d'arsenic au 19^{ème} siècle. La frise temporelle en annexe 11 résume cette évolution (231). La prise en charge des patients a été révolutionnée dans les années 2000 avec l'introduction des Inhibiteurs de la Tyrosine Kinase (ITK). Le gène responsable de la maladie produit une enzyme appelée tyrosine kinase qui entraîne la prolifération anormale des cellules. Les chercheurs se sont donc intéressés à inhiber cette enzyme. L'imatinib a été le premier ITK commercialisé pour le traitement de la LMC en 2001 (232). D'autres ITK ont été synthétisés afin d'améliorer l'efficacité de la prise en charge et de lutter contre certaines résistances et intolérances au traitement. On distingue les ITK de deuxième génération représentés par le dasatinib et le nilotinib et les ITK de troisième génération représentés par le bosutinib et le ponatinib. En France, le dasatinib, nilotinib, bosutinib et le ponatinib ont reçu leur Autorisation de Mise sur le Marché (AMM) respectivement en 2006, 2007, 2013 et 2014. De plus, les ITK de deuxième génération ont obtenu leur AMM en 2010 pour traiter des cas nouvellement diagnostiqués. Avec l'arrivée des ITK sur le marché, la survie globale des patients atteints de LMC est passée de 30% dans les années 1990 à 85% en 2014 (233). La seconde figure en annexe 11 schématise les avancées en terme de survie depuis la commercialisation des ITK (231).

L'espérance de vie des patients atteints de LMC se rapprochant de l'espérance de vie de la population générale du même âge, la maladie est devenue une maladie chronique avec une exposition aux ITK sur le long terme, ce d'autant que ces traitements ne sont que suspensifs (234).

De nos jours, un patient reçoit généralement de l'imatinib en première intention de traitement, encore considéré comme le « gold standard », mais peut également recevoir du nilotinib et dasatinib (235) (236). Cependant, la réponse à l'imatinib n'est pas absolue et certains patients développent des résistances ou des intolérances à cet ITK de première génération. En l'absence de réponse optimale, un patient pourra bénéficier d'une forte dose de l'ITK originellement prescrit ou des ITK de deuxième ou troisième générations (234). De plus, des études ont montré la supériorité des ITK de deuxième génération en 1^{ere} intention de traitement par rapport à l'imatinib (237) (238) (239) (240) (241). Ces traitements entraînent des taux de rémission plus élevés et une baisse du pourcentage de progression vers la phase accélérée ou blastique. Ils engendrent un nombre d'EI sensiblement identique à ceux associés à l'imatinib. Cependant ces EI sont parfois plus graves. Ces ITK peuvent alors être considérés comme des traitements en première intention mais les patients jeunes ayant de bons pronostics et peu de comorbidités seront privilégiés car supportant mieux les EI. Le dasatinib et le nilotinib n'ont pas encore été comparés entre eux, le choix entre l'un de ces traitements se fait alors selon le profil d'EI à éviter selon les caractéristiques du patient. Les ITK de troisième génération sont toujours en cours d'évaluation pour des dispensations en 1^{ere} intention de traitement et sont déjà utilisés lors de résistance ou d'intolérance à l'imatinib, notamment lorsque les ITK de seconde génération ne sont pas adaptés. Des études ont récemment démontrées la possibilité d'arrêt du traitement au bout d'une certaine période (242). Cependant, il n'existe encore aucun consensus autour de cette problématique.

VI.2.6.1 Effets indésirables des traitements

Les ITK provoquent des EI qu'il est d'autant plus essentiel d'appréhender que les traitements sont dispensés sur le très long terme. Le tableau suivant résume les principaux EI dus aux ITK de première et deuxième génération (237) (239).

Tableau 6. Effets indésirables rencontrés la 1ère année après l'initiation des Inhibiteurs de Tyrosine

Kinase de première ou seconde génération

El, toutes sévérités comprises	Imatinib % (239) - % (237)	Dasatinib % (239)	Nilotinib (300 mg) % (237)
Hématologiques			
Neutropénie	58 – 68	65	43
Thrombopénie	62 – 56	70	48
Anémie	84 – 47	90	38
Non-hématologiques			
Diarrhées	17 – 21	17	8
Nausées	20 – 31	8	11
Vomissements	10 – 14	5	5
Rash cutané	17 – 11	11	31
Maux de tête	10 – 8	12	14
Prurit	Na – 5	Na	15
Inflammation musculaire	17 – 24	4	7
Douleur osseuse	14 – Na	11	Na
Myalgie	12 – 10	6	10
Fatigue	10 – 8	8	11
Alopécie	Na – 4	Na	8
Rétention de fluide	42 – Na	19	Na
Effusion pleurale	0 – Na	10	Na
Œdème périphérique	Na – 14	Na	5
Œdème superficiel	36 – Na	9	Na

EI : Effets Indésirables ; 'Na' : Pas de données

Les EI associés aux traitements de la LMC sont généralement signalés comme peu fréquents, mineurs, tolérables et gérables. Il est tout de même primordial de les considérer étant donné le traitement à vie. Différentes recommandations sont alors proposées par l'« European Leukemia Network » afin d'aider les PS à faire face aux EI rencontrés (243) :

- La gestion des EI ne doit pas compromettre le traitement anti-leucémique.

- De par sa nature chronique, la plupart des patients auront des EI majoritairement faciles à prendre en charge.
- La réduction ou l'arrêt du traitement ne doit être effectué qu'en dernier recours, si la gestion optimale des EI nécessite cette interruption. Une surveillance est alors préconisée pour détecter précocement ces EI.
- Une attention particulière est nécessaire quant à l'étude des comorbidités, des interactions médicamenteuses et à la survenue d'autres événements inévitables, étant donné le traitement prolongé par ITK.

L'Imatinib, le dasatinib et le ponatinib présentent dans l'ensemble une bonne innocuité au long terme. Ils entraînent des EI qui sont, pour la majorité, faciles à prendre en charge. Les ITK de deuxième génération peuvent néanmoins entraîner des problèmes inattendus qui peuvent devenir irréversibles (i.e. effusion pleurale) (243).

Les innovations thérapeutiques dans le traitement de la LMC ont révolutionné sa prise en charge en permettant aux patients d'avoir une espérance de vie similaire à la population générale. Néanmoins, les traitements sont suspensifs et nécessitent un traitement au long cours. Cette chronicisation de la prise en charge de la LMC apporte un coût important pour l'AM comme pour la société.

VI.2.7. Coût de la Leucémie Myéloïde Chronique en France

La LMC est considérée comme une maladie rare, touchant un individu pour cent mille habitants. Néanmoins, les ITK ayant permis une hausse spectaculaire de la survie des patients atteints de LMC, la prévalence de la maladie a subi une augmentation importante (232). Les ITK sont des traitements innovants et onéreux et le traitement sur la vie entière engendre des coûts très importants pour l'AM.

Malgré cela, aucune étude en France n'a étudié les coûts en lien avec la prise en charge de la LMC. Une étude américaine a évalué le coût de la prise en charge des patients traités par imatinib sur la vie entière à 320 100\$ (244). Les ITK de deuxième et troisième génération sont par ailleurs plus coûteux que l'ITK de première génération (245). De plus, les possibles EI engendrés par les traitements peuvent être très coûteux. A titre d'exemple, le surcoût engendré par une effusion pleurale chez un patient traité par ITK a été estimé à 56 594\$ sur 6 mois (246).

La LMC a un bon pronostic de survie grâce à des traitements efficaces avec une bonne innocuité sur le long terme. La qualité de vie des patients est ainsi moins impactée par la maladie en comparaison à la majorité des lymphomes. Une diminution de la qualité de vie est néanmoins constatée pour les patients de moins de 60 ans et particulièrement ceux de moins de 40 ans (247). Cette détérioration provient notamment des impossibilités de travailler liées aux effets indésirables rencontrés, d'autant plus que les patients sont traités sur le très long terme (248). Les coûts associés à la perte de productivité pourraient ainsi être non négligeables, étant donné qu'une part importante des patients sont dans la vie active. Néanmoins, aucune étude n'a évalué le fardeau économique de la LMC d'un point de vue plus global dans le cadre d'une étude COI, en considérant les coûts indirects. De plus, à notre connaissance, les coûts associés à l'aide informelle n'ont jamais été étudiés. Étant donné que la majorité des patients ne voient pas leur qualité de vie impactée par la maladie, les coûts associés à l'aide informelle pourraient être moins élevés que dans le cadre des lymphomes.

De par sa prévalence en augmentation, son traitement pris sur le très long terme et la prise en charge des EI associés, la LMC présente un fardeau économique pour le système de santé. Il est primordial d'évaluer le coût de prise en charge de la LMC, d'identifier les postes de dépenses les plus importants en termes financiers et les traitements les plus efficaces en fonction du profil des patients afin d'optimiser la prise en charge médicale de cette pathologie chronique.

Les hémopathies malignes avec notamment les lymphomes et la LMC, représentent une entité complexe rassemblant un ensemble hétérogène de maladies et de prises en charge. Les traitements associés sont innovants et permettent de bons pronostics de survie pour la majorité des patients. Néanmoins, ces traitements efficaces sont lourds, coûteux et engendrent des EI non négligeables. Ces pathologies représentent ainsi un fardeau économique pour l'AM qui les prend en charge à 100% et pour la société. Néanmoins, pas ou peu d'études ont évalué les coûts associés à leur prise en charge en France, à partir de données de vie réelle.

PARTIE 3. RESULTATS : EVALUATION DES COÛTS DES LYMPHOMES ET DE LA LEUCEMIE MYELOÏDE CHRONIQUE

VII. Evaluation des coûts des lymphomes

Ce chapitre présente les travaux effectués portant sur l'évaluation des coûts de la prise en charge thérapeutique des lymphomes en France. L'incidence de ces pathologies et l'introduction de traitements innovants de plus en plus coûteux incitent fortement à une meilleure connaissance des stratégies de prise en charge et de leurs conséquences économiques. L'article ci-dessous soumis à la revue « Journal of Medical Economics » est actuellement en cours de révision. Un résumé en français précède l'article.

VII.1. Résumé de l'article "COSTS OF HODGKIN AND NON-HODGKIN LYMPHOMAS IN FRANCE: RESULTS BASED ON THE FRENCH NATIONAL HEALTH INSURANCE DATABASE"

Introduction : Les lymphomes sont des hémopathies malignes se plaçant au 6ème rang des cancers les plus fréquents dans le monde en 2011. Ils se divisent en deux grands types de lymphomes : les Lymphomes Hodgkiniens (LH) et Non Hodgkiniens (LNH) avec respectivement une incidence estimée à 1757 et 11512 nouveaux cas en France en 2012. Les lymphomes constituent une entité vaste de pathologies. Ils sont majoritairement traités par de la chimiothérapie en association ou non avec du rituximab, un anticorps monoclonal onéreux. La phase active du traitement est caractérisée par une forte toxicité qui entraîne un nombre important d'effets indésirables. Les prises en charge diagnostiques, thérapeutique ainsi que les effets indésirables consomment des moyens conséquents, notamment dans le cadre des soins hospitaliers, générant des coûts importants. Toutefois, aucune étude n'a évalué à ce jour en France le coût des lymphomes en population générale.

Objectif : L'objectif de l'étude était d'évaluer et de comparer les coûts directs et indirects de la phase active du traitement des LH et principaux sous type de LNH à cellules B (B-LNH) : les Lymphomes Folliculaire (LF) et les Lymphomes B diffus à Grandes Cellules (LBDGC). Les coûts des patients

présentant un LH ou un B-LNH ont été comparés à des individus témoins, indemnes de la maladie et appariés selon l'âge et le sexe avec les cas.

Méthode : L'étude a été réalisée à partir d'un échantillon aléatoire représentatif de la population française couverte par l'AM, l'EGB. Les coûts pris en compte étaient les coûts directs et indirects. Les coûts intangibles n'ont pas été évalués. Le point de vue adopté était celui du payeur, l'AM. L'étude des coûts et de leurs composantes a porté sur la phase active du traitement. L'utilisation d'un modèle linéaire généralisé a permis d'évaluer les différences de coûts entre les types de lymphomes.

Résultats : L'analyse a porté sur 224 cas de lymphomes et 896 témoins. Les coûts totaux différentiels estimés pendant la phase active de traitement étaient respectivement de 32 832€, 44 539€ et 46 708€ pour les LH, LF et LBDGC. Après ajustement sur les covariables d'intérêt, les LF et les LBDGC étaient les groupes de lymphomes les plus coûteux. Les principaux postes de coûts étaient représentés par les séjours hospitaliers, les médicaments onéreux (rituximab), la consommation médicamenteuse en soins de ville et à la perte de productivité (indemnités journalières et pensions).

Conclusion : Dans un contexte, marqué par des ressources de plus en plus contraintes et la nécessité pour les autorités de santé de disposer de résultats observés en vie réelle, cette étude permet de disposer d'une évaluation détaillée des coûts associés à la prise en charge thérapeutique, en phase active de traitement, des lymphomes en France.

VII.2. REAL WORLD COSTS OF HODGKIN AND THE MAIN B-CELL NON-HODGKIN LYMPHOMAS IN FRANCE

REAL WORLD COSTS OF HODGKIN AND THE MAIN B-CELL NON-HODGKIN LYMPHOMAS IN FRANCE

Michael Mounié^{1, 2, 3}, Nadège Costa^{1, 2}, Cécile Conte⁴, Dominique Petiot⁵, Didier Fabre⁵, Fabien Despas^{4, 6}, Maryse Lapeyre-Mestre^{3, 4, 6}, Guy Laurent⁷, Nicolas Savy^{3, 8}, Laurent Molinier^{1, 2, 3, 5}

1 : *Unité d'Evaluation Médico-Economique, Centre Hospitalier Universitaire, Toulouse F-31059, France*

2 : *Institut National de la Santé et de la Recherche Médicale –INSERM, Unité Mixte de Recherche UMR 1027, F-31000 Toulouse, France*

3 : *Université Toulouse III Paul Sabatier, Toulouse, France*

4 : *Laboratoire de Pharmacologie Médicale, facultés de Médecine, 37 allées Jules Guesde, Université Toulouse III Paul Sabatier 31000 Toulouse*

5 : *Département d'Information Médicale, Centre Hospitalier Universitaire, Toulouse F-31059, France*

6 : *Service de Pharmacologie Clinique, CIC 1436, CHU Toulouse, F-31000 Toulouse, France*

7 : *Service d'hématologie CHU Toulouse, Institut Universitaire du Cancer-Oncopôle de Toulouse, 31100 Toulouse*

8 : *Institut Mathématiques de Toulouse, UMR 5219, CNRS, F-31062, Toulouse, France.*

Corresponding author:

Mr Mounié Michael

Unité d'Evaluation Médico-Economique, Département d'Information Médicale, Centre Hospitalier

Universitaire, 2 Rue Viguerie, 31300 Toulouse Cedex 09, France

Tel. +33677708149

E-mail: mounie.m@chu-toulouse.fr

ORCID: 0000-0003-0298-267X

Abstract

Background: Lymphomas are costly diseases that suffer from a lack of detailed economic information, notably in a real world setting. Decision makers are increasing search for Real World Evidence (RWE) to assess the impact, in real life, of healthcare management and to support their public decisions. Thus, we aimed to assess the real world net costs of active treatment phase of adult Hodgkin Lymphoma (HL), Follicular Lymphoma (FL) and Diffuse Large B Cell Lymphoma (DLBCL).

Methods: We performed a retrospective cohort study using population-based data from a national representative sample of the French population covered by the health insurance system. Cost analysis was performed from the French health insurance perspective and took into account direct and indirect costs (€2018). Healthcare costs were studied over the active treatment phase. We used multivariate modelling to adjust cost differences between lymphoma subtypes.

Results: Analyses were performed on 224 lymphoma patients and 896 controls. The mean additional costs due to HL, FL and DLBCL patients were respectively €32,832, €44,539 and €46,708 for the active treatment phase. The main additional cost driver was principally inpatient stay (hospitalization costs and costly cancer related drugs), followed by outpatient medication and productivity loss. When adjusted, DLBCL remains significantly the most costly lymphoma subtype.

Conclusion: This study provides an accurate assessment of the main lymphoma subtypes related cost with high magnitude of details in a real world setting. It supplies useful economics information for a rational healthcare spending while identifying where potentials cost saving could be realized.

Keys words: *Lymphoma; Cost evaluation; French health insurance; Population based database*

1. Introduction

Lymphomas are malignant hemopathies separated into two main subtypes: Hodgkin Lymphoma (HL) and Non Hodgkin Lymphoma (NHL). They account together for approximately 3–4% of cancers worldwide that placing them to the 6th rank of cancers in 2011 [1]. In France, the annual incidence of HL and NHL were respectively estimated to 1,757 and 11,512 new cases in 2012 [2]. HL are B-cell neoplasms and NHL are divided into B-cell and T/NK-cell neoplasms where B-cell NHL (B-NHL) are the most frequent, notably represented by Follicular Lymphoma (FL) and Diffuse Large B-Cell Lymphoma (DLBCL) [3]. The 5-years net survival of HL is equal to 85 % and vary widely according sub-types for B-NHL with 87% for FL and 57% for DLBCL [4].

Type and duration of treatment depend on lymphoma stages and sub-types. Treatment of lymphoma is mainly based on drug strategies that combine conventional cytotoxic chemotherapy with expensive medication as, for example, rituximab for B-NHL. Lymphoma patients may be treated by immunotherapy, radiation therapy and surgery [5, 6]. The active treatment phase is characterized by high toxicity with important magnitude of Adverse Events (AEs) [7]. The diagnosis, treatments and AEs may lead to important physical and psychological vulnerability [7, 8]. Thus, the high consumption of care and the resulting number

of hospitalizations lead to a significant cost supported by the social health insurance which should be assessed through a Cost-Of-Illness (COI) study.

The aim of a COI study is to identify and measure the costs of a particular disease. COI outcome tells how much society is spending on a particular disease and the amount that would be saved if the disease was eradicated [9]. They reveal the different cost components and their relative societal burden. COI studies help health policy makers to rationalize health care expenditure by identifying areas where potential economic savings can be realized [10]. Likewise, these studies supplies useful economic information to assess innovation in healthcare management.

Nowadays, decision makers are increasing interest for Real-World Evidence (RWE) [11]. It's allows to assess the impact, in real life, of healthcare management and to guide health authority in their public decisions. Nevertheless, only few studies assessing the cost of lymphoma in real world setting were found in Europe [12] or in North America [13-17]. They were focused on special treatments, particular line of treatments or AEs and they did not finely describe cost components of the active treatment phase. Only two studies have assessed lymphoma's costs in France, but not in real world setting [18, 19].

In this context, our aims were to assess direct and indirect net costs during the active treatment phase of HL, FL and DLBCL patients according to a population-based real

world database and to identify the most

2. Method

2.1. Study design, setting and population

We performed a population-based, retrospective, cohort study using a representative random sample of the French national health insurance database (“Système National des Données de Santé” (SNDS)), called the “Echantillon Généraliste des Bénéficiaires” (EGB).

The SNDS is a national medical and administrative database, allowing the access to health care consumption and corresponding reimbursements for 98.8 % of the French population [20]. The EGB database is a representative sample according to age and sex corresponding to 1/97 of SNDS population and includes demographic data, ambulatory care reimbursement (including drug dispensing), inpatient care data, medical data (Long-Term Diseases (LTD) diagnoses), and characteristics of healthcare providers [21]. Patients may be enrolled in a special scheme called LTD which allows the reimbursement of 100% of disease-related costs. All these anonymized data are prospectively recorded and individualized. Therefore, in accordance with French legislation, approval by an ethics committee was not required.

important costs component.

2.2. Identification of the study population

Study population was identified using inpatient care data according to a validated incident identification lymphoma case algorithm [22]. HL (ICD code: C81), FL (ICD code: C82) and DLBCL (ICD code: C833) patients were considered. Exclusion period of prevalent lymphomas and inclusion period of incident lymphomas was ranged respectively from 01/03/2007 to 28/02/2009 and from 01/03/2009 to 28/02/2013. Follow up continued until 31/12/2015.

Study period was the active treatment phase identified between 01/03/2009 and 31/12/2015 and defined by: 1) an index date per patient as the first ICD-10 lymphoma discriminant code minus 21 days. 2) The end of the active treatment phase defined by the date of the last ICD-10 code of lymphoma management if there is no more discriminant ICD-10 codes of lymphoma management associated during the following 12 weeks. Index date was chosen 21 days before the first ICD-10 lymphoma code to take into account all potential clinical examination related to diagnosis of lymphoma before the first lymphoma related hospitalization. The 12 weeks without ICD-10 codes of lymphoma management were chosen to take into account the different healthcare courses of lymphoma sub-types, notably the maintenance therapy for FL, and the variability of management according to patients.

We did not consider T-cell lymphoma because algorithm cannot identify them correctly given their particular healthcare management. Beside, we have not considered others B-Cell lymphoma to avoid misclassification bias and because FL and LBDGC are globally representative of B-NHL management (i.e. indolent and aggressive lymphomas). For each case, we have randomly selected from the EGB 4 controls among patients not suffering from lymphoma to estimate the net cost of lymphoma. Controls were matched on gender and age. We did not use a propensity score to match our control according the low number to socio-demographic data we had. Controls have the same follow up duration of their associated cases.

2.3. Costs estimates

Cost analysis was performed from the French health insurance perspective and included direct and indirect costs. Direct medical costs corresponded to inpatient stays, outpatient care (i.e. visits, medical and paramedical procedures), drugs and medical equipment. Direct non-medical costs were limited to transportation costs, and indirect costs to daily allowance and disability pension. Costs were estimated by multiplying the number of resources used by the corresponding reimbursement tariff given by the French health insurance (Online Resource 1).

In France, public and private hospital payment is based on DRG system. Inpatient stays were valued using DRG tariffs and the price of

additional expensive drugs. Outpatient cares were valued according to the French Common Classification of Medical Acts (CCAM). Medication and medical equipment were valued with the French health insurance tariffs. Transportations, paramedical acts and visits were valued using the General Classification of Professional Act (NGAP). Costs linked to productivity loss were valued using the daily benefits given by the health insurance for short absences in the workplace and using disability pension for long-term sick leave.

Additional expensive drugs were divided into rituximab and others expensive drugs. DRG were split into 4 classes: lymphoma's diagnoses, lymphoma's treatments, lymphoma's complications and other hospitalizations. All inpatient stays combining both lymphoma's treatments and lymphoma's complications or diagnoses are grouped in lymphoma's treatments category. Outpatient medications were categorized using the main ATC (Anatomical Therapeutic Chemical) classification groups [23].

All costs were inflated to corresponding 2018 prices using the French Consumer Price Index (CPI) from the Organisation for Economic Co-operation and Development (OECD) website [24].

2.4. Others data

We have derived baseline comorbidities and Charlson Comorbidity Index (CCI) from Bannay and al. paper during the year before index date [25]. They use medical procedures,

drugs and discharge diagnosis in hospital in addition to LTD scheme to identify the comorbidities and the CCI. Finally, they apply new weights to a better assessment of CCI. The related comorbidities identification algorithm and the new weights are described in Bannay and al. paper and were implemented in the EGB for cases and controls. Length of the active treatment phase and place of management (i.e. private vs public) were also described for lymphoma patients. Deaths were not taken into account in analysis because of: 1) the exact day of death was unknown and 2) the difficulty to consider the death through the construction of the active treatment phase. Indeed, death occurred after the active treatment phase because we considered the last lymphoma related ICD-10 code as the end of the active treatment phase.

2.5. Statistical analysis

Descriptive statistics were used to summarise the data and included mean \pm Standard Deviation (SD) or quantiles, for continuous variables and occurrences with percentages for qualitative variables. Baseline characteristics were compared between lymphoma patients and controls using the Z-test for quantitative variables and Fisher's exact test for qualitative variables.

Lymphoma patient's costs were compared to those of the controls to estimate the additional cost of different lymphoma subgroups. Costs were described in terms of mean and their Bias-Corrected and accelerated (BCa)

bootstrap 95% Confidence Intervals (CI) which adjusts for skewness distribution of cost. Depending on cost distribution, cost differences between lymphoma patients and their controls were tested using a statistical Z-Test or a Mann Whitney Wilcoxon non-parametric test.

A Generalized Linear Model (GLM) with gamma distribution and log link was implemented to adjust the cost differences between lymphoma subgroups with covariates. Age in quantile, gender, CCI and the place of care were used as adjustments covariables. Age and gender were maintained in the model despite if they were non-significant. Statistical analyses were performed using R software (version 3.1.2).

3. Results

3.1. Patient characteristics

Lymphoma patients and controls identification are summarized in figure 1 and descriptive statistics according to lymphoma sub-types are synthesized in table 1. All comorbidities are summarized in the supplementary files (Online Resource 2).

The mean age of HL, FL and DLBCL patients was respectively 46.9 ± 19 , 63 ± 13.2 and 65.9 ± 16.1 years with 40.4%, 46.9% and 55.3% of women. HL and DLBCL patients had not significantly higher CCI than controls ($p=0.172$, $p=0.084$) contrary to FL patients

($p < 0.01$). This significant difference principally came from the important number of others cancer types in FL group (Online Resource 2). Furthermore, DLBCL controls have higher proportion of diabetes than cases. The mean follow-up with the first and third quantile were respectively 199 days [155; 253], 419 days [104; 860] and 186 [126; 232] for HL, FL and DLBCL patients. Management by private hospital structures accounted between 9% for FL and 13.5% for DLBCL.

3.2. Cost analysis

Mean additional costs (Table 2) due to HL, FL and DLBCL patients were respectively €32,832, €44,539 and €46,708 for the active treatment phase. The main cost drivers were inpatient stay, medication, and productivity loss. All significant cost differences between lymphomas and controls are detailed in the online supplementary material (Online Resource 3).

Inpatient stay was the most important cost driver amounted to a significant additional cost ($p < 0.01$) of respectively €22,929, €36,180 and €36,285 for HL, FL and DLBCL patients. These additional costs were mainly led by: 1) hospitalization costs arising from the diagnosis and the treatment, especially for DLBCL with respectively €3,614 and €13,140; 2) expensive drug costs coming from rituximab for FL and DLBCL patients with respectively €22,723 and €14,303 and from other expensive drugs for HL patients with €7,661.

The second most important cost driver was outpatient medication which amounted to a significant additional cost ($p < 0.01$) of respectively €4,153, €3,662 and €5,778 for HL, FL and DLBCL patients. These additional costs were firstly led by antineoplastic and immunomodulators drug costs which were mainly composed by the three following growth factor drugs (ATC code: L03AA): filgrastim, lenograstim and pegfilgrastim. Secondly, drug costs were explained by blood and hematopoietic organs drugs which are driven by the antianemic preparations (ATC code: B03).

The third main cost driver come from productivity loss which amounted to a significant additional cost ($p < 0.01$) equal to €3,468, €2,103 and €1,592 for respectively HL, FL and DLBCL and led by the daily allowance.

Table 3 shows the cost variations associated to lymphoma subtypes. The CCI was not significant in both univariate and multivariate analysis and was excluded from the model. LBDGC patients are associated with a significant higher cost of healthcare management than HL patients (RR=1.47; 95% CI: 1.04; 2.08, $p=0.030$) while FL patients tend to amount to a higher cost than HL patients (RR=1.42; 95% CI: 0.98; 2.07, $p=0.068$). Beside, patients who are managed in private center are associated with a 2-times significant cost decrease than patients who are managed in public center (RR=0.53; 95% CI: 0.35; 0.79, $p=0.003$).

4. Discussion

Only few data are available on the COI of lymphoma sub-types using population based data. Five studies were found in North America and one in Europe. American studies were focused on HL, DLBCL or FL patients according to special treatments or line of treatments [13-17]. Moreover, the healthcare system in the US or Canada is considerably different than the French healthcare system especially because of health care organization, reimbursement conditions and different unit prices used to value healthcare resources which makes harder the comparison between studies. According to the European context, Wang and al. used decision model to predict cost of DLBCL patient based on UK's population based database [12]. They were focused on costs of diagnosis, treatment, supportive care, follow-up and end-of-life care. They have estimated life time costs of patient treated with first line treatment to £22,122. Our results are higher despite the lower follow up we used. That may be explained by the larger part of cost component we took into account. Unfortunately, authors did not detail all inpatient cost components and we cannot compare our results.

There is a lack of economic evaluation of lymphoma in France. We reviewed only two studies on the topic [18, 19]. Best and al. in 2005 assessed the cost of DLBCL patients from the French payer perspective and treated by rituximab to €41,952 over 15 years with 33.6% due to rituximab. Our results are quite

similar considering that we took into account larger magnitude of cost component while they considered a larger follow up period. Deconinck and al. in 2010 estimated the cost of FL treated with Rituximab during maintenance therapy. Nowadays maintenance therapy with rituximab is considered as the gold standard and we consider it as a part of active treatment phase [26]. Furthermore, both of these studies have considered French healthcare perspective but they did not only use French data.

Our study is the first which assesses HL, FL and DLBCL related cost in France, using real world population-based data. In a context where decision makers are increasing interest for RWE, our study provides detailed economic information about HL and the main NHL subtype's management. Our finding spotlight where management could be improved in terms of quality of care and cost saving.

We noticed that FL and DLBCL are significantly the most costly lymphoma despite management differences, notably according follow-up duration. Related expenditures are mainly lead by rituximab. This cost could be largely reduced in the future, especially with the recent availability of two biosimilars in France, truxima® and rixathon® in 2017. According to the European Society for Medical Oncology (ESMO), biosimilar medications present a necessary and timely opportunity for physicians, patients and healthcare systems [27]. This medication could lead to an important cost saving but naturally depending

on the extent of biosimilar adoption [28]. Physicians and patients express reservations regarding biosimilar efficacy, safety and cost saving [29]. In this context, RWE provided in our study will be helpful to assess the budget impact or cost-effectiveness associated with these biosimilars. Findings associated may help to convince stakeholders about biosimilar benefit in real world setting and to enhance product uptake [30].

We have observed a high inpatient's DRG cost related to hospitalization stay for treatment, notably for DLBCL patients (€13,140). This cost is mainly determined by the duration of hospital stay that is longer than one night (€10,564), while chemotherapy is generally administered in day hospitalization setting. Chemotherapy is toxic and the resulting complications occur frequently during a cycle of chemotherapy. Nevertheless, they are managed at the following admission for chemotherapy administration, which explains these results and justifying the low cost related to complications alone. That highlights the need to prevent avoidable complications to reduce this cost with promoting telemedicine programs or improving therapeutic patient education [31-33]. In this context, a better management of AEs could leads to improve of patient quality of life, reduce day absence from work and decrease productivity loss related cost, the third cost driver in our study [34]. In addition, subcutaneous rituximab instead intravenous rituximab could reduce treatment burden for B-NHL patient and improve resource utilization [35].

The main strength of our study is the use of real life data with a population-based reimbursement database. The French health insurance databases are the best population based data source for performing economic studies in France [21]. It provides exhaustive reimbursement data of the healthcare consumption of a large part of French population with detailed inpatient and outpatient financial data which allows thorough analysis. The second main strength of our work is the use of a validated algorithm to select our population. The validation study shows a great sensibility of the algorithm allowing a correct identification of cases. In addition, false negative may concern patients never hospitalized for their lymphoma because different disease management or a gap between diagnosis and treatment [22]. Thus, according to the exhaustiveness of health care consumption in the EGB, the use of validated algorithm to these data is a great of interest to conduct economic evaluation.

Our results present some limitations. The first one is due to the high misclassifications rate on lymphoma sub-types: approximately 20% of diagnoses change after an expert review and the most frequent discrepancies were among lymphoma sub-types (i.e. 41.3%) [36]. We have defined which subgroups lymphoma patient belongs according to type, number and recency of ICD-10 code found to avoid it. In addition, we have arbitrary chosen 12 weeks without ICD-10 codes of lymphoma management after the last ICD-10 code related to lymphoma to conclude about the end of

active treatment phase. That allows taking into account the large gaps between guidelines and the practice arising from patient and hospital characteristics and affecting treatment adherence [37]. That could lead to consider few patients which early relapsed and started a second active treatment phase and an overestimation of our results. However, overestimation is small according to the limited number of relapsed lymphoma patients during 12 weeks after the end of active treatment phase [38-40].

In a context where decision maker increasing interest for RWE, our study provide important information with high magnitude of details on lymphoma related costs that will help health policy to better understand the healthcare management of lymphoma [11]. This work underlines the strength to work with a powerful tool as the health insurance database using validated algorithm to supply useful RWE information for assessing novelties in healthcare management.

References

1. Roman E, Smith AG. Epidemiology of lymphomas. *Histopathology*. 2011 Jan;58(1):4-14
2. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, Forman D, Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer*. 2013 Apr;49(6):1374-403. doi: 10.1016/j.ejca.2012.12.027. Available at <http://eco.iarc.fr>. Accessed on march 1, 2018.
3. Chihara D, Nastoupil LJ, Williams JN, Lee P, Koff JL, Flowers CR. New insights into the epidemiology of non-Hodgkin lymphoma and implications for therapy. *Expert Rev Anticancer Ther*. 2015 May;15(5):531-44.
4. Monnereau A, Uhry Z, Bossard N, Cowppli-Bony A, Voirin N, Delafosse P, Remontet L, Troussard X, Maynadié M. Survie des personnes atteintes de cancer en France métropolitaine, 1989-2013. Partie 2 – Hémopathies malignes. Saint-Maurice : Institut de veille sanitaire ; 2016. 144 p.
5. Ansell SM. Hodgkin Lymphoma: Diagnosis and Treatment. *Mayo Clin Proc*. 2015 Nov;90(11):1574-83.
6. Ansell SM. Non-Hodgkin Lymphoma: Diagnosis and Treatment. *Mayo Clin Proc*. 2015 Aug;90(8):1152-63.
7. Kasi PM, Tawbi HA, Oddis CV, Kulkarni HS. Clinical review: Serious adverse events associated with the use of rituximab - a critical care perspective. *Crit Care*. 2012 Aug 31;16(4):231.
8. Mojs E., Warchol-Biedermann K., Sanborski W. What Do We Know About Psychological Outcomes of Lymphoma in Adults? *European Psychologist* (2017), 22, pp. 121-131. <https://doi.org/10.1027/1016-9040/a000285>.
9. Byford S, Torgerson DJ, Raftery J. Economic note: cost of illness studies. *BMJ* 2000;320:1335.
10. Koopmanschap MA. Cost-of-illness studies. Useful for health policy? *PharmacoEconomics* 1998;14:143–148.
11. Romio S, Sturkenboom M, Corrao G. Real-world data from the health decision maker perspective. What are we talking about? *Epidemiology Biostatistics and Public Health* - 2013, Volume 10, Number 3.
12. Wang, H. I., Roman, E., Crouch, S., Aas, E., Burton, C., Patmore, R., & Smith, A. (2018). A Generic Model for Follicular Lymphoma: Predicting Cost, Life Expectancy, and Quality-Adjusted-Life-Year Using UK Population-Based Observational Data. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 21(10), 1176-1185.
13. Morrison VA, Bell JA, Hamilton L & al. Economic burden of patients with diffuse large B-cell and follicular lymphoma treated in the USA. *Future Oncol*. 2018 Oct;14(25):2627-2642.
14. Changxia Shao, Jinan Liu, Wei Zhou, Monika K. Raut, Matthew Monberg, Xiting Cao, Alejandro D. Ricart & Arun Balakumaran (2018) Treatment patterns, health care resource utilization, and costs in patients with relapsed/refractory Hodgkin lymphoma treated with brentuximab vedotin, *Leukemia & Lymphoma*, DOI: 10.1080/10428194.2018.1508665
15. Bonafede, M., Feliciano, J., Cai, Q., Noxon, V., Princic, N., Richhariya, A., & Straus, D. J. (2018). Real-world analysis of cost, health care resource utilization, and supportive care in Hodgkin lymphoma patients with frontline failure. *ClinicoEconomics and outcomes research : CEOR*, 10, 629-641. doi:10.2147/CEOR.S178649
16. Richard T. Maziarz, Yanni Hao, Annie Guerin, Genevieve Gauthier, Marjolaine Gauthier-Loiselle, Simu K. Thomas & Lamis Eldjerou (2018) Economic burden following allogeneic hematopoietic stem cell transplant in patients with diffuse large B-cell lymphoma, *Leukemia & Lymphoma*, 59:5, 1133-1142, DOI: 10.1080/10428194.2017.1375100
17. Khor S, Beca J, Krahn M and al. Real world costs and cost-effectiveness of Rituximab for diffuse large B-cell lymphoma patients: a population-based analysis. *BMC Cancer*. 2014 Aug 12;14:586. doi: 10.1186/1471-2407-14-586.
18. Best JH, Hornberger J, Proctor SJ, Omnes LF, Jost F. Cost-effectiveness analysis of rituximab combined with chop for treatment of diffuse large B-cell lymphoma. *Value Health*. 2005 Jul-Aug;8(4):462-70.
19. Deconinck E, Miadi-Fargier H, Pen CL, Brice P. Cost effectiveness of rituximab maintenance therapy in follicular lymphoma: long-term economic evaluation. *Pharmacoeconomics*. 2010;28(1):35-46
20. Bezin J , Duong M, Lassalle R, Droz C, Pariente A, Blin P, Moore N. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 2017 Aug;26(8):954-962
21. Moulis G, Lapeyre-Mestre M, Palmaro A, Pugnet G, Montastruc J-L, Sailler L. French health insurance databases: What interest for medical research? *Rev Med Interne*. 2015 Jun;36(6):411-7
22. Conte C, Palmaro A, Grosclaude P, Daubisse-Marliac L, Despas F, Lapeyre-Mestre M. A novel approach for medical research on lymphomas: A study validation of claims-based algorithms to identify incident cases. *Medicine (Baltimore)*. 2018 Jan;97(2):e9418
23. Anatomical Therapeutic Chemical Classification. Available at: <http://ec.europa.eu/health/documents/community-register/html/atc.htm>. [Accessed June 1, 2017].

24. Organisation for Economic Co-operation and Development (OECD). Anon. 4. PPPs and exchange rates. Available at: https://stats.oecd.org/Index.aspx?DataSetCode=SNA_TABLE4. Accessed June 1, 2017.
25. Bannay A, Chaignot C, Blotière PO, Basson M, Weill A, Ricordeau P, Alla F. The Best Use of the Charlson Comorbidity Index With Electronic Health Care Database to Predict Mortality. *Med Care*. 2016 Feb;54(2):188-94.
26. Salles G, Seymour JF, Offner F, and al. Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial. *Lancet*. 2011;377:42-51
27. Taberero J, Vyas M, Giuliani R, and al. Biosimilars: a position paper of the European Society for Medical Oncology, with particular reference to oncology prescribers. *ESMO Open* 2016;1:e000142. doi: 10.1136/esmoopen-2016-000142
28. Ronnebaum, S. and Atzinger C. Enhancing Biosimilar Adoption With Real-World Evidence. *Health Economics and Outcomes Research (HEOR) article*. July/August 2018 Value & Outcomes Spotlight
29. Dolan, C. Opportunities and Challenges in Biosimilar Uptake in Oncology. *Am J Manag Care*. 2018 Jun;24(11):-S0
30. Simoens, S., Jacobs, I., Popovian, R., Isakov, L., & Shane, L. G. (2017). Assessing the Value of Biosimilars: A Review of the Role of Budget Impact Analysis. *PharmacoEconomics*, 35(10), 1047-1062.
31. Compaci G, Ysebaert L, Obéric L, Derumeaux H, Laurent G. Effectiveness of telephone support during chemotherapy in patients with diffuse large B cell lymphoma: the Ambulatory Medical Assistance (AMA) experience. *Int J Nurs Stud*. 2011;48(8):926–32. doi: 10.1016/j.ijnurstu.2011.01.008
32. Sirintrapun SJ, Lopez AM. Telemedicine in Cancer Care. *Am Soc Clin Oncol Educ Book*. 2018 May 23;(38):540-545.
33. World Health Organization. Therapeutic Patient Education Continuing Education Programmes for Health Care Providers in the Field of Prevention of Chronic Diseases. 1998
34. Khalid M, Kamal, Jordan R, Covvey, Ankur Dashputre, Somraj Ghosh, Surbhi Shah, Monali Bhosle, and Christopher Zacker. A Systematic Review of the Effect of Cancer Treatment on Work Productivity of Patients and Caregivers. *Journal of Managed Care & Specialty Pharmacy* 2017 23:2, 136-162
35. Andrew Davies, Claude Berge, Axel Boehnke, Subcutaneous Rituximab for the Treatment of B-Cell Hematologic Malignancies: A Review of the Scientific Rationale and Clinical Development. *Adv Ther*. 2017; 34(10): 2210–2231.
36. Laurent C, Baron M, Amara N, Haioun C, Dandoit M, Maynadié M. Impact of Expert Pathologic Review of Lymphoma Diagnosis: Study of Patients From the French Lymphopath Network. *J Clin Oncol*. 2017 Jun 20;35(18):2008-2017
37. Stienen, J. J., Hermens, R. P., Wennekes, L., van de Schans, S. A., van der Maazen, R. W., Dekker, H. M., ... PEARL study group (2015). Variation in guideline adherence in non-Hodgkin's lymphoma care: impact of patient and hospital characteristics. *BMC cancer*, 15, 578. doi:10.1186/s12885-015-1547-8.
38. Feldman, T., Gromko, L., Protomastro, E. A., Starovoitova, L., Mato, A. R., Glisch, C., Sadekh, M., Schoener, B., Stives, S., Valentinetti, M., Timberg, M., Bhattacharyya, P. K., Chow, K. F., Yang, X., Skarbnik, A. P., & Goy, A. (2014). Relapses of Diffuse Large B Cell Lymphoma in Rituximab Era Are Limited to the First Two Years after Frontline Therapy. *Blood*, 124(21), 1307. Accessed December 28, 2018. Retrieved from <http://www.bloodjournal.org/content/124/21/1307>.
39. Casulo C, Byrtek M, Dawson KL, Zhou X, Farber CM, Flowers CR and al. Early Relapse of Follicular Lymphoma After Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone Defines Patients at High Risk for Death: An Analysis From the National LymphoCare Study. *J Clin Oncol*. 2015 Aug 10;33(23):2516-22
40. Montanari F, Diefenbach C. Relapsed Hodgkin lymphoma: management strategies. *Curr Hematol Malig Rep*. 2014 Sep;9(3):284

Table 1. Baseline characteristics according to lymphoma subgroups (HL and B-NHL)

	HL			FL			DLBCL		
	Cases (n=57)	Controls (n=228)	P	Cases (n=64)	Controls (n=256)	P	Cases (n=103)	Controls (n=412)	P
Age (mean, (sd))	46.9 (19.0)	46.9 (19.0)	1	63 (13.2)	63 (13.2)	1	65.9 (16.1)	65.9 (16.1)	1
Gender (women, (%))	23 (40.4)	92 (40.4)	1	30 (46.9)	120 (49.6)	1	57 (55.3)	228 (55.3)	1
CCI (n, (%)) :									
0	47 (82.5)	201 (88.2)	0.172	41 (64.1)	204 (79.7)	0.009	75 (72.8)	337 (81.8)	0.084
1-2	7 (12.3)	23 (10.1)		18 (28.1)	43 (16.8)		20 (19.4)	57 (13.8)	
3-4	1 (1.8)	3 (1.3)		3 (4.7)	9 (3.5)		5 (4.9)	15 (3.6)	
≥5	2 (3.5)	1 (0.4)		2 (3.1)	0 (0)		3 (2.9)	3 (0.7)	
Management by Private	6 (10.5)			6 (9)			14 (13.5)		
Follow-up (mean, [Q1 - Q3])	199 [155 ; 253]			419 [104 ; 860]			186 [126 ; 232]		

HL: Hodgkin Lymphoma, FL: Follicular Lymphoma, DLBCL: Diffuse Large B Cells Lymphoma CCI: Charlson Comorbidity Index,

P: P value

Table 2. Mean costs of lymphoma cases, controls, and additional costs according to lymphoma subgroups (€, 2018).

		Inpatient stays	Medical visits	Medical acts	Paramedical acts	Medication	Medical equipments	Transportation	Productivity loss	TOTAL
HL	Cases (n=57)	23737 [15382; 48817]	147 [119; 177]	1103 [844; 1496]	283 [177; 501]	4408 [3335; 6341]	131 [73; 360]	1965 [1364; 2800]	3736 [2539; 5515]	35537 [25852; 61952]
	Controls (n=228)	808 [317; 2367]	71 [55; 112]	120 [86; 219]	41 [24; 89]	255 [164; 504]	42 [21; 109]	66 [18; 281]	268 [168; 453]	1685 [1084; 3680]
	ΔCost	22929	76	983	242	4153	89	1899	3468	32832
	P	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
FL	Cases (n=64)	37062 [29216; 53748]	305 [218; 411]	1394 [1035; 1736]	385 [209; 766]	4437 [3288; 5462]	382 [128; 1748]	967 [660; 1284]	2589 [1329; 4601]	47568 [37391; 61989]
	Controls (n=256)	882 [551; 1603]	211 [156; 398]	355 [289; 485]	130 [85; 234]	775 [495; 1748]	100 [67; 151]	34 [20; 68]	486 [299; 902]	3029 [2248; 4608]
	ΔCost	36180	94	1039	256	3662	282	934	2103	44539
	P	<0.001	0.164	<0.001	0.056	<0.001	0.229	<0.001	0.011	<0.001
DLBCL	Cases (n=103)	36794 [32768; 41827]	168 [141; 202]	1110 [942; 1323]	399 [289; 686]	6087 [5263; 7253]	198 [139; 310]	1501 [1188; 1982]	1764 [1190; 2515]	48143 [42784; 53941]
	Controls (n=412)	508 [311; 1013]	78 [70; 88]	133 [115; 155]	139 [84; 256]	309 [266; 374]	60 [39; 128]	24 [15; 37]	172 [99; 290]	1434 [1137; 1903]
	ΔCost	36285	90	977	260	5778	138	1477	1592	46708
	P	<0.001	<0.001	<0.001	0.005	<0.001	0.001	<0.001	<0.001	<0.001

HL: Hodgkin Lymphoma, FL: Follicular Lymphoma, DLBCL: Diffuse Large B Cells Lymphoma, P ; P value

Table 3. Gamma multivariate regression model on additional cost of lymphoma

Variable		RR	CI (95%)	P
Age	[18 ; 55]	1		
	[56 ; 70]	0.99	[0.7 ; 1.38]	0.934
	[71 ; 89]	0.8	[0.57 ; 1.12]	0.2
Gender	Female	1.05	[0.81 ; 1.37]	0.707
Place of management	Private	0.53	[0.35 ; 0.79]	0.003**
Lymphoma subtype	HL	1		
	FL	1.42	[0.98 - 2.07]	0.068.
	DLBCL	1.47	[1.04 - 2.08]	0.030*

HL : Hodgkin Lymphoma ; FL : Follicular Lymphoma ; DLBCL : Diffuse Large B-Cell Lymphoma ; RR : Relative Risk ; P : P Value.

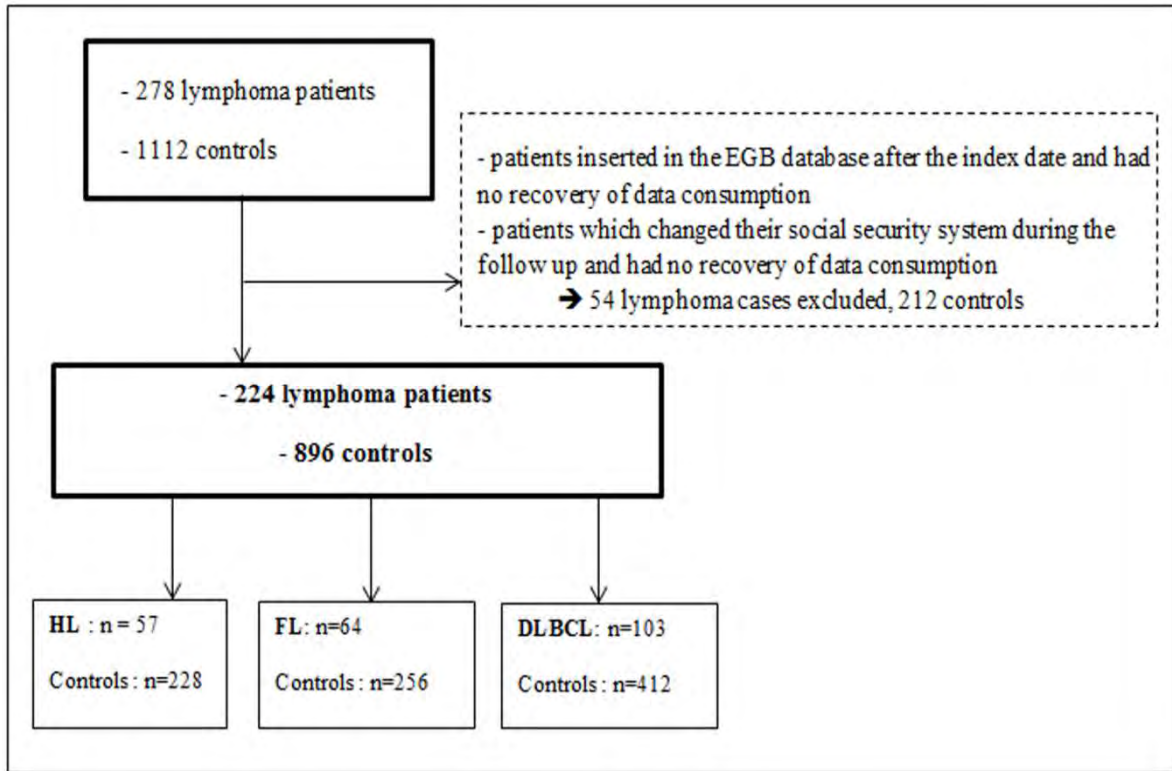


Figure 1. Identification of lymphoma's and control's population.

HL: Hodgkin Lymphoma, FL: Follicular Lymphoma, DLBCL: Diffuse Large B-Cell Lymphoma

Supplementary material

Online Resource 1. Unit costs - type and sources

<i>Consumed resources</i>	<i>Unit costs types</i>	<i>Unit cost sources</i>
Hospitalizations		
Inpatient stays	Disease Related Groups	Technical Agency of Hospital information ¹
Day hospital	Disease Related Groups	Technical Agency of Hospital information ¹
Hospitalization for rehabilitation	Per diem cost	Technical Agency of Hospital information ²
Psychiatric hospitalisation	Per diem cost	Technical Agency of Hospital information ³
General practitioner and specialist visits	General Classification of Professional Acts	Ameli (Health insurance) ⁴
Medical acts		
Imaging	Comon Classification of Medical Acts	Ameli (French Health Insurance) ⁵
Biology	Classification of acts of medical biology	Ameli (French Health Insurance) ⁶
Others	Comon Classification of Medical Acts	Ameli (French Health Insurance) ⁵
Paramedical acts	General Classification of Professional Acts	Ameli (French Health Insurance) ⁴
Medication	Medication database and tariffs information	Ameli (French Health Insurance) ⁷
Medical equipment	List of product and services	Ameli (French Health Insurance) ⁸
Transportation	General Classification of Professional Acts	Ameli (French Health Insurance) ⁹

1: <http://www.atih.sante.fr/manuel-des-ghm-version-provisoire-11g>; 2: <http://www.atih.sante.fr/financement-des-etablissements/ssr>; 3: <http://www.atih.sante.fr/financement-des-etablissements/psy>; 4: http://www.ameli.fr/fileadmin/user_upload/documents/NGAP.pdf; 5: <http://www.ameli.fr/accueil-de-la-ccam/telechargement/historique/ccam-version-43/index.php>; 6: http://www.codage.ext.cnamts.fr/codif/nabm/index_presentation.php?p_site=AMELI; 7: http://www.codage.ext.cnamts.fr/codif/bdm_it/index_presentation.php?p_site=AMELI; 8: http://www.codage.ext.cnamts.fr/codif/tips/index_presentation.php?p_site=AMELI; 9: <http://www.ameli.fr/professionnels-de-sante/transporteurs/votre-convention/tarifs/index.php>

Online Resource 2. Baseline comorbidities according to lymphoma subgroups

Comorbidities (n, (%))	HL			FL			DLBCL		
	Cases	Controls	P	Cases	Controls	P	Cases	Controls	P
Myocardial Infarction	1 (1.8)	3 (1.3)	1	2 (3.1)	3 (1.2)	0.262	1 (1)	6 (1.5)	1
Congestive Heart Failure	2 (3.5)	0 (0)	0.039	0 (0)	4 (1.6)	0.587	4 (3.9)	7 (1.7)	0.243
Periphral Vascular Disease	0 (0)	1 (0.4)	1	2 (3.1)	7 (2.7)	1	2 (1.9)	12 (2.9)	0.746
Cerebrovascular Disease	1 (1.8)	3 (1.3)	1	1 (1.6)	2 (0.8)	0.489	3 (2.9)	7 (1.7)	0.427
Dementia	0 (0)	2 (0.9)	1	1 (1.6)	2 (0.8)	0.489	0 (0)	10 (2.4)	0.224
Chronic Pulmonary Disease	1 (1.8)	8 (3.5)	0.693	1 (1.6)	15 (5.9)	0.21	4 (3.9)	28 (6.8)	0.363
Connective Tissue Disease	2 (3.5)	1 (0.4)	0.103	0 (0)	1 (0.4)	1	3 (2.9)	3 (0.7)	0.098
Ulcer Disease	0 (0)	0 (0)	1	1 (1.6)	0 (0)	0.2	0 (0)	0 (0)	1
Mild Liver Disease	1 (1.8)	2 (0.9)	0.489	1 (1.6)	2 (0.8)	0.489	2 (1.9)	1 (0.2)	0.103
Diabetes	1 (1.8)	14 (6.1)	0.318	3 (4.7)	24 (9.4)	0.316	0 (0)	42 (10.2)	<0.001
Hemiplegia	0 (0)	1 (0.4)	1	0 (0)	1 (0.4)	1	1 (1)	1 (0.2)	0.36
Moderate Or Severe Renal Disease	0 (0)	1 (0.4)	1	1 (1.6)	0 (0)	0.2	1 (1)	5 (1.2)	1
Diabete With End Organ Damage	0 (0)	1 (0.4)	1	0 (0)	1 (0.4)	1	0 (0)	3 (0.7)	1
Any Tumor (except lymphoma)	5 (8.8)	12 (5.3)	0.347	16 (25)	28 (10.9)	0.007	18 (17.5)	31 (7.5)	0.004
Moderate Or Severe Liver Disease	1 (1.8)	0 (0)	0.2	0 (0)	1 (0.4)	1	0 (0)	0 (0)	1
Metastatic Solid Tumor	2 (3.5)	1 (0.4)	0.103	2 (3.1)	0 (0)	0.039	3 (2.9)	1 (0.2)	0.027
HIV.AIDS	1 (1.8)	0 (0)	0.2	1 (1.6)	0 (0)	0.2	1 (1)	0 (0)	0.2

HL : Hodgkin Lymphoma, FL: Follicular Lymphoma, DLBCL: Diffuse Large B Cell Lymphoma, P: P-value

Online Resource 3. Detailed costs (€, 2018) according to cost drivers and lymphoma subgroups

	HL			FL			DLBCL		
	Cases (n=57)	Controls (n=228)	ΔCost	Cases (n=64)	Controls (n=256)	ΔCost	Cases (n=103)	Controls (n=412)	ΔCost
Inpatient stays	23737 [15382;48817]	808 [317; 2367]	22929	37062 [29216; 53748]	882 [551; 1603]	36180	36794 [32768; 41827]	508 [311; 1013]	36285
Additional expensive drugs:	8040 [1457; 38245]	555 [93; 2046]	7485	25004 [17878; 39770]	-	25004	15693 [13463; 19868]	88 [0; 265]	15605
Rituximab	379 [49; 1133]	-	379	22723 [15852; 35421]	-	22723	14303 [12163; 18658]	0 [0; 0]	14303
Other expensive drugs*	7661 [1204; 31488]	555 [109; 2844]	7106	2281 [1261; 4028]	-	2281	1390 [818; 2541]	88 [0; 265]	1302
DRG tariffs :	14512 [11607; 20490]	210 [128; 327]	14302	11359 [8819; 13968]	614 [396; 1297]	10746	19737 [16271; 22048]	327 [212; 530]	19411
Diagnosis	3878 [2540; 6242]	-	3878	1971 [1198; 3207]	-	1971	3614 [2613; 4732]	-	3614
Treatment	8704 [6946; 12376]	-	8704	7332 [5881; 10265]	-	7332	13140 [10700; 15507]	-	13140
Adverse event	529 [199; 1295]	-	529	258 [84; 738]	-	258	1049 [593; 1748]	-	1049
Other hospitalizations	1401 [735; 2784]	210 [121; 325]	1191	1792 [1085; 2978]	614 [393; 1237]	1178	1743 [1131; 2467]	327 [213; 540]	1416
Productivity loss	3736 [2539; 5515]	268 [168; 453]	3468	2589 [1329; 4601]	486 [299; 902]	2103	1764 [1190; 2515]	172 [99; 290]	1592
Daily allowance	3302 [2258; 5101]	133 [69; 247]	3169	2390 [1229; 4376]	106 [45; 289]	2284	1404 [877; 2155]	45 [18; 125]	1359
Medical acts	1103 [844; 1496]	120 [86; 219]	983	1394 [1035; 1736]	355 [289; 485]	1039	1110 [942; 1323]	133 [115; 155]	977
Laboratory tests	487 [391; 645]	30 [24; 42]	457	614 [458; 809]	93 [75; 124]	521	509 [444; 585]	39 [32; 45]	470
Imaging Acts	136 [103; 183]	22 [16; 30]	114	272 [188; 359]	89 [70; 113]	182	175 [135; 224]	37 [30; 46]	137
Technical Flat rates	445 [278; 740]	7 [3; 12]	438	426 [283; 630]	32 [21; 59]	394	370 [266; 513]	11 [7; 18]	359
Transportation	1965 [1364; 2800]	66 [18; 281]	1899	967 [660; 1284]	34 [20; 68]	934	1501 [1188; 1982]	24 [15; 37]	1477
Ambulance	95 [39; 233]	41 [2; 229]	54	114 [55; 250]	7 [4; 13]	106	392 [263; 607]	13 [7; 28]	379
Taxi	870 [534; 1403]	9 [0; 40]	861	527 [321; 794]	19 [7; 49]	835	677 [501; 951]	4 [2; 7]	673
VSL	878 [470; 1692]	5 [1; 15]	873	283 [158; 500]	5 [2; 19]	847	378 [268; 577]	7 [3; 12]	371
Medications (**)	4408 [3335; 6341]	255 [164; 504]	4153	4437 [3288; 5462]	775 [495; 1748]	3662	6087 [5263; 7253]	309 [266; 374]	5778
anti-infectious for systemic use (J)	349 [188; 934]	9 [6; 12]	340	350 [228; 568]	19 [13; 30]	331	267 [197; 412]	9 [7; 12]	258
Antineoplastics and immuno-modulators (L)	2609 [1860; 3608]	23 [0; 78]	2586	2405 [1708; 3169]	220 [15; 1398]	2184	3720 [3184; 4360]	34 [11; 101]	3686
Growth Factors (L03AA)	2521 [1827; 3512]	-	2521	2235 [1629; 3065]	1 [0; 3]	2234	3546 [3010; 4155]	1 [0; 3]	3545
Filgrastim	488 [233; 994]	-	488	511 [212; 1274]	-	511	546 [382; 757]	-	546
Lenograstim	1572 [1048; 2285]	-	1572	571 [323; 953]	-	571	1476 [1122; 1990]	-	1476
Pegfilgrastim	461 [184; 1008]	-	461	1153 [634; 1873]	1 [0; 3]	1152	1523 [1067; 2110]	1 [0; 3]	1522
Blood and hematopoietic organs (B)	616 [322; 1204]	10 [4; 34]	606	799 [473; 1278]	43 [25; 79]	756	1405 [1068; 2021]	14 [10; 20]	1391
Antithrombotic (B01)	68 [26; 207]	10 [4; 35]	56	288 [130; 790]	20 [15; 27]	268	268 [171; 413]	20 [15; 27]	249
Antianemic preparations (B03)	483 [234; 1086]	-	483	466 [237; 873]	3 [1; 9]	463	991 [708; 1544]	3 [1; 9]	988
Blood substitutes and infusion solutions (B05)	65 [7; 351]	-	65	21 [8; 72]	0 [0; 1]	20	102 [31; 361]	0 [0; 1]	101
Nervous system	95 [67; 139]	59 [31; 155]	36	158 [88; 373]	95 [68; 150]	62	112 [85; 160]	47 [37; 63]	65
Digestive tracks and metabolism (A)	529 [395; 711]	26 [19; 37]	503	270 [203; 332]	97 [68; 148]	173	272 [226; 330]	53 [39; 81]	220
Total	35537 [25852; 61952]	1685 [1084; 3680]	33852	47568 [37391; 61989]	3029 [2248; 4608]	44539	48143 [42784; 53941]	1434 [1137; 1903]	46708

All cost differences presented are significant ($p < 0.01$). HL: Hodgkin Lymphoma, FL: Follicular Lymphoma, DLBCL: Diffuse Large B Cells Lymphoma, Other expensive drugs *: All medication not taken into account in DRG as expensive drugs except Rituximab, (**): related ATC code, -: No consumption found

VIII. Evaluation du coût de la leucémie myéloïde chronique

Ce chapitre présente les travaux portant sur l'évaluation des coûts de la prise en charge thérapeutique des patients atteints de LMC dans l'ex-région Midi-Pyrénées, et particulièrement sur l'impact économique des différents Inhibiteurs de Protéines Kinases.

L'article ci-dessous a été publié dans le vol 36, n°2-3, 109-12 de la revue « Journal de gestion et d'économie médicale » (249). Le résumé de l'article figure ci-dessous.

**EVALUATION DU COUT DE LA PRISE EN CHARGE ES PATIENTS ATTEINTS DE LEUCEMIE MYELOIDE
CHRONIQUE (LMC) EN MIDI-PYRENEESS : IMPACT ECONOMIQUE DES DIFFERENTS INHIBITEURS DE
PROTEINES KINASES**

Michael Mounié^{1,2,3}, Nadège Costa^{1, 2}, Martin Gauthier^{2,4},
Aurore Palmaro^{2,3}, Cyrille Delpierre^{2,3}, Fabien Despas^{2,3,5},
Maryse Lapeyre Mestre^{2,3,5}, Guy Laurent^{2,4}, Nicolas Savy^{3,6}, Laurent Molinier^{1,2,3}

1 : Unité d'Evaluation Médico-Economique, CHU Toulouse F-31059, France

2 : UMR 1027, INSERM, facluté de médecine, Toulouse, France

3 : Université de Toulouse III – Paul Sabatier, Toulouse, France

4 : Département d'Hématologie et de médecine Interne, Institut Universitaire du Cancer-
Oncopole, 31100 Toulouse

5 : Service de Pharmacologie Clinique, CIC 1436, CHU Toulouse, F-31000 Toulouse, France

6 : Institut Mathématiques de Toulouse, UMR 5219, CNRS, F-31062, Toulouse, France.

Auteur correspondant : Mr Mounié Michael

Unité d'Evaluation Médico-Economique, Département d'Information Médicale, Centre
Hospitalier Universitaire, 2 Rue Viguerie, 31300 Toulouse Cedex 09, France

Tel. +33677708149

E-mail: mounie.m@chu-toulouse.fr

Résumé

La Leucémie Myéloïde Chronique (LMC), syndrome myéloprolifératif fréquent, représente 15 à 20% des cas de leucémie. L'introduction des Inhibiteurs de Protéine Kinase (IPK) a révolutionné le pronostic de survie des patients, transformant la LMC en maladie chronique avec une prise en charge sur le long terme et particulièrement coûteuse.

L'objectif de l'étude est d'estimer le coût de la prise en charge des patients atteints de LMC en Midi-Pyrénées et particulièrement l'impact économique des différents Inhibiteurs de Protéines Kinases. L'analyse économique a été réalisée du point de vue de l'Assurance Maladie à partir des données du Système National d'Information Inter Régime de l'Assurance Maladie (SNIIRAM). Les coûts directs médicaux ont été étudiés de 3 mois avant jusqu'à 12 mois après l'initiation du traitement. Ces coûts ont été ajustés au moyen d'un modèle multivarié longitudinal.

Le coût moyen de la prise en charge, pour l'année suivant l'initiation des IPK, a été évalué à 35 521€ (+/- 12 984). Le poste de dépense le plus important était les IPK, représentant 87% du coût total de la prise en charge. Une prise en charge par IPK de seconde génération entraînait une augmentation du coût de 51% à 84%.

Cette étude est la première étude évaluant le coût de la prise en charge de la LMC en France. Elle met en évidence le caractère onéreux de la LMC et renforce la nécessité de mettre en œuvre des travaux complémentaires en prenant en considération une population d'étude plus importante et en intégrant des postes de dépenses supplémentaires.

Mots clés : Leucémie Myéloïde Chronique, coût, économie de la santé, Inhibiteurs de Protéines Kinases

Abstract

Chronic Myeloid leukemia (CML) is a common myeloproliferative disorder, accounting for 15 to 20% of all leukemia cases. The introduction of Protein Kinase Inhibitors (PKI) has significantly changed the CML patient's survival and turned it into a chronic disease with a long term and costly care management. The aim of this study is to estimate the total annual cost of CML patients in Midi Pyrenees and assess the economic impact of different Tyrosine Kinase Inhibitors.

Economic analysis was carried out from the French health insurance perspective. Direct medical costs were taken into account and assessed from 3 months prior to 12 month after the treatment start. Longitudinal multivariate model was used to adjust mean total costs.

Total cost of CML patient's management amounted to 35 211€ (+/- 15 573) for the first year after PKI's initiation. PKI are the main cost driver, accounting for 87% of total cost. Second-generation IPK bring a significantly cost increase ranged from 51% to 84%.

This study is the first dealing with the cost of CML patients in France. It shows the costly nature of this disease and highlights the need to conduct additional works considering a larger sample size and integrating all costs not included here.

Keys words: Chronic Myeloid Leukemia, cost, health economics, Protein Kinase Inhibitors

Introduction

La Leucémie Myéloïde Chronique (LMC) est un syndrome myéloprolifératif fréquent, représentant 15 à 20% des cas de leucémie. En France, son incidence a été estimée entre 800 et 1000 nouveaux cas par an [1]. La LMC est liée à une anomalie provenant d'une translocation réciproque entre les chromosomes 9 et 22 (formation du chromosome de Philadelphie) aboutissant à la fusion des gènes *BCR* (Breakpoint Cluster Region) et *ABL* (Abelson), formant une protéine chimérique BCR-ABL. Cette protéine de fusion présente une activité tyrosine kinase constitutivement active et indépendante des mécanismes de régulation, responsable du phénotype prolifératif. L'introduction, dans les années 2000, des Inhibiteurs de Protéine Kinase (IPK) a révolutionné la prise en charge de la maladie. En effet, depuis l'introduction de l'imatinib en 2001, la survie globale à 5 ans des patients atteints de LMC n'a cessé d'augmenter et est passée de 30% dans les années 90 à

85% en 2014 [2]. L'espérance de vie des patients atteints de LMC se rapprochant de l'espérance de vie de la population générale du même âge, la maladie est alors devenue une maladie chronique avec une exposition aux IPK sur le long terme [3].

Il existe à présent 5 types d'IPK dans le traitement de la LMC : l'IPK de première génération, l'imatinib ; les IPK de deuxième génération, le dasatinib et le nilotinib, et depuis peu, les IPK de troisième génération, le bosutinib et le ponatinib [4]. Un patient reçoit généralement de l'imatinib en première intention de traitement, encore considéré comme le « gold standard ». En l'absence de réponse optimale ou en cas d'intolérance à l'imatinib, il pourra bénéficier d'une forte dose d'imatinib ou d'IPK de deuxième et troisième générations [3, 4]. Tous ces traitements innovants sont onéreux. Ward MA et al en 2015 ont montré une augmentation significative de 30% du coût de la prise en charge des patients atteints de LMC

lorsque ceux-ci sont traités par des IPK de deuxième ou troisième génération [5]. Reed SD et al évaluent le coût de la prise en charge des patients atteints de LMC en 2004 traités par imatinib à 320 100\$ sur un horizon temporel correspondant à la vie entière [6]. Hoyle M et al ont montré en 2011 que le coût de la prise en charge des patients résistants ou intolérants à l'imatinib et traités par fortes doses d'imatinib était de 172 415£, tandis que ceux traités par du dasatinib coutaient 221 325£ [7]. Bien que les IPK aient démontré leur efficacité sur la survie, ces traitements causent un nombre important d'effets indésirables tels que des épanchements pleuraux, des cytopénies, des céphalées, du prurit, des rashes cutanés et d'autres phénomènes biochimiques anormaux [8, 9]. Guérin A et al ont estimé le coût relatif à une effusion pleurale chez des patients atteints de LMC et traités par IPK à 56 594\$ sur 6 mois [10]. En France, le coût des traitements en lien avec la LMC est pris en charge à 100% par l'Assurance

Maladie au titre de son inscription aux Affections de Longue Durée (ALD). Cependant, il n'existe pas de données françaises sur le coût de la prise en charge de la LMC.

Nous faisons l'hypothèse que le coût de la prise en charge de la LMC est principalement dû aux IPK et est particulièrement important lorsque des IPK de seconde génération sont prescrits.

L'objectif principal de cette étude est d'évaluer le coût de prise en charge, la première année de traitement, des patients atteints de LMC en Midi-Pyrénées et d'étudier les facteurs jouant un rôle sur les variations de ce coût, notamment le type d'IPK prescrit. L'objectif secondaire de l'étude est d'évaluer le coût des patients atteints de LMC 3 mois avant l'initiation des IPK.

Matériel et méthodes

Données

Nous avons réalisé une étude rétrospective observationnelle portant sur la population

de Midi-Pyrénées. Les données ont été extraites à partir du Système National d'Information Inter Régime de l'Assurance Maladie (SNIIRAM). Le SNIIRAM est une base médico-administrative rassemblant actuellement toute la consommation de soins et les remboursements associés de 94.5% de la population Française [11]. Cette base de données comprend également des données démographiques portant sur les individus (âge, sexe, statut marital, date de décès) ; les caractéristiques des prescripteurs, les comorbidités pour les individus pris en charge lors d'une hospitalisation et les ALD. Lors de la réalisation de l'étude, le SNIIRAM couvrait environ 86% de la population Française [11].

Critère de jugement

Les données du SNIIRAM étaient disponibles du 01/01/2011 au 31/12/2014. Les patients ont été inclus dans l'étude entre le 01/10/2011 et le 31/12/2014 et la période du 01/01/2011 au 01/10/2011 a permis l'identification des cas prévalents et

leurs exclusions. Les patients ont été identifiés à l'aide d'un algorithme considérant tous les patients majeurs comme ayant une LMC s'ils recevaient au moins un remboursement d'IPK relatif à la LMC (imatinib, dasatinib, nilotinib, bosutinib, ponatinib) et aucune indication de ces IPK pour d'autres pathologies. Les autres pathologies pouvant se traiter par ces IPK ont été identifiées à partir des actes médicaux, des séjours hospitaliers et des prescriptions médicamenteuses. L'algorithme au complet est décrit dans le travail de Gauthier M et al [12].

La période sur laquelle nous avons travaillé était trois mois avant le début du traitement jusqu'à un an après. Elle a ensuite été séparée en deux phases : la période « pré-cancer » correspondant aux 3 mois précédant le début du traitement par IPK et la période « cancer » correspondant à l'année suivant le début du traitement par IPK. Nous avons choisi d'étudier le coût des patients atteints de LMC sur la période « pré-cancer » afin d'appréhender la

consommation supplémentaire de médicaments avant le début du traitement [13] et les différents examens diagnostics relatifs à la LMC.

Calcul des coûts

L'analyse des coûts a été réalisée selon le point de vue de l'Assurance Maladie. Les coûts directs médicaux ont été pris en considération. Ils correspondaient aux coûts relatifs aux hospitalisations, aux actes médicaux (i.e. consultations et actes externes) et aux médicaments. Les séjours hospitaliers ont été valorisés selon les tarifs en vigueur associés aux Groupes Homogènes de séjour (GHS). La valorisation des actes médicaux et des médicaments s'est faite selon les tarifs appliqués par l'Assurance Maladie. Les coûts ont été estimés en multipliant le nombre de ressources utilisées par leurs tarifs respectifs. Les coûts sont présentés en euros 2014 et ont été ajustés à l'aide de l'indice français des prix à la consommation [14].

Analyses statistiques

Les statistiques descriptives utilisées pour décrire notre population étaient le nombre d'individus, la proportion par rapport à la population totale étudiée pour les variables discrètes et la médiane avec le premier et troisième quartiles (Q1-Q3) pour les variables continues.

Les coûts ont été discrétisés selon les différents postes de dépenses et par périodes de 3 mois. Ils ont été décrits par leurs moyennes sur la population entière et leurs intervalles de confiance à 95%.

Des boîtes à moustaches ont été utilisées pour décrire les coûts des IPK au cours du temps et selon les différents types d'IPK.

Un modèle multivarié longitudinal, avec une distribution gamma et une fonction de lien logarithmique, utilisant les « Generalised Estimating Equation » (GEE) a été mis en œuvre, sur la période « cancer », afin d'évaluer l'impact du type d'IPK prescrit sur le coût de la prise en charge. Ces modèles permettent de prendre

en compte des données longitudinales et de considérer ainsi les possibles changements de traitement au cours du temps. Cette modélisation permet d'obtenir des approximations proches de la réalité en prenant en compte différents paramètres de confusion par rapport à la simple analyse descriptive. Les variables prises en compte dans le modèle étaient toutes les variables d'intérêts ($p < 0,20$) ainsi que l'âge et le sexe qui ont été forcées jusqu'au modèle final. Le modèle final a été sélectionné à l'aide d'une méthode pas à pas descendante. En fonction de la relation entre le coût et les variables, et pour une meilleure interprétation, les variables du modèle ont été décomposées selon les quartiles où leurs modalités ont pu être rassemblées. Les analyses ont été effectuées sur le logiciel R (v3.1.2).

Éthique / consentement

Ce travail a été réalisé sur des données anonymes. Par conséquent, conformément à la législation française, l'approbation par un comité d'éthique n'était pas requise.

L'analyse de ces données est autorisée à des fins de recherche (loi nationale sur la protection des données personnelles: Commission nationale des technologies de l'information et décision sur la liberté n° 89-117) et a été réalisée sur la base d'un accord conclu entre notre équipe de recherche et le département médical régional de la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés.

Résultats

Au total, 103 patients atteints de LMC pris en charge dans l'ex région Midi-Pyrénées ont été identifiés par l'algorithme de sélection entre le 01/10/2011 et le 31/12/2014. Dix-neuf patients ont été exclus pour cause d'un suivi trop court (< 1 an à partir de l'initiation aux IPK). Finalement, 84 patients atteints de LMC ont été inclus dans l'étude. Les caractéristiques de ces patients sont présentées dans le tableau I. L'âge médian de ces patients était de 62 ans (Q1-Q3 : 50-73) et 47,6% étaient des femmes. Une

grande majorité des patients identifiés ont reçu en première intention de l'imatinib (72,6%) tandis que 11,9% ont bénéficié du dasatinib et 15,5% du nilotinib. Aucun patient au cours de la période d'étude n'a reçu de bosutinib ou ponatinib. Seulement 5 patients ont changé d'IPK au cours du suivi. Les patients présentaient peu de comorbidités en plus de la LMC, avec pour la grande majorité (78,6%) un score de comorbidités de Charlson faible (1 ou 2). Des antécédents d'autres cancers ont été retrouvés chez 17,9 % des patients, des antécédents de maladie cardiovasculaire chez 9,5% et des antécédents de maladie respiratoire chez 6%. Les patients atteints de LMC ne sont en général pas ou peu hospitalisés.

Le coût moyen de la prise en charge des patients atteints de la LMC sur la période totale de l'étude était de 37 756€ (+/- 13451) dont 2 235€ (+/-3515) pour la période « pré-cancer » et 35 521€ (+/- 12 984) pour la période « cancer ». Les coûts déclinés par poste de dépenses sont

présentés dans le tableau II. En se focalisant sur la période « cancer », nous constatons que les médicaments constituent le poste de dépenses le plus important dans la prise en charge des patients, avec 91% du coût total et 87% dus aux IPK seuls. Le coût lié aux IPK est plus important lors de la phase d'initiation en comparaison aux autres périodes de temps. On observe également une augmentation du coût relatif au dasatinib entre 6 et 9 mois après le début du traitement. Les postes de dépense les plus importants sont ensuite les hospitalisations et les actes médicaux représentant chacun 4% du coût total de la prise en charge effectuée lors de la période « cancer ». Le coût des actes médicaux est porté par les consultations chez le généraliste et le spécialiste, par les actes de biologie ainsi que les actes médicaux codés selon la Classification Commune des Actes Médicaux.

Nous observons un coût important porté par les hospitalisations durant la période

« pré-cancer », représentant 64,5% du coût des hospitalisations sur la période totale de l'étude. Ce coût élevé est lié à des séjours pour diagnostics de leucémie (tableau III). Sur les 38 patients ayant eu au moins une hospitalisation lors de la phase « pré cancer », plus de la moitié d'entre eux ont eu une hospitalisation pour bilan de leucémie, lymphomes, autres syndromes myéloprolifératif ou de la chimiothérapie. Parmi ces patients, 14 étaient hospitalisés dans le cadre d'un bilan pour « autre leucémie » avec un coût moyen de 432 €, représentant 30% du coût total des hospitalisations en phase diagnostique.

La figure 1 présente les boîtes à moustache avec les coûts des différents IPK prescrits au cours du temps sur la période « cancer ». Le coût lié aux IPK est plus important lors de la phase d'initiation et les coûts moyens sont plus élevés pour les patients pris en charge avec des IPK de seconde génération.

Les résultats du modèle multivarié final sont présentés dans le tableau IV. Les 3

premiers mois suivant le début du traitement étaient associés à une augmentation du coût significative de 64% (RR=1,64, IC: 1,44-1,87, $p<0,001$) par rapport au reste de la période de suivi. Les patients bénéficiant respectivement d'un traitement par nilotinib ou par dasatinib ont eu une prise en charge plus coûteuse de 51% (RR=1,51, IC: 1,23-1,82, $p<0,001$) et 84% (RR=1,84, IC : 1,50-2,25, $p<0,001$) par rapport à ceux recevant de l'imatinib. L'âge, le sexe, le score de comorbidité de Charlson et le nombre d'hospitalisations ne présentent pas d'effets significatifs sur le coût de prise en charge des patients atteints de LMC.

Discussion

Cette étude est la première à évaluer le coût de la prise en charge de la LMC avec des données françaises. Le coût total moyen de la prise en charge des patients a été estimé à 35 521 € pour la première année de suivi et 2 235 € pour les 3 mois précédant le début du traitement. Le principal poste de dépense pendant la

période « cancer » était la médication représentant 91% du coût total de la prise en charge et 87% sont dus aux IPK. Les patients recevant du nilotinib ou du dasatinib ont une prise en charge plus coûteuse de 50% et 84% respectivement, comparativement à ceux bénéficiant de l'imatinib. Cela s'explique en grande partie par le prix moyen des traitements. Le coût journalier moyen du traitement, est estimé à 77€ pour l'imatinib, 97€ pour le nilotinib et 110€ pour le dasatinib [15, 16, 17]. De plus, deux études ont relevé un nombre important d'effets indésirables tels que des épanchements pleuraux et des cytopénies chez les patients traités par dasatinib ou des céphalées, du prurit, des rashs cutanés et une élévation de la bilirubine et du glucose chez les patients traités par nilotinib [8, 9]. Ces effets indésirables plus fréquents chez les patients traités par IPK de seconde génération pourraient également contribuer à l'augmentation des coûts de la prise en charge des patients.

Le coût important observé au cours des trois premiers mois de la prise en charge thérapeutique s'explique par une prescription d'IPK supplémentaire au commencement du traitement en plus des prescriptions à chaque fin de mois.

L'augmentation du coût du dasatinib durant la période]6 ; 9] mois découle des 5 patients changeant de traitement lors de cette période et passant de l'imatinib au dasatinib. Ces patients débutent un nouveau traitement et on retrouve l'augmentation du coût due à la prescription supplémentaire en début de traitement.

Nous observons un coût hospitalier important pendant la période « pré-cancer », représentant 64,5% du coût hospitalier total observé sur la période totale d'étude. En effet, les séjours hospitaliers sont majoritairement concentrés dans la phase diagnostique car les patients atteints de LMC sont parfois présumés atteints d'une autre leucémie. Ces patients sont alors hospitalisés et des

compléments de bilan sont réalisés afin de confirmer le diagnostic. Nous retrouvons ainsi dans les tables du Programme de Médicalisation du Système d'Information (PMSI) pendant la phase diagnostique un grand nombre de séjours hospitaliers pour « autres leucémies » (n=14).

La comparaison des résultats de notre étude aux autres études de la littérature constitue un exercice délicat. La plupart des travaux ont été conduits sur un horizon temporel plus important que celui de notre étude. De plus, les différents types de systèmes de santé, le point de vue pris en compte, le périmètre des coûts retenus, l'année de l'étude, la nature du traitement et la ligne de traitement considérée jouent un rôle primordial sur les résultats d'étude économique [18]. Néanmoins, nous retrouvons dans ces différents travaux un coût de prise en charge significativement plus important lorsqu'un IPK de deuxième génération est prescrit [5, 7].

La principale force de notre étude est d'être bâtie sur une approche en

population, basée sur l'utilisation de données d'observation. En effet, ces données sont fondées sur des pratiques réelles et permettent d'obtenir des estimations proches de la réalité. Le SNIIRAM est une des bases de données les mieux adaptée aux études économiques et pharmaco-épidémiologiques en France [11]. Il permet l'accès à des informations démographiques, à la consommation de soins et aux remboursements extrêmement détaillés de 94.5% de la population française et ainsi de réaliser des analyses approfondies sur les pathologies observées en population générale.

Nous avons fait le choix de découper la période de suivi en intervalle de temps de 3 mois. Nous étions donc confrontés à la nécessité de travailler avec des données longitudinales pour lesquelles l'hypothèse d'indépendance entre les valeurs de notre variable d'intérêt n'était pas respectée. Nous avons utilisé une approche en population générale utilisant les GEE qui offrent la possibilité de corriger les

estimations des coefficients de régression à l'aide d'une matrice de corrélation préalablement choisie. Ces modèles permettent de considérer le temps comme une variable explicative et de prendre en compte les changements de traitement au cours de la période de suivi. L'utilisation de ce type de modèle apporte des estimations précises des effets des différents paramètres considérés sur les variations globales du coût de la prise en charge des patients.

Notre étude présente toutefois un certain nombre de limites. La première tient à l'absence de certains postes de dépenses. Malgré l'utilisation du SNIIRAM, nous n'avons pas eu accès aux actes paramédicaux, aux équipements médicaux et aux indemnités journalières. Nous pouvons émettre l'hypothèse que les coûts présentés dans cet article sont sous-estimés, ce d'autant que les patients atteints de LMC perçoivent des indemnités journalières de manière plus fréquente que la population générale [19]. La deuxième

limite est en lien avec la base utilisée qui a été construite à partir d'une extraction régionale des données de la base nationale du SNIIRAM. En plus des différences de pratiques entre les régions, le nombre de patients identifiés dans l'étude reste assez limité mais permet néanmoins une première estimation du coût de la LMC. De plus, l'horizon temporel considéré dans l'étude est également réduit. En effet, le SNIIRAM nous a permis d'accéder aux données sur les trois années précédant la demande jusqu'à l'année en cours. Toutefois seulement 84 des patients identifiés par l'algorithme présentaient un suivi de 1 an ou plus par rapport à l'initiation des IPK et 56 avaient un suivi de 2 ans ou plus. Nous avons alors restreint notre étude à 1 an après le début du traitement afin de disposer d'une taille de population suffisante pour assurer une puissance convenable aux tests statistiques réalisés. La troisième limite est liée au fait que nous n'avons pas évalué le coût propre de la LMC mais seulement le coût de la

prise en charge des patients atteints de LMC. En effet, nous ne disposions pas d'individus témoins permettant de comparer le coût des patients atteints de LMC aux individus issus de la population générale indemnes de cette pathologie et donc d'évaluer le surcoût propre à la LMC.

Cette étude constitue un travail pilote qui devra être poursuivi et complété par des futurs travaux menés à l'échelon national ayant pour objectif principal l'évaluation du coût de la LMC et répondant aux limites précédemment évoquées. Les coûts de prise en charge de cette maladie étant essentiellement expliqués par les remboursements d'IPK, l'apparition de génériques modifiera grandement les dépenses de l'assurance-maladie.

Conclusion

Cette étude permet de disposer d'une première estimation du coût de la prise en charge en France des patients atteints de

LMC en population générale. Elle met toutefois en évidence, la nécessité de réaliser des travaux complémentaires palliant aux limites présentes dans cette étude. Une analyse des différentes trajectoires de soins des patients, selon le type de traitement considéré couplé à l'étude d'un critère d'efficacité clinique en complément des données économiques, permettrait d'évaluer l'efficience de chacun des traitements, et ainsi d'apporter des informations utiles aux décideurs en les aidant à faire des choix dans l'allocation des ressources.

Remerciements

Ce travail a reçu le soutien de l'Agence Nationale de la recherche (ANR) dans le cadre des investissements d'avenir (ANR-11-PHUC-001, Programme CAPTOR).

Références

- 1- Sant M, Allemani C, Tereanu C, et al. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood* 2010; 116(19):3724-34.
- 2- Sacha T. Imatinib in chronic myeloid leukemia: an overview. *Mediterr J Hematol Infect Dis*. 2014; 6(1):e2014007.
- 3- O'Brien S, Radich JP, Abboud CN, et al. Chronic Myelogenous Leukemia, Version 1.2014. *J Natl Compr Cancer Netw JNCCN*. 2013 Nov;11(11):1327-40.
- 4- Baccarani M, Deininger MW, Rosti G, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. *Blood*. 2013 Aug 8;122(6):872-84
- 5- Ward MA, Fang G, Richards KL, et al. Comparative evaluation of patients newly initiating first-generation versus second-generation tyrosine kinase inhibitors for chronic myeloid leukemia and medication adherence, health services utilization, and healthcare costs. *Curr Med Res Opin*. 2015 Feb;31(2):289-97.
- 6- Reed SD, Anstrom KJ, Ludmer JA, et al. Cost-effectiveness of imatinib versus interferon-alpha plus low-dose cytarabine for patients with newly diagnosed chronic-phase chronic myeloid leukemia. *Cancer*. 2004 ;101(11):2574-83.
- 7- Hoyle M, Rogers G, Moxham T, et al. Cost-effectiveness of dasatinib and nilotinib for imatinib-resistant or -intolerant chronic phase chronic myeloid leukemia. *Value Health*. 2011 Dec;14(8):1057-67
- 8- Kantarjian H, Shah NP, Hochhaus A, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med*. 2010 Jun 17;362(24):2260-70.
- 9- Saglio G, Kim DW, Issaragrisil S, et al. Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. *N Engl J Med*. 2010 Jun 17;362(24):2251-9.
- 10- Guérin A, Wu EQ, Bollu VK, et al. The economic burden of pleural effusions in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors. *J Med Econ*. 2013;16(1):125-33
- 11- Moulis G, Lapeyre-Mestre M, Palmaro A, Pugnet G, Montastruc J-L, Sailler L. French health insurance databases: What interest for medical research? *Rev Med Interne Rev Med Interne*. 2015 Jun;36(6):411-7
- 12- Gauthier M, De Barros S, Palmaro A, Lapeyre-Mestre M, Despas F. Psychotropic drug initiation in patients having a CML diagnosed: a population-based study. Poster session presented at the 20th Annual Meeting of French Society of Pharmacology and Therapeutics, 19-21 April 2016, Nancy, France. [Abstract disponible : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/fcp.12190>]
- 13- Pearson S-A, Abrahamowicz M, Srasuebkul P, Buckley NA. Antidepressant therapy in cancer patients: initiation and factors associated with treatment. *Pharmacoepidemiol Drug Saf*. 2015 Jun;24(6):600-9
- 14- PPPs and exchange rates [Internet]. [cité 9 juin 2016]. Disponible sur: https://stats.oecd.org/Index.aspx?DataSetCode=SNA_TABLE4
- 15- Vidal 2016: https://www.vidal.fr/Medicament/glivec-66355-prescription_delivrance_prise_en_charge.htm
- 16- Vidal 2016: https://www.vidal.fr/Medicament/sprycel-77851-prescription_delivrance_prise_en_charge.htm
- 17- Vidal 2016: https://www.vidal.fr/Medicament/tasigna-84033-prescription_delivrance_prise_en_charge.htm
- 18- Drummond M, Sculpher M. Common methodological flaws in economic evaluations. *Med Care*. juill 2005;43(7 Suppl):5-14.
- 19- Horsboel TA, Nielsen CV, Nielsen B, Andersen NT, De Thurah A. Wage-subsidised employment as a result of permanently reduced work capacity in a nationwide cohort of patients diagnosed with haematological malignancies. *Acta Oncol*. 2015 May;54(5):743-9

Tableau I. Description de la population

Caractéristiques	Population totale
Nombre de patients	84
Sexe féminin, n (%)	40 (47,6)
Âge : médiane [Q1-Q3]	62 [50-73]
Type d'IPK prescrit à l'initiation, n (%) :	
Imatinib	61 (72,6)
Dasatinib	10 (11,9)
Nilotinib	13 (15,5)
Changement d'IPK durant le traitement, n(%)	5 (6)
Nombre d'hospitalisation, médiane [Q1-Q3]	1 [0-2]
Comorbidités, n (%) :	
Cancer	15 (17,9)
Maladie cardiovasculaire	8 (9,5)
Maladie respiratoire	5 (6,0)
Diabète	1 (1,2)
Score de Charlson, n (%) :	
2	66 (78,6)
3-4	16 (19,1)
>4	2 (2,4)

IPK : Inhibiteur de la Protéine Kinase, Q1-Q3: 1er et 3ieme quartile

Tableau II. Coût (€, 2014) détaillés selon les postes de dépenses et le temps

Temps	-3 ; 0] mois]0 ; 3] mois]3 ; 6] mois]6 ; 9] mois]9 ; 12] mois
	Moyenne [IC 95%]	Moyenne [IC 95%]	Moyenne [IC 95%]	Moyenne [IC 95%]	Moyenne [IC 95%]
Hospitalisations	1 443 [852 ; 2 033]	299 [57 ; 541]	448 [175 ; 721]	243 [58 ; 429]	548 [0 ; 1 277]
Actes médicaux	613 [445 ; 782]	439 [370 ; 509]	408 [309 ; 506]	286 [224 ; 348]	354 [183 ; 525]
Consultation	185 [157 ; 214]	195 [169 ; 222]	146 [125 ; 168]	141 [116 ; 166]	139 [112 ; 167]
Consultation généraliste	115 [97 ; 133]	110 [89 ; 131]	76 [63 ; 89]	82 [64 ; 100]	90 [69 ; 111]
Consultation spécialiste	70 [52 ; 88]	86 [72 ; 100]	71 [56 ; 85]	59 [44 ; 74]	49 [36 ; 62]
Acte CCAM	256 [104 ; 409]	87 [30 ; 143]	113 [25 ; 202]	58 [5 ; 112]	129 [-35 ; 293]
Actes d'imagerie	76 [37 ; 114]	12 [2 ; 23]	22 [7 ; 36]	10 [3 ; 17]	52 [-8 ; 111]
Autres actes CCAM	181 [59 ; 302]	74 [21 ; 128]	92 [6 ; 177]	48 [-5 ; 101]	77 [-37 ; 191]
Actes de biologie	172 [130 ; 214]	157 [132 ; 182]	148 [123 ; 174]	87 [71 ; 102]	85 [70 ; 101]
Médicaments	179 [107 ; 250]	11 453 [10 113 ; 12 794]	7 132 [5 988 ; 8 276]	7 272 [5 794 ; 8 750]	6 639 [5 500 ; 7 778]
IPK seuls	0 [0 ; 0]	11254 [9927 ; 12581]	6737 [5659 ; 7814]	6887 [5392 ; 8382]	6043 [4905 ; 7182]
Imatinib	0 [0 ; 0]	6826 [5641 ; 8011]	3981 [3048 ; 4913]	2696 [1837 ; 3556]	3176 [2314 ; 4039]
Nilotinib	0 [0 ; 0]	2387 [1068 ; 3705]	1627 [656 ; 2598]	1223 [480 ; 1966]	1377 [527 ; 2226]
Dasatinib	0 [0 ; 0]	2042 [675 ; 3408]	1129 [398 ; 1860]	2968 [1450 ; 4486]	1490 [561 ; 2419]
Autres	179 [108 ; 249]	199 [97 ; 301]	395 [-117 ; 907]	384 [46 ; 722]	596 [160 ; 1032]
Coût total	2 235 [1 483 ; 2 987]	12 192 [10 847 ; 13 537]	7 988 [6 794 ; 9 181]	7 801 [6 284 ; 9 318]	7 541 [6 067 ; 9 015]

IC : Intervalle de confiance, CCAM : Classification Commune des Actes Médicaux, IPK : Inhibiteur de Protéine Kinase

Tableau III. Détails des coûts des hospitalisations 3 mois avant le début du traitement (€, 2014)

GHM retrouvés	Nombre Individus]-3 ; 0] mois
	n	Moyenne [IC 95%]
Toutes hospitalisations	38	1 443 [852 ; 2 033]
Hospitalisation pour bilan de leucémie, lymphomes, autres syndromes myéloprolifératifs ou de la chimiothérapie (i)	20	517 [211 ; 823]
Hospitalisation pour bilan « autre Leucémie » (ii)	14	432 [131 ; 734]
Hospitalisation pour bilan: lymphome, leucémie ou syndrome myéloprolifératif (iii)	5	42 [3 ; 81]
Hospitalisation pour chimiothérapie (iv)	2	43 [-25 ; 110]
Autres hospitalisations (v)	21	926 [381 ; 1471]

GHM : Groupe Homogène de Malade ; GHM retrouvés : i : 17M111, 17M112, 17M113, 17M11T, 28Z07Z, 17C051, 17C03J, 17M091, 17M12T, 17K07J ; ii : 17M111, 17M112, 17M113, 17M11T ; iii: 17C051, 17C03J, 17M091, 17M12T, 17K07J ; iv: 28Z07Z ; v: 02M072, 03C162, 03M111, 04M033, 05K061, 05K101, 05K10J, 05M092, 05M13T, 06C044, 06K04J, 08C461, 08c473, 08C492, 08M25T, 11M043, 11M12T, 12M041, 13C062, 16M091, 16M10

Tableau IV. Résultats du modèle multivarié

Variables		RR	P-value	IC (95%)
Sexe	Féminin	0,99	0,935	[0,82 ; 1,20]
Age (années)	<54	1		
	[54 ; 70]	1	0,99	[0,81 ; 1,23]
	>70	0,86	0,2500	[0,66 ; 1,11]
Intervalles de Temps (mois)]3 ; 12]	1		
]0 ; 3]	1,64	<0,001	[1,44 ; 1,87]
Principe actif IPK	Imatinib	1		
	Dasatinib	1,84	<0,001	[1,5 ; 2,25]
	Nilotinib	1,5	<0,001	[1,23 ; 1,82]

RR : Risque Relatif, IC : Intervalle de confiance, IPK : Inhibiteur de Protéine Kinase

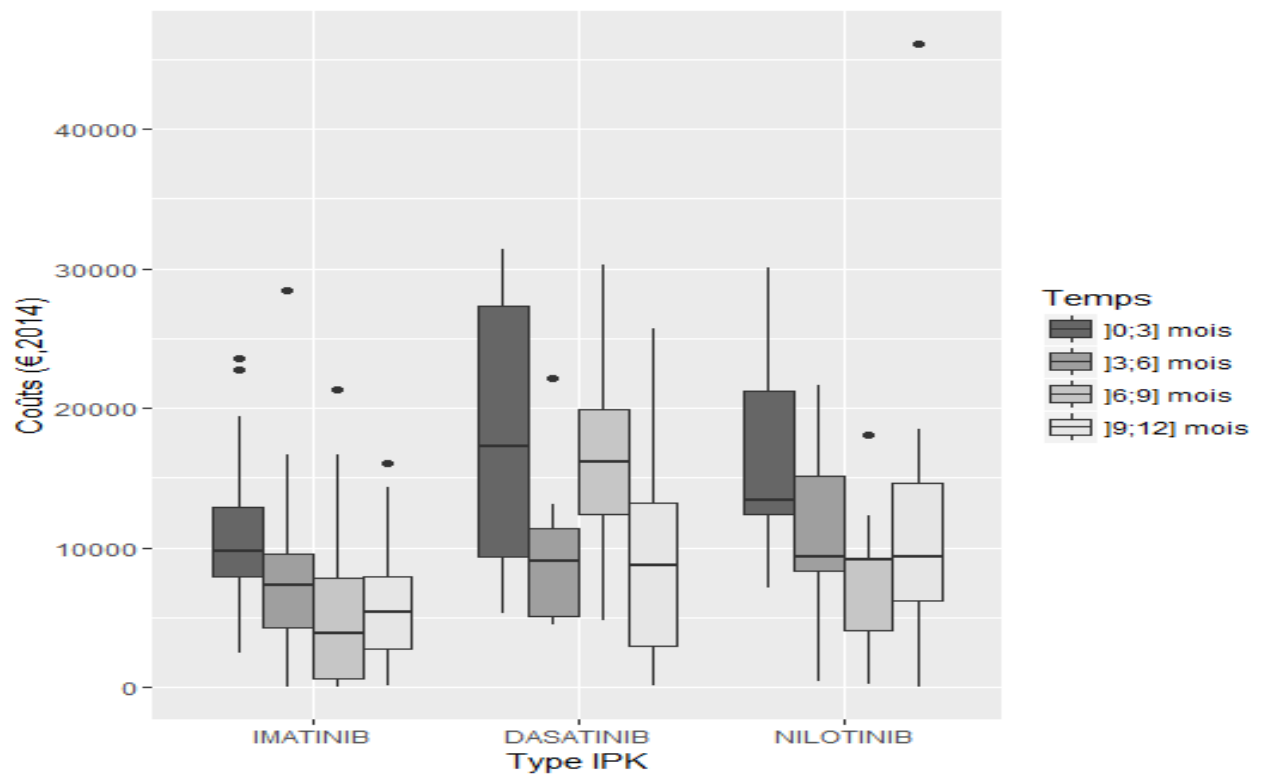


Figure 1. Boîtes à moustache des coûts des IPK selon le temps et le type d'IPK prescrit

IX. Discussion générale et perspectives

IX.1. Un contexte en pleine mutation

IX.1.1. Evolutions des prises en charge en cancérologie

Notre système de soins connaît actuellement des évolutions majeures avec une réorganisation de la prise en charge des patients en lien avec les progrès technologiques et thérapeutiques (5). L'organisation des soins rencontre ainsi de profondes transformations notamment provoquées par la place croissante de la médecine personnalisée, par l'arrivée de nouvelles formes de chimiothérapie orale et par le développement de l'ambulatoire (250).

La médecine personnalisée est en plein essor. Les caractéristiques physiques, sociales, psychologiques et culturelles jouent des rôles importants dans les résultats des stratégies de traitement (251). De plus, les progrès récents en génomique ont montré que la réussite d'un traitement peut dépendre des caractéristiques biologiques du patient (252). Ces avancés tant au niveau des connaissances que des techniques d'analyses et de gestion des « big data » associés ont également permis le développement de la médecine prédictive (253). Cette branche naissante de la médecine s'attache à identifier les individus à risque aggravé, nécessitant des prises en charge particulières en termes de prévention, de dépistage, de suivi ou de traitement (5). Les diagnostics personnalisés doivent ainsi être réalisés de manières coordonnées entre les différents professionnels de la santé. La prise en charge est ensuite définie selon les paramètres génotypiques et socio-économiques du malade afin d'optimiser son efficacité (254). Un suivi personnalisé et pluridisciplinaire du patient durant la phase thérapeutique et l'après-cancer constitue un élément primordial afin d'éviter les ruptures dans les parcours de santé des patients. Ce suivi adapté permet de prendre en considération les besoins des patients, de leur entourage et de favoriser la réinsertion du malade dans sa vie professionnelle et personnelle. L'ensemble de ces actions nécessitent des évolutions organisationnelles majeures et une grande coordination entre les différents

professionnels de la santé. Elles sont essentielles pour assurer une efficacité optimale des soins et une bonne qualité de vie aux malades.

Les traitements et leurs modes d'administration sont également en pleine évolution. La médecine orale et les soins ambulatoires connaissent un fort développement. On parle ainsi de « virage ambulatoire » (255). Le développement majeur de l'hospitalisation de jour comme de l'hospitalisation à domicile en lien avec la médecine de ville est de nature à améliorer la réponse aux besoins de santé. Les pouvoirs publics ont ainsi l'ambition d'atteindre un taux de 70% en chirurgie ambulatoire en 2022 (256). Selon l'étude UNICANCER mobilisant un panel d'experts, près de 50% des patients atteints d'un cancer pourraient bénéficier des anti-cancéreux oraux d'ici 2020 (257). Plus confortables pour les patients et plus économiques pour l'AM, ces modes de prises en charge ambulatoires ont cependant pour conséquence une gestion plus compliquée de la toxicité des traitements. Des indications précises couplées à une bonne organisation des professionnels de santé intervenant dans la prise en charge du patient sont ainsi essentielles afin de conserver l'efficacité des traitements (5). L'éducation thérapeutique a notamment un rôle fort dans ce contexte. Elle doit poursuivre son développement afin de lutter contre la mauvaise observance des traitements et de limiter les EI dus à leur toxicité. Le patient tient ainsi un rôle nouveau dans la prise en charge. Suite au développement des technologies de communication et d'information, il ne se place plus comme spectateur mais comme acteur de sa santé. C'est la médecine participative (258). Des progrès majeurs sont ainsi nécessaires en matière de coordination des prises en charge, de renforcement de l'accès aux soins primaires et à l'hospitalisation à domicile et de communication avec les patients (256).

Plusieurs « plans cancers » ont été mis en place par le gouvernement français depuis 2003 afin d'organiser la lutte contre cette maladie et d'améliorer la qualité des prises en charge. Le plan cancer actuel (2014 – 2019) s'articule autour de quatre axes principaux : guérir le plus de personnes malades, préserver la continuité et la qualité de la vie, investir dans la prévention et la recherche et

optimiser le pilotage et l'organisation des soins (5). Il est structuré autour de 17 grands objectifs et se déclinent en 208 actions. Il insiste particulièrement sur l'importance d'adapter les modalités d'organisation et de financement aux enjeux de la cancérologie. Le 4^{ème} axe souligne la nécessité d'évaluer l'efficacité des actions afin de garantir l'accès aux innovations et à des soins de qualité pour l'ensemble des patients. Cette recherche d'efficacité nécessite une connaissance approfondie des stratégies de prise en charge et un partage d'informations de nature économique. Il est primordial que les décideurs et les acteurs du système de santé puissent disposer d'informations robustes et pertinentes leur permettant d'améliorer la réponse aux besoins, tant des patients que des professionnels de la santé.

IX.1.2. Place de l'évaluation économique dans le système de santé français

Dans ce contexte marqué par de profondes mutations et des ressources contraintes, des choix budgétaires s'imposent. En 2017, la CSBM s'élevait à 8.7% du PIB, correspondant à 199.3 milliards d'euros (259). Son évolution en constante progression demeure supérieure à celle du PIB. L'estimation du coût des stratégies de prise en charge des patients est devenue primordiale, car elle peut aider à faire des choix dans l'allocation des ressources et permet un suivi économique des pathologies.

La cancérologie n'échappe pas à cette évolution. Les dépenses remboursées par l'AM pour des prises en charge en lien avec les cancers s'élevaient à 14.7 milliards d'euros en 2016 (260). Les dépenses en cancérologie sont vouées à progresser en raison du progrès technologique et médical, de la croissance démographique, du vieillissement de la population et d'une évolution de la demande (261) (262). Il est essentiel de prendre en compte ces évolutions afin de préserver la qualité de la réponse apportée aux besoins de la population.

L'évaluation économique des stratégies de prise en charge apporte des informations complémentaires à l'évaluation médicale. Elle a vocation à favoriser une allocation optimale des

ressources limitées dans le domaine de la santé. Ce type d'évaluation peut répondre à trois objectifs distincts :

- Estimer le fardeau économique d'une prise en charge ou d'une pathologie,
- Estimer le bilan économique net pour la société d'un programme de santé. Réaliser ce travail ambitieux suppose d'estimer dans une unité commune les coûts et les bénéfices générés par ce programme, sachant que ces coûts ne sont que partiellement et imparfaitement exprimables en unités monétaires,
- Choisir parmi plusieurs stratégies tendant à un même but, celles qui auront une efficacité médicale maximale pour un budget fixé ou un coût minimal pour une efficacité constante. Cette approche aboutit à classer les stratégies selon leur rapport coût-résultat respectif et à les mettre en œuvre successivement par ordre de rapports croissants jusqu'à atteindre les limites du budget disponible. Contrairement à la première approche, il ne s'agit pas de se prononcer sur le bien-fondé des programmes de santé et sur l'importance des budgets qu'il faut consentir, mais d'établir un ordre de priorité pour leur mise en œuvre respective.

En France, la place accordée aux évaluations économiques dans le domaine de la santé est encore modeste. Dans le cadre des missions de la HAS, la Commission Evaluation Economique et de Santé Publique (CEESP) a été créée pour contribuer à ce que la mesure de l'intérêt pour la société d'une stratégie ou d'un produit soit prise en compte dans les décisions les concernant, notamment celles de leur prix et de leur remboursement (263). La mesure de l'intérêt pour la société d'une nouvelle stratégie ou d'un produit, comparé à l'existant, est établie par l'étude de son efficacité. Cette analyse permet d'articuler de façon adéquate les objectifs de santé publique et les moyens qui y sont consacrés. La CEESP publie des recommandations en santé publique portant sur la qualité de la prise en charge, des rapports d'évaluation technologique/économique et des avis d'efficacité lorsqu'un produit de santé revendique une amélioration du service attendu ou une amélioration du service médical rendu élevée (niveau I, II, III) et est susceptible d'avoir un impact significatif sur les dépenses

de l'AM (263). Ainsi, dans ces cas-là, les résultats des EME peuvent participer à la négociation du prix des produits de santé (264). En opposition à d'autres systèmes de santé tel qu'au Royaume-Uni où l'EME est décisionnelle et détermine l'entrée d'un produit ou d'une technologie de santé sur la liste de remboursement, en France l'EME est seulement informationnelle (24). Cependant, ces évaluations se focalisent sur les médicaments ou DM innovants et coûteux. Le comité technique des vaccinations commande également des EME afin d'établir des stratégies vaccinales sur les nouveaux vaccins. Néanmoins, étant donné la nature informationnelle de l'EME, elle devrait pouvoir s'extrapoler à toute décision publique.

IX.1.3. Evolution des modes de financement : le financement au parcours de soins

La mise en place d'un dispositif de financement au parcours de soins constitue un axe majeur de la réforme de l'organisation et du financement du système de santé français. Les modalités de financement des établissements de santé ont évolué au cours des dernières décennies. Actuellement plusieurs modalités de financement coexistent en France : le financement des soins de ville, le financement hospitalier en partie fondé sur la T2A et le financement du secteur médico-social. A l'heure où les parcours de santé se mettent en place progressivement pour assurer une prise en charge globale des patients, les modes de tarification doivent s'adapter à l'évolution du système de santé (265).

Si la T2A dans son principe de mise en œuvre est relativement simple et qu'elle a permis un financement au regard de l'activité médicale, elle présente néanmoins plusieurs effets néfastes. Compte tenu du fait que la T2A ne prend pas en compte l'intégralité d'un parcours de soins, elle peut entraîner une augmentation du volume de soins. En effet, la pertinence des actes mis en jeu n'est pas toujours évaluée (266). En complément, la T2A, se focalisant sur la répartition des ressources contraintes au prorata de l'activité, freine les perspectives de développement de nouvelles activités,

d'innovations ou d'investissements, souvent bénéfiques pour le patient (267). C'est particulièrement le cas en cancérologie ou les modes de prise en charge évoluent très rapidement, parallèlement aux progrès technologiques.

Dans ce cadre, le concept de parcours de soins, de santé ou de vie a vu le jour afin que les patients reçoivent « les bons soins, par les bons professionnels dans les bonnes structures et au bon moment » (267). Ce concept de parcours s'explique par un enchaînement de soins ordonné, logique et coordonné entre tous les professionnels de soins, médico-sociaux ou sociaux (268). C'est précisément dans ce cas que réside tout l'intérêt du regroupement d'activités et des services en un seul paiement que les anglo-saxons appellent « bundling » ou « bundles payments » (269). Les « bundles payments » sont des modes de rémunération intermédiaires entre des paiements atomisés, comme la T2A ou la rémunération au prix de journée, et des paiements globalisés comme la dotation globale de financement ou le financement à la capitation. Les trois types de paiement regroupés les plus couramment retrouvés dans la littérature sont (270):

- « A l'épisode » qui vise majoritairement les actes chirurgicaux lourds et les soins médicaux aigus,
- « Au parcours de soins » qui vise notamment les maladies chroniques dont les cancers,
- « Shared savings programme » qui se focalise sur une population donnée et visent à inciter à la coordination dans le cadre d'une approche territoriale.

Ces modes de rémunération regroupés favorisent la coordination des différents professionnels de santé et la conception d'une offre de soins organisée autour du patient (267). Les professionnels de santé sont incités à se concerter et à prendre en considération l'ensemble des étapes. Ces modes de rémunération conduisent ainsi à une meilleure efficacité du système, une amélioration de la qualité et de la sécurité des prises en charge, notamment au moyen d'une réduction des taux de complication, de ré-hospitalisation et d'admission en urgence, et une réduction des coûts. Ils

présentent également l'intérêt de favoriser l'intégration des stratégies de prévention au parcours de soins (270).

Les évolutions apportées par ces modes de financement conduisent à un changement dans le positionnement des professionnels de santé en général, et de l'hôpital en particulier. Etant donné la diminution des ré-hospitalisations et de complications, une réduction du volume de l'activité hospitalière est attendue. En contrepartie, l'hôpital peut acquérir l'opportunité de jouer un rôle central comme coordonnateur de soins et détenteur de fonds (270).

IX.2. Evaluation des coûts des stratégies de prise en charge

IX.2.1. Fardeau économique des hémopathies malignes

Les pathologies cancéreuses représentent un enjeu de santé publique majeur en raison de leur forte incidence, de leur impact en termes de mortalité et morbidité et de leurs conséquences économiques. La majorité des hémopathies malignes est concernée par ces problématiques.

Les pronostics de survie associés à ces pathologies se sont fortement améliorés, notamment grâce aux avancées thérapeutiques majeures réalisées depuis 20 ans ainsi qu'aux prises en charge pluridisciplinaires (202). Ce gain d'efficacité ne va pas sans des coûts en forte progression supportés par la collectivité. D'une part, les innovations et les nouveaux modes de prise en charge consomment des moyens de plus en plus importants, et d'autre part, les améliorations de la survie globale engendrent des prises en charge nécessitant des traitements sur un plus long terme. La LMC et certains lymphomes (i.e. LF) sont ainsi traités sur des périodes qui peuvent aller de plusieurs années jusqu'à la vie entière. Les autorités de santé doivent s'adapter à ces évolutions et relever le défi financier qu'elles engendrent tout en préservant la qualité des soins.

Dans ce cadre, l'évaluation des coûts de ces pathologies est fondamentale. Elle apporte des informations détaillées sur les composantes de coûts associées aux prises en charge, à leurs poids et

leurs survenues, globalement et selon les caractéristiques sociodémographiques (i.e. âge, sexe, catégorie socio-professionnelle...) ou médicales (i.e. traitement utilisé, comorbidités, facteurs de risque, complications...) des individus. Les acteurs de la santé ont la possibilité d'utiliser ces éléments afin d'identifier les domaines où des économies potentielles sont réalisables et prioriser les recherches et financements en fonction (18). Ce type d'évaluation facilite également la réalisation d'évaluations de l'efficacité mettant au regard les éléments obtenus avec des informations complémentaires dans l'implémentation de modèles statistiques permettant l'extrapolation des résultats.

Toutefois, peu de données économiques sont disponibles aujourd'hui sur ces pathologies en France. Aucune étude n'a estimé le coût de la prise en charge du LH ou de la LMC. Seulement quelques études se sont penchées sur le coût des prises en charge des LNH (218) (220). Les données utilisées pour produire les résultats provenaient pour partie de la littérature internationale, non représentative du contexte français et des pratiques médicales. A ce titre, la comparaison des résultats d'évaluations économiques menées dans des pays différents, et donc des contextes différents, est difficile et largement discutée, notamment en raison des différences de disponibilité des traitements, de pratiques cliniques, de prix relatifs et de mesures incitatives concernant les professionnels de santé et les organisations (271). Dans le 4ème axe de l'actuel plan cancer, l'objectif 15 « Appuyer les politiques publiques sur des données robustes et partagées », insiste sur la nécessité d'utiliser des données produites de qualité dans les évaluations (5). Les informations transmises aux professionnels de santé et aux décideurs se doivent d'être proches des pratiques de terrain et représentatives de la population afin de guider la prise de décisions de façon pertinente et objective en France, au niveau populationnel. Afin d'accompagner les pouvoirs publics dans ce processus de décision, l'usage de données de vraie vie est essentiel.

IX.2.2. Estimation des coûts de prise en charge en population générale : un enjeu majeur pour le système de santé

Les évaluations économiques des stratégies de prise en charge en cancérologie reposent dans bien des cas sur l'utilisation de populations sélectionnées. Elles sont réalisées dans le cadre plus général d'essais cliniques, randomisés ou non. Les ECR constituent la référence dans le domaine de l'évaluation en médecine et leur usage est recommandé par la HAS (272). Mis en œuvre sur la base d'un schéma expérimental rigoureux, les ECR ont vocation à maximiser la validité interne des résultats. Les patients inclus sont recrutés selon des critères stricts, excluant les individus aux caractéristiques particulières. Ils sont randomisés et font l'objet d'une surveillance accrue. Les ECR évaluent ainsi l'impact d'un produit de santé ou d'une stratégie de prise en charge dans les conditions optimales de comparaison. Ils fournissent des résultats essentiels à son introduction ou son déploiement sur le marché. Néanmoins, ces essais se focalisent majoritairement sur des séquences thérapeutiques ou des prises en charge particulières réalisées dans un nombre restreint d'établissements spécialisés et sur un périmètre de coûts limité lorsqu'ils comportent un volet économique. De plus, ils sont réalisés sur des courtes périodes qui ne permettent pas toujours d'appréhender la totalité des effets et des coûts des options testées (129). Les principes méthodologiques des ECR entravent ainsi les possibilités de généralisation des résultats à la population générale et leur usage dans le cadre de la prise de décisions publiques (128). Ils ne permettent pas de rendre compte de l'impact réel des recommandations ou des innovations sur les pratiques.

L'approche en population, fondée sur l'utilisation de données d'observation, est utile car elle permet de produire des estimations non biaisées correspondant à des résultats fondés sur les pratiques réelles (8). Les décideurs souhaitent disposer d'informations portant sur l'utilisation des produits et services de santé lors de leur mise en œuvre en « vie réelle » (7) (120). Ces informations permettent notamment de confirmer ou d'infirmer ceux préalablement obtenus dans le cadre des ECR. Les estimations obtenues peuvent alors être utilisées pour réévaluer les choix associés aux

remboursements du produit en vue d'une éventuelle remise à niveau (24). Dans ce cadre, plusieurs sources de données d'observation peuvent être mobilisées. Les données issues d'études épidémiologiques partagent un inconvénient majeur avec les ECR qui est la difficulté à les mettre en place en raison de leurs délais de mise en œuvre et de leurs coûts (128). En comparaison, la mutualisation des coûts de création et de gestion des registres ou des bases de données médico-administratives du SNDS constitue un réel avantage. Les registres, en recensant de façon exhaustive l'ensemble des cas atteints d'une pathologie dans une zone géographique déterminée, représentent une base exhaustive de sondage en population générale (133). En France, ils ont été créés dans le but de disposer de données d'incidence fiables (134). Plus récemment dans le cadre des registres de cancer, un programme de recueil ponctuel et spécifique de certaines localisations s'est développé au sein du réseau FRANCIM (Réseau Français des Registres de Cancer) avec un objectif d'évaluation des pratiques de soins. Il s'appuie sur la mise en place de cohortes de patients présentant une même localisation, diagnostiqués une année donnée dans une zone géographique définie et suivis pendant une période de cinq ans. Ces enquêtes, dites de « hautes résolutions », permettent d'obtenir des informations en population générale sur l'efficacité médicale, mesurée en termes de survie, des différents traitements réalisés. Toutefois, elles ne permettent que très difficilement d'évaluer le coût associé à une pathologie ou une prise en charge.

Dans ce contexte, l'exploitation des bases de données médico-administratives du SNDS peut être très utile. Ces bases renseignent de manière continue et détaillée la totalité de la consommation des soins remboursés de 98.8 % de la population française. Fondées sur l'observation des pratiques réelles, elles permettent de produire des résultats dotés d'une très bonne validité externe et apportent des informations pertinentes et objectives aux autorités de santé. Une attention particulière portant sur les étapes de sélection des cas étudiés est néanmoins essentielle afin de conserver la validité des conclusions. De par la taille de la population incluse, la nature et le volume des données présentes, son alimentation continue et l'ouverture de ses accès, le SNDS est voué à être au cœur de la recherche médicale en France. Il constitue un entrepôt de données

particulièrement adapté à l'évaluation des parcours de soins et de leurs conséquences économiques, notamment selon les profils de la population et les catégories d'acteurs (273). Une connaissance approfondie des possibilités offertes, des avantages, des limites et des niveaux de complexité méthodologiques associés à l'exploitation de ces données est donc indispensable (274). Le SNDS appartient à l'ère du « Big data » en matière de santé. L'exploitation et l'analyse de ces bases de données fondées sur des méthodologies adaptées peut fournir des résultats avec des niveaux de preuves équivalents et complémentaires à ceux des ECR (7). A titre d'exemple, dans le cadre d'une étude comparative, la mise en œuvre de méthodes d'appariements telles que les scores de propension, la régression sur discontinuité ou à variables instrumentales ou des méthodes d'ajustement à partir des modèles prédictifs peuvent ainsi être utilisées afin de corriger les biais de sélection présents lorsqu'il n'y a pas de randomisation (275).

IX.2.3. Modélisation des coûts et prise en compte de l'incertitude

Les études menées en population générale exploitent des données représentatives des pratiques courantes qui nécessitent l'usage de méthodologies adaptées. Les individus inclus dans ces études peuvent présenter des caractéristiques particulières qui auront un impact sur la mesure des coûts et entachant la validité des résultats (8). Sans randomisation, les biais de sélection ou de confusion peuvent fausser les résultats de l'étude. Dans ce contexte, la modélisation statistique des données est essentielle car elle peut permettre de contrôler l'hétérogénéité des populations ou des pratiques (276).

La modélisation est essentielle pour l'acquisition des connaissances en jouant un rôle moteur lors de la conception de stratégies de prise en charge et de l'interprétation des phénomènes. Dans le cadre de l'évaluation économique, elle peut aider à clarifier les conséquences de processus décisionnels et à simuler l'effet de programmes alternatifs. Sans se substituer aux données d'observation,

l'utilisation de modèles statistiques peut compléter ces dernières lorsqu'elles sont trop partielles pour pouvoir répondre aux préoccupations des acteurs du système de santé. C'est notamment le cas lorsqu'on ne dispose que de données à court terme pour évaluer les conséquences d'un processus sur une plus longue période, lorsqu'on n'a accès qu'à des données d'efficacité sur un critère intermédiaire pour évaluer un effet final, ou encore, lorsque l'observation ne porte que sur des populations particulières et non sur l'ensemble de la population éligible au traitement (163).

Les modèles de prédiction considèrent dans les inférences les caractéristiques de la population à disposition, prédisent les valeurs de la variable d'intérêt conditionnellement à ces caractéristiques et estiment leurs effets. Ils ajustent ainsi les estimations réalisées selon les paramètres de l'étude afin de contrôler les différents biais rencontrés. Cette étape est d'autant plus importante que les résultats obtenus peuvent alimenter dans un second temps les modèles de décision utilisés et recommandés par la HAS dans le cadre des évaluations d'efficience (3). La qualité des résultats d'un modèle de décision dépend de celle des données qui l'alimentent. Dans ce cadre, les modèles de prédiction permettent d'obtenir des informations non biaisées et agrégées conditionnellement aux caractéristiques individuelles utilisées. Ces résultats peuvent ainsi permettre d'extrapoler les résultats sur un horizon temporel plus important ou à des populations plus larges.

Selon le rapport de l'ISPOR portant sur les bonnes pratiques dans le cadre de la modélisation décisionnelle des stratégies de soins, « les conclusions d'une étude sont conditionnelles aux hypothèses et aux données sur lesquelles le modèle s'est construit » (161). Les données économiques présentent des particularités en raison de leur distribution et de leur nature. D'une part, les coûts sont positifs, leur distribution étirée vers la droite et d'autre part, ils peuvent s'analyser au cours du temps notamment lorsque la pathologie étudiée est chronique, ou selon diverses zones géographiques ou entités. Les hypothèses fondamentales des modèles de prédiction classiques ne sont plus respectées et des méthodologies appropriées doivent être mises en œuvre afin de proposer des résultats valides.

Les choix effectués par les décideurs fondés sur la base de résultats produits dans le cadre d'évaluations économiques peuvent être déterminants pour la population. En effet, ils peuvent affecter directement la répartition des ressources et par voie de conséquences la santé de la population. Les décideurs ont ainsi un rôle de « porteurs de risques » qui requiert des précautions quant à l'interprétation des résultats obtenus (84). Dans cette perspective, la variabilité des résultats d'une évaluation économique devrait systématiquement être testée dans le cadre d'analyses de sensibilité. Ces analyses complémentaires ont pour but d'explicitier les conséquences des variations de paramètres incertains sur l'estimation des coûts et/ou des effets de santé. Elles jouent ainsi un rôle majeur dans le cadre des évaluations économiques. Elles peuvent conforter les décideurs dans leurs choix et permettent une meilleure négociation des prix relatifs aux traitements efficaces (83) (24).

Les usages abusifs des outils statistiques dans les recherches biomédicales sont fréquents. Il convient de maîtriser ces outils afin de continuer à approfondir nos connaissances et d'améliorer les résultats dans le domaine de la santé (277).

X. Conclusion et perspectives

Nous avons développé une approche analytique permettant de déterminer les coûts associés à deux groupes d'hémopathies malignes en France. A partir de données extraites du SNDS permettant d'appréhender l'intégralité des prises en charge hospitalières et extrahospitalières en population, nous avons évalué et analysé les coûts des stratégies de prise en charge de ces cancers.

Concernant la prise en charge des lymphomes, les résultats les plus marquants portent sur l'importance des coûts des séjours hospitaliers et des molécules onéreuses (i.e. rituximab), ainsi que sur la forte toxicité des traitements et la prise en charge tardives des effets indésirables. La place importante occupée par le coût en lien avec les pertes de productivité, notamment pour les LH, confirme le besoin d'accompagner les patients dans leur vie professionnelle pendant et après la phase de traitement. Notre travail a également porté sur l'étude des coûts de la prise en charge des patients atteints de LMC et sur l'impact économique des traitements par ITK. Les résultats de cette première étude française portant sur les coûts de la LMC en Midi-Pyrénées soulignent le caractère onéreux des traitements médicamenteux, mais aussi la nécessité de réaliser des travaux complémentaires en population générale, sur une période plus importante et adoptant un point de vue plus global.

Ce travail constitue une base de réflexion pour le futur et offre plusieurs perspectives. Il donne la possibilité de quantifier l'impact économique de nouveaux médicaments dans le traitement de ces hémopathies malignes. Il permet également d'appréhender le coût de l'introduction de nouvelles modalités de prise en charge de ces pathologies. Il propose des modèles adaptés aux caractéristiques des patients et directement exploitables, qui permettront dans l'avenir un suivi économique de ces pathologies. Le travail engagé dans cette thèse nous incite à poursuivre et développer plus largement notre réflexion dans le champ de l'évaluation médicale et économique des stratégies de prise en charge en population.

Notre projet suit différents objectifs. Dans un premier temps, il va consister à compléter les résultats préalablement obtenus et de mettre à disposition des éléments économiques pertinents sur les parcours de soins, et de vie, associés aux hémopathies malignes étudiées. Dans ce cadre, une extraction nationale des cas de LMC effectuée à partir du SNDS comprenant l'ensemble des consommations de soins accessible sur une période de 5 ans est d'ores et déjà disponible. Son exploitation permettra de décrire de manière exhaustive et représentative les coûts associés aux parcours de soins couvrant l'ensemble des patients atteints de LMC en France. Des travaux complémentaires portant sur la période d'après cancer des patients atteints de lymphomes vont être mis en œuvre. Les résultats produits pourront être utiles à l'amélioration de la coordination des soins et du suivi des patients à chaque étape de leur maladie, tant dans une perspective professionnelle que personnelle.

Dans un second temps, des travaux seront menés dans le champ des génériques et des bio-similaires, source d'économies potentielles. Des modèles centrés sur les patients et alimentés majoritairement à partir des informations préalablement estimées dans le cadre des travaux menés durant ce travail de thèse ainsi qu'à partir des données de la littérature seront mis en œuvre. Ils auront pour objectif d'estimer les sources d'économies potentielles en lien avec l'introduction sur le marché futur des ITK génériques (i.e. LMC) et des bio-similaires du rituximab (i.e. B-LNH) dans le traitement de ces hémopathies. Ces travaux fourniront des résultats utiles aux décideurs dans leurs prises de décisions.

Bibliographie

- 1- Caillol, M., Le Coz, P., Aubry, R. & Bréchat, P. (2010). Réformes du système de santé, contraintes économiques et valeurs éthiques, déontologiques et juridiques. Santé Publique, vol. 22,(6), 625-636. doi:10.3917/spub.106.0625.
- 2- Benamouzig D. (2013)., « L'économie de la santé sous surveillance médicale (1960-1990) », Histoire, médecine et santé, 4 | 2013, 41-67.
- 3- Haute autorité de santé. Choix méthodologiques pour l'évaluation économique. Accessible [Octobre 2018] : https://www.has-sante.fr/portail/plugins/ModuleXitiKLEE/types/FileDocument/doXiti.jsp?id=c_1120708
- 4- Sanchez, CP. & Saout, C. (2017). Prix et accès aux traitements médicamenteux innovants. Avis du Conseil économique, social et environnemental. Accessible [Octobre 2018] : http://www.lecese.fr/sites/default/files/pdf/Avis/2017/2017_04_acces_medicaments_innovants.pdf.
- 5- Institut national du Cancer. Plan cancer 2014-2019. Accessible [Octobre 2018]: <http://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Plan-Cancer-2014-2019>
- 6- Bourret, P. & Le Moigne, P. (2014). Essais cliniques, production de la preuve et mutations de la biomédecine. Sciences sociales et santé, vol. 32,(3), 5-11. doi:10.3917/sss.323.0005.
- 7- Maisonneuve, H. Babany, G. Données de vie réelle et recherche clinique: le complément des données des essais randomisés? PresseMed.(2015), <http://dx.doi.org/10.1016/j.lpm.2015.02.005>
- 8- Begaud, B., Polton D., von Lennep F., Les données de vie réelle, un enjeu majeur pour la qualité des soins et la régulation du système de santé – l' exemple du médicament. 2017. Rapport remis à la ministre de la santé. Accessible [Octobre 2018]: http://solidarites-sante.gouv.fr/IMG/pdf/rapport_donnees_de_vie_reelle_medicaments_mai_2017vf.pdf
- 9- LOI n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé. Accessible [Octobre 2018]: <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000031912641&categorieLien=id>
- 10- Assurance maladie. Finalité du SNIIRAM. Accessible [Octobre 2018]: <https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/sniiram/finalites-du-sniiram.php>

- 11- Moulis, G., Lapeyre-Mestre, M., Palmaro, A., Pugnet, G., Montastruc, J-L., Sailler L. French health insurance databases: What interest for medical research? *Rev Med Interne Rev Med Interne*. 2015 Jun;36(6):411-7
- 12- Maunoury, F. (2016). Aide à la décision des modèles médico-économiques. *Journal de gestion et d'économie médicales*, vol. 34,(2), 177-186. doi:10.3917/jgem.162.0177.
- 13- Collège des économistes de la santé. Guide méthodologique pour l'évaluation économique des stratégies de santé. Paris :s.n., 2003 Accessible [Octobre 2018]: www.ces-asso.org/docs/Rapport_final-CES.pdf
- 14- Baleyrier F., Domenech C., Thomas X. Nouvelles thérapeutiques conventionnelles en oncologie. (2011). *Bulletin du Cancer* Vol 98 - N° 8, P. 901-913
- 15- Drummond M., O'Brien B., Stoddart Greg et al., 1998, Méthodes d'évaluation économique des programmes de santé . Paris : Economica
- 16- Jefferson T, Demicheli V, Mugford M. Cost-of-illness studies, elementary economic evaluation in health care. 2nd Ed. London: BMJ Publishing Group; 2000. pp. 17–29.
- 17- Byford S, Torgerson DJ, Raftery J. Economic note: cost of illness studies. *BMJ* 2000;320:1335.
- 18- Jo, C. (2014). Cost-of-illness studies: concepts, scopes, and methods. *Clinical and Molecular Hepatology*, 20(4), 327–337. <http://doi.org/10.3350/cmh.2014.20.4.327>
- 19- Hassan Serrier. (2011). Théories et méthodes d'évaluation du coût social de facteurs de risque professionnels en France : application au cas des cancers d'origine professionnelle. *Economies et finances*. Université de Bourgogne.
- 20- Riveline, C. Evaluation des coûts. Éléments d'une théorie de la gestion. Ecole Nationale des Mines de Paris. 1980
- 21- Shiell, A., Gerard, K., Donaldson, C. Cost of illness studies: An aid to decision-making? *Health Policy*. Volume 8, Issue 3, December 1987, Pages 317-323
- 22- Auweiler, PW. Müller, D. Stock, S. Gerber, A. Cost effectiveness of rituximab for non-Hodgkin's lymphoma: a systematic review. *Pharmacoeconomics*. 2012 Jul 1;30(7):537-49.
- 23- Hornberger, J., Chien, R., Friedmann M. Cost-effectiveness of rituximab as maintenance therapy in patients with follicular non-Hodgkin lymphoma after responding to first-line rituximab plus chemotherapy. *Leuk Lymphoma*. 2012 Dec;53(12):2371-7.
- 24- Briggs, AH. Handling uncertainty in economic evaluations of healthcare interventions *BMJ* 1999 Sep 4
- 25- M. Jeantet et A. Lopez, avec le concours de N.Destais. Evaluation médico-économique en santé. Accessible [Octobre 2018]: <http://www.igas.gouv.fr/spip.php?article436>

- 26- De Jager, P. Difficulties associated with calculating the variance of the ICER statistic. 2017
<https://www.researchgate.net/publication/315619378>
- 27- Diamond, GA, Kaul, S. Cost, effectiveness, and cost-effectiveness. *Circ Cardiovasc Qual Outcomes*. 2009;2
- 28- Chamot, E. et al. Analyse économique appliquée à la santé: un tour d'horizon. In: *Bulletin des médecins suisses*, 2001, vol. 82, n° 39, p. 2060-2064. <https://archive-ouverte.unige.ch/unige:28070>
- 29- Robinson, R. (1993). Cost-utility analysis. *BMJ : British Medical Journal*, 307(6908), 859–862.
- 30- Pettitt, D. A., Raza, S., Naughton, B., Roscoe, A., Ramakrishnan, A., Ali, A., ... Brindley, D. A. (2016). The limitations of QALY: a literature review. *Journal of Stem Cell Research and Therapy*, 6(4), Article: 1000334.
- 31- McGregor, M. (2003). Cost–utility analysis: Use QALYs only with great caution. *CMAJ: Canadian Medical Association Journal*, 168(4), 433–434.
- 32- Meunier, V. (2009). Analyse coût-bénéfice: guide méthodologique. Number 2009-06 of the *Cahiers de la Sécurité Industrielle*, Institute for an Industrial Safety Culture, Toulouse, France (ISSN 2100-3874). Available at http://www.icsi-eu.org/francais/dev_cs/cahiers/.
- 33- Eisenberg, JM. Clinical economics: a guide to the economic analysis of clinical practices. *JAMA* 1989;262:2879-86
- 34- Abraham, C., Thedie, J. Le prix d'une vie humaine dans les décisions économique. *Revue française de recherche opérationnelle*, 1960, n°16.
- 35- Haute autorité de santé. Choix méthodologiques pour l'analyse de l'impact budgétaire à la HAS. 2016. Accessible [Octobre 2018]: https://www.has-sante.fr/portail/upload/docs/application/pdf/2016-12/guide_methodologique_choix_methodologiques_pour_lanalyse_de_limpact_budgetaire_a_la_has.pdf
- 36- Mauskopf, J. Prevalence - based economic evaluation. *Value in Health* 1998;1(4):251 – 9
- 37- Trueman, P., Drummond, M., Hutton, J. Developing guidance for budget impact analysis. *Pharmacoeconomics* 2001;19(6):609 - 21
- 38- Collège des Economistes de la Santé. Guide méthodologique pour la mise en place d'une analyse d'impact budgétaire. Paris : s.n. 2008
- 39- Ghabri, S., Poullié, A., Autin, E. & Josselin, J. (2017). Le guide d'analyse d'impact budgétaire de la HAS : un nouvel outil d'aide à la décision. *Santé Publique*, vol. 29,(4), 585-588. doi:10.3917/spub.174.0585.

- 40- Levy, P. L'analyse d'impact budgétaire et l'évaluation économique: des approches substituables ou complémentaires. Lettres du Collège des Economistes de la Santé. 2008
- 41- International Society for Pharmacoeconomics and Outcomes Research, Sullivan, SD., Mauskopf, JA., Augustovski, F., Jaime Caro, J., Lee, KM., et al. Budget impact analysis- principles of good practice : report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. Value Health. 2014 ;17(1) :5-14.
- 42- Grignon, M., Midy F. La notion de coût en économie de la santé. Journal d'économie médicale 2002, Vol. 20, 2, pp. 105-117.
- 43- Williams, A., 1970. Cost-benefit analysis. The managed economy. Cairncross A. (éd.). Oxford : Blackwell.
- 44- Posnett, J., Jan S., 1996, Indirect cost in economic evaluation : the opportunity cost of unpaid inputs, Health Economics , vol 5, pp. 13-23
- 45- Launois, R., 1999, Un coût, des coûts, quels coûts ?, Journal d'Economie Médicale , vol 17, n°1, pp. 77-82.
- 46- Green, DI. Pain-Cost and Opportunity-Cost. The Quarterly Journal of Economics, Volume 8, Issue 2, 1 January 1894, Pages 218–229, <https://doi.org/10.2307/1883711>
- 47- Davenport, Value and Distribution, Chicago, 1908, chap. 7.
- 48- Lassegue, P. Gestion de l'entreprise et comptabilité, 2^e ed, Paris : Dalloz 1965.
- 49- Andréani, E. Le coût d'opportunité. In: Revue économique, volume 18, n°5, 1967. pp. 840-858;
- 50- Marshall, A. Principes d'économie politique, t. II, Paris, Giard et Brière, 1909, p. 47.
- 51- collectif, 1997b, Recommandations de bonnes pratiques des méthodes d'évaluation économique des stratégies thérapeutiques. La lettre du collège, vol n° spécial, avril 1997.
- 52- Gold, M., R., Siegel J. E., Russell, L. B. et al., 1996, Cost effectiveness in health and medicine. Oxford: Oxford University Press.
- 53- Costa, N., Ferlicq, L., Derumeaux-Burel, H., Rapp, T., Garnault, V., Gillette-Guyonnet, S., Andrieu, S., Vellas, B., Lamure, M., Grand, A., Molinier, L. (2012). Comparison of informal care time and costs in different age-related dementias: a review. BioMed research international, 2013, 852368.
- 54- Berad, A., Gervès, C., Fontaine, R., Aquino, J.P., Plisson, M. Combien coûte la maladie d'Alzheimer ? Rapport, fondation Médéric Alzheimer. Accessible [Octobre 2018]: <http://www.fondation-mederic-alzheimer.org/Nos-Travaux/Nos-etudes>
- 55- Costa N. (2013). Analyse médico-économique de différentes stratégies de prévention dans les maladies neurodégénératives : Application à la maladie d'Alzheimer et à la maladie de Parkinson. (Thèse de doctorat inédite). Université de Claude Bernard Lyon 1.

- 56- Cunnama, L., Sinanovic, E., Ramma, L. & al. Using Top-down and Bottom-up Costing Approaches in LMICs: The Case for Using Both to Assess the Incremental Costs of New Technologies at Scale. *Health Econ.* 2016 Feb;25 Suppl 1:53-66
- 57- Assurance Maladie. Principe généraux de la classification commune des actes médicaux. Accessible [Octobre 2018]: https://www.ameli.fr/fileadmin/user_upload/documents/CCAM-Principes_generaux_mars2011.pdf.
- 58- Site Assurance Maladie. Liste des médicaments remboursables. [Octobre 2018]: https://www.ameli.fr/medecin/exercice-liberal/remuneration/nomenclatures-codage/medicaments#text_11322
- 59- Site Assurance Maladie. Liste des produits et des prestations. [Octobre 2018]: <https://www.ameli.fr/medecin/exercice-liberal/remuneration/nomenclatures-codage/liste-produits-prestations-lpp>
- 60- Haute autorité de santé. Tarifs des Dispositifs médicaux. Accessible [Octobre 2018]: https://www.has-sante.fr/portail/jcms/c_412129/fr/les-procedures-d-inscription-et-de-tarifification-des-dispositifs-medicaux
- 61- Haute autorité de santé. Les coûts de la qualité et de la non-qualité des soins dans les établissements de santé : état des lieux et propositions. 2004. Accessible [Octobre 2018] : https://www.has-sante.fr/portail/jcms/r_1499622/fr/couts-de-la-qualite-et-de-la-non-qualite-des-soins-dans-les-etablissements-de-sante-etat-des-lieux-et-propositions
- 62- Burlaud, A. & Simon, C. (2013). Lexique. Dans : Alain Burlaud éd., *Le contrôle de gestion* (pp. 112-114). Paris: La Découverte.
- 63- Mercier, G., Costa, N., Dutotc, C. et Riched V-P. Sources de données, données utilisées et modalité de recueil. Calcul et analyse des coûts hospitaliers : guide méthodologique. *Revue d'Épidémiologie et de Santé Publique*. Volume 66, Supplement 2, March 2018, Pages S73-S91
- 64- Guerre, P., Hayes, N., Bertaux, A-C. Estimation du coût hospitalier : approches par « micro-costing » et « gross-costing ». *Revue d'Épidémiologie et de Santé Publique*. Volume 66, Supplement 2, March 2018, Pages S65-S72
- 65- Robert, B., Fetter, J.D., Thompson, R.E. Mills. A System for Cost and Reimbursement Control in Hospitals. *Yale J Biol Med*. 1976 May
- 66- Agence technique de l'information sur l'hospitalisation. Présentation MCO. Accessible [Octobre 2018]: <https://www.atih.sante.fr/mco/presentation>
- 67- Assurance Maladie. Prise en charge transport. Accessible [Octobre 2018]: <https://www.ameli.fr/assure/remboursements/rembourse/transport/transport>

- 68- Service public. Tarifs conventionnel des ambulances et véhicules sanitaires légers Accessible [Octobre 2018]: <https://www.service-public.fr/professionnels-entreprises/vosdroits/F22121>
- 69- Assurance Maladie. Référentiel de prescription des transports. Accessible [Octobre 2018]: <https://www.ameli.fr/content/referentiel-de-prescription-des-transports>
- 70- Sultan-Taïeb, H., Tessier, P. & Béjean, S. (2009). Capital humain et coûts de friction: Quels critères de choix pour l'évaluation des pertes de production ?. *Revue économique*, vol. 60,(2), 293-306. doi:10.3917/reco.602.0293.
- 71- Becker, S.G. *A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press. Published March 1994.
- 72- Assurance Maladie. Remboursement indemnité journalière. Accessible [Octobre 2018]: <https://www.ameli.fr/assure/remboursements/indemnites-journalieres/arret-maladie>.
- 73- Assurance Maladie. Remboursement pension d'invalidité. Site de l'AM. Accessible [Octobre 2018]: <https://www.ameli.fr/employeur/demarches/invalidite/pension-invalidite>.
- 74- O'shea, E. *The Costs of Caring for People with Dementia and Related Cognitive Impairments*. National council on ageing and older people, 2000. Report 60.
- 75- Van den Berg, B., Brouwer, W.B., Koopmanschap, M.A. Economic valuation of informal care. An overview of methods and applications. *Eur J Health Econ*. 2004 Feb;
- 76- Koopmanschap, MA., van Exel, J.N., Van den Berg, B., Brouwer, W.B. An overview of methods and applications to value informal care in economic evaluations of healthcare. *Pharmacoeconomics*. 2008;26(4):269-80.
- 77- Launois, R. Un coût, des coûts, quels coûts ? *Journal d'économie médicale*. 1999. Vol. 17, 1, pp. 77-82
- 78- Torgerson, D.J., & Raftery, J. (1999). Discounting. *BMJ : British Medical Journal*, 319(7214), 914–915.
- 79- Gafni A, Torrance GW. Risk attitude and time preference in health. *Management Science*. 1984;30(4):440-451.
- 80- Lebègue, D. Commissariat général du plan. Révision du taux d'actualisation des investissements publics. 2005. Accessible [Octobre 2018]: <http://www.plan.gouv.fr/intranet/upload/actualite/Rapport%20Lebegue%20Taux%20actualisation%2024-01-05.pdf>.
- 81- Brouwer, W.B.F., Niessen, L.W., Postma, M.J., Rutten, F.F.H. Need for differential discounting of costs and health effects in cost effectiveness analyses. *BMJ*. 2005;331: 446–448.

- 82- Bilcke, J., Beutels, P., Brisson, M., Jit, M. Accounting for methodological, structural, and parameter uncertainty in decision-analytic models: a practical guide. *Med Decis Making*. 2011 Jul-Aug;31(4):675-92.
- 83- Briggs, A.H., O'Brien, B.J., Blackhouse, G. Thinking Outside the Box : Recent Advances in the Analysis and Presentation of Uncertainty in Cost-Effectiveness Studies. *Annu Rev Public Health* 2002. 23:377-401.
- 84- Arrow, K.J., Lind, R.C. Uncertainty and the Evaluation of Public Investment Decisions. *The American Economic Review*. Vol. 60, No. 3 (Jun., 1970), pp. 364-378
- 85- Zaugg, V., Savoldelli, V., Sabatier, B. & Durieux, P. (2014). Améliorer les pratiques et l'organisation des soins : méthodologie des revues systématiques. *Santé Publique*, vol. 26,(5), 655-667. doi:10.3917/spub.145.0655.
- 86- The Health Science. Graphique de Tornado, illustration. Accessible [Octobre 2018]: <https://www.thehealthscience.com/topics/tornado-diagram>
- 87- Doubilet, P., Begg, C.B., Weinstein, M.C., Braun, P., McNeil, B. J. Probabilistic Sensitivity Analysis Using Monte Carlo Simulation: A Practical Approach. *Medical Decision Making*, 1985, Vol 5, Issue 2, pp. 157 – 177.
- 88- Baltussen, R.M., Hutubessy, R.C., Evans, D.B., Murray, C.J. Uncertainty in cost-effectiveness analysis. Probabilistic uncertainty analysis and stochastic league tables. *Int J Technol Assess Health Care*. 2002 Winter;18(1):112-9.
- 89- Adalsteinsson, E., Toumi, M. Benefits of probabilistic sensitivity analysis – a review of NICE decisions. *Journal of Market Access & Health Policy*. 2013;1:10.3402/jmahp.v1i0.21240. doi:10.3402/jmahp.v1i0.21240.
- 90- Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: Not an optional extra. *Health Econ*. 2005;14(4):339–47.
- 91- Briggs A., Gray A., 1998, The distribution of health care costs and their statistical analysis for economic evaluation, *Journal of Health Services Research and Policy*, vol 3, n° 4, pp. 233-245.
- 92- Théorème Central Limite (TCL) Lucien Le Cam, « The central limit theorem around 1935 », *Statistical Science*, vol. 1, no 1, 1986, p. 78–91 [Accessible Juin 2018: <https://projecteuclid.org/euclid.ss/1177013818>]
- 93- Van Hout, B. A., Al, M. J., Gordon, G. S. and Rutten, F. F. H. (1994), Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ.*, 3: 309–319. doi:10.1002/hec.4730030505
- 94- Fieller, EC. (1954). "Some problems in interval estimation.". *Journal of the Royal Statistical Society, Series B*. 16 (2): 175–185. JSTOR 2984043.

- 95- Willan, A. R. and O'Brien, B. J. Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. *Health Econ.* 1996; 5: 297-305
- 96- Chaudhary, M. A. and Stearns, S. C. Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. *Statistics and Medicine* 1996; 15: 1447-1458
- 97- Efron, B. Bootstrap Methods: Another Look at the Jackknife. *Ann. Statist.* Volume 7, Number 1 (1979), 1-26.
- 98- Mullahy, J., Manning, W. *Statistical issues of cost-effectiveness analysis.* [Auteur du livre] Sloan F. *Valuing health care.* Cambridge: Cambridge University Press, 1994, pp, 149-184.
- 99- O'Brien, B.J., Drummond, M.F., Labelle R.J., and Willan, A. "In Search of Power and Significance: Issues in the Design and Analysis of Stochastic Cost-Effectiveness Studies in Health Care." *Medical Care* 32, no. 2 (1994): 150-63. <http://www.jstor.org/stable/3766311>.
- 100- Carpenter, J., Bithell, J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Stat Med.* 2000 May 15;19(9):1141-64.
- 101- Hall, P. *The Bootstrap and Edgeworth Expansion.* Springer Series in Statistics 1992. Disponible (2017) : <https://people.eecs.berkeley.edu/~jordan/sail/readings/edgeworth.pdf>
- 102- Canty, A., Ripley, B. Package boot R. July 2017, disponible : <https://cran.r-project.org/web/packages/boot/boot.pdf>
- 103- Polsky, D., Glick, H.A., Willke, R., Schulman, K. Confidence intervals for cost-effectiveness ratios: a comparison of four methods. *Health Econ.* 1997 May-Jun;6(3):243-52.
- 104- Siani, C., Peretti, C., Moatti, J.P. Revisiting methods for calculating confidence region for ICERs ? Are Fieller's and bootstrap methods really equivalent ? 2003, Document de travail-Grequam. Disponible (2017) <http://www.ces-asso.org/sites/default/files/C17-Siani-article.pdf>
- 105- Mircioiu, C., Atkinson, J. A Comparison of Parametric and Non-Parametric Methods Applied to a Likert Scale. Shaw N, ed. *Pharmacy: Journal of Pharmacy, Education and Practice.* 2017;5(2):26. doi:10.3390/pharmacy5020026.
- 106- Costa, N., Mounié, M, Bernard, G. et al.. Cost-effectiveness Analysis of TESPAL vs Arterial Embolization for Intractable Epistaxis : a long-term analysis. Under review: clinical otolaryngology.
- 107- Tarricone, R. Cost-of-illness analysis. What room in health economics? *Health Policy.* 2006 Jun;77(1):51-63

- 108- Caillol, H. (2015). Ouverture des données de santé : l'expérience de l'Assurance maladie. *Informations sociales*, 191,(5), 60-67. Accessible [Octobre 2018] : <https://www.cairn.info/revue-informations-sociales-2015-5-page-60.htm>.
- 109- Cour des comptes. (2016). Les données personnelles de santé gérées par l'assurance maladie. Accessible [Octobre 2018]: <https://www.ccomptes.fr/fr/publications/les-donnees-personnelles-de-sante-gerees-par-lassurance-maladie>
- 110- Ancelle T. *Statistique. Épidémiologie*. Paris : Maloine; 2002. ISBN 2 224 027060
- 111- Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. *Lancet*. 2002 Jan 5;359(9300):57-61.
- 112- Arnaud C. *Enseignement du Deuxième Cycle des Etudes Médicales. La médecine fondée sur les preuves. LES ENQUETES EPIDEMIOLOGIQUES*. Disponible : http://www.medecine.ups-tlse.fr/DCEM2/module1/sous_module1/003_methodologie_generale_CA.pdf
- 113- Nissen, T., Wynn, R. (2014). The clinical case report: a review of its merits and limitations. *BMC Research Notes*, 7, 264. <http://doi.org/10.1186/1756-0500-7-264>
- 114- Chakra CAN, Pariente A, Pinet M, Nkeng L, Moore N, Moride Y. Case series in drug safety. A review to determine characteristics and quality. *Drug Saf*. 2010;33:1081–1088. doi: 10.2165/11539300-000000000-00000
- 115- Setia, M. S. (2016). *Methodology Series Module 3: Cross-sectional Studies*. *Indian Journal of Dermatology*, 61(3), 261–264. <http://doi.org/10.4103/0019-5154.182410>
- 116- Goldberg, M., Zins, M. Principe et intérêt des cohortes épidémiologiques. *Actualité et dossier en santé publique n° 78*. Haut conseil de la santé publique. Accessible [Octobre 2018] : <https://www.hcsp.fr/explore.cgi/Adsp?clef=118>
- 117- Merito, M., Bonaccorsi, A., Pammolli, F. et al. Economic evaluation of HIV treatments: the I.CO.N.A. cohort study. *Health Policy*. 2005 Nov;74(3):304-13.
- 118- Setia, M. S. (2016). *Methodology Series Module 2: Case-control Studies*. *Indian Journal of Dermatology*, 61(2), 146–151. <http://doi.org/10.4103/0019-5154.177773>
- 119- Mounié, M., Costa, N., Sailler, L. et al. Incremental Costs in Giant Cell Arteritis. *Arthritis Care Res*. 2017 Oct 3;
- 120- Drummond, M.F. *Experimental versus Observational Data in the Economic Evaluation of Pharmaceuticals*. *Medical Decision Making*. Vol 18, Issue 2_suppl, pp. S12 - S18. First Published April 1, 1998
- 121- Mahajan, A. (2015). Limitations of CBA study: Controlled before after study. *Lung India*. 2015 Nov-Dec; 32(6): 670–671.

- 122- Haute autorité de Santé. 2007. Méthodes quantitative pour évaluer les interventions visant à améliorer les pratiques. Accessible [Octobre 2018] : https://www.has-sante.fr/portail/upload/docs/application/pdf/eval_interventions_ameliorer_pratiques_guide.pdf
- 123- Mahajan, A. (2015). Limitations of CBA study: Controlled before after study. *Lung India*. 2015 Nov-Dec; 32(6): 670–671.
- 124- Site de l'Agence National de Sécurité du Médicaments. Définition essais cliniques. Accessible [Octobre 2018]: [https://www.ansm.sante.fr/Activites/Essais-cliniques/Repertoires-des-essais-cliniques-de-medicaments/\(offset\)/1](https://www.ansm.sante.fr/Activites/Essais-cliniques/Repertoires-des-essais-cliniques-de-medicaments/(offset)/1)
- 125- Marc Righini, Grégoire Le Gal. Principes de la médecine fondée sur les preuves. *Médecine thérapeutique*. 2005;11(6):424-427.
- 126- Ayerbe, C., Missonier, A. (2007). Validité interne et validité externe de l'étude de cas: principes et mise en oeuvre pour un renforcement mutuel, *Revue Finance Contrôle Stratégie*, 10, (2), 37-62
- 127- Drummond M.F. Economic analysis alongside controlled trials. *Departments of Health* 1994
- 128- Jain, S., Arora, N.K. Randomized controlled trial and economic evaluation. *Indian J Pediatr*. 2000 May;67(5):363-8
- 129- Sanson-Fisher, R.W., Bonevski, B., Green, L.W., D'Este, C. Limitations of the randomized controlled trial in evaluating population-based health interventions. *Am J Prev Med*. 2007 Aug;33(2):155-61.
- 130- De Ketele, JM., Roegiers, X. *Methodologie du recueil d'informations. Méthodes en sciences humaines*. Sieme édition.
- 131- Bellary, S., Krishnankutty, B., & Latha, M. S. (2014). Basics of case report form designing in clinical research. *Perspectives in clinical research*, 5(4), 159-66.
- 132- Institut National du Cancer. Source. Les registres du cancer. Accessible [Octobre 2018] : <http://lesdonnees.e-cancer.fr/Themes/Les-registres-de-cancers>
- 133- Source FRANCIM. Dr P. Grosclaude – communication personnelle. 2018.
- 134- INSERM. Les registres de morbidité. Accessible [Octobre 2018] : <https://www.inserm.fr/recherche-inserm/recherche-en-sante-publique/registres-morbidite>
- 135- Service public - site de l'administration Française. Dossier Médical. <https://www.service-public.fr/particuliers/vosdroits/F12210>
- 136- Conte, C., Palmaro, A., Grosclaude, P., Daubisse-Marliac, L., Despas, F., & Lapeyre-Mestre, M. (2018). A novel approach for medical research on lymphomas: A study

- validation of claims-based algorithms to identify incident cases. *Medicine*, 97(2), e9418.
<http://doi.org/10.1097/MD.00000000000009418>
- 137- Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L et al. Value of a national administrative database to guide public decisions: From the système national d'informaton interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. *Rev Epidemiol Sante Publique*. 2017 Oct;65 Suppl 4:S149-S167
- 138- Assurance maladie. Présentation du SNIIRAM. Document téléchargeable : Accessible [Octobre 2018] :
https://www.ameli.fr/fileadmin/user_upload/documents/Presentation_du_Sniiram.pdf.
- 139- Bezin J , Duong M, Lassalle R, Droz C, Pariente A, Blin P, Moore N. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 2017 Aug;26(8):954-962
- 140- Institut National des Données de Santé. Composantes du SNDS. Accessible [Octobre 2018] : <https://www.indsante.fr/fr/les-composantes-du-snds>
- 141- Décret n° 2016-1871 du 26 décembre 2016 relatif au traitement de données à caractère personnel dénommé « système national des données de santé ». Accessible [Octobre 2018] :
<https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000033702840&categorieLien=id>
- 142- Agence technique de l'information sur l'hospitalisation. Présentation générale du PMSI. Accessible [Octobre 2018] :
<http://www.atih.sante.fr/sites/default/files/public/content/1096/presgen.pdf>
- 143- Source Département d'Information Médicale Hotel Dieu St-Jacques. Pr L. Molinier – communication personnelle
- 144- Rey, G., Rican, S., Jouglu, E. (2011). Mesure des inégalités de mortalité par cause de décès. Approche écologique à l'aide d'un indice de désavantage social. *Bulletin Epidémiologie Hebdomadaire* n°8-9. Accessible [Octobre 2018] :
http://opac.invs.sante.fr/index.php?lvl=notice_display&id=9263
- 145- Centre d'épidémiologie sur les causes médicales de Décès. Description CépiDC. Accessible [Octobre 2018] : <http://www.cepidc.inserm.fr/index.php?p=accueil>
- 146- Epidémiologie France. Détails BCMD. Accessible [Octobre 2018] :
[https://epidemiologie-france.aviesan.fr/ccontent/pdf/\(ObjectId\)/88181](https://epidemiologie-france.aviesan.fr/ccontent/pdf/(ObjectId)/88181)
- 147- Système National des Données de Santé. Description des données en open data. Accessible [Octobre 2018] : <https://www.snds.gouv.fr/SNDS/Open-Data>.

- 148- De Roquefeuil, L., Studer, A., Neumann, A. & Merlière, Y. (2009). L'échantillon généraliste de bénéficiaires : représentativité, portée et limites. *Pratiques et Organisation des Soins*, vol. 40,(3), 213-223. doi:10.3917/pos.403.0213.
- 149- Séminaire INSERM. 2017. Lapeyre-Mestre, M. SNDS - Quoi de neuf dans l'accès aux grandes bases de données? Accessible [Octobre 2018] : http://www.u1027.inserm.fr/medias/fichier/seminaire-snds_1499065607062-pdf?ID_FICHE=190363&INLINE=FALSE
- 150- Système National des Données de Santé. Description SNDS. Accessible [Octobre 2018] : <https://www.snds.gouv.fr/SNDS/Qu-est-ce-que-le-SNDS>.
- 151- Bossi Malafosse, J. (2016). Les nouvelles règles d'accès aux bases médico-administratives : quel effet sur l'ouverture des données ?. *I2D – Information, données & documents*, volume 53,(3), 61-64. <https://www.cairn.info/revue-i2d-information-donnees-et-documents-2016-3-page-61.htm>
- 152- Legifrance. Décret n° 2016-1871 du 26 décembre 2016 relatif au traitement de données du SNDS. Accessible [Octobre 2018] : <https://www.legifrance.gouv.fr/eli/decret/2016/12/26/AFSE1625287D/jo/texte>.
- 153- Quantin, C., Benzenine, E., Hägi, M., Auverlot, B., Abrahamowicz, M., Cottenet, J. et al. Estimation of national colorectal-cancer incidence using claims databases. *JCancer Epidemiol* 2012;2012:298369
- 154- Assurance maladie. Algorithme d'identifications des pathologies dans le SNIIRAM. Accessible [Octobre 2018] : <https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/etudes-en-sante-publique/cartographie-des-pathologies-et-des-depenses/methodologie.php> .
- 155- Bannay, A., Chaignot, C., Blotière, P.O., Basson, M., Weill, A., Ricordeau, P., Alla, F. The Best Use of the Charlson Comorbidity Index With Electronic Health Care Database to Predict Mortality. *Med Care*. 2016 Feb;54(2):188-94.
- 156- Tuppin P, Pestel L, Samson S, Cuerq A, Rivière S, Tala S & al. [The human and economic burden of cancer in France in 2014, based on the Sniiram national database]. *Bull Cancer*. 2017 Jun;104(6):524-537
- 157- INSERM. Démarche réglementaires en fonction du projet. Accessible [Octobre 2018] : <https://www.inserm.fr/professionnels-recherche/recherche-sur-personnes/soumission-projets-impliquant-personne-humaine>
- 158- Ministère des solidarités et de la santé. Réglementation des recherches impliquant la personne humaine. Accessible [Octobre 2018] : <https://solidarites-sante.gouv.fr/systeme->

[de-sante-et-medico-social/recherche-et-innovation/recherches-impliquant-la-personne-humaine/](https://www.indsante.fr/fr/actualite/publication-des-nouvelles-methodologies-de-reference-mr)

- 159- Institut National des Données de Santé. Les nouvelles méthodologies de référence. Accessible [Octobre 2018] : <https://www.indsante.fr/fr/actualite/publication-des-nouvelles-methodologies-de-reference-mr>
- 160- Système National des Données de Santé. Procédure d'accès au SNDS. Accessible [Octobre 2018] : <https://www.snds.gouv.fr/SNDS/Processus-d-acces-aux-donnees>
- 161- Weinstein, M.C., O'Brien, B., Hornberger, J. et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices—Modeling Studies. *Value Health* 2003;6:9–17.
- 162- Kuntz, K., Sainfort, F., Butler, M. et al. Decision and Simulation Modeling in Systematic Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013 Feb.
- 163- Stavros, P. et Alastair, G. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting *BMJ* 2011; 342 :d1766
- 164- Liu, L., Cowen M.E., R.L. Strawderman R.L. and Shih, YCT. A Flexible Two-Part Random Effects Model for Correlated Medical Costs *J Health Econ.* 2010 Jan; 29(1): 110–123.
- 165- Martinez M.J. Modèles linéaires généralisés à effets aléatoires : contributions au choix de modèle et au modèle de mélange. *Mathématiques [math]. Université Montpellier II - Sciences et Techniques du Languedoc*, 2006. Français. <tel-00388820>
- 166- Wikipédia. Densité de la loi gamma. Accessible [Octobre 2018] : https://fr.wikipedia.org/wiki/Loi_Gamma
- 167- Wikipédia. Densité de la loi log normale. Accessible [Octobre 2018] : https://fr.wikipedia.org/wiki/Loi_log-normale
- 168- Nixon, R. M. and Thompson, S. G. (2004), Parametric modelling of cost data in medical studies. *Statist. Med.*, 23: 1311–1331. doi:10.1002/sim.1744
- 169- Fu, L., Moncher, B.R. (2004). Severity Distributions for GLMs: Gamma or Lognormal? Evidence from Monte Carlo Simulations. FCAS, MAAA. Accessible [Octobre 2018] : <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.559.6085&rep=rep1&type=pdf>
- 170- Polgreen, L. A., & Brooks, J. M. (2012). Estimating incremental costs with skew: a cautionary note. *Applied health economics and health policy*, 10(5), 319-29.
- 171- Moran, J. L., Solomon, P. J., Peisach, A. R. and Martin, J. (2007), New models for old questions: generalized linear models for cost prediction. *Journal of Evaluation in Clinical Practice*, 13: 381–389. doi:10.1111/j.1365-2753.2006.00711.x

- 172- Wikipédia. Densité de la loi inverse gaussienne. Accessible [Octobre 2018] : https://fr.wikipedia.org/wiki/Loi_inverse-gaussienne
- 173- Tweedie, M. C. K. (1984). An index which distinguishes between some important exponential families, in J. K. Ghosh and J. Roy (eds), *Statistics: Applications and New Directions*, Proceedings of the Indian Statistical Institute Golden Jubilee International Conference, Calcutta: Indian Statistical Institute.
- 174- Kurz, C. F. (2017). Tweedie distributions for fitting semicontinuous health care utilization cost data. *BMC Medical Research Methodology*, 17, 171. <http://doi.org/10.1186/s12874-017-0445-y>
- 175- Bortoluzzo, A.B., Claro, D.P., Caetano, M.A.L. et Artes, R. (2011). Estimating total claim size in the auto insurance industry: a comparison between tweedie and zero-adjusted inverse gaussian distribution. *BAR - Brazilian Administration Review*, 8(1), 37-47. <https://dx.doi.org/10.1590/S1807-76922011000100004>
- 176- Noriszura, I. (2013). ESTIMATION OF CLAIM COST DATA USING ZERO ADJUSTED GAMMA AND INVERSE GAUSSIAN REGRESSION MODELS. *Journal of Mathematics and Statistics*. 9. 186-192. 10.3844/jmssp.2013.186.192.
- 177- Zeger, S.L., Liang, K.Y., Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics* 1988;44:1049–1060.
- 178- Ballinger, G.A. Using Generalized Estimating Equations for Longitudinal Data Analysis. *Organizational Research Methods*. Vol 7, Issue 2, pp. 127 – 150. First Published April 1, 2004. <https://doi.org/10.1177/1094428104263672>
- 179- Seck, B. Estimation pour les modèles linéaires généralisés : Approche marginale, approche conditionnelle et application (2006). Faculté des sciences et de génie. Université Laval Quebec. Accessible [Octobre 2018] : http://archimede.mat.ulaval.ca/theses/B-Seck_06.pdf
- 180- Package GeePack, 2016. Accessible [Octobre 2018] : <https://cran.r-project.org/web/packages/geepack/geepack.pdf>
- 181- Grivord, P. et Guillermin, M. Les modèles multiniveaux. *Methodologie statistique*. INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES (2016) Accessible [Octobre 2018] : <https://www.insee.fr/fr/statistiques/2022152>
- 182- Laird, N.M., Ware, J.H. Random-effects models for longitudinal data. *Biometrics*. 1982 Dec;38(4):963-74.
- 183- Université de Bristol. Figure 17-18-19 extraite d'un cours relatif aux modèles mixte par l'université de Bristol. Accessible [Octobre 2018] : <http://www.bristol.ac.uk/cmm/learning/videos/random-slopes.html>

- 184- Commengues, D., Jacqmin-Gadda, H. (2015). Modèles biostatistiques pour l'épidémiologie. Editeur : de Boeck. Code ISBN : 9782807300262
- 185- Hubbard, A.E., Ahern, J., Fleischer, N.L. et al, To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. *Epidemiology*. 2010 Jul;21(4):467-74.
- 186- Jacqmin-Gadda H, Sibillot S, Proust C, Molina JM, Thiebaut R, Robustness of the linear mixed model to misspecified error distribution, *Computational Statistics & Data Analysis* 51. 5142-5154
- 187- J. M. Neuhaus, J. D. Kalbfleisch and W. W. Hauck. A Comparison of Cluster-Specific and Population-Averaged Approaches for Analyzing Correlated Binary Data. *International Statistical Review* Vol. 59, No. 1 (Apr., 1991), pp. 25-35
- 188- Muff S., Held L., Keller L.F. Marginal or conditional regression models for correlated non-normal data? *Methods in Ecology and Evolution* 2016. doi: 10.1111/2041-210X.12623
- 189- Pan, W., Louis, T., & Connett, J. (2000). A Note on Marginal Linear Regression with Correlated Response Data. *The American Statistician*, 54(3), 191-195. doi:10.2307/2685589
- 190- Crouchley, R., & Davies, R. B. (1999). A comparison of population average and random-effect models for the analysis of longitudinal count data with base-line information. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 162(3), 331-347. DOI: [10.1111/1467-985X.00139](https://doi.org/10.1111/1467-985X.00139)
- 191- Leyrat, C., Morgan, K.E., Leurent, B., Kahan, B.C. Cluster randomized trials with a small number of clusters: which analyses should be used? *Int J Epidemiol*. 2018 Feb 1;47(1):321-331. doi: 10.1093/ije/dyx169.
- 192- Zhang, H., Yu, Q., Feng, C., Gunzler, D., Wu, P. & Tu, X.M. (2012) A new look at the difference between the GEE and the GLMM when modeling longitudinal count responses, *Journal of Applied Statistics*, 39:9, 2067-2079, DOI: 10.1080/02664763.2012.700452
- 193- Swan, Taryn (2006) Generalized estimating equations when the response variable has a Tweedie distribution: an application for multi-site rainfall modelling. Honours thesis, University of Southern Queensland. (Unpublished). Accessible [Octobre 2018] : https://www.researchgate.net/publication/267376353_Generalized_Estimating_Equations_When_the_Response_Variable_Has_a_Tweedie_Distribution_An_Application_for_Multi-Site_Rainfall_Modelling.
- 194- Yang Y., Demyabix S., Lio Y., Wang J., Inferring Tweedie Compound Poisson Mixed Models with Adversarial Variational Bayes. American International Group Inc. (Submitted on 16 Jun 2017 (v1), last revised 14 Feb 2018

- 195- Liu, L., Cowen, M. E., Strawderman, R. L., & Shih, Y.-C. T. (2010). A Flexible Two-Part Random Effects Model for Correlated Medical Costs. *Journal of Health Economics*, 29(1), 110–123. <http://doi.org/10.1016/j.jhealeco.2009.11.010>
- 196- Foucart, T. « Colinéarité et régression linéaire », *Mathématiques et sciences humaines* [En ligne], 173 | Printemps 2006, mis en ligne le 22 mai 2006, consulté le 13 septembre 2018. URL : <http://journals.openedition.org/msh/2963>
- 197- Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11
- 198- Wilking N.E., Hofmarcher T., Lindgren P., Jönsson B. The burden and direct cost of cancer in Europe (EU-28). DOI: 10.1200/JCO.2016.34.15_suppl.6618. *Journal of Clinical Oncology* 34, no. 15_suppl (May 20 2016) 6618-6618.
- 199- Jönsson B, Hofmarcher T, Lindgren P, Wilking N. The cost and burden of cancer in the European Union 1995-2014. *Eur J Cancer*. 2016 Oct;66:162-70. doi: 10.1016/j.ejca.2016.06.022. Epub 2016 Aug 31.
- 200- Institut national du Cancer. Incidence cancer. Accessible [Octobre 2018] : <http://www.e-cancer.fr/Professionnels-de-sante/Les-chiffres-du-cancer-en-France/Epidemiologie-des-cancers/Donnees-globales>
- 201- Monnereau A, Remontet L, Maynadié M, Binder-Foucard F, Belot A, Troussard X, Bossard N. Estimation nationale de l'incidence des cancers en France entre 1980 et 2012. Partie 2 – Hémopathies malignes. Saint-Maurice : Institut de veille sanitaire ; 2013. 88 p.
- 202- Livre blanc L'organisation, de la prise en charge des patients sous thérapies orales en hématologie. 2016. Accessible [Octobre 2018] : <https://www.oncorif.fr/2017/06/livre-blanc-lorganisation-de-la-prise-en-charge-des-patients-sous-therapies-orales-en-hematologie/>
- 203- Chihara D, Nastoupil LJ, Williams JN, Lee P, Koff JL, Flowers CR. New insights into the epidemiology of non-Hodgkin lymphoma and implications for therapy. *Expert Rev Anticancer Ther*. 2015 May;15(5):531-44
- 204- Comprendre les lymphomes non hodgkiniens. Rapport de l'institut national du cancer. Accessible [Octobre 2018] : www.e-cancer.fr/content/download/63214/569021/file/GUILYMPH11.pdf
- 205- Cours 2009 - 2010 lymphomes malins. Université Médicale Virtuelle Francophone Accessible [Octobre 2018] : http://campus.cerimes.fr/hematologie/enseignement/hematologie_164/site/html/cours.pdf

- 206- Monnereau A, Uhry Z, Bossard N, Cowppli-Bony A, Voirin N, Delafosse P, Remontet L, Troussard X, Maynadié M. Survie des personnes atteintes de cancer en France métropolitaine, 1989-2013. Partie 2 – Hémopathies malignes. Saint-Maurice : Institut de veille sanitaire ; 2016. 144 p
- 207- Haute autorité de santé. Rapport ALD 30, guide du médecin sur le lymphome non hodgkinien Accessible [Octobre 2018] : https://www.has-sante.fr/portail/plugins/ModuleXitiKLEE/types/FileDocument/doXiti.jsp?id=c_1254553
- 208- Laurent C, Baron M, Amara N, Haioun C, Dandoit M, Maynadié M. Impact of Expert Pathologic Review of Lymphoma Diagnosis: Study of Patients From the French Lymphopath Network. *J Clin Oncol.* 2017 Jun 20;35(18):2008-2017
- 209- Lister TA, Crowther D, Sutcliffe SB et al. Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting. *J Clin Oncol.* 1989 Nov;7(11):1630-6.
- 210- Borel, C., Lamy, S., Compaci, G., Récher, C., Jeanneau, P., Nogaro, J. C., ... Laurent, G. (2015). A longitudinal study of non-medical determinants of adherence to R-CHOP therapy for diffuse large B-cell lymphoma: implication for survival. *BMC Cancer*, 15, 288. <http://doi.org/10.1186/s12885-015-1287-9>
- 211- Jozette J.C. Stienen, Rosella P.M.G. Hermens, Lianne Wennekes et al. Variation in guideline adherence in non-Hodgkin's lymphoma care: impact of patient and hospital characteristics. *BMC Cancer.* 2015; 15: 578
- 212- Oncomip treatment guideline of adult lymphomas ("Référentiel Oncomip 2017 Traitement des Lymphomes de l'Adulte"). Accessible [Octobre 2018] : <http://www.oncomip.org/fr/dldoc/?t=recommandations&f=doc1&d=541&h=2916a3fda676eac6842c411081a7dc5c>
- 213- Compaci G, Obéric L, Ysebaert L, Borel C, Laurent G. AMA (Assistance des malades ambulatoires) : une procédure de télémedecine appliquée à la prise en charge des lymphomes agressifs. *Oncologie* (2011) 13: 585. <https://doi.org/10.1007/s10269-011-2069-x>
- 214- Haute autorité de santé. Guide parcours de soins Lymphome de Hodgkin classique de l'adulte. Accessible [Octobre 2018] : https://www.has-sante.fr/portail/jcms/c_1650372/fr/lymphome-de-hodgkin-classique-de-l-adulte-guide-du-parcours-de-soins
- 215- Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al. WHO classification of tumours of haematopoietic and lymphoid tissues, Fourth edition. Lyon:IARC Press;2008

- 216- Institut Nation du Cancer. Comprendre le lymphome hodgkinien. Accessible [Octobre 2018] : <http://www.e-cancer.fr/content/download/96834/1035844/file/Comprendre%20le%20lymphome%20hodgkinien-2014.pdf>
- 217- Le lymphome hodgkinien - La Société de leucémie et lymphome du Canada. Rapport. Disponible : http://www.sllcanada.org/sites/default/files/National/PNT4497-PS57_Hogkin_Lymphoma_10_16_FR_V7_LR.pdf
- 218- Best JH, Hornberger J, Proctor SJ, Omnes LF, Jost F. Cost-effectiveness analysis of rituximab combined with chop for treatment of diffuse large B-cell lymphoma. *Value Health*. 2005 Jul-Aug;8(4):462-70.
- 219- Coiffier B, Lepage E, Briere J, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med* 2002;346: 235–42.
- 220- Deconinck E, Miadi-Fargier H, Pen CL, Brice P. Cost effectiveness of rituximab maintenance therapy in follicular lymphoma: long-term economic evaluation. *Pharmacoeconomics*. 2010;28(1):35-46
- 221- Clinicaltrials.gov. Essais EORTC. Détails centres d'inclusion. https://clinicaltrials.gov/ct2/show/study/NCT00004179?show_locs=Y#locn
- 222- Mojs E., Warchol-Biedermann K., Sanborski W. What Do We Know About Psychological Outcomes of Lymphoma in Adults? *European Psychologist* (2017), 22, pp. 121-131. <https://doi.org/10.1027/1016-9040/a000285>
- 223- Smith, S. K., Zimmerman, S., Williams, C. S., & Zebrack, B. J. (2009). Health status and quality of life among non-Hodgkin lymphoma survivors. *Cancer*, 115(14), 3312-23.
- 224- Yabroff, K. R. and Kim, Y. (2009), Time costs associated with informal caregiving for cancer survivors. *Cancer*, 115: 4362-4373. doi:10.1002/cncr.24588
- 225- European Society for Medical Oncology (ESMO). Guide pour les patients LMC. Accessible [Octobre 2018] : <http://www.esmo.org/content/download/6600/115001/file/FR-LMC-Guide-pour-les-Patients.pdf>
- 226- Pr Arsène Burny. Lettre éditoriale du Fond National de la Recherche Scientifique. 2006. La cellule souche sous tous ses aspects. Accessible [Octobre 2018] : <http://www.fnrs.be/docs/Lettre/lettre64.pdf>
- 227- A. Hochhaus, S. Saussele, G. Rosti, F.-X. Mahon, J. J. W. M. Janssen, H. Hjorth-Hansen, J. Richter, C. Buske, ; Chronic myeloid leukaemia: ESMO Clinical Practice Guidelines for

- diagnosis, treatment and follow-up, *Annals of Oncology*, Volume 28, Issue suppl_4, 1 July 2017, Pages iv41–iv51.
- 228- Sant M, Allemani C, Tereanu C, et al. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood* 2010; 116(19):3724-34.
- 229- T. Leguay, F.-X. Mahon. Leucémie myéloïde chronique. 2005 EM consulte. Doi : 10.1016/S1155-1984(05)39940-7
- 230- Société Française d'Hématologie. Leucémie myéloïde chronique. *Hématologie*. 2007;13(5):328-330. doi:10.1684/hma.2007.0174
- 231- Sorel N, Cayssials E, Brizard F, & al. Actualisation des traitements et du suivi moléculaire dans la prise en charge de la leucémie myéloïde chronique. *Ann Biol Clin* 2017 ; 75 (2) : 129-45
- 232- Nasr R, Bazarbachi A. Chronic myeloid leukemia: "archetype" of the impact of targeted therapies. *Pathol Biol (Paris)*. 2012 Aug;60(4):239-45
- 233- Sacha T. Imatinib in chronic myeloid leukemia: an overview. *Mediterr J Hematol Infect Dis*. 2014; 6(1):e2014007
- 234- O'Brien S, Radich JP, Abboud CN, et al. Chronic Myelogenous Leukemia, Version 1.2014. *J Natl Compr Cancer Netw JNCCN*. 2013 Nov;11(11):1327–40.
- 235- Baccarani M, Deininger MW, Rosti G, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. *Blood*. 2013 Aug 8;122(6):872-84
- 236- Baccarani, M., Castagnetti, F., Gugliotta, G. et al. *Ann Hematol* (2015) 94(Suppl 2): 141. <https://doi.org/10.1007/s00277-015-2322-2>
- 237- Saglio G, Kim DW, Issaragrisil S, et al. Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. *N Engl J Med*. 2010 Jun 17;362(24):2251-9.
- 238- Cortes JE, Jones D, O'Brien S, Jabbour E, Konopleva M, Ferrajoli A, et al. Nilotinib as front-line treatment for patients with chronic myeloid leukemia in early chronic phase. *J Clin Oncol* 2010;28(3):392–7.
- 239- Kantarjian H, Shah NP, Hochhaus A, Cortes J, Shah S, Ayala M, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med* 2010;362(24):2260–70.
- 240- Larson RA, Hochhaus A, Hughes TP, Clark RE, Etienne G, Kim DW, et al. Nilotinib vs imatinib in patients with newly diagnosed Philadelphia chromosome-positive chronic myeloid leukemia in chronic phase : ENESTnd 3-year follow-up. *Leukemia* 2012 ; 26 : 2197-203.

- 241- Jabbour E, Kantarjian HM, Saglio G, Steegmann JL, Shah NP, Boqué C, et al. Early response with dasatinib or imatinib in chronic myeloid leukemia : 3-year follow-up from a randomized phase 3 trial (DASISION). *Blood* 2014 ; 123 : 494-500.
- 242- (2018). When to Stop Tyrosine Kinase Inhibitors for the Treatment of Chronic Myeloid Leukemia. *Current treatment options in oncology*, 19(3), 15. doi:10.1007/s11864-018-0532-2
- 243- Steegmann, J. L., Baccarani, M., Breccia, M., Casado, L. F., García-Gutiérrez, V., Hochhaus, A., Clark, R. E. (2016). European LeukemiaNet recommendations for the management and avoidance of adverse events of treatment in chronic myeloid leukaemia. *Leukemia*, 30(8), 1648–1671. <http://doi.org/10.1038/leu.2016.104>
- 244- Reed SD, Anstrom KJ, Ludmer JA, et al. Cost-effectiveness of imatinib versus interferon-alpha plus low-dose cytarabine for patients with newly diagnosed chronic-phase chronic myeloid leukemia. *Cancer*. 2004 ;101(11):2574-83.
- 245- Ward MA, Fang G, Richards KL, et al. Comparative evaluation of patients newly initiating first-generation versus second-generation tyrosine kinase inhibitors for chronic myeloid leukemia and medication adherence, health services utilization, and healthcare costs. *Curr Med Res Opin*. 2015 Feb;31(2):289-97.
- 246- Guérin A, Wu EQ, Bollu VK, et al. The economic burden of pleural effusions in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors. *J Med Econ*. 2013;16(1):125-33
- 247- Efficace, F., Baccarani, M., Breccia, et al. (2011). Health-related quality of life in chronic myeloid leukemia patients receiving long-term therapy with imatinib compared with the general population. *Blood*, 118(17), 4554-4560. Accessed January 15, 2019. <https://doi.org/10.1182/blood-2011-04-347575>.
- 248- Flynn, K. E., & Atallah, E. (2016). Quality of Life and Long-Term Therapy in Patients with Chronic Myeloid Leukemia. *Current hematologic malignancy reports*, 11(2), 80-5.
- 249- Mounié, M., Costa, N., Gauthier, M., Palmaro, A., Delpierre, C., Despas, F .. & Molinier, L. (2018). Évaluation du coût de la prise en charge des patients atteints de Leucémie Myéloïde Chronique (LMC) en Midi-Pyrénées : impact économique des différents Inhibiteurs de Protéines Kinases. *Journal de gestion et d'économie médicales*, vol. 36,(2), 109-122. doi:10.3917/jgem.182.0109.
- 250- Boucher S., Systèmes de financement et évolution de la prise en charge en cancérologie. 2016. *Innovations & Thérapeutiques en Oncologie*. Volume 2, numéro 5, Septembre-Octobre 2016

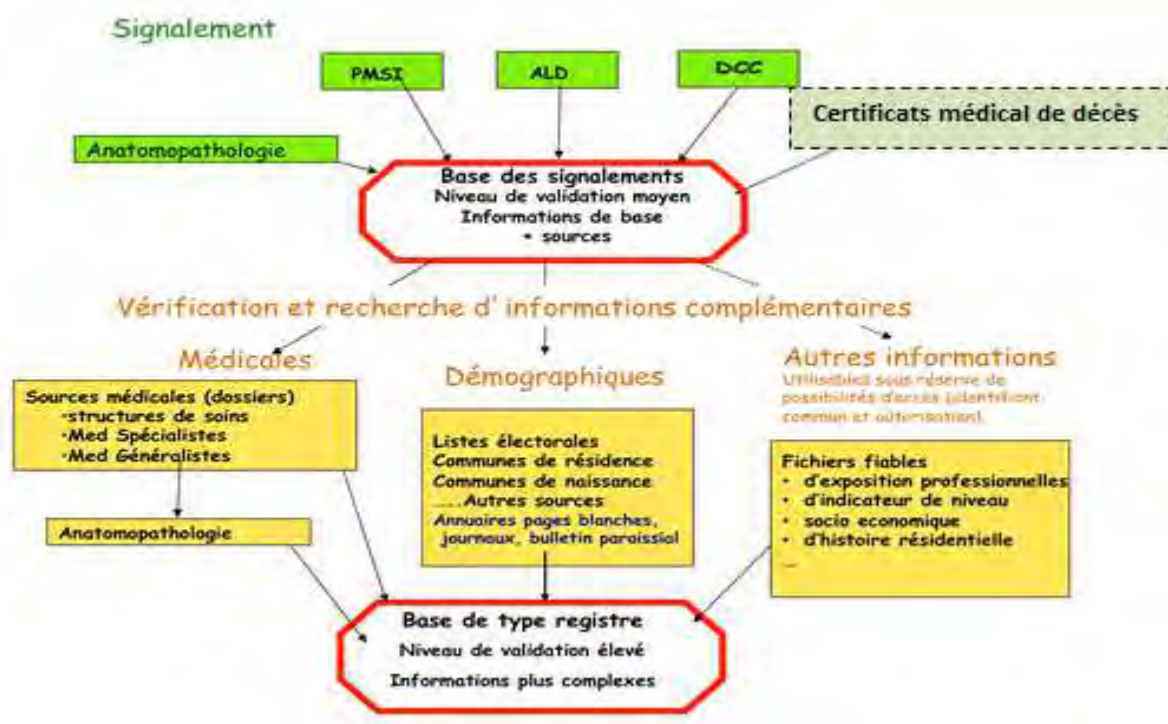
- 251- Vogenberg, F. R., Isaacson Barash, C., & Pursel, M. (2010). Personalized medicine: part 1: evolution and development into theranostics. *P & T : a peer-reviewed journal for formulary management*, 35(10), 560-76. Brittain HK, Scott R, Thomas E. The rise of the genome and personalised medicine. *Clin Med (Lond)*. 2017 Dec;17(6):545-551. doi: 10.7861/clinmedicine.17-6-545.
- 252- Brittain HK, Scott R, Thomas E. The rise of the genome and personalised medicine. *Clin Med (Lond)*. 2017 Dec;17(6):545-551. doi: 10.7861/clinmedicine.17-6-545.
- 253- Société française de médecine prédictive et personnalisée. Accessible [Octobre 2018] : <http://sfmpp.org/expertises/medecine-predictive/>
- 254- Ken Redekop W, Deirdre Mladi. 2013. The Faces of Personalized Medicine: A Framework for Understanding Its Meaning and Scope. *Value in Health*, <https://doi.org/10.1016/j.jval.2013.06.005>.
- 255- Assurance Maladie. Cnamts, rapport sur l'évolution des charges et produits de l'assurance maladie pour 2018, juillet 2017. Accessible [Octobre 2018] : <https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/rapports-et-periodiques/rapports-charges-produits-de-l-assurance-maladie/rapports-charges-et-produits-pour-2018-et-2019/rapport-charges-et-produits-pour-l-annee-2018.php>
- 256- Rapport cour des comptes. Sécurité sociale 2018. Accessible Octobre 2018 : <https://www.ccomptes.fr/fr/publications/securite-sociale-2018>
- 257- Site du réseau UNICANCER. Quelle prise en charge des cancers en 2020 ?. Accessible [Octobre 2018] : <http://www.unicancer.fr/patients/quelle-prise-charge-cancers-2020>
- 258- Silber, D. Médecine 2.0 : les enjeux de la médecine participative. *Press Med* 2009. Doi : 10.1016/j.lpm.2009.06.011
- 259- Direction de la recherche, des études, de l'évaluation et des statistiques. 2017. Dépenses de santé en 2017. Accessible [Octobre 2018] : <https://drees.solidarites-sante.gouv.fr/etudes-et-statistiques/publications/panoramas-de-la-drees/article/les-depenses-de-sante-en-2017-resultats-des-comptes-de-la-sante-edition-2018>
- 260- Assurance maladie. Une analyse médicalisée des dépenses remboursées par l'assurance maladie sur la période 2012 -2016. Accessible [Octobre 2018] : <https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/etudes-en-sante-publique/cartographie-des-pathologies-et-des-depenses/presentation-et-principaux-resultats.php>
- 261- Savage, P., & Mahmoud, S. (2015). Development and economic trends in cancer therapeutic drugs: a 5-year update 2010–2014. *British Journal of Cancer*, 112(6), 1037–1041. <http://doi.org/10.1038/bjc.2015.56>

- 262- 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. Partie 1 – Tumeurs solides. Saint-Maurice (Fra) : Institut de veille sanitaire, 2013. 122 p.
- 263- Haute Autorité de Santé. Missions de la Commission Évaluation Économique et de Santé Publique. Accessible [Octobre 2018] : https://www.has-sante.fr/portail/jcms/c_419565/fr/commission-evaluation-economique-et-de-sante-publique
- 264- Haute Autorité de Santé. Présentation La Commission évaluation économique et de santé publique. Accessible [Octobre 2018] : https://www.has-sante.fr/portail/upload/docs/application/pdf/2012-11/quest-ce_que_la_ceesp-brochure2clics_2012-11-22_10-13-34_631.pdf
- 265- Emery G, Le Fur C, Epis-de-Fleurian AA, Josserand A. Parcours de soins, vers la mise en place d'un financement adapté : raisons et conditions de réussite. *Thérapie*. Volume 73, Issue 1, February 2018, Pages 25-32
- 266- Assurance Maladie. Cnamts, rapport sur l'évolution des charges et produits de l'assurance maladie pour 2017, juillet 2016. Accessible [Octobre 2018] : <https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/rapports-et-periodiques/rapports-charges-produits-de-l-assurance-maladie/rapports-charges-et-produits-pour-2013-a-2017/rapport-charges-et-produits-pour-l-annee-2017.php>
- 267- Véran, O. L'évolution des modes de financement des établissements de santé. Rapport du ministère des solidarités et de la santé. 2017. Accessible [Octobre 2018] : <https://solidarites-sante.gouv.fr/ministere/documentation-et-publications-officielles/rapports/sante/article/l-evolution-des-modes-de-financement-des-etablissements-de-sante>
- 268- Ministère des solidarités et de la santé. Parcours de santé, de soins et de vie : une approche globale au plus près des patients. Accessible [Octobre 2018] : <https://solidarites-sante.gouv.fr/systeme-de-sante-et-medico-social/parcours-des-patients-et-des-usagers/article/parcours-de-sante-de-soins-et-de-vie>
- 269- Haute autorité de santé. Définition Parcours de santé (2017). FAQ. Accessible [Octobre 2018] : http://www.has-sante.fr/portail/jcms/c_1247635/questions-reponses-parcours-de-soins
- 270- Press MJ, Rajkumar R, Conway PH. Medicare's New Bundled Payments Design, Strategy, and Evolution. *JAMA*. 2016;315(2):131–132. doi:10.1001/jama.2015.18161

- 271- Beija, M., Georges-Picot, A. Quelles leçons tirer du financement au parcours de santé à l'étranger ? 2018. FINANCES HOSPITALIERES, n° 123, 2018/04/01, pages 15-20. Accessible [Octobre 2018] : <http://www.bdsp.ehesp.fr/Base/520239/>
- 272- Maynard A. Economic evaluation techniques in healthcare: reinventing the wheel? *Pharmacoeconomics* 1997;11(2):115-8
- 273- Haute autorité de santé. Méthodes quantitatives pour évaluer les interventions visant à améliorer les pratiques. Accessible [Octobre 2018] : https://www.has-sante.fr/portail/jcms/c_597750/fr/methodes-quantitatives-pour-evaluer-les-interventions-visant-a-ameliorer-les-pratiques
- 274- Site SNDS. Mobiliser les données du SNDS pour mieux évaluer le parcours de santé des aînés (PAERPA). Accessible [Octobre 2018] : <https://www.snds.gouv.fr/SNDS/Actualites/Actu-4>
- 275- Garrison LP Jr, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report.. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report. *Value Health*. 2007 Sep-Oct;10(5):326-35.
- 276- Josselin, JM. et Le Maux, B. *Statistical Tools for Program Evaluation. Methods and Applications to Economic Policy, Public Health, and Education*. Heidelberg: Springer.
- 277- Gillaizeau F, Grabar S. Modèles de régression multiple. *Sang Thrombose Vaisseaux* 2011 ; 23 (7) : 360-70 doi:10.1684/stv.2011.0632
- 278- Thiese, M. S., Arnold, Z. C., & Walker, S. D. (2015). The misuse and abuse of statistics in biomedical research. *Biochimica medica*, 25(1), 5-11.

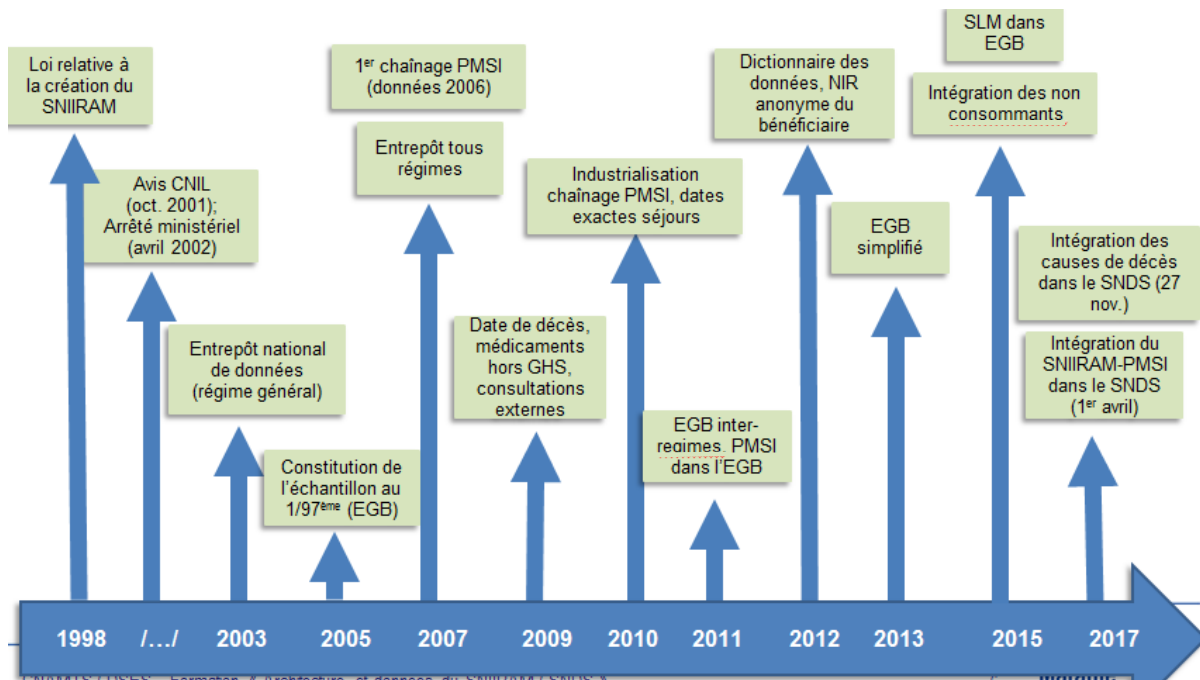
ANNEXE

Annexe 1. Structure générale de la collecte des informations dans un registre (133).



PMSI : Programme de Médicalisation des Systèmes d'Information, ALD : Affection Longue Durée. DCC : Dossiers Communicant de Cancérologie.

Annexe 2. Evolution du SNIIRAM depuis sa création jusqu'à son intégration dans le SNDS (138).

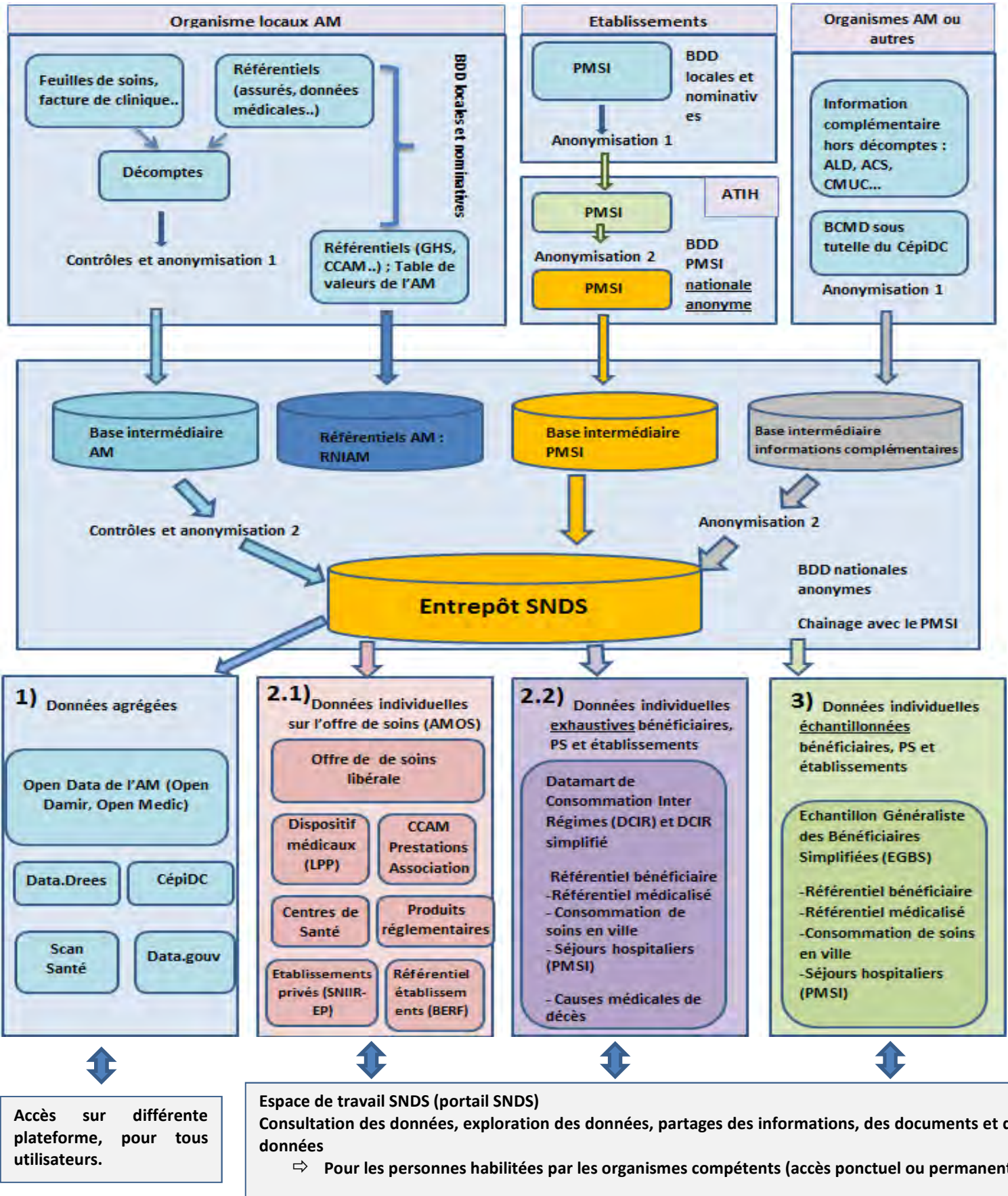


SNIIRAM : Système national d'information inter-régimes de l'Assurance maladie ; CNIL : Commission nationale de l'informatique et des libertés ; EGB : Échantillon généraliste des bénéficiaires ; PMSI : Programme de médicalisation des systèmes d'Information

Annexe 3. Principales informations contenues dans le PMSI

Les données contenues dans les bases du PMSI sont celles des Résumé de Sortie Anonyme (RSA), recueillies directement au niveau des établissements avant d'être centralisés au niveau national. On retrouve alors des informations sur :

- les séjours facturés directement à l'AM (clinique privé ou secteur médico-social handicap) et les séjours du PMSI pour les séjours en MCO, Hospitalisation A Domicile (HAD), les Soins de Suite et de Réadaptation (SSR) et la psychiatrie (PSY)
- les médicaments et dispositifs facturés en supplément (en « sus ») des forfaits GHS
- les dates des séjours, les codes diagnostics (principaux, reliés, associés), les coûts associés et des informations sur l'établissement (i.e. FINESS, statut juridique)



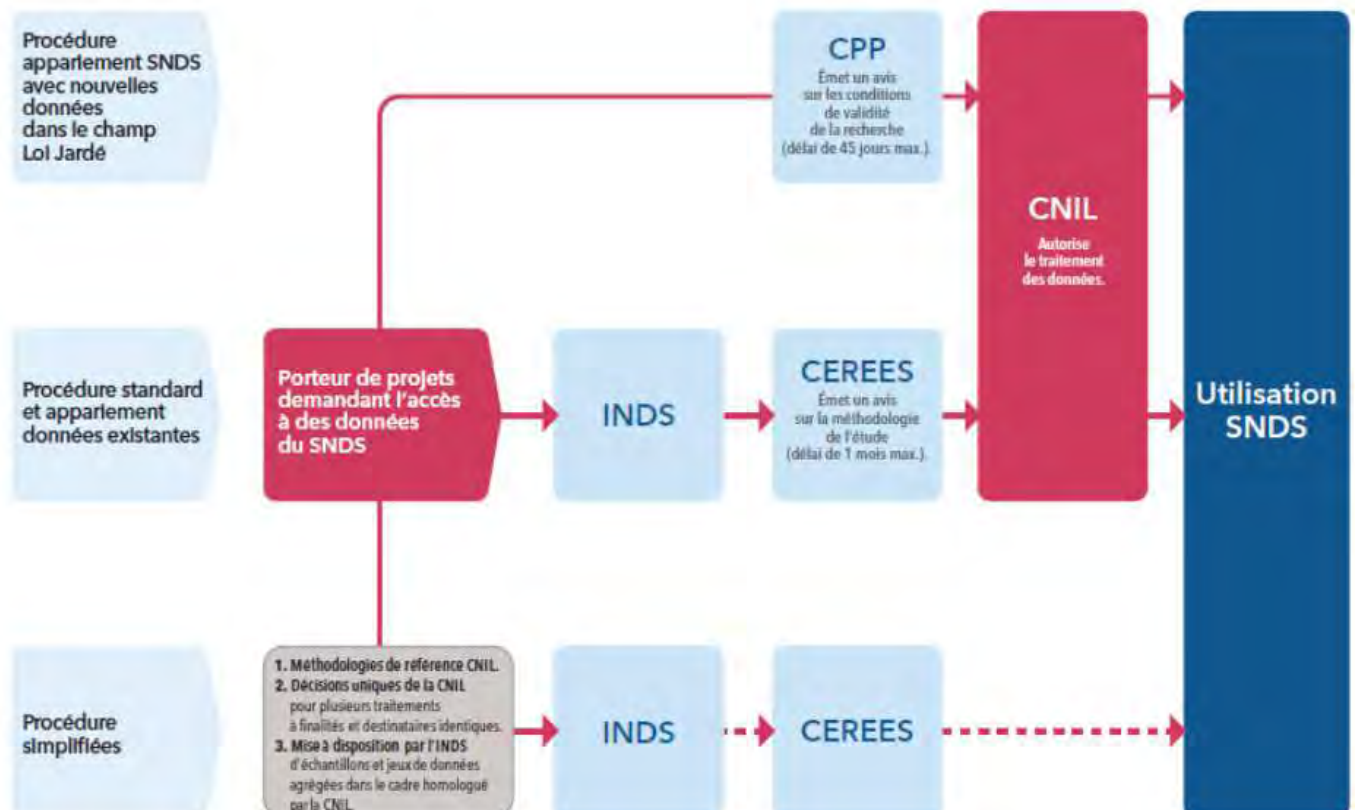
AM : Assurance Maladie, GHS : Groupe Homogène de Séjours, CCAM : Classification Commune des Actes Médicaux, PMSI : Programme Médicalisé des Systèmes d'Information, BDD : Base De Données, ATIH : Agence Technique d'Information Hospitalière, ALD : Affection Longue Durée, ACS : Aide à la Complémentaire Santé, CMU-C : Couverture Maladie Universelle Complémentaire, BCMD : Base des Cause Médicales de Décès, CépiDC : Centre d'épidémiologie sur les causes médicales décès, RNIAM : Répertoire National Inter Régime de l'AM, DAMIR : Dépenses de l'AM Inter Régimes, UCD : Unité Commune de Dispensation, PS : Professionnels de Santé, LPP : Liste des Produits et Prestations, SNIIREP : Système National Inter-Régime des Etablissements Privés, BERF : Base Etablissement Référentielle, EGBS : Echantillon Généraliste des Bénéficiaires Simplifiées, DCIR : Datamart de Consommation Inter Régimes; SNDS: Système national des données de santé

Annexe 5. Organisme ayant droit à un accès permanent au SNDS (152)

N°	Organisme ayant droit	Profondeur historique maximale (années)
1	La direction de la recherche, des études, de l'évaluation et des statistiques, la direction générale de la santé et la direction générale de l'offre de soins, la direction de la sécurité sociale, la direction du budget et le service de santé des armées ;	9
2	Les agences régionales de santé ;	9
3	Les caisses nationales des régimes de l'assurance maladie obligatoire, les organismes locaux et régionaux de l'assurance maladie obligatoire ;	9
4	La Caisse nationale de solidarité pour l'autonomie ;	9
5	La Haute Autorité de santé ;	19
6	L'Autorité de sûreté nucléaire ;	5
7	L'Agence nationale de santé publique ;	19
8	L'Agence nationale de sécurité du médicament et des produits de santé ;	19
9	L'Agence de biomédecine ;	19
10	L'Institut de radioprotection et de sûreté nucléaire ;	5
11	L'Institut national du cancer ;	19
12	L'Etablissement français du sang ;	19
13	L'Agence technique de l'information sur l'hospitalisation ;	9
14	L'Agence nationale d'appui à la performance des établissements de santé et médico-sociaux ;	5
15	L'Institut national des données de santé ;	5
16	L'Institut de recherche et documentation en économie de la santé ;	9
17	L'Institut national d'études démographiques ;	5
18	L'Observatoire français des drogues et toxicomanies ;	5
19	Le Haut Conseil pour l'avenir de l'assurance maladie ;	5
20	Le Fonds de financement de la couverture maladie universelle ;	5
21	Les observatoires régionaux de la santé ;	5
22	Les unions régionales de professionnels de santé ;	5
23	Les équipes de recherche de l'Institut national de la santé et de la recherche médicale ;	5
24	Les équipes de recherche des centres hospitaliers universitaires et des centres de lutte contre le cancer ;	5
25	Les équipes de recherche et de formation de l'Ecole des hautes études en santé publique.	5

**Il faut rajouter l'année en cours à la profondeur historique maximale indiqué.*

Annexe 6. Synthèse du circuit d'accès ponctuel au SNDS (160)



SNDS : Système National des Données de Santé. CPP : Comité de Protection de la Personne. CNIL : Commission Nationale de l'Informatique et des Libertés. INDS : Institut National des Données de Santé. CEREES : Comité d'Expertise pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé.

Annexe 7. Ecriture et résolution des modèles de population générale (179).

Soient $Y_{i,j}$ la variable à expliquer, $X_{i,j}$ les variables explicatives et $\beta_{i,j}$ les coefficients associés, $\varepsilon_{i,j}$ les résidus du modèle. On a :

$$g(Y_{i,j}) = \beta_{i,j}X_{i,j} + \varepsilon_{i,j}$$

Où:

$$g(E[Y_{i,j}]) = g(\mu_{i,j}) = x_{i,j}^T \beta$$

Avec $\mu_{i,j}$ la moyenne de la distribution représentant la variable à expliquer, g la fonction de lien, x le vecteur de régression, β le vecteur des coefficients de régression et avec $i = 1 \dots n_i$ et $j = 1 \dots n_j$

Soit :

$$\varphi a_{i,j} = \text{Var}(Y_{i,j})$$

Où φ est un paramètre de dispersion et $a_{i,j} = a(\mu_{i,j})$ est une fonction de variance connue.

Soit:

$$\vec{Y}_i = (Y_{i,1}, \dots, Y_{i,n_i})$$

Soient $\vec{R}_i(\alpha)$, la matrice de corrélation de travail complètement décrite par n vecteur α à spécifier et \vec{V}_i la matrice de covariance de travail de \vec{Y}_i :

$$\vec{V}_i = \varphi \vec{A}_i^{1/2} \vec{R}_i(\alpha) \vec{A}_i^{1/2}$$

Avec \vec{A}_i la matrice diagonale des $a_{i,j}, j = 1, \dots, n_i$

Pour les estimations $(\hat{\varphi}, \hat{\alpha})$ de (φ, α) , l'estimation $\hat{\beta}$ est solution des équations appelées GEE :

$$\sum_{i=1}^K \frac{\partial \vec{\mu}_i^T}{\partial \beta} \vec{V}_i^{-1} (\vec{Y}_i - \vec{\mu}_i) = \vec{0}$$

Afin d'estimer les paramètres de la matrice de variance \vec{V}_i et de vérifier l'ajustement du modèle, on définit les résidus :

$$e_{i,j} = \frac{Y_{i,j} - \mu_{i,j}}{\sqrt{b''(X_{i,j}^T \beta)}}$$

Que l'on évalue en $\beta = \hat{\beta}$. Pour estimer le paramètre de dispersion φ , on pose :

$$\hat{\varphi} = \frac{1}{N - p} \sum_{i=1}^K \sum_{j=1}^{n_i} e_{i,j}^2$$

Où $N = \sum_{i=1}^K n_i$ est le nombre total d'observation dans l'échantillon et $p = \dim(\beta)$

Ainsi, les estimations de β se calculent à partir de l'algorithme suivant :

- Estimer β sous l'hypothèse d'indépendance (i.e. comme un modèle linéaire généralisé simple) et appeler l'estimateur obtenu $\hat{\beta}_0$.
- Estimer α et φ à partir de $\hat{\beta}$ (i.e. $\hat{\beta}_0$ à la 1ere itération) et des $e_{i,j}$
- Poser $\vec{V}_i = \hat{\varphi} \vec{A}_i^{1/2} \vec{R}_i(\hat{\alpha}) \vec{A}_i^{1/2}$
- Mettre à jour la valeur de $\hat{\beta}$:

$$\hat{\beta}_{m+1} = \hat{\beta}_m + \left(\sum_{i=1}^K \vec{D}_i^T \vec{V}_i^{-1} \vec{D}_i \right) \left(\sum_{i=1}^K \vec{D}_i^T \vec{V}_i^{-1} [\vec{Y}_i - \vec{\mu}_i(\hat{\beta}_m)] \right)$$

Avec $\vec{D}_i^T = \frac{\partial \vec{\mu}_i^T}{\partial \beta}$

- Répéter les étapes de l'algorithme jusqu'à convergence : différence entre $\hat{\beta}_{m+1}$ et $\hat{\beta}_m$ plus petite qu'une tolérance spécifiée

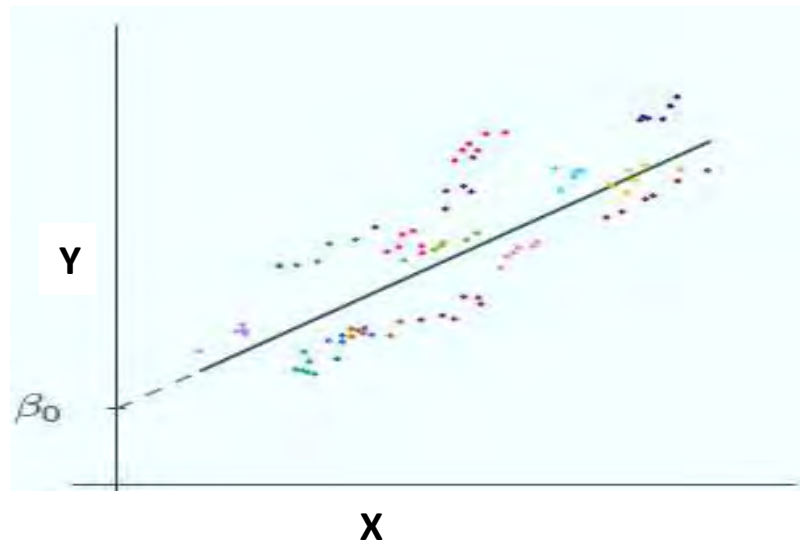
La matrice de corrélation $\vec{R}_i(\alpha)$ utilisée est une matrice de travail et donc possiblement fautive. La variance de $\hat{\beta}$, \vec{V} est alors estimée par un estimateur de variance sandwich robuste.

Annexe 8. Les différents types de modèles mixtes.

Les figures utilisées dans cette annexe sont extraites d'un cours sur les modèles mixtes donné par l'université de Bristol en Angleterre (183).

Soit la représentation graphique d'un jeu de données (Y en fonction de X) en cluster ou les modèles mixtes pourraient être utilisés.

Figure 1, annexe 8. Représentation de Y en fonction de X avec des données en clusters et la droite de régression estimée dans le cadre des modèles linéaires.



Comme on le voit sur le nuage de point, et contrairement à la figure 10 dans le corps du manuscrit, la relation entre X et Y n'est plus uniformément répartie autour de la droite de régression et on distingue clairement des groupes. Ainsi la droite de régression (i.e. équation relative au modèle linéaire) seule n'est plus apte à fournir des estimations valides et il est nécessaire de prendre en compte les effets aléatoires introduits dans l'équation des modèles mixtes.

On peut alors distinguer deux types de modèles mixtes : les modèles mixtes à intercept aléatoire et les modèles mixtes à intercept et pente aléatoire.

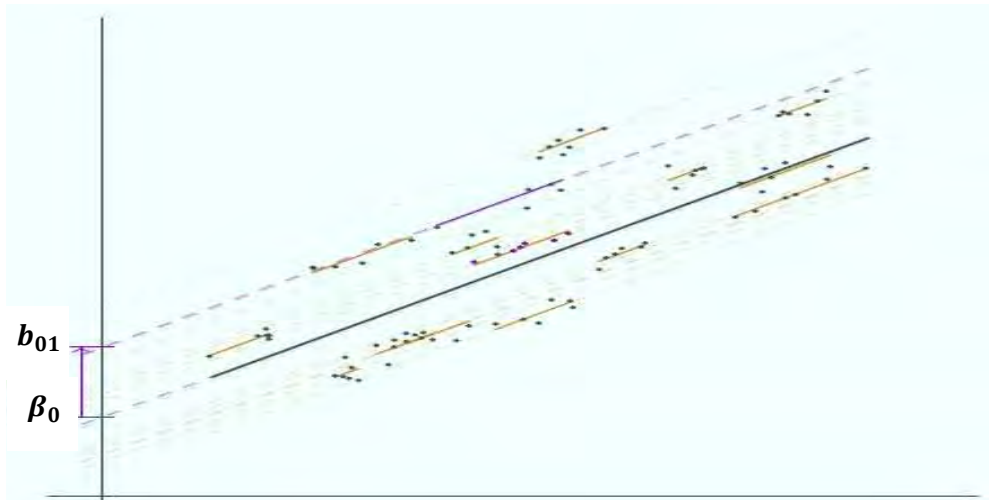
Les modèles à intercept aléatoire sont les modèles mixtes les plus simples n'incluant qu'un seul effet aléatoire. Ils supposent que seul le niveau initial est spécifique au cluster. On alors $Z_{i,j} = 0 \forall i, j$ et

$$Y_{i,j} = \beta X_{i,j} + b_{0i} + \varepsilon_{i,j}$$

$$Y_{i,j} = \beta_0 + b_{0i} + \beta_1 X_{i,j} + \varepsilon_{i,j}$$

Avec $\varepsilon_{i,j} \sim N(0, \sigma_\varepsilon^2)$ et $b_i \sim N(0, \sigma_{b_0}^2)$. En reprenant la typologie de la figure ci-dessus avec des données en clusters, on aurait alors la figure suivante.

Figure 2, annexe 8. Représentation des droites de régression selon les intercepts aléatoires



Ainsi, si le modèle était seulement à effets fixes, l'intercept serait toujours égal à β_0 . Ici, il est égal à $\beta_0 + b_{0,i}$, selon le groupe i . De plus, étant donné que $b_{0,i}$ suit une loi normale de moyenne nulle, la moyenne de l'intercept global sera toujours égal à β_0 , tandis que pour chaque cluster son intercept sera égal à $\beta_0 + b_{0,i}$. Chaque intercept dans chacun des groupes sera ainsi aléatoire (car $b_{0,i}$ aléatoire). La variance des $b_{0,i}$ (i.e. $\sigma_{b_0}^2$) expliquera l'hétérogénéité inobservée apportée par les différents clusters tandis que la variance des $\varepsilon_{i,j}$ (i.e. σ_ε^2) expliquera l'hétérogénéité inobservée apportée par l'ensemble de la population.

Les modèles à intercept et pente aléatoires s'adaptent généralement mieux aux données. Comme dans le modèle précédemment, le niveau initial (i.e. intercept) variera d'un cluster à l'autre mais les pentes mesurées dans chaque cluster seront également différentes pour chaque groupe. On revient alors à l'équation générale des modèles mixtes.

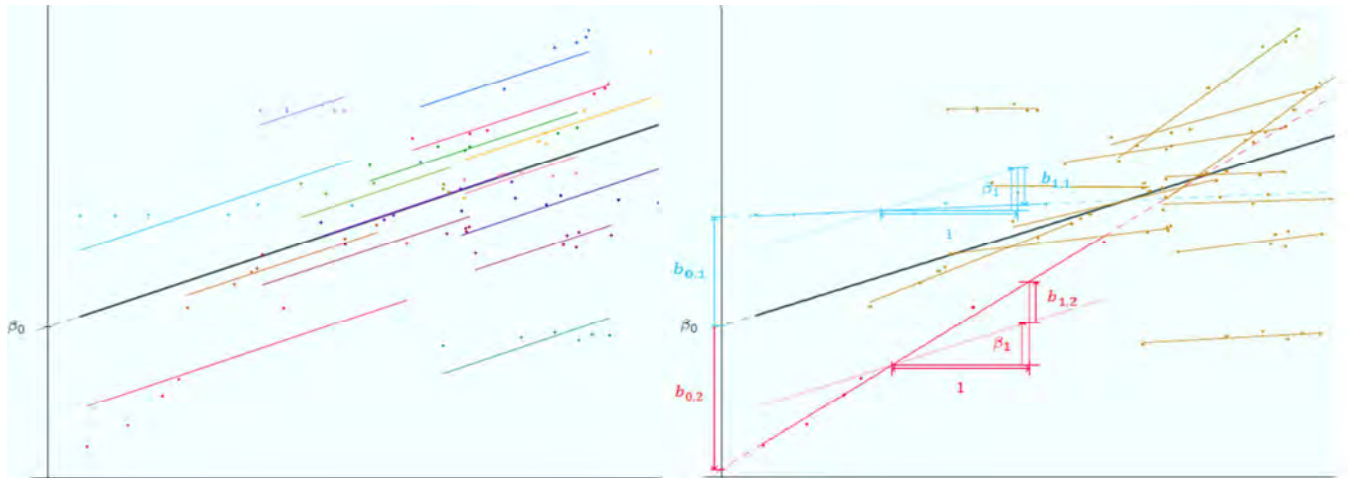
$$Y_{i,j} = \beta X_{i,j} + b_i Z_{i,j} + \varepsilon_{i,j}$$

$$Y_{i,j} = \beta_0 + \beta_1 X_{i,j} + b_{0,i} + b_{1,i} Z_{i,j} + \varepsilon_{i,j}$$

Avec $\varepsilon_{i,j} \sim N(0, \sigma_\varepsilon^2)$ et $b_i = (b_{0,i}, b_{1,i})^T \sim N\left(0, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix}\right)$.

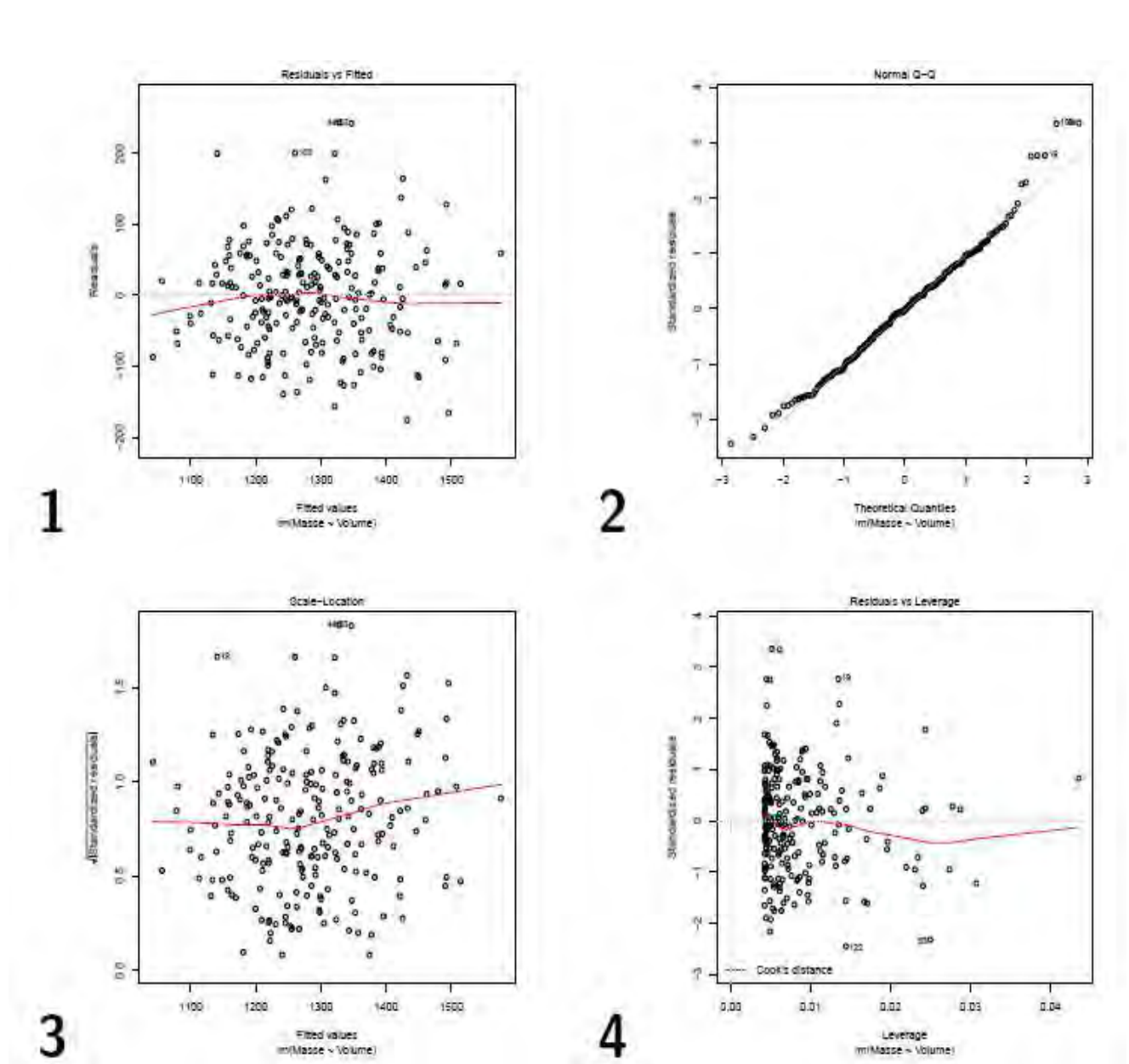
De la manière identique et pour une meilleure visualisation, on représente les intercepts aléatoires et les intercepts et pentes aléatoire sur la figure suivante.

Figure 3, annexe 8. Comparaison des droites de régression entre un modèle mixte à intercept aléatoire (gauche) et intercept et pente aléatoire (droite).



Ainsi chaque droite de régression s'adapte directement aux données du cluster lui étant associées. Identiquement, en moyenne sur toute la population l'intercept et la pente ne changent pas (i.e. effets aléatoires de moyenne nulle). Néanmoins, pour chaque cluster la droite de régression est égale à $\beta_0 + b_{0i} + \beta_1 X_{i,j} + b_{1i} Z_{i,j}$.

Annexe 9. Etude des résidus



Les 4 graphiques ci-dessus sont des outils permettant d'étude la qualité d'un modèle de régression linéaire.

- 1) Le premier graphique représente les résidus estimés en fonction des valeurs prédites. Si les résidus ne présentent pas d'organisation particulière, l'hypothèse de moyenne nulle est validée : $E(\varepsilon_i) = 0$. Si les résidus ont une dispersion à peu près constante, l'hypothèse d'homoscédasticité est validée : $V(\varepsilon_i) = \sigma^2$.

- 2) Le second graphique représente les quantiles des résidus standardisés en fonction de ceux de la loi normale $N(0,1)$. Si les points se répartissent selon une droite, l'hypothèse de normalité de Y ou des résidus est respectée.
- 3) Le troisième graphique représente $\sqrt{\varepsilon_i}$ en fonction des valeurs prédites \hat{Y}_i . Si l'amplitude des $\sqrt{\varepsilon_i}$ est constante l'hypothèse d'homoscédasticité est validée : $V(\varepsilon_i) = \sigma^2$.
- 4) Le quatrième graphique représente les résidus standardisés en fonction du levier. La ligne rouge représente alors la distance de Cook. Si elle dépasse la valeur 1 pour une observation i , il faut envisager de retirer le point i des données.

Annexe 10. Les stades des lymphomes

L'évolution des LNH se divise en 4 stades. Les stades I et II sont dit des LNH localisés alors que les stades III et IV sont considérés comme étendus ou disséminés. On retrouve :

- Stade I : début de la maladie.
 - o Atteinte d'un seul groupe ganglionnaire.
- Stade II : maladie localement avancée.
 - o Atteinte de plusieurs groupes ganglionnaires du même côté du diaphragme.
- Stade III : maladie avancée.
 - o Atteinte d'aires ganglionnaires des deux côtés du diaphragme.
- Stade IV : extension de la maladie
 - o Atteinte d'un ou plusieurs organes tels que la moelle osseuse, le poumon ou le foie.

Il peut ensuite y être rajouté des lettres :

- E : ajouté aux stades I, II ou III, s'il y a une atteinte viscérale contiguë
- B : s'il y a amaigrissement inexpliqué de plus de 10 % du poids du corps en moins de 6 mois ou fièvre inexpliquée >38 °C de plus de 8 jours ou sueurs nocturnes profuses.
- A : s'il n'y a aucun signe de B.
- X : si la masse tumorale est volumineuse.
- S : en cas d'atteinte splénique¹⁵

¹⁵ Splénique : qui appartient, est relatif à la rate

Annexe 11. Evolution des traitements et de la survie des patients atteints de la LMC (231)

Figure 1, annexe 11. Evolution des traitements de la LMC depuis la moitié du 19ieme siècle jusqu'à nos jours

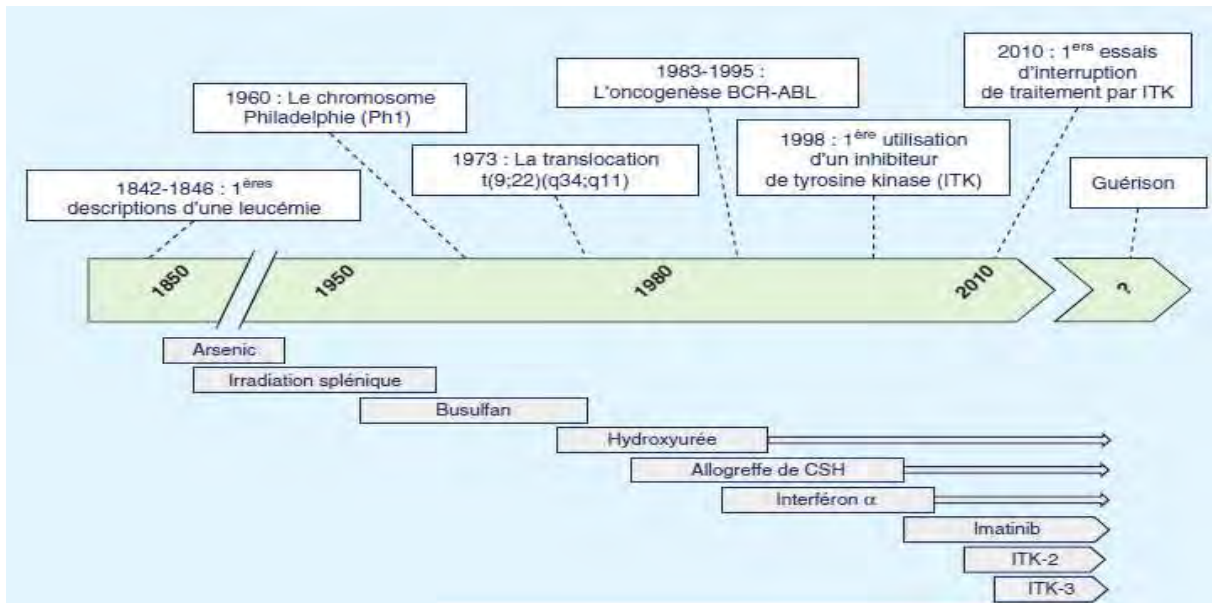
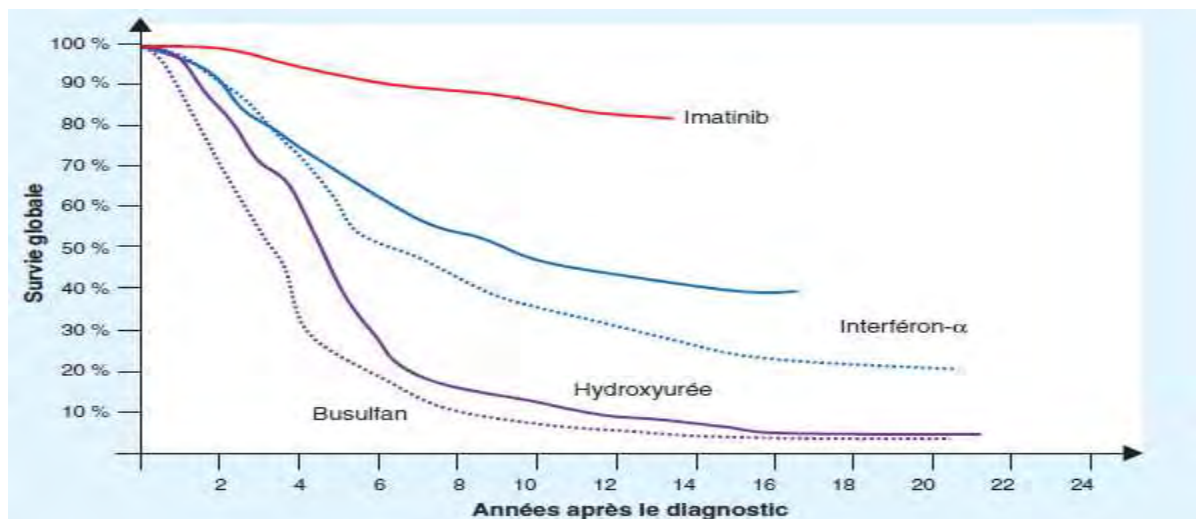


Figure 2, annexe 11. Evolution de la survie des patients atteints de LMC selon les principales prises en charge présentées ci-dessus



L'introduction de l'imatinib, le premier ITK sur le marché, a révolutionné le pronostic de survie des patients atteints de LMC.

Titre en Français

Les évaluations économiques en population dans les prises en charge des cancers : évaluation des coûts des lymphomes et de la leucémie myéloïde chronique

Résumé en Français

Le système de soin français rencontre des difficultés de financement des dépenses de santé qui placent les décideurs face à des contraintes budgétaires rendant nécessaire la priorisation des stratégies en termes de santé publique. Dans un tel contexte, les évaluations économiques des stratégies médicales, et notamment les Evaluations Médico-Economiques (EME), peuvent aider à une allocation optimale des ressources. La cancérologie notamment n'échappe pas à cette problématique. Elle constitue un enjeu national en santé de par la forte prévalence, mortalité et le caractère onéreux de ses traitements. De plus, les constantes évolutions des prises en charges des pathologies cancéreuses, parallèlement aux moyens accrus qu'elles consomment, prodiguent une importance primordiale aux évaluations économiques dans ce domaine. Les hémopathies malignes sont directement concernées par ces questions avec des avancées thérapeutiques majeures réalisées ces dernières années notamment dans la Leucémie Myéloïde Chronique (LMC), la leucémie lymphoïde, les lymphomes et les myélomes.

Les EME des prises en charge des pathologies cancéreuses sont généralement issues d'Essais Cliniques Randomisés (ECR) sur des populations sélectionnées, sur-contrôlées et centrées sur un traitement spécifique. L'approche en population basée sur des données d'observations est alors intéressante car elle permet la production de résultats reflétant les pratiques réelles. L'accès désormais facilité aux bases de données médico-administratives nationales du SNDS offre des possibilités intéressantes au regard de cet enjeu. Néanmoins, sans la rigueur méthodologique des ECR, l'information contenue dans les bases est dite « imparfaite », et le recours à la modélisation statistique est essentiel afin de structurer et d'appréhender au mieux l'information contenue dans les données. Différents modèles théoriques existent et peuvent être transposés aux problématiques des évaluations économiques des stratégies de soins des pathologies cancéreuses. La connaissance des hypothèses sous-jacentes à ces modèles permet l'usage des plus pertinents le cas échéant et ainsi l'obtention de résultats approchant au mieux la réalité.

Réalisée dans le cadre du programme de recherche CAPTOR (Cancer Pharmacology of Toulouse-Oncopole and Region) sur les médicaments anti-cancéreux, cette thèse avait pour objectifs de décrire et d'analyser les coûts des lymphomes et de la LMC en France. Les travaux effectués ont permis de mettre en évidence les différents postes de dépenses dans la prise en charge des lymphomes selon les pathologies considérées et d'identifier où des économies potentielles peuvent être réalisées. De plus, les recherches mises en œuvre ont permis une première description des coûts de la prise en charge de la LMC en France, de ces inducteurs de coûts et des variations dues aux différents traitements dispensés. Ces travaux mettent en avant le potentiel et l'importance d'utiliser les bases du SNDS pour l'évaluation économique des pathologies en France tout en utilisant des modèles adaptés aux problématiques soulevées.

Mots clefs : Evaluation économique en santé, hémopathie maligne, modélisation, bases de données médico-administratives, approche populationnelle.

Titre en Anglais

Real-world economic evaluation of cancer management: cost assessment of lymphomas and chronic myeloid leukemia

Résumé en Anglais

The French health care system has difficulties to manage health expenditure, which places decision-makers in front of budgetary constraints making necessary to prioritize strategies in public health. In this context, Economic Evaluation and Health Economic Evaluation (HEE) help to an optimal allocation of resources. Cancers are particularly concern with this problem. These pathologies are a national health issue due to their high prevalence, mortality and the costly nature of its treatments. In addition, the large changes in the management of cancer pathologies and the increased resource they consume bring a major place to economic evaluation in this area. Hematological malignancies are directly concerned by these issues with significant therapeutic advances made in recent years, particularly in Chronic Myeloid Leukemia (CML), lymphoid leukemia, lymphoma and myeloma.

HEEs for the management of cancer pathologies use generally data from Randomized Clinical Trials (RCTs). These studies are focused on a specific treatment related to strictly selected and over-controlled population. The population based approach using observational data is interesting because it allows the production of results reflecting actual practices. The recent ease access to the national health insurance databases offers interesting possibilities with regard to this issue. Nevertheless, without the methodological rigor of RCTs, the information contained in the databases could be called "imperfect". Therefore, the use of statistical modeling is essential in order to structure and improve understanding the information contained in the data. Various models exist and can be transposed to the economic evaluation of cancer management context. Knowledge of the underlying assumptions allows to use the most relevant when it's needed and to obtain results near to reality.

Conducted as part of the CAPTOR (Cancer Pharmacology of Toulouse-Oncopole and Region) research program on anti-cancer drugs, this thesis aimed to describe and analyze the costs of lymphoma and CML in France. The work carried out the various expenditure items in the management of lymphomas according to pathologies considered. In addition, our works have allowed a first description of the costs of the management of CML in France, of these cost drivers and the variations due to the different treatments dispensed. This work highlights the potential and the importance of health insurance databases to realize economic evaluation of diseases in France using suitable models according to the different problematics met.

Key words: Health economic evaluation, malignant hemopathy, modelization, administrative medical databases, population-based.
