



HAL
open science

Apport des méthodes de survie nette dans le pronostic des lymphomes malins non hodgkiniens en population générale

Morgane Mounier

► **To cite this version:**

Morgane Mounier. Apport des méthodes de survie nette dans le pronostic des lymphomes malins non hodgkiniens en population générale. Santé. Université Claude Bernard - Lyon I, 2015. Français. NNT : 2015LYO10120 . tel-01267575

HAL Id: tel-01267575

<https://theses.hal.science/tel-01267575>

Submitted on 4 Feb 2016

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 DE SCIENCES

Présentée et soutenue publiquement

Le 17 septembre 2015

Par Morgane MOUNIER

Née le 14 mars 1983 à Saint-Martin-d'Hères (38)

Apport des méthodes de survie nette dans le pronostic des lymphomes malins non hodgkiniens en population générale

Directeur de thèse : Professeur Roch GIORGI

Codirecteur de thèse : Docteur Nadine BOSSARD

Pour obtenir le grade de DOCTEUR de l'UNIVERSITÉ de Lyon 1
SPÉCIALITÉ : Biostatistiques

École doctorale : Évolution Écosystèmes Microbiologie Modélisation
Laboratoire de Biométrie et Biologie Évolutive - CNRS UMR 5558

JURY

Mme Nadine BOSSARD (PH)	Codirecteur de thèse
M. Roch GIORGI (PU-PH)	Directeur de thèse
M. Marc MAYNADIÉ (PU-PH)	Examineur
M. Bernard RACHET (DR)	Examineur
Mme Virginie RONDEAU (DR)	Rapporteur
M. Pascal ROY (PU-PH)	Examineur
Mme Catherine THIEBLEMONT (PU-PH)	Rapporteur

Au (excess) hasard de la vie...

UNIVERSITÉ CLAUDE BERNARD - LYON 1

Président de l'Université

Vice-président du Conseil d'Administration
Vice-président du Conseil des Etudes et de la Vie Universitaire
Vice-président du Conseil Scientifique
Directeur Général des Services

M. François-Noël GILLY

M. le Professeur Hamda BEN HADID
M. le Professeur Philippe LALLE
M. le Professeur Germain GILLET
M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard
Faculté de Médecine et de Maïeutique Lyon Sud – Charles
Mérieux
Faculté d'Odontologie
Institut des Sciences Pharmaceutiques et Biologiques
Institut des Sciences et Techniques de la Réadaptation
Département de formation et Centre de Recherche en Biologie
Humaine

Directeur : M. le Professeur J. ETIENNE
Directeur : Mme la Professeure C. BURILLON

Directeur : M. le Professeur D. BOURGEOIS
Directeur : Mme la Professeure C. VINCIGUERRA
Directeur : M. le Professeur Y. MATILLON
Directeur : Mme. la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies
Département Biologie
Département Chimie Biochimie
Département GEP
Département Informatique
Département Mathématiques
Département Mécanique
Département Physique
UFR Sciences et Techniques des Activités Physiques et Sportives
Observatoire des Sciences de l'Univers de Lyon
Polytech Lyon
Ecole Supérieure de Chimie Physique Electronique
Institut Universitaire de Technologie de Lyon 1
Ecole Supérieure du Professorat et de l'Education
Institut de Science Financière et d'Assurances

Directeur : M. F. DE MARCHI
Directeur : M. le Professeur F. FLEURY
Directeur : Mme Caroline FELIX
Directeur : M. Hassan HAMMOURI
Directeur : M. le Professeur S. AKKOUCHE
Directeur : M. le Professeur Georges TOMANOV
Directeur : M. le Professeur H. BEN HADID
Directeur : M. Jean-Claude PLENET
Directeur : M. Y. VANPOULLE
Directeur : M. B. GUIDERDONI
Directeur : M. P. FOURNIER
Directeur : M. G. PIGNAULT
Directeur : M. le Professeur C. VITON
Directeur : M. le Professeur A. MOUGNIOTTE
Directeur : M. N. LEBOISNE

Remerciements

C'est avec beaucoup d'émotion que je souhaite exprimer ma reconnaissance à tous ceux et celles qui, par leur enseignement, leur soutien ou simplement leur présence, m'ont aidée à la réalisation de ce travail de longue haleine.

Je tiens tout d'abord à remercier *M. le Professeur Marc Maynadié*. Il y a maintenant huit ans, vous m'avez accordé votre confiance avec *Mme le Professeur Paule Marie Carli*, en me donnant l'opportunité d'intégrer la petite équipe du registre des Hémopathies malignes de Côte d'Or. Petite de par sa taille mais grande de par son âme. Je vous remercie infiniment de m'avoir ouvert les portes de cette thèse. En espérant être à la hauteur de vos attentes...

Je remercie sincèrement mes directeurs de thèse, *Mme Nadine Bossard* et *M. le Professeur Roch Giorgi*. Je vous suis extrêmement reconnaissante d'avoir accepté de m'encadrer dans ce contexte en triangulaire un peu particulier. J'ai appris beaucoup à vos côtés et en prime, dans une ambiance des plus agréables. Vous avez su ne pas me tenir rigueur de mes quelques faux-pas et avez réussi à me canaliser durant ces années en m'apprenant, entre autres, les « bonnes manières » de la recherche publique. Veuillez être assurés de mon profond respect.

J'exprime toute ma gratitude à *Mme Virginie Rondeau* et à *Mme le Professeur Catherine Thieblemont*. Vous m'avez fait l'honneur d'être rapporteurs de cette thèse. Vos points de vue à la fois de statisticienne et de clinicienne, ont permis d'améliorer l'ensemble de ce manuscrit.

Je suis heureuse de compter parmi mon jury *M. Bernard Rachet* et *M. le Professeur Pascal Roy*. Je vous remercie de m'avoir fait l'honneur de juger mon travail de thèse. Je tiens également à remercier *le Professeur Roy* de m'avoir accueillie au sein de son service de biostatistique dans le cadre de ce travail.

Je tiens à remercier *M. Marc Colonna* et *M. Alain Monnereau* d'avoir suivi l'avancée de ce travail dans le cadre de mon comité de suivi de thèse. Vos conseils avisés m'ont été précieux.

Je souhaite exprimer toute mon estime et ma reconnaissance à *M. Laurent Remontet* et à *M. Aurélien Belot*. Vous avez eu le mérite de répondre à toutes mes questions, y compris les plus farfelues et tout cela sans porter aucun jugement à mon égard (ou alors avec

discrétion...). Je n'ai jamais réussi à vous coincer sur aucune de mes questions...ggrrr je les aurai un jour ! Je vous remercie pour ce transfert de savoir, promis j'en ferai bon usage. Aurais-je droit au service après-vente après la thèse ? Merci pour votre bonne humeur et pour avoir été bon public à toutes mes excentricités...Merci, Merci, Merci !

Je remercie chaleureusement *M. Jean Iwaz*. Thanks for your patience and your availability. I was very happy to work with you. A bit magician, you accepted impossible missions like to reduce by two the number of words in our papers.

Merci également à toute l'équipe du service de biostatistique des Hospices Civils de Lyon et à l'équipe du SESSTIM de Marseille. Vous m'avez accueillie avec beaucoup de gentillesse à chacune de mes visites.

Je souhaite également remercier :

Les membres du groupe MESURE/CENSUR. J'ai appris beaucoup durant nos réunions. *Amel, Coraline, Isabelle, Nathalie* et tous les membres du groupe *YROS*, je vous remercie pour nos échanges et votre soutien durant cette dernière ligne droite. Je clôture la saison.

Le réseau FRANCIM et le groupe européen EUROCARE, qui ont accepté nos deux études collaboratives.

Tous les collègues des différentes équipes de la Faculté de médecine de Dijon. *Valérie*, à mes débuts au registre tu as été là pour moi et je t'en remercie !

Les membres des registres de FRANCIM, pour nos échanges constructifs et pour les bons moments passés durant les congrès.

Les membres du laboratoire d'hématologie de Dijon, les hématologues, les étudiants de médecine et de pharmacie qui m'apprennent des mots toujours plus savants.

Quid d'une thèse sans soupe de sécurité ? Je souhaite exprimer toute mon amitié à :

Mes camarades de jeu au registre, Stéphanie et Sophie. Merci les filles pour votre compréhension infinie et pour m'avoir préservée durant ces années. Entre confiance et respect, je n'aurais pu imaginer meilleure équipe. *Alix et Francis*, nos petits nouveaux, j'ai dû vous paraître bizarre mais ce n'est rien c'est l'effet thèse, je redeviens normale dans très peu de temps... *Aurélie, Inès et Mylène*, travailler à vos côtés fut un immense plaisir. *Mylène* je te remercie pour ton soutien, ton coaching toujours aussi efficace et pour ton amitié fidèle.

Nos copains du village, pour ces soirées mémorables ! Cela va être difficile d'expliquer plus tard à nos enfants que l'abus d'alcool est dangereux pour la santé !

Mes amis de fac de Grenoble et Bordeaux. Une thèse... moi ? Qui l'aurait cru ?

Mes amis de presque toujours, Mélanie & Fabien, Christelle & Johan, Stéphanie & Marc, Carmen & Fanfan. Ma famille bourguignonne ! Je déclare être déchargée de toute responsabilité en cas de dérapage incontrôlé de leur part durant ma soutenance.

Mes amies d'enfance, mes racines, Céline, Clothilde, Justine, Maud. Merci les filles d'être toujours là, dans les bons comme dans les mauvais moments.

Parce que trop rarement exprimé chez les Mounier, je profite de ces quelques lignes pour témoigner de tout l'amour que je porte à ma famille :

A mes parents, Gilles & Dominique, qui m'ont apporté des bases solides me permettant d'en arriver là.

A mon frère, Arnaud, et ma sœur, Lorraine, ainsi que leur conjoint, *Valérie et Benoit.* Vous avez toujours pris soin de moi, la petite dernière, et je vous en remercie.

A mes neveux et nièces, Amory, Eléa, Flavie, Lucie, et Maxime. Tatie Momo vous remercie pour tous vos câlins et votre joie de vivre.

A mes oncles, tantes, cousins et cousines, Jean-François, Lydie & Gérard, Nicole & Henri, Anne-Sophie, Guillaume, Sébastien et Elodie. Merci de m'avoir hébergée pendant mes déplacements sur Lyon et pour votre soutien. Je sais que si j'ai besoin, vous êtes là pour moi.

A mon grand-père, Roger.

A ma cousine Flo, partie trop tôt et *à mes grands-mères, Lolo et Louise,* qui auraient été fières, j'en suis sûre.

A ma belle-famille. Claudine & Olivier, Christelle & Firmin. Vous m'avez accueillie dans la famille avec beaucoup de gentillesse et en toute simplicité, je vous en remercie.

Je terminerai en dédiant ce travail de thèse à mon essentiel :

Mes deux princesses, Marlène et Cassandra. Sans véritablement le savoir, vous m'avez apporté toute l'affection et le réconfort dont j'ai eu besoin durant ces années de thèse : des petits tracas du quotidien balayés par un sourire, un éclat de rire ou encore une douce « phrase magique » murmurée au creux de mon oreille chaque soir.

Grégory, « mon homme sweet homme ». Le plus heureux de tous que cette thèse se termine ! Merci d'avoir été, d'être, et de toujours rester à mes côtés. Tu m'as montré un autre regard sur les choses et sur la vie. Pour cela et pour tout le reste, je te suis extrêmement reconnaissante. Bien rares se font les êtres de ton envergure. Etant une espèce en voie d'extinction, tu as besoin qu'à mon tour, je prenne soin de toi. Nous allons avoir plus de temps désormais. Ah ! J'allais oublier... et si tu sens venir l'envie... ma réponse est Oui !

Table des matières

<i>Valorisation scientifique</i>	<i>vii</i>
<i>Table des illustrations</i>	<i>x</i>
<i>Abréviations</i>	<i>xii</i>
<i>Introduction générale</i>	<i>1</i>
<hr/>	
Chapitre I : Le contexte : l'épidémiologie descriptive du lymphome malin non hodgkinien	6
<hr/>	
I.1. Les lymphomes malins non hodgkiniens : l'essentiel	7
I.1.1. Aspects cliniques des LMNH	8
I.1.1.1. Diagnostic	9
I.1.1.2. Prise en charge thérapeutique.....	10
I.1.2. Évolution des classifications	11
I.1.3. Épidémiologie des LMNH	13
I.1.3.1. Incidence	13
I.1.3.2. Pronostic.....	15
I.2. Mesures des progrès en population générale	16
I.2.1. Progrès dans la prise en charge des LMNH	16
I.2.2. Données de population	19
I.2.2.1. Essais cliniques versus données de population.....	19
I.2.2.2. Les registres de cancer : source des données de population	20
I.2.2.2.1. Historique	20
I.2.2.2.2. Bases collaboratives de données de population	21
I.2.3. Discussion	22
<hr/>	
Chapitre II : Méthodes statistiques pour l'étude de la survie des patients en population générale	24
<hr/>	
II.1. Pré-requis	25
II.1.1. Notion de données censurées	26
II.1.2. Les fonctions de base en analyse de survie	27
II.1.3. Les concepts de survie.....	30
II.1.3.1. La survie toutes causes	30
II.1.3.2. La survie nette	31
II.2. Les approches d'estimation de la survie nette	31
II.2.1. Usage de la cause de décès.....	32
II.2.2. Non usage de la cause de décès.....	33
II.2.2.1. Utilisation de la mortalité attendue.....	33
II.2.2.2. Estimation de la mortalité due au cancer.....	34
II.2.3. Domaines d'application	36
II.3. Approche non paramétrique pour estimer la survie nette	36
II.3.1. Contexte	37
II.3.2. Apport de l'estimateur non paramétrique proposé par Pohar-Perme <i>et al.</i>	38

II.3.3.	Discussion	40
II.4.	Modélisation paramétrique de la survie nette	41
II.4.1.	Contexte	41
II.4.2.	Le modèle flexible du taux en excès proposé par Remontet <i>et al.</i>	42
II.4.2.1.	Écriture du modèle	43
II.4.2.2.	Modélisation de l'effet complexe des covariables.....	45
II.4.2.2.1	Modélisation des effets non linéaires.....	45
II.4.2.2.2	Modélisation des effets dépendants du temps.....	46
II.4.2.2.3	Modélisation conjointe des effets non linéaires et dépendants du temps.....	48
II.4.2.3.	La stratégie de construction du modèle flexible	49
II.4.3.	Discussion.....	52
<hr/>		
Chapitre III : Application 1 : Étude des tendances temporelles sur données françaises.....		54
<hr/>		
III.1.	Rationnel	55
III.2.	Données de survie en population française	56
III.3.	Méthodologie	56
III.4.	Article	58
III.5.	Résultats complémentaires	71
III.6.	Discussion.....	75
<hr/>		
Chapitre IV : Application 2 : Tendances et comparaisons internationales.....		77
<hr/>		
IV.1.	Rationnel	78
IV.2.	Données de survie en population européenne.....	79
IV.3.	Méthodologie	80
IV.4.	Article	83
IV.5.	Résultats complémentaires	120
IV.6.	Discussion.....	122
<hr/>		
Chapitre V : Conclusion & Perspectives		125
<hr/>		
V.1.	Conclusion	126
V.2.	Perspectives	130
V.2.1.	Évolution du cadre méthodologique de la survie nette.....	130
V.2.2.	Étude du pronostic des LMNH.....	132
<hr/>		
Annexes.....		135
<hr/>		
Références bibliographiques.....		148

Valorisation scientifique

Publications

- **Mounier M.** Bossard N, Belot A, Remontet L, Iwaz J, Dandoit M, Girard-Boulanger S, Herry A, Woronoff AS, Casasnovas RO, Maynadié M, Giorgi R; FRANCIM Network and MESURE Working Survival Group. *Trends in excess mortality in follicular lymphoma at a population level.* Eur J Haematol. février 2015;94(2):120-9.
- **Mounier M.** Bossard N, Remontet L, Belot A, Minicozzi P, De Angelis R, Capocaccia R, Iwaz J, Monnereau A, Troussard X, Sant M, Maynadié M, Giorgi R; EURO CARE-5 Working group and CENSUR Working survival group. *Changes in the dynamics of the excess mortality rates after diagnosis of follicular or diffuse large B-cell lymphoma: comparisons between European population-based data.* (Article en cours de publication dans la revue *The Lancet Haematology*).
- Sant M, Minicozzi P, **Mounier M.** Anderson LA, Brenner H, Holleczeck B, Marcos-Gragera R, Maynadié M, Monnereau A, Osca-Gelis G, Visser O, De Angelis R; EURO CARE-5 Working Group. *Survival for haematological malignancies in Europe between 1997 and 2008 by region and age: results of EURO CARE-5, a population-based study.* Lancet Oncol. août 2014;15(9):931-42.

Contribution à d'autres publications dans le domaine :

- Dandoit M, **Mounier M.** Guy J, Petrella T, Girard S, Casasnovas RO, Martin L, Bonnetain F, Maynadié M. *The heterogeneity of changes in incidence and survival among lymphoid malignancies in a 30-year French population-based registry.* Leuk Lymphoma. avril 2015;56(4):1050-7.
- Rossi C, Jégu J, **Mounier M.** Dandoit M, Colonna M, Daubisse-Marliac L, Trétarre B, Ganry O, Guizard A-V, Bara S, Bouvier V, Woronoff A-S, Monnereau A, Casasnovas O,

Velten M, Troussard X, Maynadié M. *Risk assessment of second primary cancer according to histological subtype of non-Hodgkin lymphoma*. *Leuk Lymphoma*. mars 2015;1-7.

Communications orales

- **Mounier M**, Bossard N, Remontet L, Belot A, Minicozzi P, De Angelis R, Capocaccia R, Iwaz J, Monnereau A, Troussard X, Sant M, Maynadié M, Giorgi R; FRANCIM Network and CENSUR Working Survival Group. *Net survival and excess mortality rate after follicular or diffuse large B-cell lymphoma diagnosis in Europe: trends and dynamics by age, year and area*. - Groupe des registres de langue latine. Reus (Espagne, 2015).
- **Mounier M**. *Question de survie*. Atelier ACADM – Groupe des Data managers académiques. Dijon (France, 2014).
- **Mounier M**. *Excess mortality rate: a useful indicator in cancer epidemiology: practical use*. Workshop GRELL- Groupe des registres de langue latine. Genève (Suisse, 2014).
- **Mounier M**, Bossard N, Belot A, Remontet L, Dandoit M, Girard S, Herry A, Casasnovas O, Maynadié M, Giorgi R for FRANCIM group. *Changes in the therapeutic management of follicular lymphoma: Which impact on the excess mortality risk in the general population?* - Groupe des registres de langue latine. Siracuse (Sicile, 2013).
 - ➔ Prix Enrico Anglesio décerné par la Fondo Anglesio Moroni per l’Oncologia, Torino, Italy.

Communications affichées

- **Mounier M**, Bossard N, Remontet L, Belot A, Minicozzi P, De Angelis R, Capocaccia R, Iwaz J, Monnereau A, Troussard X, Sant M, Maynadié M, Giorgi R; EURO CARE-5 Working group and CENSUR Working survival group. *Trends in net survival and dynamics of the excess mortality rate after diagnosis of follicular or diffuse large B-cell*

lymphoma in five European areas. European Hematology Association. Vienne (Autriche, 2015).

- **Mounier M.** Bossard N, Belot A, Remontet L, Dandoit M, Girard S, Herry A, Casasnovas O, Maynadié M, Giorgi R. *Evolution de la prise en charge thérapeutique des lymphomes malins non hodgkiniens : quel impact sur la mortalité en excès des patients en population générale en France ?* Société Française d'Hématologie. Paris (France, 2013).

- **Mounier M.** Le Guyader-Peyrou S, Dufour V, Collignon A, Giorgi R, Bossard N, Belot A, Remontet L, Woronoff A-S, Monnereau A, Troussard X, Maynadié M pour le réseau REPIH. *Evolution de la prise en charge thérapeutique dans les hémopathies malignes: Quel impact sur la survie des patients en population générale ?* Colloque registre INCA. Paris (France, 2012).

Table des illustrations

Figure I.1 - Classification schématique des hémopathies malignes.....	9
Figure I.2 - Evolutions majeures des classifications utilisées en hématologie et en épidémiologie.....	13
Figure I.3 - Taux d'incidence standardisé pour l'âge (à la population mondiale) en France en 2012, par sous-type de LMNH. Source : FRANCIM (Monnereau et al. 2013a).....	14
Figure I.4 - Probabilité de survie nette (et IC _{95%}) à 5 ans, standardisée sur l'âge en France en 2012, par sous-type de LMNH et par sexe. Source : FRANCIM (Monnereau et al. 2013b).....	15
Figure I.5 - Interventions dans la lutte contre les LMNH et mesures épidémiologiques associées.....	18
Figure II.1 - Probabilité de survie et probabilité de décès en fonction du temps de suivi.....	28
Figure II.2 - Fonction de risque instantané et de risque cumulé en fonction du temps de suivi.....	29
Figure II.3 - Effet linéaire (A) ou non linéaire (B) de l'âge sur le taux relatif de mortalité en excès, l'âge de référence étant fixé à l'âge moyen (60 ans).....	45
Figure II.4 - Effet proportionnel (A) ou non proportionnel (B) de l'âge sur le taux relatif de mortalité pour des patients de 30 ans par rapport à l'âge de référence, 60 ans.....	47
Figure II.5 - Effet non-linéaire et non proportionnel de l'âge sur le taux relatif de mortalité à la suite d'un diagnostic de LF, l'âge de référence étant fixé à l'âge moyen 60 ans.....	48
Figure II.6 - Stratégie de modélisation des effets complexes d'une variable continue (l'âge par exemple) proposée par Abrahamowicz & MacKenzie, en considérant le modèle du taux en excès de Remontet et al.	50
Figure III.1 - Probabilité de survie nette (et IC _{95%}) à 5 ans standardisée sur l'âge à la suite d'un diagnostic de LF, en fonction de la période de diagnostic et du sexe.....	72

Figure III.2 - Probabilité de survie nette à 5 ans (et IC _{95%}) par classe d'âge, à la suite d'un diagnostic de LF en fonction de la période de diagnostic.	73
Figure III.3 - Taux de base de mortalité en excès (sans covariables) en fonction du temps de suivi dont l'effet est modélisé par une fonction en escalier ou par une fonction spline cubique (avec un nœud à 1 an de suivi).(A) Résultat de la cohorte 1995-2004, (B) Résultat de la cohorte 1995-2010.	74
Figure III.4 - Ratio des taux de mortalité en fonction de l'année de diagnostic dont l'effet est modélisé par une fonction en escalier ou par une fonction linéaire (année de référence =1998). (A) Résultat de la cohorte 1995-2004 (ajusté sur un effet NL-NP de l'âge, période de référence 1997-98 pour la fonction en escalier). (B) Résultat de la cohorte 1995-2010 (ajusté sur un effet NL-NP de l'âge et sur un effet PP du stade, période de référence 1998-2000 pour la fonction en escalier).	74
Figure IV.1 - Ratio des taux de mortalité en excès en fonction de l'année de diagnostic, issu du modèle retenu par zone et par sous-type de lymphome. (Pour les LBDGC, le modèle retenu supposait un effet NP de l'année en Ecosse/Pays de Galles et en Europe Centrale. Aussi pour ces deux zones, le ratio est présenté à 1 an et 5 ans après le diagnostic)	121

Abréviations

AMM	Autorisation de Mise sur le Marché
ANCR	Association of the Nordic Cancer Registries
CENSUR	Challenges in the Estimation of Net SURvival
CER	Centre d'Évaluation des Registres
CHOP	Cyclophosphamide, Hydroxyadriamycine, Oncovin et Prédnisone
CIM	Classification Internationale des Maladies
CIM-O-3	Classification Internationale des Maladies en Oncologie, troisième édition
CIRC	Centre International de Recherche sur le Cancer
EUROCARE	EUROpean Cancer REgistry based study on survival and care of cancer patients
FLIPI2	Follicular Lymphoma International Prognostic Index 2
FRANCIM	FRAnce-Cancer-Incidence et Mortalité
GELA	Groupe d'Etude des Lymphomes de l'Adulte
HAEMACARE	CAnceR REgistry based Project on HAEmatologic MAlignancies
HCL	Hospices Civils de Lyon
IACR	International Association of Cancer Registries
IC _{95%}	Intervalle de Confiance à 95%
INCA	Institut National du CAncer
INSEE	Institut National de la Statistique et des Études Économiques
INTERLYMPH	INTERnational LYMPHoma epidemiology consortium
InVS	Institut national de Veille Sanitaire
IPI	Index Pronostique International
ICSS	International Cancer Survival Standard
LBDGC	Lymphome B Diffus à Grandes Cellules

LDH	Lactates DésHydrogénases
LF	Lymphome Folliculaire
LH	Lymphome Hodgkinien
LL	Log-Linéaire
LMNH	Lymphome Malin Non Hodgkinien
LYSA	Lymphoma Study Association
MESURE	Méthodes d'Estimation de la SURvie RELative
NL	Non log-Linéaire
NORDCAN	NORDic CANcer
NP	Non Proportionnel
OMS	Organisation Mondiale de la Santé
PP	Proportionnel
PRIMA	PRImary Rituximab and Maintenance
R-CHOP	Rituximab, Cyclophosphamide, Hydroxyadriamycine, Oncovin et Prédnisone
REAL classification	Revised European American Lymphoma classification
RELEVANCE	Rituximab LEnalidomide Versus ANy Chemotherapy
RNIPP	Répertoire National d'Identification des Personnes Physiques
SEER	Surveillance, Epidemiology, and End Results
SUDCAN	SUD CANcer : Survival in patients with cancer in European Latina countries
TEP	Tomographie par Émission de Positons



Introduction générale

La structuration de la lutte contre le cancer, première cause de mortalité en France (Binder-Foucard *et al.* 2014), s'est renforcée dans notre pays à partir de 2003 avec la mise en place d'un premier plan cancer qui s'est inscrit dans la politique de santé publique nationale et qui a permis de hiérarchiser les actions à mener. La France connaît actuellement son troisième plan cancer lancé en 2014 et qui s'étend jusqu'en 2019 (Plan Cancer 2015). Discipline clé de la santé publique, l'épidémiologie permet aux institutions sanitaires, représentées en première ligne par l'Institut de Veille Sanitaire (InVS) et l'Institut National du CAncer (INCA), de disposer d'indicateurs scientifiques du cancer en population générale. La surveillance épidémiologique des cancers est rendue possible grâce à l'activité des registres de cancer en population. Ces études observationnelles en population se différencient des études menées dans le cadre d'essais cliniques ou sur séries hospitalières, ces dernières se basant sur une sélection des cas à l'inclusion empêchant ainsi la généralisation des résultats à l'ensemble du collectif.

L'étude de la survie des patients en population générale est l'indicateur le plus pertinent pour apprécier l'efficacité globale du système de soins d'un pays. Afin de coordonner les efforts en matière de lutte contre le cancer, des bases collaboratives de données de population au niveau national avec le réseau FRANCIM (FRAnce Cancer Incidence et Mortalité) ou européen avec EUROCare (EUROpean CAncer REgistry) permettent de disposer d'indicateurs de survie à plus grande échelle et offrent des comparaisons géographiques entre pays ou des comparaisons temporelles en fonction des périodes de diagnostic. Le concept de survie nette apparaît comme central car il permet de distinguer la mortalité liée au cancer à la mortalité due aux autres causes. En effet, la survie nette se définit comme la survie qui serait observée si la seule cause de décès possible était le cancer. On ne s'intéresse donc dans ce cadre qu'à la mortalité induite par le cancer. Ce concept de survie nette sera défini de façon précise dès 1950 par Berkson et Gage (Berkson et Gage 1950). Deux alternatives s'offrent alors pour ne considérer que la mortalité induite par la maladie : l'approche cause spécifique qui nécessite la connaissance de la cause de mortalité principale de l'individu et l'approche basée sur l'utilisation de tables de mortalité en population générale.

Les premiers outils statistiques de la survie nette ont été développés sous forme non paramétrique afin de fournir des estimations ponctuelles. Mais devant les besoins cliniques et épidémiologiques d'étudier les effets de certaines variables sur la mortalité induite par la maladie, des outils paramétriques ont ensuite été proposés (Breslow 1975 ; Buckley 1984; Hakulinen et Tenkanen 1987 ; Andersen et Vaeth 1989 ; Estève *et al.* 1990 ; Dickman *et al.* 2004).

Depuis les années 2000, de nouveaux développements méthodologiques ont permis d'améliorer les outils existants d'une part en corrigeant le biais induit par la censure dépendante due aux autres causes de décès dans le cadre des estimateurs non paramétriques (Perme *et al.* 2012) et d'autre part en levant les hypothèses sous-jacentes à l'écriture des modèles paramétriques (Giorgi *et al.* 2003 ; Nelson *et al.* 2007 ; Remontet *et al.* 2007 ; Mahboubi *et al.* 2011). Ces avancées méthodologiques marquent un tournant dans les outils disponibles en analyse de survie. Afin de faire encore évoluer le cadre méthodologique de l'estimation de la survie nette, un groupe de travail international réunissant biostatisticiens et épidémiologistes a vu le jour en 2009 sous la coordination du Professeur Roch Giorgi (projet MESURE : Méthodes d'estimation de la SURvie Relative, ANR 09-BLAN-0357 ; 10/2009 – 04/2013) et se poursuit actuellement avec un deuxième projet nommé CENSUR : Challenges in the Estimation of Net SURvival (ANR-12-BSV1-0028). C'est dans le cadre de ces deux projets méthodologiques basés sur l'amélioration des outils pour l'estimation de la survie nette que s'intègre ce travail de thèse.

Relativement peu connus du grand public, les Lymphomes Malins Non Hodgkiniens (LMNH) représentent les hémopathies les plus répandues et restent un enjeu de santé publique de par l'accroissement de leur taux d'incidence jusque dans les années 2000 (Belot *et al.* 2008) et par l'insuffisance des connaissances sur leurs facteurs étiologiques. Des grands groupes collaboratifs en hématologie, tels que HAEMACARE ou INTERLYMPH ont fait évoluer les connaissances épidémiologiques de ces lymphomes en mettant en évidence l'importance de dissocier les sous-types de LMNH dans toutes les études. Ainsi, les classifications en hématologie n'ont cessé de se préciser permettant une meilleure définition des différentes entités dans les évolutions successives des classifications internationales. Les progrès des actions dans la lutte contre le cancer (prévention, meilleur accès aux soins, etc.) et le développement de nouveaux outils diagnostiques et de la thérapie ciblée (anticorps

monoclonaux anti-CD20) ont permis de faire considérablement évoluer la prise en charge des LMNH.

Des essais cliniques d'ampleur mondiale ont montré des résultats positifs sur la survie toutes causes des patients avec l'utilisation de cette nouvelle thérapeutique (Coiffier *et al.* 1998 ; Salles *et al.* 2008). Bien que seules les études cliniques permettent de tester l'efficacité d'une nouvelle thérapeutique, il apparaît cependant utile de compléter ces résultats par des résultats issus de bases de données populationnelles afin de déceler l'impact réel de ces avancées sur l'ensemble des sujets atteints de LMNH. Dans ce travail, nous nous sommes orientés sur l'étude de la survie à la suite d'un diagnostic de Lymphome Folliculaire (LF) ou de Lymphome B Diffus à Grandes Cellules (LBDGC). Ayant bénéficié en premier lieu des récentes innovations, ces deux sous-types servent de référentiel dans la prise en charge des lymphomes indolents pour les LF et des lymphomes plus agressifs pour les LBDGC.

Notre travail se propose d'illustrer l'apport des nouvelles méthodes d'estimation de la survie nette dans le pronostic des LMNH en population générale, dès lors qu'il s'agit de produire des comparaisons dans le temps ou entre pays. Il repose essentiellement sur l'approche paramétrique avec la modélisation du taux en excès. La nécessité de prendre en compte les effets du temps, de l'année de diagnostic et de l'âge sur la mortalité en excès dans le cadre de telles comparaisons, a guidé l'essentiel de notre travail. Par ailleurs, nous avons également souhaité restituer les effets de ces variables, non seulement sur la survie nette, mais aussi sur le taux de mortalité en excès représentant ainsi un élément nouveau dans la panoplie des indicateurs en survie.

Le Chapitre I de ce rapport est consacré au cadre de ce travail en rappelant d'une part les aspects cliniques et épidémiologiques des LMNH et d'autre part les mesures des progrès dans la prise en charge des patients en population générale en rappelant le rôle des registres de population dans l'observation du cancer.

Après un bref rappel sur les fonctions de base et concepts fondamentaux de l'analyse de survie, nous exposons dans le Chapitre II deux approches fondées sur l'estimation de la survie nette. La première approche correspond à une modélisation non paramétrique proposée par Pohar-Perme *et al.* (Perme *et al.* 2012) qui apparaît actuellement comme le seul estimateur non paramétrique qui en théorie fournit des estimations non biaisées de la survie nette. La deuxième approche se base sur une modélisation paramétrique du taux en excès proposée par

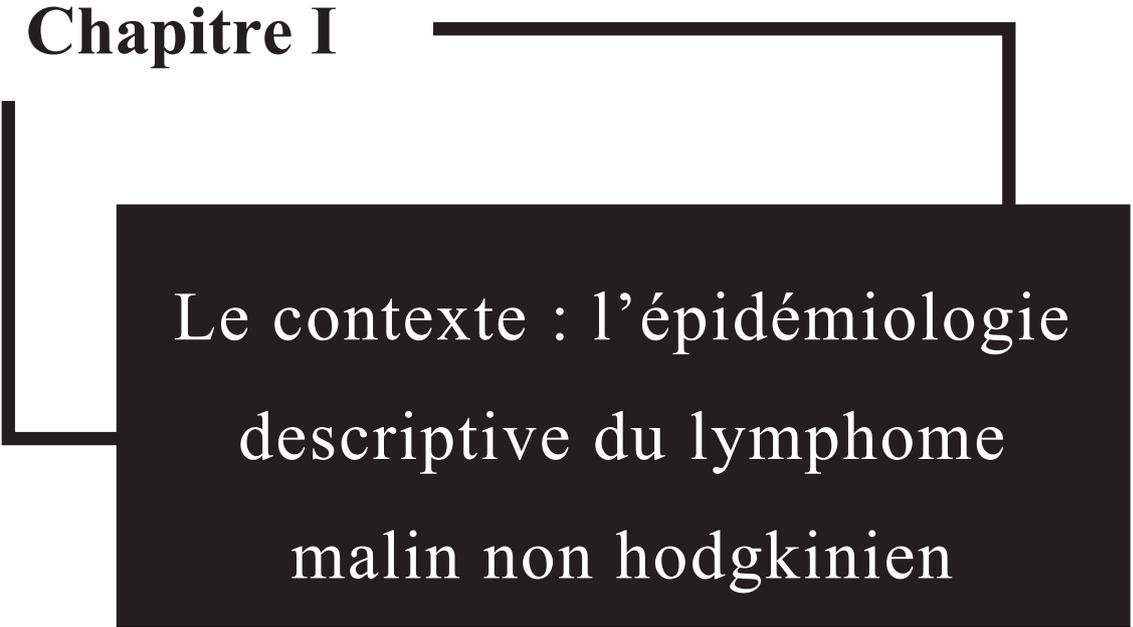
Remontet *et al.* (Remontet *et al.* 2007) qui permet de tenir compte des effets complexes des covariables sur la mortalité en excès. La stratégie de modélisation des effets complexes des variables est par ailleurs discutée en fin de chapitre.

Une application de ces méthodes sur les données du réseau des registres de cancer français FRANCIM est proposée au Chapitre III et porte sur l'étude des changements dans la dynamique du taux de mortalité en excès à la suite d'un diagnostic de LF entre le 01 janvier 1995 et le 31 décembre 2010 en fonction de l'âge, de l'année de diagnostic et du stade du lymphome.

Une deuxième application de ces outils actuels porte sur la base de population européenne EUROCORE-5 et fait l'objet du Chapitre IV en exposant les tendances de la dynamique du taux de mortalité en excès et de la survie nette après un diagnostic de LF ou de LBDGC dans cinq zones européennes en fonction de l'âge et de l'année de diagnostic entre le 01 janvier 1996 et le 31 décembre 2004.

Enfin, ce rapport s'achève au Chapitre V par les conclusions et les perspectives de ce travail.

Chapitre I



Le contexte : l'épidémiologie
descriptive du lymphome
malin non hodgkinien

Les hémopathies malignes qui regroupent les maladies malignes du sang, de la moelle osseuse et des organes lymphoïdes se situent au sixième rang des cancers en France avec 35 000 nouveaux cas estimés en 2012 (Monnereau *et al.* 2013a). Les LMNH sont des proliférations de lignées lymphoïdes et représentent les hémopathies les plus fréquentes. Afin de tenir compte des avancées dans les connaissances sur les lymphomes, les classifications internationales des maladies se sont précisées en différenciant désormais les différents types de LMNH, ce qui a également eu pour effet d'homogénéiser les données entre pays (Campo *et al.* 2011).

Des collaborations internationales en hématologie, telles que HAEMACARE ou le consortium INTERLYMPH ont mis en évidence l'importance d'étudier séparément les différentes entités de LMNH du fait de caractéristiques extrêmement diverses à la fois sur le plan physiopathologique, clinique, thérapeutique ou épidémiologique (Morton *et al.* 2007 ; Sant *et al.* 2010). Notamment, le groupe HAEMACARE a été le premier à publier des résultats de survie par sous-type de LMNH en Europe (Marcos-Gragera *et al.* 2011).

Parallèlement à une meilleure connaissance de la maladie due à l'amélioration des outils diagnostiques, la prise en charge des LMNH a connu des avancées majeures dans le traitement des patients avec entre autres le développement des anticorps monoclonaux anti-CD20 dans le traitement des lymphomes B (Coiffier 2006). L'utilisation de nouveaux traitements couplés à d'autres améliorations dans la prise en charge des LMNH, ont permis d'améliorer le pronostic des patients. Nous verrons que des mesures épidémiologiques en population générale, telles que l'incidence, la mortalité ou la survie, permettent d'évaluer la variation de la qualité globale des soins.

I.1. LES LYMPHOMES MALINS NON HODGKINIENS : L'ESSENTIEL

Le terme « hémopathie maligne » regroupe un nombre conséquent d'entités se distinguant à la fois par leurs caractéristiques cliniques et par leurs pronostics. Les hémopathies myéloïdes et les hémopathies lymphoïdes sont les proliférations développées à partir des deux lignées hématopoïétiques. Chaque lignée peut être à l'origine de plusieurs

dizaines de sous-types d'hémopathies. Les LMNH sont des hémopathies de lignées lymphoïdes.

I.1.1. ASPECTS CLINIQUES DES LMNH

Le système immunitaire, qui regroupe l'ensemble des cellules, tissus et organes qui assurent la défense de l'organisme contre les agents extérieurs, s'appuie sur le système lymphatique. Ce dernier, en charge de produire, stocker et transporter les lymphocytes, est constitué à la fois des ganglions lymphatiques et des organes lymphoïdes (tels que la rate, la moelle osseuse, les amygdales et le thymus) entre lesquels circule la lymphe via un réseau de vaisseaux microscopiques. Les lymphomes sont des cancers du système lymphoïde. Ils sont dus à une multiplication incontrôlée de lymphocytes anormaux ; les trois types de lymphocytes B, T ou NK pouvant être touchés. L'accumulation de cellules anormales conduit à la formation de masses tumorales à différents endroits du système lymphatique, le plus souvent au niveau des ganglions. Les lymphocytes circulant dans tout l'organisme, le lymphome peut donc potentiellement toucher tous les organes.

Nous distinguons deux grands types de lymphomes : *les Lymphomes Hodgkiniens (LH) et les LMNH*. Ces derniers sont les plus fréquents et représentent plus de 65% de l'ensemble des hémopathies malignes ; cependant ils restent mal connus du grand public. L'analyse microscopique permet de différencier le type de lymphomes, les LH se caractérisant par la présence de la cellule Reed Sternberg (Hodgkin 1832; Wilks 1856). Les LMNH forment quant à eux un vaste groupe d'environ 80 entités qui sont classées en deux sous-types principaux (Morton *et al.* 2007) :

- *Les lymphomes de type B* se développent à partir des lymphocytes B. Ils représentent environ 85% des cas de LMNH (les lymphomes B diffus à grandes cellules, les lymphomes folliculaires, les lymphomes de la zone marginale, le lymphome du manteau, etc.).
- *Les lymphomes de type T* se développent à partir des lymphocytes T. Ils représentent environ 15% des LMNH en particulier dans le monde occidental (les lymphomes T cutanés et non cutanés, le lymphome T périphérique, etc.).

Les deux lignées principales ainsi que les sous-types majeurs des LMNH sont rappelés sur la Figure I.1.

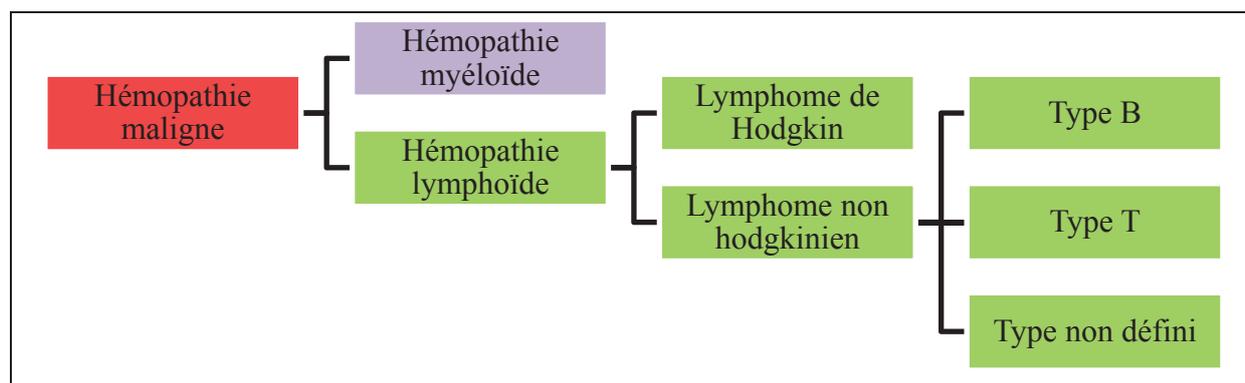


Figure I.1 - Classification schématique des hémopathies malignes.

I.1.1.1. Diagnostic

La variété des symptômes et leur faible spécificité (fièvre, sueurs nocturnes, perte de poids, augmentation de la taille des ganglions) rendent le diagnostic de lymphome délicat et dans une part non négligeable de cas, la découverte du lymphome reste fortuite. Le diagnostic d'un LMNH repose sur un examen clinique et sur une numération de formule sanguine complétée d'un immunophénotypage en cas d'hyperlymphocytose ou de présence de formes lymphoïdes anormales sur le frottis (Friedberg *et al.* 2008). Ce premier bilan est le plus souvent accompagné d'un *examen histologique en anatomopathologie* d'un échantillon de tissu prélevé sur un ganglion. Cet examen morphologique va permettre de préciser l'architecture du lymphome et la taille des cellules tumorales. Un *examen cytogénétique et moléculaire* permet ensuite de rechercher des anomalies acquises, clonales et non aléatoires. A ce stade, le diagnostic de lymphome est avéré et son type histologique connu.

Un *bilan d'extension* permet ensuite d'évaluer l'extension et l'évolutivité du lymphome. Ce bilan d'extension repose sur un bilan topographique avec imagerie (scanner, imagerie par résonance magnétique, Tomographie par Émission de Positons (TEP)) et permet de déterminer le nombre de sites atteints par le lymphome. D'autres examens sont aussi prescrits afin de mieux cibler l'extension du lymphome (myélogramme, bilan hépatique, etc.). Au terme du bilan d'extension, pourront être définis le type morphologique du lymphome et le stade qui définit son degré d'extension. Le stade du lymphome est calculé sur le nombre de localisations de la maladie et se définit par la classification Ann-Arbor en quatre catégories (stade I à IV, le stade IV correspondant au stade le plus avancé) (Armitage 2005). La notion

de grade, si elle a disparu des classifications, reste utilisée pour distinguer *les lymphomes de bas grade* (ou indolents) qui connaissent une évolution lente avec un risque de rechute élevé et *les lymphomes de haut grade* (ou agressifs) dont l'évolutivité est plus agressive.

I.1.1.2. Prise en charge thérapeutique

La prise en charge thérapeutique d'un LMNH est variable en fonction du sous-type (Zelenetz *et al.* 2014). Elle peut aller d'une simple surveillance avec abstention thérapeutique dans le cas de lymphomes de bas grade n'évoluant pas ou de faibles masses tumorales, à l'utilisation d'une polychimiothérapie en association à d'autres outils thérapeutiques tels que l'immunothérapie, la radiothérapie ou des greffes dans le cadre de lymphomes plus agressifs (Abramson et Zelenetz 2013). Le nombre de cures et l'intensité du traitement vont dépendre non seulement de la nature agressive du lymphome mais aussi d'autres facteurs tels que l'âge, le bilan d'extension et de facteurs pronostiques spécifiques.

Afin de guider le clinicien dans ses décisions thérapeutiques, des index pronostiques ont été établis. Pour les lymphomes agressifs, l'Index Pronostique International (IPI) (The International Non-Hodgkin's Lymphoma Prognostic Factors Project 1993) se base sur l'association de cinq facteurs cliniques au diagnostic : l'âge (supérieur à 60 ans), le stade clinique (avancé III ou IV), un index de performance (égal ou supérieur à 2), le taux de Lactate DésHydrogénase (LDH) (élevé), l'atteinte d'au moins deux sites extra nodaux. Des index pronostiques ont également été publiés pour les lymphomes indolents comme le Follicular Lymphoma International Prognostic Index 2 (FLIPI2) (Federico *et al.* 2009).

La prise en charge des LMNH a considérablement évolué ces vingt dernières années. Alors que jusqu'à la fin des années 1990, le traitement consistait essentiellement en une polychimiothérapie basée sur le régime Cyclophosphamide, Hydroxyadriamycine, Oncovin et Prédnisone (CHOP), le développement de la thérapie ciblée avec les anticorps monoclonaux anti-CD20 a marqué un tournant majeur dans la prise en charge thérapeutique des lymphomes B. L'utilisation de cette nouvelle thérapie s'est faite de façon progressive.

Ce sont tout d'abord les LF en situation de rechute qui ont été les premiers à bénéficier de cette nouvelle thérapie avec l'obtention en 1998 de l'autorisation de mise sur le marché (AMM) de l'anti CD-20 de première génération, le Rituximab. Des essais cliniques ont

rapporté des résultats positifs en termes de survie chez les patients qui recevaient du Rituximab en association avec la polychimiothérapie (R-CHOP) par rapport aux patients ne recevant que le traitement classique de l'époque (CHOP) (Coiffier *et al.* 1998). En 2002, le Rituximab obtient son AMM en traitement de première ligne pour les LF puis pour le traitement des LBDGC, ayant alors fait ses preuves dans le cadre d'essais cliniques menés par le Groupe d'Etude des Lymphomes de l'Adulte (GELA) (Coiffier *et al.* 2002 ; Salles *et al.* 2008).

De plus, sur la base des résultats de l'essai PRIMA (PRImary Rituximab and Maintenance) (Salles *et al.* 2011), le Rituximab a été étendu en traitement de maintenance de première ligne pour les LF, chez les patients répondeurs à l'issue de six cures de R-CHOP. C'est ce schéma thérapeutique qui est actuellement utilisé dans la prise en charge des LF.

La prise en charge des LMNH va connaître bientôt un nouvel essor avec l'utilisation de nouveaux anticorps monoclonaux anti CD-20 de deuxième génération de type I, i.e. Veltuzumab ou Ofatumumab et de type II, i.e GA101 ou LFB-R603. L'efficacité de ces nouvelles molécules est en cours d'étude dans le cadre notamment de l'essai RELEVANCE (Rituximab Lenalidomide Versus ANy Chemotherapy) mis en place par le groupe LYmphoma Study Association (LYSA). Au-delà de ces essais cliniques, il restera à en connaître le réel impact en population générale. Des études sur données de registre de cancer de population pourront permettre d'approcher les tendances chronologiques de la survie sur des bases sans sélection au départ.

I.1.2. ÉVOLUTION DES CLASSIFICATIONS

Les études menées par sous-type histologique ont été rendues possibles grâce à l'évolution des classifications. Cela a permis d'homogénéiser les procédés de diagnostic mais aussi de prendre en compte toutes les caractéristiques clinico-biologiques connues. Plusieurs sous-types de LMNH ont ainsi pu être différenciés dans les classifications internationales. La classification des hémopathies lymphoïdes a été modifiée de façon progressive au fil du temps. Quatre étapes essentielles sont à retenir au niveau de l'évolution des classifications hématologiques (Figure I.2) :

- La classification Rappaport (Rappaport, H. 1966) dans laquelle il n'y avait pas de différenciation entre cellule T et cellule B.
- La classification Kiel (Lennert *et al.* 1975) utilisée en Europe (et la classification Lukes et Collins (Lukes et Collins 1974) davantage utilisée aux USA) dans lesquelles la différenciation entre les lymphomes B des lymphomes T a été prise en compte.
- La Working formulation (Robb-Smith 1982) qui pour la première fois associe à la fois les observations histologiques, la symptomatologie et l'évolution naturelle de la maladie.
- La classification « REAL » (Revised European American Lymphoma) (Harris *et al.* 1994) qui a pris en compte les avancées majeures dans les domaines de l'immunologie, de la cytogénétique et de la biologie moléculaire.
- La classification OMS (Jaffe *et al.* 2001) : qui se positionne comme la classification internationale consensuelle. Elle a fait l'objet d'une révision en 2008.

Au niveau épidémiologique, la Classification Internationale des Maladies (CIM) a été publiée en 1900 par le gouvernement français puis reprise et publiée par l'OMS en 1948 à des fins de classement de données de mortalité et de morbidité. Cette classification regroupe toutes les maladies existantes de l'époque, allant de maladies infectieuses à des maladies tumorales malignes. Devant les recommandations de tenir compte à la fois de la morphologie et de la topographie de la tumeur, la première édition de la Classification Internationale des Maladies pour l'Oncologie (CIM-O) est publiée en 1976 (World Health Organization 1976). Cette classification a été révisée par la suite en 1990 avec la deuxième édition (Percy *et al.* 1990), puis en 2001 avec la troisième édition (Fritz 2001). La CIM-O-3 a été légèrement révisée en 2008 (Campo *et al.* 2011) afin de tenir compte des avancées biologiques et cliniques. La troisième édition de la classification internationale des maladies en oncologie (CIM-O-3) de 2008 a appliqué de très près, pour les lymphomes, les recommandations de la nouvelle classification OMS qui succède à la REAL. Aussi depuis 2001, les classifications en hématologie et en épidémiologie se sont rapprochées, la CIM-O-3 étant à ce jour utilisée à la fois par les registres des cancers et par les pathologistes.

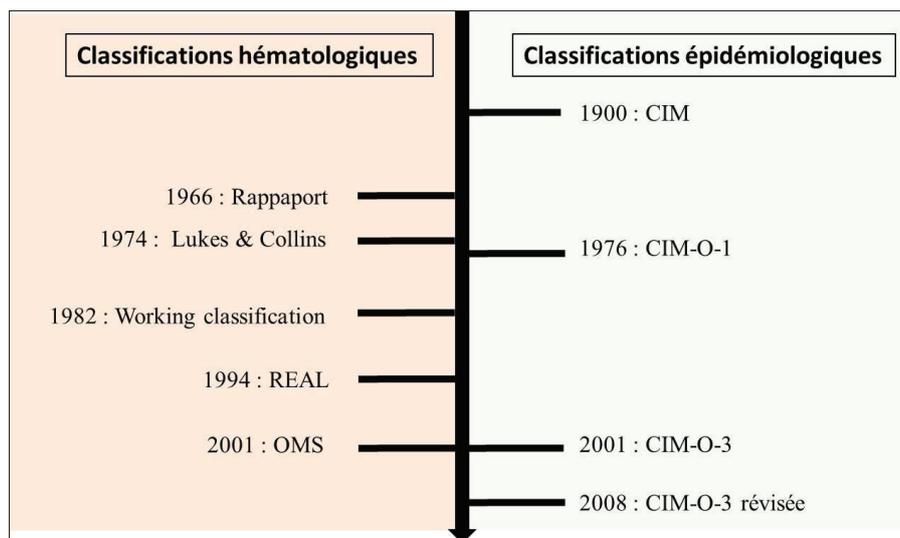


Figure I.2 - Evolutions majeures des classifications utilisées en hématologie et en épidémiologie.

A des fins d'études épidémiologiques, l'utilisation d'une seule et même classification pour l'oncologie au niveau international a permis d'homogénéiser les enregistrements d'un pays à un autre, permettant ainsi des études plus fines par sous-groupe de LMNH. Même si des diagnostics anciens ont pu être recodés en se basant sur ces nouvelles classifications, d'autres n'ont pas pu être reclassés par manque d'informations sur les comptes rendus anatomopathologiques (dus à des outils de diagnostic ne permettant pas à l'époque de définir avec précision les sous entités). Alors que certains LMNH étaient déjà référencés, par exemple les lymphomes folliculaires, d'autres entités, comme les LBDGC, ont été assez impactées par ces changements de classifications. Aussi pour certaines entités de LMNH, l'homogénéité des cas ne peut être assurée sur des périodes de diagnostic trop anciennes.

I.1.3. ÉPIDÉMIOLOGIE DES LMNH

I.1.3.1. Incidence

En France, l'incidence des LMNH a doublé depuis les années 1980 pour se stabiliser ensuite après les années 2000 (Belot *et al.* 2008). Ce doublement du nombre de cas a été perceptible partout dans le monde et il n'existe pas véritablement d'explication à la stagnation

de l'incidence des LMNH ces dernières années (Sant *et al.* 2010). Ces maladies malignes touchent essentiellement les patients âgés de plus de 60 ans.

En France, la dernière étude d'incidence par localisation de cancer, effectuée par le réseau FRANCIM en collaboration avec les Hospices Civils de Lyon, a permis pour la première fois de fournir des estimations d'incidence pour neuf sous-types de LMNH à l'échelle nationale (Monnereau *et al.* 2013a). Ces résultats d'incidence sont présentés dans la figure I.3 suivante.

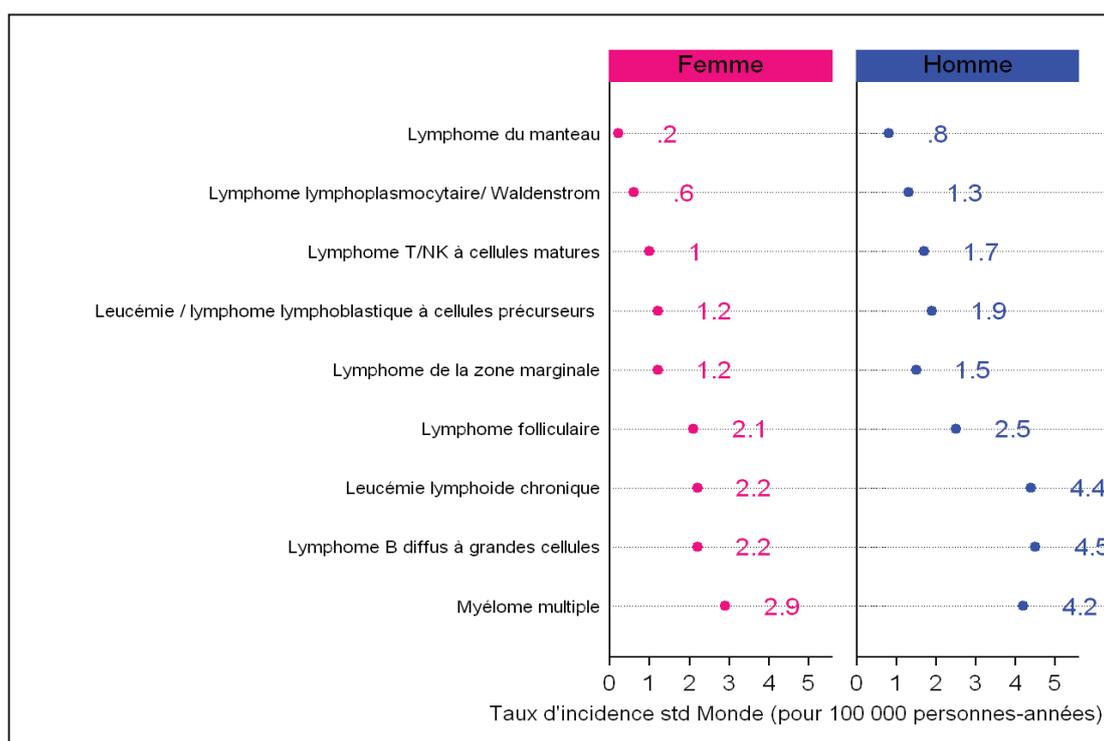


Figure I.3 - Taux d'incidence standardisé pour l'âge (à la population mondiale) en France en 2012, par sous-type de LMNH. Source : FRANCIM (Monnereau *et al.* 2013a).

Chez la femme, l'incidence standardisée sur la population mondiale varie de 0,2 à 2,9 pour 100 000 personnes-années respectivement pour les lymphomes du manteau et les myélomes multiples. Chez les hommes, elle passe de 0,8 à 4,5 pour 100 000 personnes-années respectivement pour les lymphomes du manteau et les LBDGC. Il existe une prédominance masculine avec une incidence doublée chez les hommes par rapport aux femmes, pour la leucémie lymphoïde chronique, les LBDGC et les myélomes multiples. La répartition des LMNH par type morphologique est sensiblement la même en Europe avec pour les trois sous-types les plus fréquents, un taux standardisé égal à 4,6 puis 3,8 et 3,1 pour 100 000 personnes-

années respectivement pour les myélomes multiples, les leucémies lymphoïdes chroniques et les LBDGC entre 2000 et 2002 et pour tous sexes confondus (Sant *et al.* 2010).

I.1.3.2. Pronostic

L'hétérogénéité des LMNH se perçoit également au niveau du pronostic des patients qui est variable d'une entité à une autre. De la même manière que pour l'incidence, les probabilités de survie par sous-type de LMNH font l'objet de publications récentes. Les estimations de survie nette standardisée à 5 ans en France sont présentées dans la figure I.4 (Monnereau *et al.* 2013b).

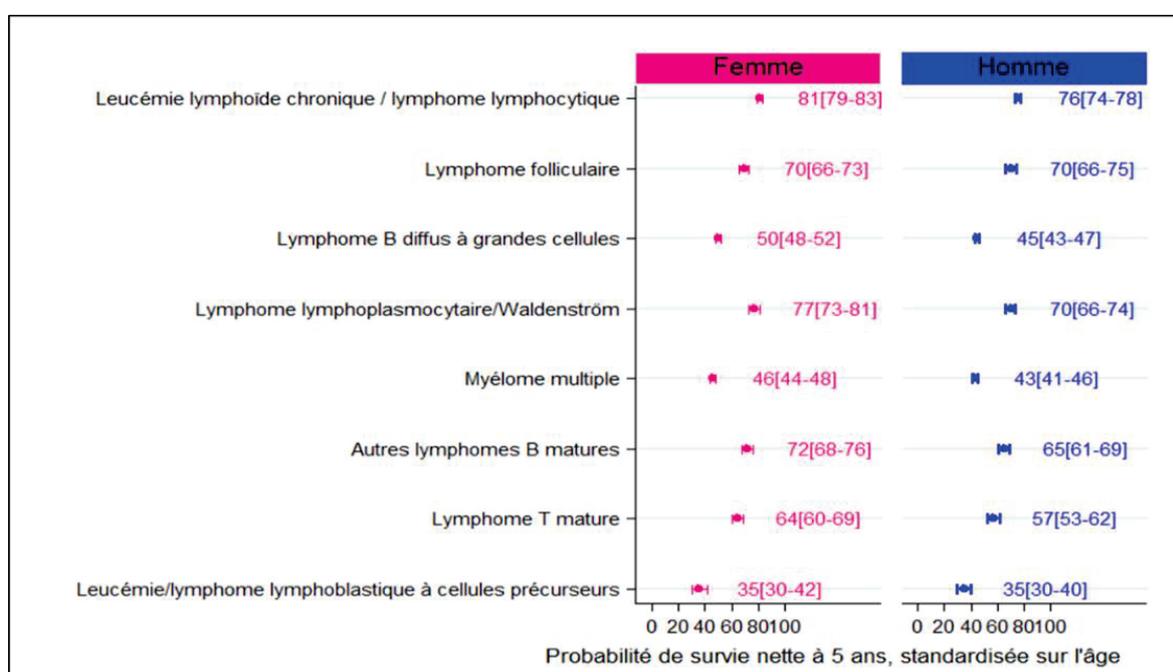


Figure I.4 - Probabilité de survie nette (et $IC_{95\%}$) à 5 ans, standardisée sur l'âge en France en 2012, par sous-type de LMNH et par sexe. Source : FRANCIM (Monnereau *et al.* 2013b).

La probabilité de survie nette à 5 ans varie de 35% (30-42) à 81% (79-83) entre le lymphome lymphoblastique et la leucémie lymphoïde chronique chez la femme et de 35% (30-40) à 76% (74-78) respectivement pour les mêmes sous-types chez l'homme. Le LBDGC, le myélome multiple et le lymphome lymphoblastique sont de mauvais pronostics avec une médiane de survie nette déjà atteinte à 5 ans (Figure I.4). A l'exception du lymphome lymphoblastique et du LF, les femmes ont une survie légèrement plus élevée que les hommes.

Plusieurs études ont mis en évidence une variation de la survie liée à l'âge en faveur des sujets plus jeunes (Sant *et al.* 2009 ; Monnereau *et al.* 2013b)¹.

En Europe, les résultats de survie par sous-type de lymphomes ont été récemment réactualisés et ont fait l'objet d'une publication à laquelle nous avons collaboré (Sant *et al.* 2014). L'article est présenté en annexe du rapport.

I.2. MESURES DES PROGRÈS EN POPULATION GENERALE

I.2.1. PROGRÈS DANS LA PRISE EN CHARGE DES LMNH

Les LMNH ont bénéficié ces dernières années de progrès majeurs liés à la fois à une meilleure connaissance de la maladie mais également à des interventions menées dans le cadre des politiques de santé publique. Ces interventions peuvent être distinguées selon plusieurs volets (Cancer 1990) (Figure I.5) :

- La prévention. Plusieurs facteurs de risque de survenue d'un LMNH ont été mis en évidence : des facteurs intrinsèques à l'individu (personnes immuno-déprimées, infections virales telles que le virus du sida, l'hépatite C ou la présence du virus Epstein-Barr (Dal-Maso et Franceschi 2003 ; Becker *et al.* 2009 ; Peveling-Oberhag *et al.* 2013) ; des facteurs liés aux habitudes de vie (obésité (Willett *et al.* 2008), consommation de tabac (Morton *et al.* 2005)) ; des facteurs familiaux (Casey *et al.* 2006 ; Wang *et al.* 2007) ou encore des facteurs environnementaux (exposition à des produits chimiques tels que pesticides, solvants ou engrais (Agopian *et al.* 2009 ; Cocco *et al.* 2013)). Malgré d'importantes études épidémiologiques portant sur le sujet notamment dans le cadre du consortium INTERLYMPH, l'étiologie des LMNH reste encore mal connue (Morton *et al.* 2014).
- La détection. Un meilleur accès aux soins avec des délais réduits de consultations auprès de médecins spécialisés permet une détection plus précoce du LMNH. Au diagnostic,

¹ Les données américaines du SEER programme (Surveillance, Epidemiology and End Results Program) ne permettent pas pour le moment de comparaisons par sous-type de lymphomes entre les données européennes et américaines, les estimations étant encore effectuées pour la totalité du groupe LMNH (Schans *et al.* 2011).

l'évolution du lymphome en est ainsi limitée avec des stades moins avancés, ce qui peut améliorer par la suite la survie des patients (Al-Hamadani et Go 2014). Une action de détection de grande ampleur pour certains cancers est la mise en place d'un programme de dépistage organisé (ex : sein, cancer du côlon) (Grech *et al.* 2015). Il n'existe pour le moment pas de programme de dépistage pour les lymphomes.

- Le diagnostic. Le diagnostic des LMNH s'est profondément amélioré avec des outils diagnostiques plus performants (biologie moléculaire, phénotypage) et des outils d'imagerie avec l'arrivée du TEP-scan qui permet d'affiner l'extension de la maladie (Sattar *et al.* 2006).

- Le traitement. La thérapeutique des LMNH a connu un profond changement avec l'utilisation de nouvelles molécules de chimiothérapie telles que l'interféron dans les années 1990 (Solal-Celigny *et al.* 1993), ou plus récemment avec la bendamustine (Flinn *et al.* 2014), mais aussi le développement de la thérapie ciblée avec les anticorps monoclonaux humanisés anti-CD20. Le rituximab est le premier anticorps monoclonal dont l'indication a été reconnue dans le traitement des lymphomes de type B. Les lymphomes folliculaires en situation de rechute ont été les premiers à bénéficier de cette nouvelle thérapeutique (Coiffier *et al.* 1998) qui a ensuite été étendue à d'autres sous-types de lymphomes, tels que les LBDGC (Coiffier *et al.* 2002). Plusieurs grands essais cliniques ont montré l'efficacité de ces nouvelles thérapeutiques qui ont amélioré considérablement le pronostic des patients (Hiddemann *et al.* 2005 ; Salles *et al.* 2008). D'autres options thérapeutiques telles que la greffe de moelle chez les patients plus jeunes en situation de rechute ont également amélioré le pronostic des sujets (Le Gouill *et al.* 2011). De plus l'inclusion dans les essais cliniques de plus en plus nombreux a permis aux patients un accès plus rapide aux nouvelles innovations. Le développement de guidelines avec des recommandations plus précises permet également une prise en charge thérapeutique plus efficace, en optimisant notamment le début de traitement pour les lymphomes indolents (Ghielmini *et al.* 2013).

- La réhabilitation. Un point majeur du dernier plan cancer en France est la préservation du bien-être des patients. Les études pour évaluer la qualité de vie des patients atteints de LMNH se développent (Jensen *et al.* 2013). L'amélioration dans la prise en charge

thérapeutique des patients avec l'utilisation de nouvelles molécules plus efficaces mais aussi moins toxiques va dans le sens d'une préservation de la qualité de vie des patients au cours de la phase thérapeutique (Smith 2015).

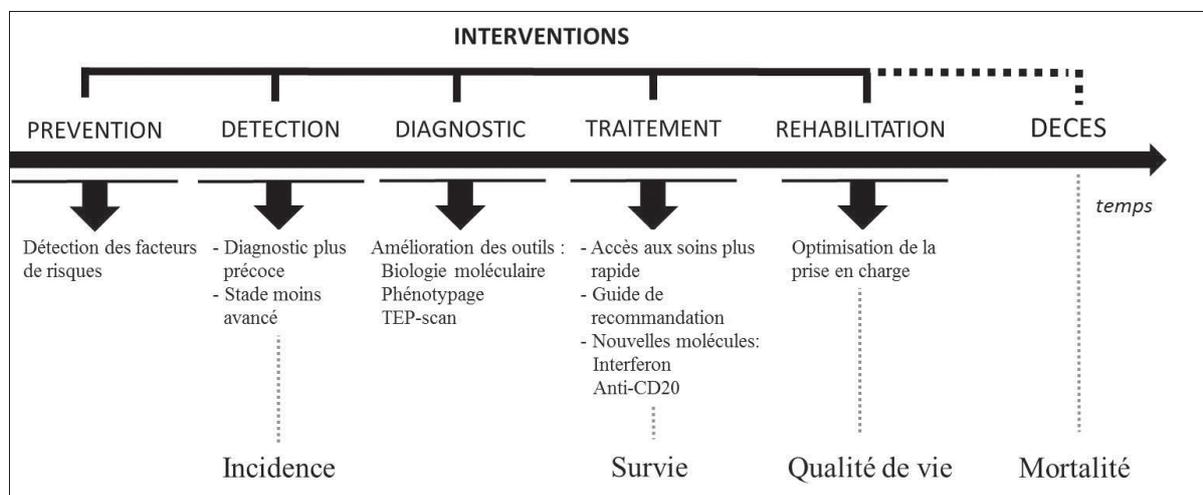


Figure I.5 - Interventions dans la lutte contre les LMNH et mesures épidémiologiques associées.

Une évaluation globale des progrès contre le cancer doit être basée sur l'étude simultanée de plusieurs indicateurs épidémiologiques qui sont l'incidence, la mortalité et la survie (Bailar et Gornik 1997; Dickman et Adami 2006).

L'étude temporelle de *l'incidence* mesure épidémiologique qui quantifie le nombre de nouveaux cas dans une population de personnes exposées au risque pendant un intervalle de temps et dans une zone géographique définis, permet de mesurer les efforts en matière de prévention.

L'analyse de la *qualité de vie* des patients mesurée à l'aide de questionnaires génériques ou spécifiques, permet de quantifier la perte de bien-être du patient pendant et à la suite de la phase thérapeutique.

L'intérêt majeur des interventions mises en place par les politiques de santé est de limiter le nombre de décès liés au cancer. L'indicateur de *mortalité*, qui représente le rapport entre le nombre de décès dus à la pathologie et la population totale pour une année et dans une zone définie, permet de mettre en avant l'efficacité de ces interventions.

Une analyse de la *survie*, qui représente une mesure du nombre de patients encore en vie à différents temps après le diagnostic, permet d'estimer le pronostic des patients et son évolution au cours des périodes de diagnostic.

I.2.2. DONNÉES DE POPULATION

I.2.2.1. Essais cliniques versus données de population

Un *essai thérapeutique* est défini comme « tout projet de recherche qui affecte de façon prospective des sujets humains à des groupes d'intervention et de comparaisons afin d'étudier la relation de cause à effet entre un acte médical et l'évolution d'un état de santé » selon le comité international des rédacteurs de revues médicales (ICMJE 2015). Après une sélection des patients via des critères d'inclusion à l'entrée de l'étude (sur l'âge, le sexe, le stade d'avancement de la maladie par exemple), l'allocation aléatoire des traitements (la randomisation) assure une comparabilité entre les groupes. En garantissant l'absence de facteurs confondants, l'essai thérapeutique randomisé est la méthode de choix pour évaluer l'impact d'une nouvelle thérapeutique sur l'évolution de la maladie. Néanmoins les résultats basés sur un essai ne peuvent pas être étendus à l'ensemble des malades en population générale du fait de la non représentativité des patients induite par des critères d'inclusion en amont de l'étude.

A l'inverse des essais thérapeutiques randomisés, les *études populationnelles* (ou études en population générale) incluent la totalité des patients atteints de la maladie d'intérêt dans une zone géographique et sur une période de temps délimitées. Cette exhaustivité sur le nombre de cas permet de disposer d'indicateurs épidémiologiques pour l'ensemble de la population à risque. Néanmoins, il est difficile, voire impossible, de mettre en exergue l'impact d'un facteur spécifique (un traitement par exemple) en se positionnant sur les résultats d'une étude populationnelle, du fait de la présence de facteurs de confusion (Dickman et Adami 2006).

Les essais thérapeutiques randomisés et les études populationnelles ne répondent pas aux mêmes objectifs : ces deux types d'études sont donc complémentaires et doivent ainsi être menées de front.

I.2.2.2. Les registres de cancer : source des données de population

Les registres de cancer constituent la source de référence pour l'épidémiologie des cancers en population générale (Parkin 2006). Au sens du Comité National des Registres (CNR), un registre est défini comme « *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* ». Il existe deux types de registres de cancer : les registres hospitaliers et les registres de population. *Les registres hospitaliers* enregistrent tous les cas de cancer qui ont été suivis dans une structure hospitalière alors que *les registres de population* recensent tous les cas de cancer. Seules les bases issues des registres de population permettent de disposer de données exhaustives.

Les sources de notification des registres sont les différentes structures médicales de la zone couverte par le registre (centres hospitaliers universitaires ou des hôpitaux périphériques, des centres anticancéreux, des cliniques privées, des laboratoires d'anatomopathologie, etc.). La plupart des registres fonctionnent en recueil retardé, c'est-à-dire qu'il existe un décalage entre l'année de recueil et l'année d'incidence². Les informations recueillies sont variables d'un registre à l'autre, dépendant de la perspective dans laquelle le registre a été fondé et des conditions locales. Néanmoins, certaines données sont indispensables pour l'enregistrement des cas et sont donc communes à tous : le sexe, la date de naissance, le lieu de résidence, la date d'incidence (ou de diagnostic), la morphologie et la topographie du cancer, la base de diagnostic, le statut vital et la date de dernière nouvelle du patient. Un suivi actif permet une mise à jour du statut vital du patient à l'aide de différentes sources d'information, principalement en interrogeant le Répertoire National d'Identification des Personnes Physiques (RNIPP).

I.2.2.2.1 Historique

Le premier registre des cancers de population a été créé à Hambourg en 1926 (Wagner 1985). Par la suite, d'autres registres de cancer ont vu le jour et il existe actuellement plus de 300 registres généraux de cancer dans le monde (Forman *et al.* 2013). Plusieurs pays

² En France, le décalage est de 2 ans

enregistrent les cas de cancer au niveau national. Le premier registre national créé est le registre danois en 1942, mais la majorité des pays disposent de registres ne couvrant qu'une zone bien définie (un département ou une région) (Wagner 1991).

En France, le premier registre des cancers a été mis en place en 1975 (Faivre *et al.* 1979) à la suite d'initiatives individuelles, puis à partir de 1986, les registres ont été créés autour d'une politique nationale de santé publique et de recherche avec la création du CNR. Afin de coordonner les efforts dans l'observation du cancer en France, le réseau *France Cancer Incidence et Mortalité (FRANCIM)* a été créé en 1997 et regroupe l'ensemble des registres des cancers français qualifiés par le CNR. Une base de données commune à l'ensemble des registres a été constituée et est hébergée dans le service de biostatistique des Hospices Civils de Lyon (HCL). Cette base commune permet la réalisation d'études sur les tendances évolutives des indicateurs épidémiologiques pour chaque cancer. En 2015, la base de données commune regroupe les bases de 28 registres de cancer dont 16 registres généraux et 12 registres spécialisés (dont 3 à vocations nationales). La qualité et l'exhaustivité des enregistrements des registres français sont vérifiées tous les quatre ans via un audit par le CNR qui depuis 2013 a été remplacé par le Centre d'Évaluation des Registres (CER) ayant les mêmes vocations. Depuis 2008, l'observation et la surveillance des cancers en France se sont progressivement structurées pour aboutir à un programme partenarial entre le réseau des registres de cancer français FRANCIM, le service de biostatistique des HCL, l'InVS et l'INCA.

1.2.2.2 Bases collaboratives de données de population

La recherche épidémiologique sur le cancer est coordonnée au niveau international par le Centre International de Recherche sur le Cancer (CIRC) créé en 1965 par l'OMS. Le CIRC est en lien étroit avec l'Association Internationale des Registres du Cancer (IACR) qui a vu le jour en 1966 et qui a pour objectif de promouvoir et de normaliser les méthodes de recueil de données des registres. La collaboration entre ces deux groupes a donné naissance à plusieurs projets d'ampleur internationale dont les deux principaux sont le projet *Cancer incidence in five continents* et le projet *GLOBOCAN*. Le projet Cancer incidence in five continents fournit les incidences par cancer provenant de 290 bases de données de registres issus de 68 pays (Matsuda et Saika 2012). Le projet GLOBOCAN fournit des estimations de l'incidence, mortalité et de la prévalence des principaux cancers pour 184 pays (Ferlay *et al.* 2015).

Au niveau européen, la base de population EUROCARE (EUROpean CANcer REgistry based study on survival and care of cancer patients) (EUROCARE 2014) a été initiée en 1989 par des institutions italiennes et représente le plus gros projet collaboratif sur l'étude de la survie des cancers en Europe. La base actuelle EUROCORE-5 compte plus de 21 millions de diagnostics de cancer provenant de 116 bases de données de registres et représente 30 pays européens.

Il existe également en Europe, le projet NORDCAN (NORDic CANcer) initié en 2002 par l'Association des registres de cancer nordique (ANCR) (Engholm *et al.* 2010). Ce projet fournit des estimations des principaux indicateurs épidémiologiques du cancer pour 50 localisations dans les pays nordiques.

Le dernier projet collaboratif en date en Europe est le projet SUDCAN (SUD CANcer : Trends in net survival in six latin speaking countries), coordonné par le Groupe des REgistres de Langue Latine (GRELL). Ce projet se rattache à la base de données européenne d'EUROCARE-5.

Au niveau américain, le SEER programme (Surveillance, Epidemiology and End Results) créé en 1973, regroupe les différents registres de cancer aux Etats-Unis. Il représente la source d'information pour l'incidence et la survie de tous les cancers aux USA (Edwards *et al.* 2014). Le programme recouvre environ 28% de la population américaine.

Au niveau international, le projet CONCORD est la continuité du projet Cancer incidence in five continents et a pour objectif l'estimation de la survie dans différents pays. Ce projet collaboratif se poursuit avec la publication récente des derniers résultats du projet CONCORD-2 qui regroupe plus de 25 millions de diagnostics de cancer entre 1995 et 2009 et provenant de 279 bases de données de registre des cancers (Allemani *et al.* 2015).

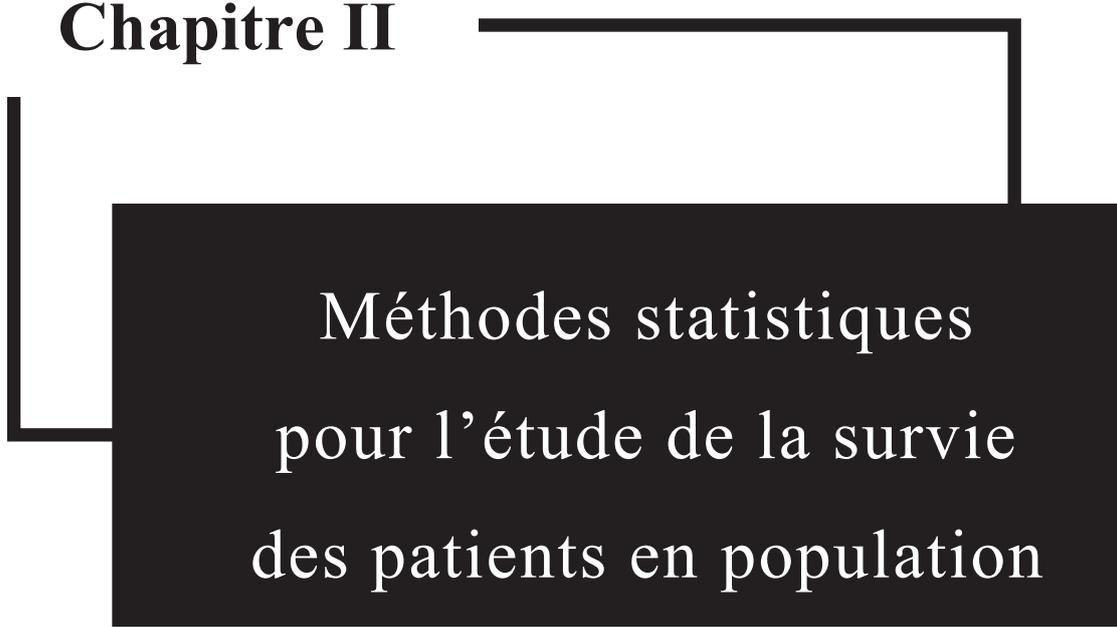
I.2.3. DISCUSSION

Même si les LMNH restent un cancer rare en termes d'incidence par rapport à d'autres cancers, ils restent un enjeu de santé publique avec le doublement du nombre de cas en 20 ans et des facteurs étiologiques encore mal connus. Alors que dans la plupart des études les

indicateurs épidémiologiques étaient estimés pour l'ensemble des LMNH, les recommandations de grands groupes collaboratifs en hématologie ont permis une prise de conscience générale sur la pertinence de dissocier les sous-types de LMNH (Morton *et al.* 2007 ; Sant *et al.* 2010). Les LMNH ont bénéficié d'avancées majeures liées à la fois à l'amélioration des connaissances de la maladie mais également à l'utilisation de nouvelles thérapeutiques. Même si la quasi-totalité des patients atteints d'un LMNH ne peuvent être guéris, de nouveaux espoirs liés aux progrès de ces dernières années résident dans l'allongement de l'espérance de vie des patients. L'efficacité du système de soins des lymphomes peut être visible par le biais d'études portant sur des données des registres de cancer qui permettent des estimations d'incidence, de mortalité et de survie en population générale. Ces estimations font régulièrement l'objet de publications et les bases collaboratives de données de population permettent des comparaisons des principaux indicateurs entre pays.

Notre travail a porté sur l'estimation de la survie nette des patients à la suite d'un diagnostic de LMNH. Cet indicateur, mesuré sur des données de population nécessite des outils adaptés afin de tenir compte des biais inhérents aux études observationnelles ou interventionnelles. C'est l'objet du projet CENSUR qui fixe le cadre de ce travail et dont le but est d'améliorer les méthodes d'estimations de la survie nette des patients en population générale.

Chapitre II



Méthodes statistiques
pour l'étude de la survie
des patients en population

Afin de comparer le poids du cancer lui-même sur la mortalité des patients entre différentes zones géographiques ou différentes périodes, il est nécessaire de disposer d'outils statistiques de survie permettant de s'affranchir des autres causes de décès que celles liées au cancer lui-même. La survie nette, indicateur théorique (car non observable) qui correspond à la situation hypothétique où seule la maladie d'étude serait la cause de mortalité dans la population suivie, apparaît alors comme le seul indicateur adéquat dans les études comparatives.

Il existe plusieurs approches pour estimer la survie nette d'un groupe. Toutefois, Danieli *et al.* (Danieli *et al.* 2012) ont montré qu'en réalité seules deux d'entre elles fournissent des estimations non biaisées de la survie nette : l'estimateur non paramétrique proposé par Pohar-Perme (Perme *et al.* 2012) et l'approche paramétrique basée sur une modélisation ajustée sur les variables démographiques de la table de mortalité.

II.1. PRÉ-REQUIS

L'analyse de données de survie qui mesure la probabilité de survenue d'un événement dans une population en tenant compte de son délai d'apparition a deux particularités majeures. La première est qu'elle ne concerne que des variables aléatoires positives en s'intéressant aux durées de vie. Dans ce cadre, la loi normale ne peut être utilisée comme distribution et d'autres lois doivent alors être choisies. La deuxième particularité est que la variable aléatoire d'intérêt n'est pas toujours observée chez tous les sujets, ces observations incomplètes sont alors considérées comme des données censurées.

De par ces spécificités, les méthodes standard d'analyse statistique sont inappropriées pour l'étude des données de survie et de nouveaux outils statistiques ont donc été proposés.

Pour conduire une analyse de survie il est nécessaire de définir une fenêtre d'observation qui va dépendre de l'objectif clinique du travail. La borne gauche de cette fenêtre correspond à ce que l'on appelle la *date d'origine*, c'est-à-dire le moment à partir duquel les patients vont entrer dans l'étude pour être suivis (la date d'origine pouvant être différente pour chaque individu). La borne de droite de la fenêtre de suivi va correspondre à ce que l'on

nomme la *date de point*, qui est la date au-delà de laquelle nous ne tenons plus compte de l'apparition de l'événement.

En épidémiologie du cancer, la date de naissance ou la date de diagnostic sont souvent utilisées comme date d'origine et la durée d'intérêt correspond au délai écoulé jusqu'au décès du patient.

II.1.1. NOTION DE DONNÉES CENSURÉES

Comme indiqué précédemment, une des particularités essentielle de l'analyse de survie est que l'événement d'intérêt peut ne pas être observé pour tous les individus à la date de point, la donnée étant alors considérée comme incomplète ou censurée. Cette notion de censure peut prendre plusieurs formes.

La *censure à droite* est observée lorsque l'individu n'a pas subi l'événement d'intérêt à la fin de l'étude dans le cas où :

- le sujet n'a pas encore subi l'événement d'intérêt à la date de point. Dans ce cas, le patient sera censuré à la date de point. Le sujet est défini comme exclu-vivant.
- le sujet a quitté l'étude avant la date de point. Le patient sera alors censuré à sa date de dernière nouvelle. Le sujet est défini comme perdu de vue.
- il existe des risques concurrents, c'est-à-dire si un autre événement empêche l'observation de l'événement d'intérêt dans la fenêtre d'étude. Dans la situation où l'événement d'intérêt est le décès lié à la maladie, le décès du patient dû à une autre cause peut empêcher la survenue de l'événement d'intérêt. Le patient sera alors censuré à sa date de décès dû à une autre cause.

D'autres types de censures peuvent être observés, même s'ils sont moins répandus : la *censure à gauche* si l'individu a déjà subi l'événement à la date d'origine et la *censure par intervalle* si l'individu subit l'événement entre deux observations, sans qu'on ne puisse en connaître la date précise.

Les méthodes classiques d'analyse de survie supposent l'indépendance entre le temps de censure et le temps d'événement du patient afin de disposer d'estimations non biaisées de survie. La survie du groupe des individus censurés est alors considérée comme identique après

le temps de censure à celle des individus non censurés et l'estimation de la survie peut ainsi être généralisée à tous les individus de la cohorte. En d'autres termes, les individus censurés à une date t doivent être représentatifs des individus encore à risque au temps t .

En revanche, s'il y a dépendance entre la cause de la censure et l'événement d'étude, cela peut introduire un biais dans l'estimation de la durée de survie, avec une probable surestimation de cette dernière (Kalbfleisch et Prentice 2011).

Le nombre de patients censurés peut être minimisé en amont de l'analyse en utilisant des bonnes pratiques telles que:

- assurer un recul suffisant entre la date d'origine et la date de point de l'étude pour limiter le nombre d'exclus-vivants.
- garantir un système de suivi actif pour tous les individus de la cohorte afin de réduire le nombre de perdus de vue.

En revanche, le mécanisme de censure induit par la présence de risques concurrents est plus complexe. En effet, si le temps de décès dû au cancer est dépendant du temps de décès dû aux autres causes (en étant affecté par une même covariable comme l'âge), alors cela induit la présence de censure « dépendante » et donc un biais dans l'analyse. Seul un outil statistique adapté permettra de corriger le biais dans cette situation. Nous reviendrons plus longuement sur ce mécanisme de censure qui marque un point essentiel dans les méthodes d'estimation de la survie nette.

II.1.2. LES FONCTIONS DE BASE EN ANALYSE DE SURVIE

L'analyse de survie se base sur plusieurs fonctions dépendantes du temps. La variable aléatoire, continue positive, nommée « durée de survie » dont on étudie la distribution est notée T . La loi de probabilité de T peut être caractérisée par plusieurs fonctions liées entre elles.

- Soit $f(t)$ la fonction de densité de probabilité de T définie par :

$$f(t) = \lim_{\Delta t \rightarrow 0^+} \frac{P(t \leq T < t + \Delta t)}{\Delta t}$$

Il s'agit de façon classique de la densité d'une variable aléatoire continue.

- Soit $F(t)$ la *fonction de répartition* de T définie par :

$$F(t) = P(t \leq T) = \int_0^t f(u)du$$

$F(t)$ correspond à la probabilité qu'une personne décède entre le temps 0 et le temps t (Figure II.1). Cette fonction de répartition est monotone et croissante. Aussi, nous avons pour $F(t)$:

$$\lim_{t \rightarrow 0^+} F(t) = 0 \text{ \& } \lim_{t \rightarrow +\infty} F(t) = 1$$

De cette probabilité de décès, nous pouvons en déduire la fonction de survie.

- Soit $S(t)$ la *fonction de survie* définie par :

$$S(t) = P(T > t) = 1 - F(t)$$

La fonction de survie donne la probabilité qu'une personne survive au-delà d'un temps t (Figure II.1). Nous avons pour $S(t)$: $\lim_{t \rightarrow 0^+} S(t) = 1 \text{ \& } \lim_{t \rightarrow +\infty} S(t) = 0$

A la date de début d'étude, tous les individus sont en vie et tous les individus sont à risque de connaître l'événement d'intérêt et le connaîtront si le temps de suivi est suffisant.

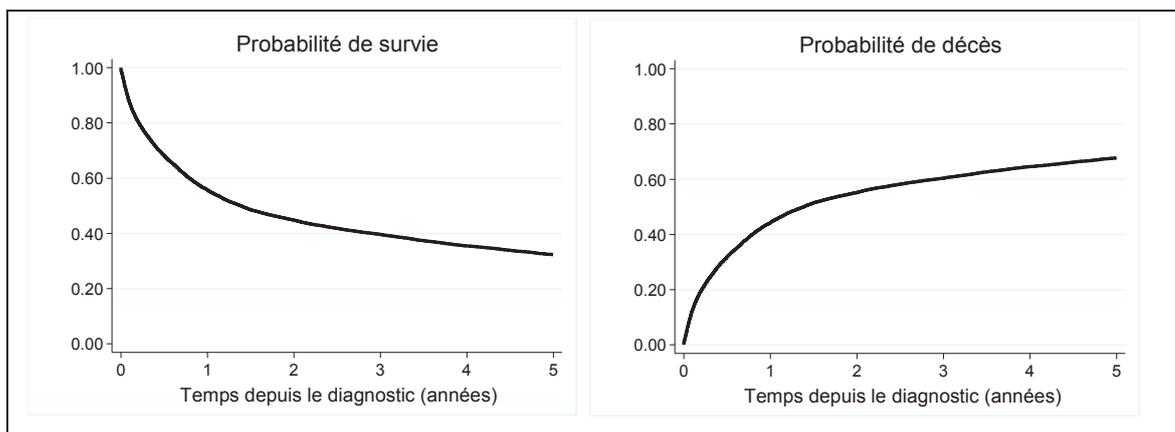


Figure II.1 - Probabilité de survie et probabilité de décès en fonction du temps de suivi.

La fonction de survie $S(t)$ et de décès $F(t)$ sont des probabilités et de ce fait ne peuvent pas prendre des valeurs supérieures à 1 et sont toujours positives.

- Soit $\lambda(t)$ la fonction de risque instantané de T définie par :

$$\lambda(t) = \lim_{\Delta t \rightarrow 0^+} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = -\frac{d \ln S(t)}{dt}$$

Cette fonction est définie comme la probabilité de décès durant un intervalle de temps Δt conditionnellement au fait que le sujet soit en vie au début de l'intervalle (Figure II.2). De par sa définition, cette fonction correspond à un taux et peut donc prendre des valeurs supérieures à 1. La valeur obtenue dépend également de l'unité de temps utilisée. Nous la nommons « *taux instantané de mortalité* » dans un intervalle de temps défini. Ce taux traduit la force de mortalité subie par les patients encore en vie. De cette fonction découle la fonction de risque cumulé.

- Soit $\Lambda(t)$ la fonction de risque cumulé de T définie par :

$$\Lambda(t) = \int_0^t \lambda(u) du = -\ln S(t)$$

Cette fonction correspond à la somme cumulée des taux instantanés $\lambda(t)$ jusqu'au temps t (Figure II.2). Elle est nommée « *taux cumulé de mortalité* » au temps t .

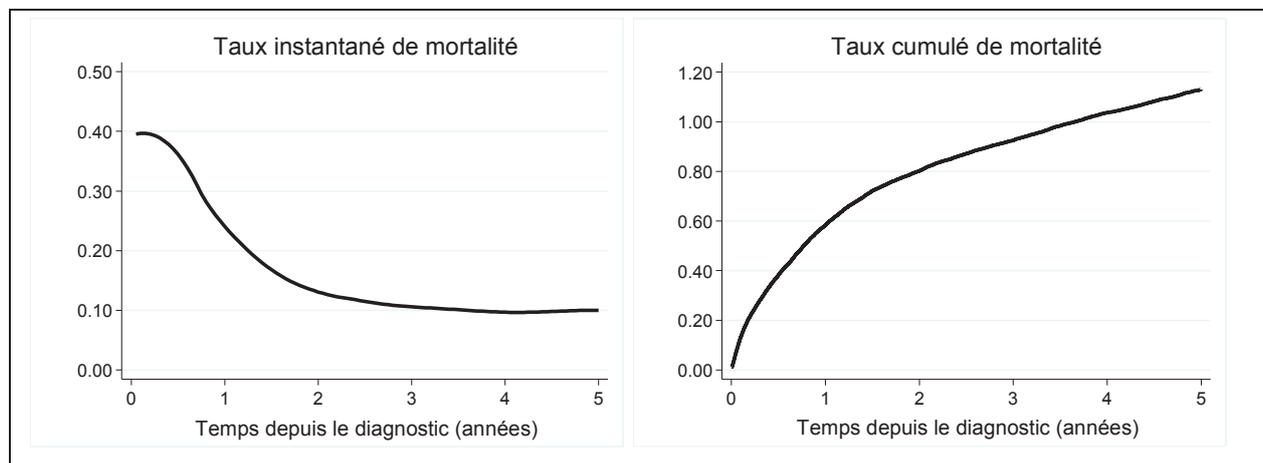


Figure II.2 - Fonction de risque instantané et de risque cumulé en fonction du temps de suivi.

Les fonctions de risque instantané et de risque cumulé ont pour unité un nombre de décès par personne-temps.

Aussi, la définition de la distribution de probabilité de T repose sur l'une des quatre fonctions suivantes : $S(t)$, $f(t)$, $\lambda(t)$ et $\Lambda(t)$. Ces quatre fonctions sont liées les unes aux autres et

l'estimation de l'une permet de déduire l'estimation des trois autres et inversement. Dès lors, nous pouvons définir la relation entre la fonction de survie $S(t)$ et la fonction de risque $\lambda(t)$.

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

Ainsi plus le risque diminue et plus la probabilité de survie est élevée.

On note aussi que lorsque le taux instantané est relativement faible, alors il correspond approximativement à la probabilité de décéder dans l'année grâce à la propriété mathématique qui permet d'approcher $\exp(-\lambda)$ par $1 - \lambda$ quand λ est petit.

A la vue de ces diverses fonctions qui définissent de façon équivalente la survie, les résultats d'une analyse de survie peuvent ainsi être exprimés soit comme une *probabilité de survie* soit comme un *taux d'événement* par personne-temps.

II.1.3. LES CONCEPTS DE SURVIE

Lorsque le décès du patient est l'événement d'intérêt dans l'analyse de survie, deux concepts fondamentaux sont à définir : le concept de survie « toutes causes » et le concept de « survie nette ». A noter que les définitions des fonctions présentées dans le paragraphe précédent sont valables dans le cadre de ces deux concepts.

II.1.3.1. La survie toutes causes

Nous parlons d'une analyse de la *survie toutes causes* lorsque l'événement d'intérêt est le décès du patient, toutes causes de décès confondues qu'elles soient liées ou non à la maladie d'étude. Aussi tous les décès observés durant la fenêtre de suivi sont pris en compte comme événement. La survie toutes causes est donc la résultante de l'ensemble des forces de mortalité auxquelles les individus sont exposés.

De façon générale, les résultats issus d'une analyse de survie toutes causes permettent d'affirmer qu'il y a augmentation/stabilisation/diminution de la survie des patients mais ils ne permettent pas de mettre en évidence la part de la mortalité induite par la maladie dans l'évolution de la survie.

Remarque 1. Certains auteurs utilisent d'autres termes pour définir le concept de survie toutes causes, tels que « survie observée », « survie globale » ou encore « survie brute ». Ces trois termes peuvent être ambigus et nous préférons ici le terme « survie toutes causes ».

II.1.3.2. La survie nette

La survie nette est définie comme la survie dans la situation hypothétique où la maladie d'étude serait la seule cause possible de décès, (c'est-à-dire lorsque tous les autres risques compétitifs de décéder sont éliminés). Il s'agit d'un concept théorique. De ce fait, le temps qui découle de ce concept n'est pas observable pour un individu.

La probabilité de survie toutes causes est par définition toujours inférieure à la probabilité de survie nette. L'écart se creuse davantage entre les deux probabilités lors d'estimations à long terme puisque plus le temps s'allonge et plus le risque que le décès soit dû à la maladie diminue. Néanmoins si la maladie d'étude est de mauvais pronostic ou que le suivi est de courte durée, les décès observés auront pour la quasi-totalité une cause liée à la maladie. Dans cette situation, la survie toutes causes approchera la survie nette.

II.2. LES APPROCHES D'ESTIMATION DE LA SURVIE NETTE

Deux approches théoriques sont possibles pour estimer la survie nette : l'approche « cause spécifique », qui se base sur l'utilisation de la cause de décès des patients, et l'approche par estimation de la mortalité attendue des patients, qui ne nécessite pas de connaître la cause exacte de décès. Nous verrons par la suite que seule l'approche par estimation de la mortalité attendue est un estimateur non biaisé de la survie nette.

II.2.1. USAGE DE LA CAUSE DE DÉCÈS

Dans cette approche nommée *approche cause spécifique*, les causes de décès des patients sont supposées connues et vont être classiquement catégorisées en deux groupes : les décès dus au cancer et les décès dus aux autres causes. Les patients décédés d'autres causes sont alors considérés comme n'ayant pas subi l'événement d'intérêt et sont censurés à leur date de décès et seuls les décès dus au cancer seront considérés comme événement. Le délai d'intérêt sera alors le délai entre la date de diagnostic et la date de décès dû au cancer. Dans ce contexte, les outils d'estimation de la survie toutes causes peuvent alors être utilisés (tels que l'estimateur non paramétrique de Kaplan-Meier (Kaplan et Meier 1958) ou le modèle semi-paramétrique de Cox (Cox 1972)).

Cependant, même si les causes de mortalité sont accessibles, il apparaît difficile de déterminer avec précision la cause principale de décès du patient. Est-elle liée directement ou indirectement à la maladie ?

De plus le degré de cette information dépend de la qualité de l'enregistrement qui peut être variable en fonction des zones géographiques. Percy *et al.* ont été les premiers à quantifier l'impact d'une mauvaise classification de la cause de décès sur les statistiques de mortalité (Percy *et al.* 1981). Aussi, l'approche cause spécifique n'est pas un outil adapté pour les études comparatives entre zones géographiques ou entre périodes de diagnostic.

Remarque 2. *Une étude menée sur les données du registre des cancers genevois a néanmoins montré que les erreurs de codification de la cause de décès principale ne portaient que sur un pourcentage minime des décès, ces problèmes de codification ayant donc un impact limité sur les estimations de survie nette (Schaffar et al. 2013).*

La connaissance et l'exactitude de la cause principale de décès ne sont pas les seules difficultés rencontrées dans l'approche cause spécifique. Il a été montré que cette approche fournit la plupart du temps une estimation biaisée de la survie nette car elle ne respecte pas l'hypothèse de censure indépendante (Perme *et al.* 2012). En effet, une même variable peut affecter la mortalité « autres causes » et la mortalité liée au cancer, et rendre cette hypothèse inacceptable. En pratique ces deux causes dépendent toujours au moins d'une covariable démographique qui est l'âge. Si ces facteurs démographiques ne sont pas pris en compte en amont dans l'analyse, l'approche cause spécifique n'estime pas correctement la survie nette.

Satten et Robins (Satten *et al.* 2001) ont proposé un estimateur pondéré afin de corriger les biais induits par la présence de censure dépendante dans le cadre de cette approche.

Il n'en reste pas moins que l'approche cause spécifique suppose la connaissance exacte de la cause principale de décès.

II.2.2. NON USAGE DE LA CAUSE DE DÉCÈS

II.2.2.1. Utilisation de la mortalité attendue

Il existe une alternative pour estimer la survie nette d'un groupe de patients sans toutefois utiliser les causes de décès. Le principe de cette approche est d'approximer la mortalité « autres causes » des individus par leur mortalité attendue s'ils étaient exempts du cancer. La mortalité attendue représente la mortalité « naturelle » auquel le patient est exposé en fonction de ses caractéristiques (âge, sexe par exemple). L'approximation de la mortalité attendue est basée sur l'utilisation de tables de mortalité de la population générale. Une table de mortalité fournit les taux de décès³ sur un intervalle de temps défini (en général par année) et en fonction de certaines caractéristiques telles que l'âge, le sexe et la zone géographique. La mortalité attendue d'un patient peut alors être estimée par le taux de décès correspondant dans la table de mortalité au regard de ses caractéristiques individuelles.

En France, les tables de mortalité sont stratifiées sur l'âge⁴, le sexe, le département et l'année de décès et sont fournies par l'Institut National de la Statistique et des Études Économiques (INSEE).

Afin d'éviter de trop fortes variations du taux de mortalité aux âges élevés, induites par des faibles effectifs, il est souhaitable d'utiliser une procédure de lissage de la table de mortalité afin d'atténuer ces variabilités observées.

Remarque 3. *Les tables de mortalité en population générale tiennent compte des décès dus à la maladie d'étude puisque toutes les causes de décès sont prises en compte. Néanmoins il a été montré que cela n'affectait que peu l'estimation de la survie nette en pratique (Talback et*

³ Dans une table de mortalité, le taux de décès est égal au nombre de personnes décédées de toutes causes confondues divisé par l'effectif de la population à risque.

⁴ Par classe de 1 an (allant de 0 à 99 ans).

Dickman 2011). *L'hypothèse est donc faite que la mortalité induite par la maladie d'étude reste minimale par rapport à la mortalité toutes causes confondues. Cette hypothèse est tout de même mise en défaut dans le cas des sujets âgés atteints du cancer de la prostate (Jooste et al. 2013).*

Remarque 4. *L'utilisation de tables de mortalité stratifiées sur d'autres facteurs pronostiques pouvant avoir un impact sur la mortalité a été discutée à plusieurs reprises (Dickman et al. 1997, 1998). Une mauvaise adéquation des tables de mortalité peut être à l'origine de certaines différences retrouvées dans des comparaisons de survie nette entre plusieurs groupes. Des études de simulation ont montré que le manque de stratification des tables de mortalité sur un facteur influençant la mortalité « autres causes » biaise les estimations des effets de cette même variable sur la mortalité due au cancer et, de façon moins marquée, inhibe les estimations des effets des autres variables (Grafféo et al. 2012). Il serait souhaitable dans l'avenir de disposer en France de tables de mortalité stratifiées sur d'autres facteurs pronostiques permettant ainsi de corriger les biais induits par un manque d'information dans ces tables. Ce travail méthodologique s'intègre dans un axe de recherche du projet CENSUR.*

II.2.2.2. Estimation de la mortalité due au cancer

La première approche proposée pour estimer la survie nette d'un groupe de patients malades en se basant sur leur mortalité attendue est une approche probabiliste (Berkson 1942 ; Berkson et Gage 1950). La survie relative au temps t , $S_{relative}(t)$, se définit comme le ratio entre la survie observée au temps t , $S_o(t)$, et la survie attendue au temps t , $S_a(t)$:

$$S_o(t) = S_a(t) * S_{relative}(t)$$

$$\Leftrightarrow S_{relative}(t) = \frac{S_o(t)}{S_a(t)}$$

La survie attendue, S_a , correspond à la survie d'un groupe de patients sains ayant les mêmes caractéristiques démographiques que le groupe de patients malades. Cette survie attendue est alors calculée via les tables de mortalité en population générale.

Remarque 5. Les estimateurs basés sur ce ratio de survie sont regroupés sous le terme de « Ratio-estimate ». Pendant longtemps la survie relative a été utilisée comme outil pour estimer la survie nette. Cependant, plusieurs auteurs (Danieli *et al.* 2012 ; Perme *et al.* 2012) ont montré que cette survie relative apparaît comme un estimateur de la survie nette seulement dans le cas précis où la censure est indépendante. Nous y reviendrons dans la suite du chapitre.

De façon analogue, cette relation entre survie nette, observée et attendue peut être écrite en termes de risque instantané de mortalité. Le taux instantané de mortalité observée au temps t noté $\lambda_o(t)$, est égal à la somme du taux instantané de mortalité attendue à t noté $\lambda_a(t)$ et du taux de mortalité en excès dû au cancer à t noté $\lambda_{excès}(t)$:

$$\lambda_o(t) = \lambda_a(t) + \lambda_{excès}(t)$$

Le taux de mortalité observée peut être vu comme la résultante de deux forces de mortalité : la mortalité en excès due à la maladie d'intérêt et la mortalité « autres causes ». Si le taux de mortalité en excès est nul alors les sujets « malades » auront une mortalité identique à celle des sujets « non malades » ayant des caractéristiques similaires. Les modèles statistiques utilisant cette relation additive entre le taux de mortalité attendue et le taux de mortalité en excès sont regroupés sous le terme de modèle dit « du taux en excès » (Estève *et al.* 1990 ; Giorgi *et al.* 2003 ; Dickman *et al.* 2004 ; Nelson *et al.* 2007 ; Remontet *et al.* 2007 ; Mahboubi *et al.* 2011)

Remarque 6. Certains facteurs peuvent entraîner un taux instantané négatif de mortalité due à la maladie. Dans cette situation, les individus « malades » ont une meilleure survie que des individus « non malades » ayant les mêmes caractéristiques. Cela est plausible car les patients peuvent être suivis plus régulièrement par leur médecin, ce qui peut affecter la mortalité attendue. Même si cette situation peut être observable, plusieurs estimateurs de survie nette apportent une contrainte de positivité au taux instantané de mortalité due à la maladie.

Remarque 7. D'autres auteurs (Berkson et Gage 1950 ; Hill *et al.* 1985 ; Andersen et Vaeth 1989) se basent sur une relation multiplicative des risques de mortalité entre la mortalité

attendue λ_a , et la mortalité due à la maladie, qui est vue comme un risque relatif $\lambda_{relatif}$. La mortalité observée est alors égale à: $\lambda_o(t) = \lambda_a(t) * \lambda_{relatif}(t)$. Ces modèles apportent moins de restrictions en termes mathématiques que les modèles additifs (Stare et al. 2005). Néanmoins, une relation additive apparaît plus plausible biologiquement dans l'étude du pronostic du cancer en population (Buckley 1984 ; Estève et al. 1990 ; Bolard et al. 2001). Récemment, Elie et al. (Elie et al. 2014) ont proposé un modèle combinant à la fois un taux de mortalité en excès et un taux relatif de mortalité due à la maladie.

II.2.3. DOMAINES D'APPLICATION

En recherche clinique, la survie toutes causes reste l'approche la plus utilisée. La cause de décès étant recueillie, la survie dite spécifique, qui consiste à censurer les décès autres causes, est également souvent rapportée dans le but d'approcher la survie nette. Cependant il a été montré que la méthode d'estimation de la survie spécifique n'était pas satisfaisante pour estimer la survie nette en raison de la censure induite par les décès liés aux autres causes.

Enfin la survie sans progression ou sans événement est un indicateur majeur pour les études de recherche clinique. La survie nette estimée dans le cadre précédemment décrit, c'est-à-dire n'utilisant pas la cause de décès mais la « mortalité attendue », n'est pas ou très peu utilisée en recherche clinique. Cela fait d'ailleurs partie d'un axe de recherche du programme CENSUR.

L'estimation de la survie nette reste à ce jour utilisée essentiellement dans le cadre des études menées en population générale. Elle permet, comme indiqué précédemment, de comparer entre pays et/ou entre périodes de diagnostic, la mortalité en excès liée au cancer. En outre, ces études populationnelles permettent des estimations à long terme.

II.3. APPROCHE NON PARAMÉTRIQUE POUR ESTIMER LA SURVIE NETTE

La particularité de la statistique non paramétrique en analyse de survie réside dans le fait qu'aucune hypothèse n'est faite quant à la loi de probabilité réelle des observations pour la

fonction de survie. Ainsi, l'approche non paramétrique permet d'éviter les erreurs qui peuvent être causées par le mauvais choix de la fonction de probabilité.

II.3.1. CONTEXTE

Les premiers estimateurs de la survie nette basés sur l'utilisation des tables de mortalité sont les « ratio-estimates » qui reposent, comme vu précédemment, sur le ratio entre la survie observée et la survie attendue du collectif étudié. Ederer et Heise (Ederer F. and Heise H. 1959), Ederer (Ederer *et al.* 1961) et Hakulinen (Hakulinen 1982) ont proposé trois estimateurs non paramétriques qui se différencient les uns des autres par le calcul de la probabilité de survie attendue, la survie observée étant quant à elle estimée à l'aide d'un outil de survie toutes causes tel que la méthode actuarielle ou la méthode de Kaplan-Meier (Böhmer 1912 ; Kaplan et Meier 1958).

Tout comme pour l'approche cause-spécifique, ces « ratio-estimates » ne prennent pas en compte la présence éventuelle de censure dépendante induite par des risques de décès concurrents.

Pour expliquer ce mécanisme de censure déjà abordé au paragraphe II.1.1, prenons l'exemple d'un individu qui décède dans la fenêtre de suivi. Pour cet individu deux temps d'événements sont susceptibles d'être observables : le temps de décès dû au cancer et le temps de décès « autres causes ». Ces deux états étant irréversibles, seul le temps le plus proche du diagnostic sera observé. Considérons que pour notre individu, le premier temps de décès observé correspond à un décès « autres causes ». Ce patient va alors sortir de la population à risque, susceptible de décéder du cancer. La composition du collectif des personnes à risque se modifie donc avec le temps, car les patients ayant un risque éventuellement accru de décès par cancer sont exclus du groupe à risque. Il s'ensuit un biais dans l'estimation de la survie nette.

Remarque 8. *Il paraît absurde de vouloir à tout prix considérer qu'un patient est encore à risque de décéder de la maladie s'il est déjà décédé d'une autre cause. Rappelons au lecteur que le concept de survie nette reste une situation hypothétique et non observable en réalité. La survie nette représente néanmoins un indicateur clinique pertinent.*

Afin d'obtenir des estimations non biaisées de la survie nette, l'utilisation d'un estimateur qui puisse s'affranchir de l'hypothèse de censure indépendante est nécessaire.

II.3.2. APPORT DE L'ESTIMATEUR NON PARAMÉTRIQUE PROPOSÉ PAR POHAR-PERME ET AL.

En 2012, Pohar-Perme *et al.* (Perme *et al.* 2012) proposent un nouvel estimateur non paramétrique de la survie nette tenant compte de la présence éventuelle de censure non indépendante. Dans ce nouvel estimateur, la population à risque est corrigée au cours du suivi afin d'inclure tous les patients susceptibles d'être encore à risque de décéder du cancer, même s'ils sont déjà décédés d'une autre cause.

En survie nette, ce problème de censure dépendante a déjà été soulevé dans le cadre de l'approche cause spécifique par Robins et Satten (Robins 1993 ; Satten *et al.* 2001). Les auteurs proposent alors de corriger le groupe des patients à risque par une pondération correspondant à l'inverse de la probabilité de censure. Dans ce travail, la survie cause spécifique est estimée à l'aide de l'estimateur de Nelson-Aalen (Nelson 1972 ; Aalen 1978). Deux processus de comptage interviennent alors : le processus N qui correspond au processus de comptage du nombre d'événements et le processus Y qui correspond au processus de comptage du nombre de personnes à risque susceptibles de subir l'événement. Satten *et al.* ont proposé de pondérer simultanément ces deux processus pour chaque individu par leur temps de survie cause spécifique $S_{csi}(t)$. Pour un individu i ($i = 1, \dots, n$), le taux cumulé de mortalité cause spécifique $\hat{\Lambda}_{cs}^w$ pondéré par w , est alors défini comme :

$$\hat{\Lambda}_{cs}^w(t) = \int_0^t \frac{dN_{cs}^w(u)}{Y^w(u)}$$

$$\text{Avec } N_{cs}^w(t) = \sum_{i=1}^n N_{csi}^w(t) = \sum_{i=1}^n \frac{N_{csi}(t)}{S_{csi}(t)} \text{ Et } Y^w(t) = \sum_{i=1}^n Y_i^w(t) = \sum_{i=1}^n \frac{Y_i(t)}{S_{csi}(t)}$$

N_{cs}^w correspond au nombre de décès dus au cancer au temps t , pondéré par w . Y^w correspond au nombre de personnes à risque au temps t , pondéré par w .

La pondération par w sur la population à risque Y , augmente le nombre de patients à risque afin de tenir compte de la proportion attendue des patients qui pourraient ne pas être inclus. La pondération par w sur le nombre de décès N , augmente le nombre d'événements de la même manière. Cette pondération corrige essentiellement les problèmes liés à l'effet de l'âge sur les deux forces de mortalité. Comme cet effet est plus prononcé chez les sujets âgés, la pondération sera moindre chez les sujets jeunes. Le calcul de la pondération pouvant être obtenu directement par les données individuelles, aucune modélisation paramétrique n'est nécessaire.

Pohar-Perme *et al.* ont alors proposé de faire de même dans un cadre où il n'est pas nécessaire d'utiliser la cause de décès. Pour cela, l'expression du taux cumulé de mortalité attendue de l'estimateur Ederer II (Ederer F. and Heise H. 1959) est réécrite sous forme de processus de comptage et la même pondération par w est apportée à la fois sur le taux cumulé de mortalité observée $\widehat{\Lambda}_o$, et sur le taux cumulé de mortalité attendue $\widehat{\Lambda}_a$ afin d'estimer le taux cumulé de mortalité en excès $\widehat{\Lambda}_{\text{excès}}$. Le nouvel estimateur proposé par Pohar-Perme *et al.* se définit alors comme :

$$\widehat{\Lambda}_{\text{excès}}(t) = \widehat{\Lambda}_o(t) - \widehat{\Lambda}_a(t)$$

$$\Leftrightarrow \widehat{\Lambda}_{\text{excès}}(t) = \int_0^t \frac{dN^w(u)}{Y^w(u)} - \int_0^t \frac{\sum_{i=1}^n Y_i^w(u) d\Lambda_a(u)}{Y^w(u)}$$

Avec $N^w(t) = \sum_{i=1}^n N_i^w(t) = \sum_{i=1}^n \frac{N_i(t)}{S_{ai}(t)}$ et $Y^w(t) = \sum_{i=1}^n Y_i^w(t) = \sum_{i=1}^n \frac{Y_i(t)}{S_{ai}(t)}$

N^w correspond au nombre de décès observés au temps t , pondéré par w . Y^w correspond au nombre de personnes à risque au temps t , pondéré par w .

Remarque 9. Dans le cadre du projet CENSUR, Grafféo *et al.* ont présenté un test de comparaison de distribution de fonctions de survie attaché à l'estimateur non paramétrique de Pohar-Perme lors de la 34^{ème} conférence annuelle de l'International Society for Clinical Biostatistics à Munich en 2013.

II.3.3. DISCUSSION

Parmi les méthodes non paramétriques développées en analyse de survie, l'estimateur de Pohar-Perme reste le seul estimateur capable sur le plan théorique de fournir des estimations non biaisées de la survie nette.

Dans le cadre de la survie nette, Danieli *et al* (Danieli *et al.* 2012) se sont intéressés à la comparaison de différents estimateurs sur données simulées. L'objectif de ce travail était de comparer les estimateurs non paramétriques d'Ederer I, d'Ederer II, d'Hakulinen et l'approche cause spécifique à l'estimateur Pohar-Perme. Les auteurs montrent qu'en l'absence de censure dépendante liée à l'âge, tous les estimateurs fournissent des résultats similaires. En revanche, si une censure dépendante est introduite, alors ces estimateurs sous-estiment la mortalité en excès par rapport aux estimations obtenues avec l'estimateur Pohar-Perme.

Des études menées sur des données de population ont par la suite été publiées afin de mettre en évidence l'impact de l'utilisation de ces divers estimateurs sur la probabilité de survie nette. Roche *et al.* (Roche *et al.* 2013) ont montré des différences importantes de probabilités de survie à long terme (à 10 ans après le diagnostic) entre ces différents estimateurs, particulièrement pour les cancers de bon pronostic. Néanmoins sur une durée plus courte, Seppä *et al.* (Seppä *et al.* 2013) ont montré qu'à 5 ans après le diagnostic, les estimations n'étaient pas véritablement différentes si l'on compare les résultats obtenus avec l'estimateur Pohar-Perme et l'estimateur d'Ederer II.

Même s'il ne subsiste aucun doute quant à la consistance de l'estimateur de Pohar-Perme pour estimer la survie nette, cet estimateur reste néanmoins assez sensible à des évolutions brutales des taux de mortalité attendue. Dickman *et al.* (Dickman *et al.* 2013) ont mis en avant une variance plus élevée de l'estimateur Pohar-Perme par rapport à l'estimateur Ederer II associé à une standardisation interne⁵ sur l'âge, ce qui induit des intervalles de confiance assez larges en fin de suivi.

En France, le réseau des registres FRANCIM, en collaboration avec le service de biostatistique des HCL, ont fait le choix d'estimer la survie nette à 5 et 10 ans pour quarante-

⁵ Cette standardisation interne consiste à estimer les probabilités de survie à un temps fixé dans différentes classes d'âge prédéfinies. Des poids calculés sur la structure d'âge du collectif de départ sont ensuite utilisés pour pondérer les survies par classe d'âge et ainsi obtenir une probabilité de survie globale.

sept localisations de cancer en utilisant l'estimateur Pohar-Perme (Monnereau *et al.* 2013b ; Jooste *et al.* 2013).

Au niveau européen, les récentes estimations de survie publiées à partir d'EUROCORE-5 restent basées sur l'estimateur d'Ederer II (Sant *et al.* 2014). En revanche au niveau international, les estimations de survie publiées à partir du projet CONCORD-2 sont basées sur le nouvel estimateur proposé par Pohar-Perme (Allemani *et al.* 2015).

II.4. MODÉLISATION PARAMÉTRIQUE DE LA SURVIE NETTE

Une autre approche statistique pour l'analyse de la survie consiste en l'utilisation d'outils paramétriques. L'avantage de ces outils par rapport à des estimateurs non paramétriques est de permettre l'étude de l'association entre la mortalité du patient et des facteurs pronostiques en contrôlant l'effet d'autres variables simultanément.

II.4.1. CONTEXTE

Devant les besoins cliniques et épidémiologiques de pouvoir étudier l'effet de covariables sur la mortalité des patients, des modèles paramétriques ont été développés. Ces modèles de régression ont d'abord été développés dans un cadre de survie toutes causes, le modèle le plus utilisé étant le modèle semi-paramétrique de Cox (Cox 1972). Très vite est ensuite apparu le besoin d'étendre de tels modèles dans le concept de la survie nette (Breslow 1975; Buckley 1984; Hakulinen et Tenkanen 1987; Andersen et Vaeth 1989).

Une relation additive entre le taux de mortalité attendue et le taux de mortalité induit par la maladie étant plus plausible cliniquement, les modèles basés sur la modélisation du taux en excès se sont développés (Estève *et al.* 1990 ; Dickman *et al.* 2004). Ces modèles largement utilisés en épidémiologie reposent sur plusieurs hypothèses :

- l'hypothèse d'un risque de base⁶ constant par intervalle de temps prédéfini.

⁶ Le risque de base correspond au risque lorsque toutes les variables du modèle sont égales à la modalité de référence.

- l'hypothèse de proportionnalité des risques⁷ qui impose que le risque entre deux groupes est constant au cours du temps de suivi.
- l'hypothèse de log-linéarité pour les variables continues qui suppose un risque linéaire avec l'augmentation d'une unité de la variable.

Ces hypothèses sont assez fortes et ne sont pas toujours vérifiées en pratique clinique. Elles peuvent de ce fait impacter les résultats issus du modèle paramétrique. Aussi de nouveaux modèles, dits flexibles, ont été proposés pour lever ces hypothèses de base en survie toutes causes (Royston et Parmar 2002 ; Abrahamowicz et MacKenzie 2007) et de façon parallèle en survie nette (Giorgi *et al.* 2003 ; Nelson *et al.* 2007 ; Remontet *et al.* 2007 ; Mahboubi *et al.* 2011).

II.4.2. LE MODÈLE FLEXIBLE DU TAUX EN EXCÈS PROPOSÉ PAR REMONTET ET AL.

Dans le concept de la survie nette, Remontet *et al.* (Remontet *et al.* 2007) ont proposé un modèle flexible du taux de mortalité instantané en excès. En se basant sur l'utilisation de données individuelles⁸, ce modèle peut être vu comme une extension du modèle proposé par Estève *et al.* (Estève *et al.* 1990). Les estimations des paramètres du modèle sont obtenues en maximisant la log-vraisemblance complète⁹.

Dans les modèles du taux en excès, la modélisation intervient seulement sur le taux en excès $\lambda_{excès}$, le taux attendu λ_a étant fourni par les tables de mortalité en population générale. Nous avons alors :

$$\lambda_o(t, X) = \lambda_a(\hat{age} + t, Z) + \lambda_{excès}(t, X, \beta)$$

Avec, \hat{age} correspondant à l'âge au diagnostic, t au temps de suivi, Z étant les variables d'ajustement de la table de mortalité. β représente le vecteur des paramètres du modèle et X représente le vecteur des covariables du modèle. A noter que X inclut les covariables Z et l'âge.

⁷ Nous parlons alors de modèle de risques proportionnels.

⁸ A l'inverse des approches utilisant les modèles linéaires généralisés qui se basent sur des données groupées et qui nécessitent de catégoriser l'ensemble des covariables du modèle (Hakulinen et Tenkanen 1987; Dickman *et al.* 2004)

⁹ Dans les modèles du taux en excès, le calcul de la vraisemblance ne fait intervenir que le taux de mortalité attendue des patients décédés.

II.4.2.1. Écriture du modèle

Le modèle proposé par Remontet *et al.* (Remontet *et al.* 2007) se base sur la modélisation du logarithme népérien du taux de mortalité instantané en excès. En l'absence de covariables, le modèle s'écrit :

$$\ln[\lambda_{\text{excès}}(t)] = f(t)$$

$f(t)$ correspond au risque de base. Le choix de la fonction f est laissé libre à l'utilisateur et peut correspondre à une fonction plus ou moins souple au regard du jeu de données. En cancérologie, comme le risque de décéder peut être plus fort en début de suivi, il est préférable d'utiliser une fonction suffisamment flexible afin de modéliser une dynamique du taux de base non linéaire. L'utilisation de splines¹⁰ de régression apparaît alors comme un bon outil pour approcher des fonctions ayant un caractère non linéaire.

Dans le cadre de nos deux applications présentées aux *chapitres III et IV*, le risque de base a été modélisé de façon continue jusqu'à 5 ans de suivi, à l'aide d'une spline de régression cubique avec un nœud à 1 an. En cancérologie, une mortalité élevée est souvent observée durant la première année de suivi. Nous avons alors positionné notre nœud à l'endroit où une courbure était attendue. La modélisation portant sur 5 ans, il ne nous a pas paru nécessaire d'ajouter un nœud supplémentaire au cours du suivi. Notre choix *a priori* s'est basé sur des travaux menés sur les données du réseau FRANCIM. (Remontet *et al.* 2007). Dans ce travail portant sur 18 localisations de cancer, les auteurs ont comparé la vraisemblance de plusieurs modèles dans lesquels l'emplacement du premier nœud variait, passant à 1 an, 1 an ½, 2 ans et 3 ans de suivi. L'emplacement à 1 an de suivi apparaissait alors comme le meilleur choix et ce pour l'ensemble des localisations de cancer.

¹⁰ Une fonction spline est basée sur une régression polynômiale (de degré variant) par morceaux (dans des intervalles dont les extrémités sont appelées nœuds) en imposant des conditions de continuité, de pente et de courbure aux frontières des intervalles.

Remarque 10. La modélisation du risque de base pour nos deux applications fait appel à des splines paramétriques, i.e. dit de régression (dépendant d'un ensemble de paramètres à l'inverse des splines non paramétriques, i.e. dit de lissage). Un avantage des splines de régression est d'apporter une bonne flexibilité à la fonction afin de s'adapter aux données tout en conservant une relation linéaire entre les paramètres. L'utilisation de telles splines nécessite de fixer au préalable :

- l'ordre de la spline qui permet de déterminer les propriétés de continuité de l'estimation. Des splines fréquemment utilisées sont les splines cubiques car elles représentent un bon compromis entre souplesse et lissage.

- le nombre de nœuds qui permet de contrôler la flexibilité de la fonction et qui dépend de la complexité du phénomène étudié. L'augmentation du nombre de nœuds offre une plus grande flexibilité mais peut induire un sur ajustement des données.

- la position des nœuds. Le choix de l'emplacement peut être guidé par la structure des événements observés sur les données ou être spécifié à l'aide d'études antérieures. Afin de s'assurer de cela, plusieurs modèles se basant sur des splines de natures différentes peuvent être comparés (à l'aide du critère d'information d'Akaike par exemple).

Certains auteurs (Bolard et al. 2002 ; Giorgi et al. 2003 ; Nelson et al. 2007) s'orientent vers l'utilisation de splines naturelles qui bénéficient de l'ajout de deux nœuds supplémentaires (aux extrémités) en imposant une contrainte de linéarité au-delà de ces nœuds extrêmes. Cette contrainte permet d'éviter à la fonction d'être trop dépendante de valeurs extrêmes notamment en fin de suivi. En général, l'emplacement des nœuds pour des splines naturelles est basé sur les quantiles de la distribution des temps de décès observés afin d'avoir approximativement le même nombre d'événements dans chacun des intervalles de temps (Abrahamowicz et al. 1996).

Une fois le risque de base modélisé, des covariables peuvent ensuite être incluses dans le modèle de Remontet *et al.* dans lequel une relation additive entre les covariables est supposée. Le modèle permet de tenir compte simultanément des effets non linéaires et /ou non proportionnels des covariables.

II.4.2.2. Modélisation de l'effet complexe des covariables

II.4.2.2.1 Modélisation des effets non linéaires

L'effet d'une covariable continue est linéaire si le rapport entre deux taux évalués à X et $X + 1$, ne dépend pas de la valeur de X . Par exemple si l'on souhaite étudier l'effet de la variable âge sur le taux de mortalité en excès, l'augmentation d'une année de l'âge aura le même effet sur le risque de mortalité chez les patients passant de 30 à 31 ans que chez les patients passant de 65 à 66 ans. La progression du risque n'est cependant pas toujours linéaire avec l'augmentation de la variable, induisant ainsi un effet non linéaire sur la mortalité.

Un exemple d'effet linéaire (A) ou non linéaire (B) de l'âge sur le taux de mortalité en excès est présenté sur la Figure II.3 à l'aide du taux relatif de mortalité en fonction de l'âge au diagnostic. Ce taux relatif correspond au ratio de risque entre un âge fixé et l'âge de référence (égal dans notre exemple à 60 ans). Par exemple pour des patients de 80 ans, le ratio de risque est égal au taux de mortalité en excès des sujets de 80 ans divisé par le taux de mortalité des sujets de 60 ans.

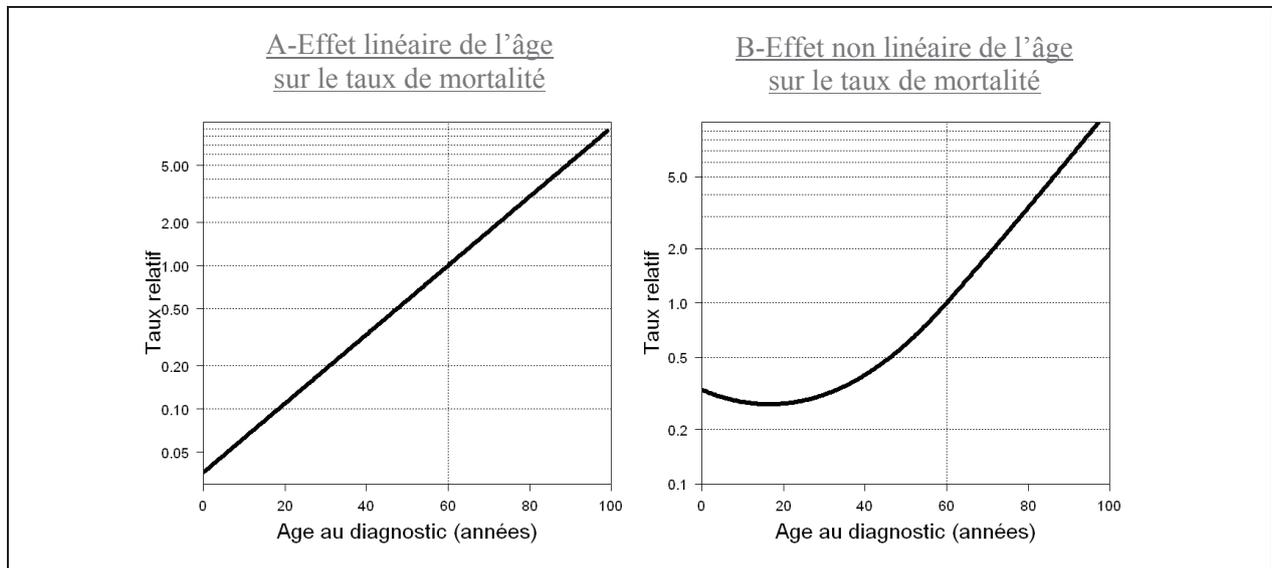


Figure II.3 - Effet linéaire (A) ou non linéaire (B) de l'âge sur le taux relatif de mortalité en excès, l'âge de référence étant fixé à l'âge moyen (60 ans).

Cette problématique de linéarité est en fait la même que celle rencontrée précédemment pour modéliser le risque de base de façon continue.

En supposant un effet non linéaire de l'âge sur le taux de mortalité en excès, le modèle de Remontet *et al.* se définit alors comme :

$$\ln[\lambda_{\text{excès}}(t, \hat{\text{age}})] = f(t) + g(\hat{\text{age}})$$

L'effet non linéaire de la variable est modélisé par la fonction $g(\hat{\text{age}})$, g correspondant à une fonction flexible, telle qu'une spline de régression, la forme de la fonction étant laissée libre.

Le ratio de risque entre deux âges sera alors défini comme:

$$\frac{\lambda_{\text{excès}}(t, \hat{\text{age}}_1)}{\lambda_{\text{excès}}(t, \hat{\text{age}}_0)} = \frac{\exp(f(t) + g(\hat{\text{age}}_1))}{\exp(f(t) + g(\hat{\text{age}}_0))} = \exp(g(\hat{\text{age}}_1) - g(\hat{\text{age}}_0))$$

La fonction du ratio de risque en excès dépendant alors des valeurs originelles de l'âge, l'effet de l'âge est dans ce cas modélisé de façon à pouvoir tenir compte d'un éventuel effet non linéaire (Figure II.3 (B)).

II.4.2.2.2 Modélisation des effets dépendants du temps

L'effet d'une covariable (qu'elle soit continue ou catégorielle) sur le taux de mortalité est proportionnel si le rapport entre deux taux évalués à X et $X + 1$ ne dépend pas du temps. En d'autres termes, si l'on s'intéresse au ratio de risque entre deux groupes alors le risque est toujours le même à un an ou à 5 ans après le diagnostic. Le risque entre ces deux groupes est alors considéré comme constant tout au long du temps de suivi. Cependant certains facteurs pronostiques peuvent avoir un effet à court terme ou bien à plus long terme induisant ainsi un effet non proportionnel sur la mortalité.

Un exemple d'effet proportionnel (A) ou non proportionnel (B) de l'âge sur le taux de mortalité en excès est présenté sur la figure II.4. Cette figure représente le taux relatif de décès entre deux âges (correspondant ici au taux de mortalité chez les sujets de 30 ans par rapport au taux de mortalité chez les sujets de 60 ans) en fonction du temps de suivi depuis le diagnostic.

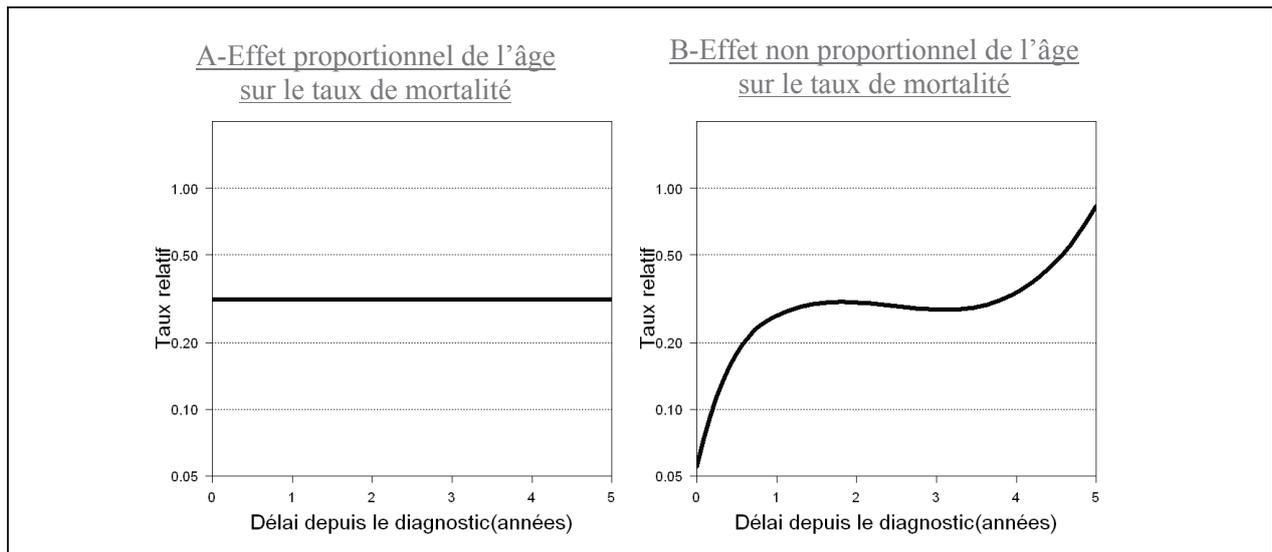


Figure II.4 - Effet proportionnel (A) ou non proportionnel (B) de l'âge sur le taux relatif de mortalité pour des patients de 30 ans par rapport à l'âge de référence, 60 ans.

Afin de considérer un effet non proportionnel de l'âge sur le taux de mortalité (Figure II.4 (B)), une interaction entre la variable et une fonction dépendante du temps est ajoutée dans le modèle qui s'écrit alors :

$$\ln[\lambda_{\text{excès}}(t, \hat{\text{age}})] = f(t) + \hat{\text{age}} + h(t) * \hat{\text{age}}$$

L'effet non proportionnel de l'âge est modélisé par l'interaction $h(t) * \hat{\text{age}}$, h correspondant à une fonction spline ayant en général les mêmes caractéristiques que la fonction f servant à modéliser le risque de base. Le ratio de risque entre deux âges sera alors :

$$\frac{\lambda_{\text{excès}}(t, \hat{\text{age}}_1)}{\lambda_{\text{excès}}(t, \hat{\text{age}}_0)} = \exp(\hat{\text{age}}_1 - \hat{\text{age}}_0 + h(t)(\hat{\text{age}}_1 - \hat{\text{age}}_0))$$

Le rapport de risque entre $\hat{\text{age}}_1$ et $\hat{\text{age}}_0$ dépendant de $h(t)$, l'effet de l'âge sur la mortalité en excès est alors considéré comme pouvant être non proportionnel (Figure II.4 (B)).

Dans nos deux applications basées sur données réelles, nous avons choisi d'utiliser une spline cubique avec un nœud à 1 an de suivi pour modéliser l'effet non proportionnel des covariables, l'âge et l'année de diagnostic, tout comme pour modéliser le risque de base.

II.4.2.2.3 Modélisation conjointe des effets non linéaires et dépendants du temps

Il arrive qu'une variable continue puisse avoir simultanément un effet non linéaire et non proportionnel sur le taux de mortalité. Une telle situation est visualisée sur la figure II.5 qui représente le taux relatif de décès en fonction de l'âge et du délai depuis le diagnostic.

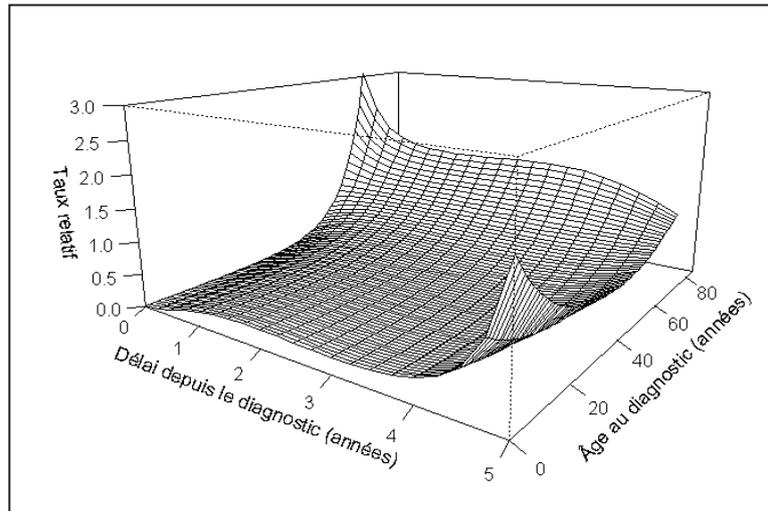


Figure II.5 - Effet non linéaire et non proportionnel de l'âge sur le taux relatif de mortalité à la suite d'un diagnostic de LF, l'âge de référence étant fixé à l'âge moyen 60 ans.

Dans cet exemple portant sur la mortalité en excès à la suite d'un diagnostic de LF, l'âge a un effet non linéaire à temps fixé (Figure II.5). A 3 ans après le diagnostic par exemple, le taux relatif de décéder est relativement stable jusqu'à 60 ans et augmente ensuite chez les patients plus âgés. L'âge a également un effet non proportionnel avec une mortalité précoce observée chez les sujets âgés et une mortalité plus importante en fin de suivi chez les sujets jeunes (Figure II.5). Dans cette situation, il apparaît alors indispensable de prendre en compte les effets complexes de l'âge dans la modélisation du taux de mortalité. En considérant conjointement un effet non linéaire et non proportionnel de l'âge sur le taux de mortalité en excès, le modèle de Remontet *et al.*, s'écrit alors :

$$\ln[\lambda_{\text{excès}}(t, \text{âge})] = f(t) + g(\text{âge}) + h(t) * \text{âge}$$

A partir de l'écriture du modèle, nous pouvons estimer le rapport de risque entre deux âges qui est égal à :

$$\frac{\lambda_{excès}(t, \hat{age}_1)}{\lambda_{excès}(t, \hat{age}_0)} = \exp(g(\hat{age}_1) - g(\hat{age}_0) + h(t)(\hat{age}_1 - \hat{age}_0))$$

Ce rapport de risque dépendant à la fois des valeurs originelles de l'âge mais également de $h(t)$, l'effet de l'âge est dans ce cas modélisé de façon à pouvoir être non linéaire et non proportionnel (Figure II.5).

Lorsque la spline est utilisée pour modéliser l'effet de facteur pronostic, comme l'âge, la nature de l'effet ne peut être connue en amont de l'étude. De ce fait, l'emplacement des nœuds doit donc être décidé arbitrairement. Aussi, afin de modéliser un effet non linéaire de l'âge et de l'année de diagnostic sur la mortalité lors de nos deux applications présentées aux *chapitres III et IV*, nous avons utilisé une spline quadratique à un nœud. Pour l'effet de l'âge, le nœud a été fixé à la valeur moyenne de l'âge pour chacune des cohortes étudiées. Pour l'effet de l'année de diagnostic, le nœud a été fixé à l'année 1998 (pour les LF) ou à 2002 (pour les LBDGC) afin de marquer l'arrivée de la thérapie ciblée dans la prise en charge des lymphomes.

II.4.2.3. La stratégie de construction du modèle flexible

L'effet d'une variable sur la mortalité peut être de différentes natures. Cette variable peut n'avoir aucun effet sur la mortalité ou bien avoir un effet proportionnel/non proportionnel et/ou avoir un effet linéaire/non linéaire pour les variables continues.

Alors que l'effet réel de la variable n'est pas connu en amont de l'analyse, chaque effet doit faire l'objet d'une étude spécifique afin d'ajuster correctement le modèle aux données observées. Il a été montré qu'un modèle peut entraîner des résultats faussés dès lors que l'effet d'une variable n'est pas correctement pris en compte dans la modélisation (Quantin *et al.* 1999).

Certains auteurs (Abrahamowicz et MacKenzie 2007) ont également mis en évidence la difficulté engendrée par le fait que la non prise en compte d'un effet non linéaire pouvait induire à tort un effet proportionnel pour cette même variable (et inversement). Ainsi,

Abrahamowicz *et al.* ont suggéré de tester simultanément l'effet non linéaire et non proportionnel d'une variable continue en s'orientant sur un arbre de décision à deux branches. La figure II.6 présente un exemple de la stratégie de modélisation d'une variable pronostique, l'âge, en se basant sur le modèle de Remontet *et al.* qui suppose une relation additive entre les effets des variables. Cinq modèles emboîtés (dans lesquels la variable continue est modélisée avec un effet différent dans chacun des modèles) sont alors comparés à l'aide d'un test statistique de comparaison (un test du rapport de vraisemblance par exemple) Pour déterminer quel modèle doit être retenu parmi les cinq, une sélection pas à pas descendante est utilisée.

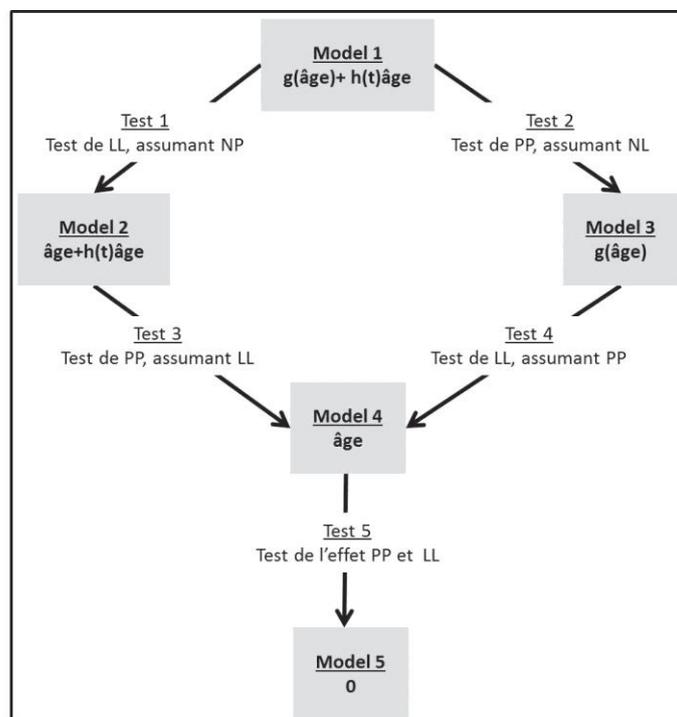


Figure II.6 - Stratégie de modélisation des effets complexes d'une variable continue (l'âge par exemple) proposée par Abrahamowicz & MacKenzie, en considérant le modèle du taux en excès de Remontet *et al.*

L'idée de cette stratégie est de débiter avec le modèle le plus complexe (Modèle 1) qui suppose un effet non linéaire et non proportionnel (NL/NP) de la variable, puis d'éliminer dans le modèle, les effets de la variable qui ne sont pas significatifs au seuil de significativité fixé en général à 0,05.

- Dans un premier temps, en partant du modèle 1 nous testons la significativité de l'effet non linéaire (NL) (Test 1) puis celle de l'effet non proportionnel (NP) (Test 3) et en enfin la significativité de l'effet proportionnel et linéaire (PP-LL) (Test 5) de la variable (Figure II.6).

- Dans un second temps, en partant du modèle 1, nous testons la significativité de l'effet NP (Test 2) puis celle de l'effet NL (Test 4) pour terminer par la significativité de l'effet PP-LL (Test 5) de la variable continue.

A la suite de cette stratégie, deux situations peuvent être observées :

- les deux branches mènent au même modèle.
- les deux branches ne permettent pas de conclure à un modèle final unique. Cette situation peut se produire si le test de la linéarité ne conduit pas au même résultat selon qu'il est évalué en considérant un effet NP ou PP ; ou si le test de la proportionnalité ne conduit pas à la même conclusion selon qu'il est évalué en supposant un effet NL ou LL.

Lors de l'application de cette stratégie de construction du modèle sur les données européennes présentée au *Chapitre IV*, les deux branches ont conduit au même modèle à l'exception de l'Ecosse/Pays de Galles pour les LBDGC pour lesquels un effet non linéaire de l'année était retrouvé significatif en supposant un effet NP alors que ce même effet était non significatif dans le modèle supposant un effet PP. Nous suggérons dans cette situation, de conserver les deux effets NP et NL de la variable dans le modèle final.

Wynant *et al.* ont également proposé une méthode bootstrap afin d'évaluer la stabilité du modèle sélectionné par la procédure pas à pas descendante (Wynant et Abrahamowicz 2014).

La stratégie se complique si l'on souhaite tenir compte dans le modèle de l'effet complexe de plusieurs variables simultanément. Wynant *et al.* (Wynant et Abrahamowicz 2014) se sont intéressés à la stratégie de construction d'un tel modèle en proposant six stratégies dans la situation où plusieurs variables pronostiques doivent être modélisées.

Nous verrons par la suite que l'application présentée au *Chapitre IV* de ce travail fait appel à l'une des stratégies proposées par Wynant et Abrahamowicz afin de modéliser simultanément les effets complexes de deux variables pronostiques, qui sont l'âge et l'année de diagnostic. La stratégie consistait alors à tester les effets complexes d'une variable (Figure II.6) tout en ajustant sur un effet complexe (NP-NL) des autres variables du modèle.

II.4.3. DISCUSSION

Danieli *et al.* (Danieli *et al.* 2012) ont montré que dans la « boîte à outils » actuelle des estimateurs de la survie nette, seules deux approches fournissent une estimation non biaisée de la survie nette : l'approche non paramétrique proposée par Pohar-Perme *et al.* et l'approche paramétrique avec ajustement sur les variables démographiques de la table de mortalité (essentiellement l'âge).

Une approche non paramétrique ne nécessite pas de stratégie de modélisation, mais ne permet pas d'étudier l'effet conjoint de variables sur la mortalité. Aussi, si l'intérêt réside dans la modélisation des effets de variables sur la mortalité alors l'utilisation d'un estimateur paramétrique est à privilégier. Comme certaines variables peuvent avoir un effet complexe sur la mortalité, il est nécessaire d'utiliser des outils paramétriques tels que les modèles flexibles qui prennent en compte de tels effets.

Les variables continues, telles que l'âge, sont souvent catégorisées dans les modèles de régression donc le ratio de risque peut évoluer de façon non linéaire entre les catégories. Néanmoins, l'hypothèse de log-linéarité reste assumée au sein de chaque catégorie mais si ces catégories sont larges alors cette hypothèse semble peu plausible. Il s'avère alors plus pertinent de conserver le caractère continu de la variable dans la modélisation tout en s'assurant de la forme linéaire ou non de son effet sur la mortalité. Une stratégie de modélisation est alors essentielle afin de tester conjointement les effets non proportionnel et/ou non linéaire de la variable pour éviter que la mauvaise spécification d'un effet entraîne une mauvaise estimation d'un autre effet.

Une des difficultés dans l'utilisation de modèles flexibles réside dans le choix des fonctions souples utilisées pour modéliser de façon continue à la fois l'effet du temps de suivi et les effets des variables sur le taux de mortalité en excès. L'utilisation de spline de régression nécessite de définir le degré de la spline, le nombre et la position des nœuds. A notre connaissance, il n'existe pas véritablement de consensus quant à l'emplacement et quant aux nombres de nœuds, la décision se base alors sur des fondements cliniques et sur des connaissances à priori.

Le risque de base étant modélisé de façon continue au cours du temps de suivi dans le modèle de Remontet *et al.* (Remontet *et al.* 2007), ce dernier permet d'estimer la dynamique du taux instantané de mortalité en excès et la probabilité de survie nette de la population d'étude.

Dans le cadre des modèles du taux en excès, un autre modèle flexible a été proposé en 2011 (Mahboubi *et al.* 2011). Ce modèle est sensiblement identique à celui proposé par Remontet *et al.* à la différence près qu'une relation multiplicative entre les covariables du modèle est préférée à une relation additive. Une étude a été conduite dans le cadre du projet MESURE afin de comparer les estimations issues de ces deux modèles sur un jeu de données réelles et a montré que les résultats obtenus étaient sensiblement identiques avec ces deux modèles. Il est à noter que le choix d'une relation additive entre les covariables permet d'éviter des problèmes d'identifiabilité qui peuvent être rencontrés lorsque la relation est définie comme multiplicative (Abrahamowicz et MacKenzie 2007 ; Mahboubi *et al.* 2011).

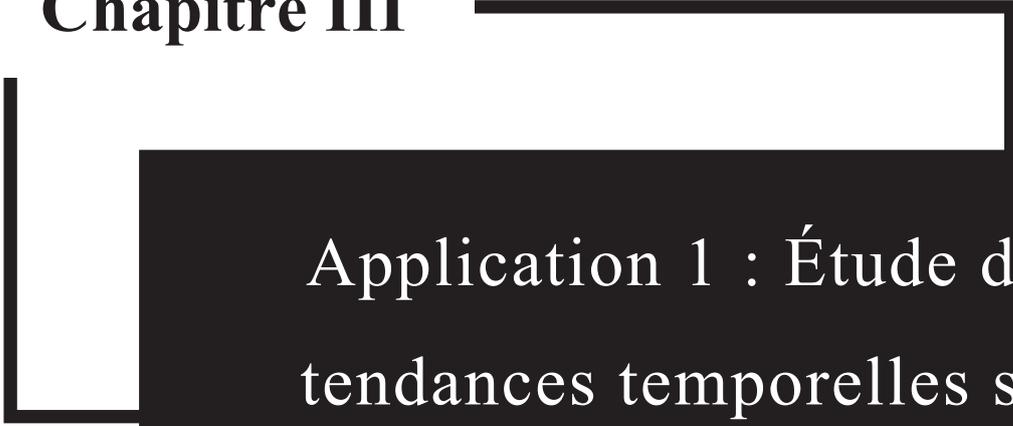
Alors que les deux modèles précités se basent sur la modélisation du taux de mortalité instantané en excès, d'autres auteurs ont proposé de modéliser la fonction de mortalité cumulée en excès à l'aide d'un modèle flexible (Nelson *et al.* 2007). Cette approche permet de faciliter les estimations des paramètres et ainsi limite le temps de calcul. Cependant l'interprétation des effets des covariables du modèle est plus complexe car la modélisation du taux cumulé engendre des interactions entre ces variables sans qu'elles ne soient clairement définies dans l'écriture du modèle (Nelson *et al.* 2007).

Dans ce travail, deux applications du modèle de Remontet *et al.* associant les effets complexes de l'âge et de l'année de diagnostic ont été réalisées sur des bases de données de population.

Le Chapitre III présente une application basée sur les données françaises du réseau FRANCIM à partir desquelles nous avons estimé les tendances de la dynamique du taux de mortalité en excès à la suite d'un diagnostic de LF en fonction de l'âge, de l'année de diagnostic et du stade de la maladie.

La deuxième application, qui fait l'objet du Chapitre IV, a porté sur les données européennes issues d'EUROCARE-5 à partir desquelles nous avons estimé les tendances de la survie nette et de la dynamique du taux de mortalité en excès à la suite d'un diagnostic de LF ou d'un LBDGC, et ceci en fonction de l'âge et de l'année de diagnostic pour cinq zones européennes.

Chapitre III



Application 1 : Étude des tendances temporelles sur données françaises

III.1. RATIONNEL

De grands groupes collaboratifs ont mis en évidence l'importance de dissocier dans les études les sous-types de LMNH étant donnée la variété de leurs caractéristiques épidémiologiques et cliniques (Sant *et al.* 2008). Pour la première fois en France, le réseau français des registres de cancers, FRANCIM, a publié des résultats de survie nette par sous-type de lymphome sur la période de diagnostic allant de 1995 à 2004 (Monnereau *et al.* 2013b).

Ces résultats ont rapporté une amélioration de la survie au cours des périodes de diagnostic, amélioration plus ou moins prononcée en fonction des sous-types de lymphome. Ces évolutions positives avec les périodes de diagnostic récentes ont été interprétées, entre autres, par le rôle des avancées thérapeutiques dans la prise en charge des LMNH avec notamment le développement des anticorps monoclonaux et l'utilisation de l'immunothérapie. Plusieurs grands essais cliniques ont rapporté des résultats positifs en termes de survie chez les patients qui recevaient ce nouveau traitement en association au traitement de référence par rapport aux patients ne recevant que le traitement de référence (Coiffier *et al.* 1998 ; Salles *et al.* 2008). Ce sont les LF qui ont été les premiers à bénéficier de cette nouvelle thérapeutique qui a obtenu son AMM en France en 1998 d'abord en traitement de consolidation et ensuite à partir de 2002 en traitement de première intention. Alors que les résultats d'essais cliniques ne peuvent être généralisés à l'ensemble des patients atteints d'un LF du fait de critères d'inclusion en amont de l'étude, il apparaît intéressant de compléter ces résultats par des études en population générale afin de mesurer l'impact réel de ces changements majeurs de prise en charge sur la mortalité des patients. Afin d'estimer l'évolution de la mortalité due au cancer, la survie nette apparaît alors comme l'outil adéquat.

Comment ces évolutions de prise en charge du lymphome ont-elles modifié la mortalité induite par le LF en population générale ? Ont-elles eu un effet à court terme ou à plus long terme sur cette mortalité ?

L'objectif de ce travail était de compléter les précédents résultats de survie nette publiés en France pour les LF (Monnereau *et al.* 2013b) en estimant la tendance de la dynamique du taux de mortalité en excès en fonction de l'année de diagnostic, ajustée sur l'âge au diagnostic et le stade du lymphome, chez des patients diagnostiqués entre 1995 à 2010. L'année de

diagnostic était considérée comme une mesure indirecte de l'évolution de la prise en charge des LF au cours du temps, l'année 1998 marquant l'arrivée de l'immunothérapie dans la prise en charge des patients.

III.2. DONNÉES DE SURVIE EN POPULATION FRANÇAISE

Deux études de survie basées sur des cohortes historiques ont été menées en utilisant la base commune du réseau FRANCIM.

- La première analyse portait sur les données de huit registres des cancers comprenant 1477 cas de LF diagnostiqués entre le 1^{er} janvier 1995 et le 31 décembre 2004 chez des patients de plus de 15 ans. Le statut vital des patients était actualisé au 1^{er} janvier 2008. Seuls les registres de cancer disposant d'un recul suffisant étaient inclus dans l'étude. Les codes ICD-O-3 retenus pour les LF étaient les codes 9690/3, 9691/3, 9695/3 et 9698/3 respectivement pour les LF sans autres informations, de grade 2, de grade 1 et de grade 3.
- La deuxième analyse a porté spécifiquement sur les données de deux registres des cancers issus de la première analyse dans lesquels le stade Ann-Arbor au diagnostic du lymphome était renseigné. Cette deuxième analyse a alors porté sur 451 LF diagnostiqués entre le 1^{er} janvier 1995 et le 31 décembre 2010 chez des patients de plus de 15 ans. Le statut vital des patients était actualisé au 1^{er} janvier 2013.

Pour ces deux analyses, les taux lissés de mortalité attendue en population générale étaient disponibles par sexe, âge, année et par département géographique français.

III.3. MÉTHODOLOGIE

Le taux de mortalité en excès ainsi que les effets des variables ont été modélisés jusqu'à 5 ans après le diagnostic en se basant sur le modèle proposé par Remonet *et al.* (Remonet *et al.* 2007). Les deux bases de données ont fait l'objet de deux analyses spécifiques :

▪ Dans la première analyse portant sur la cohorte 1995-2004 (1477 patients), la mortalité en excès a été considérée comme dépendante du temps depuis le diagnostic, de l'âge au diagnostic et de l'année de diagnostic, ces variables étant modélisées de façon continue. L'effet de l'année de diagnostic a été inclus dans le modèle avec un effet non proportionnel et non linéaire de l'âge. En effet, des explorations antérieures sur ces données de survie ont montré l'effet souvent complexe de l'âge au diagnostic.

Afin de tester les effets complexes de l'année (qui était notre variable d'intérêt), quatre modèles ont été comparés en supposant :

- un effet non linéaire et non proportionnel de l'année (NL / NP) :

Modèle 1 : $\ln[\lambda_c(t, \hat{age}, \text{année})] = f(t) + g(\hat{age}) + h(t) * \hat{age} + i(\text{année}) + j(t) * \text{année}$

- un effet non linéaire et proportionnel de l'année (NL / PP) :

Modèle 2 : $\ln[\lambda_c(t, \hat{age}, \text{année})] = f(t) + g(\hat{age}) + h(t) * \hat{age} + i(\text{année})$

- un effet linéaire et proportionnel de l'année (LL / PP) :

Modèle 3 : $\ln[\lambda_c(t, \hat{age}, \text{année})] = f(t) + g(\hat{age}) + h(t) * \hat{age} + \text{année}$

- pas d'effet de l'année :

Modèle 4 : $\ln[\lambda_c(t, \hat{age}, \text{année})] = f(t) + g(\hat{age}) + h(t) * \hat{age}$

λ_c correspond au taux instantané de mortalité en excès due au cancer, t au temps depuis le diagnostic et les fonctions f, g, h, i et j représentent :

- f : une spline cubique avec un nœud à 1 an de suivi pour l'effet du temps depuis le diagnostic.

- g : une spline quadratique avec un nœud à l'âge moyen pour l'effet non linéaire de l'âge.

- h : une spline cubique avec un nœud à 1 an de suivi pour l'effet non proportionnel de l'âge.

- i : une spline quadratique avec un nœud à l'année 1998 pour l'effet non linéaire de l'année.

- j : une spline cubique avec un nœud à 1 an de suivi pour l'effet non proportionnel de l'année.

L'année de référence fixée à 1998 marque l'arrivée de la thérapie ciblée dans la prise en charge thérapeutique des LF. En partant du modèle le plus complexe (Modèle 1), les modèles ont été comparés deux à deux à l'aide d'une stratégie pas à pas descendante en utilisant un test du

rapport de vraisemblance avec un seuil égal à 0,05. En se basant sur le modèle retenu avec la stratégie précitée, la dynamique du taux de mortalité en excès a été estimée à différents âges (30, 50, 65 et 80 ans) et pour différentes années de diagnostic (1995, 1998, 2000 et 2004). Le ratio du taux de mortalité en excès a également été estimé en fonction de l'année de diagnostic (le ratio correspondant au taux de mortalité pour une année donnée divisé par le taux de l'année de référence (1998)).

- La deuxième analyse portant sur la cohorte 1995-2010 (451 cas) a fait l'objet d'une stratégie de modélisation du taux similaire à la précédente analyse, en ajustant également l'effet de l'année de diagnostic sur le stade du lymphome. Pour cela, le stade Ann-arbor a été considéré en 2 catégories, en regroupant classiquement les stades I-II et les stades III-IV. Cette variable catégorielle a ensuite été ajoutée dans les quatre modèles précités, en supposant un effet proportionnel de cette variable sur le taux de mortalité. Nous avons exclu de l'analyse les patients pour lesquels l'information du stade était manquante (soit 6% des cas sur la période 1995-2010) en nous assurant au préalable que les caractéristiques principales de ces patients étaient similaires à celles de l'ensemble de la cohorte. En se basant sur le modèle retenu à l'aide d'une stratégie similaire à la première analyse, la dynamique du taux en excès a été estimée par stade, pour différentes années et différents âges. Le ratio des taux de mortalité par stade a également été estimé en fonction de l'année de diagnostic.

L'ensemble des analyses a été réalisé avec le logiciel R en utilisant les programmes développés par Remontet *et al.*

III.4. ARTICLE

Les résultats de ce travail ont fait l'objet d'une publication dans la revue *European Journal of Hematology* sous forme d'un article scientifique publié en ligne en juin 2014.



ORIGINAL ARTICLE

Trends in excess mortality in follicular lymphoma at a population level

Morgane Mounier¹, Nadine Bossard^{2,3,4,5}, Aurélien Belot^{2,3,4,5,6}, Laurent Remontet^{2,3,4,5}, Jean Iwaz^{2,3,4,5}, Mylène Dandoit¹, Stéphanie Girard-Boulanger¹, Aurélie Herry¹, Anne-Sophie Woronoff⁷, René-Olivier Casanovas⁸, Marc Maynadié^{1,9}, Roch Giorgi¹⁰, for FRANCIM Network and MESURE Working Survival Group

¹Registre des Hémopathies Malignes de Côte d'Or, EA 4184, Université de Bourgogne, Dijon; ²Université de Lyon, Lyon; ³Université Lyon 1, Villeurbanne; ⁴CNRS UMR5558, Laboratoire de Biométrie et Biologie Evolutive, Equipe Biostatistique Santé, Villeurbanne; ⁵Service de Biostatistique, Hospices Civils de Lyon, Lyon; ⁶Département des Maladies Chroniques et traumatismes, Institut de Veille Sanitaire, Saint Maurice; ⁷Registre des cancers du Doubs, EA 3181, Centre Hospitalier Universitaire de Besançon, Besançon; ⁸Service d'Hématologie Clinique, Centre Hospitalier Universitaire de Dijon, Dijon; ⁹Service d'Hématologie Biologique, Centre Hospitalier Universitaire de Dijon, Dijon; ¹⁰UMR S912 Sciences Economiques & Sociales de la Santé et Traitement de l'Information Médicale (SESSTIM), Marseille, France

Abstract

Background: Since the 1990s and since the development of humanised monoclonal antibodies in 1998, the treatment of non-Hodgkin lymphoma has undergone profound changes. Follicular lymphoma (FL) was the first to benefit from this treatment, and several clinical trials have shown a significant improvement in overall survival, but little information is available at a population level. *Objective:* Our objective was to estimate changes in FL-specific mortality at a population level, with an appropriate methodology. *Methods:* Two French retrospective population-based studies on FL were conducted, one from 1995 to 2004, in 1477 patients, and one from 1995 to 2010, in 451 patients. Trends in excess mortality rates (EMRs) according to age, sex, Ann Arbor stage and year of diagnosis were evaluated using the flexible model of Remontet *et al.* *Results:* Trends in the EMR differed according to age at diagnosis and was higher in advanced stage (III, IV) in patients older than 65 yr. The EMR decreased linearly from 1995 to 2010. This decrease was more marked for advanced stages. *Conclusion:* FL-specific mortality decreased over the years of diagnosis, and the difference according to the lymphoma stage diminished in more recent years. However, progress in the management of FL was not able to erase age-related differences.

Key words follicular lymphoma; immunotherapy; excess mortality; population-based study

Correspondence Roch Giorgi, UMR S912 Sciences Economiques & Sociales de la Santé et Traitement de l'Information Médicale (SESSTIM), Faculté de Médecine, 27 Bd Jean Moulin, F-13385 Marseille, France. Tel: +33 (0)491 388 089; Fax: +33 (0)491 385 749; e-mail: Roch.GIORGI@ap-hm.fr

Accepted for publication 15 June 2014

doi:10.1111/ejh.12403

Non-Hodgkin lymphoma (NHL) is the most common haematological malignancy. In Europe, its incidence doubled during the second half of the 20th century, but has recently stabilised (1, 2). NHL is a heterogeneous group, and the incidence and prognosis vary according to the disease histology (3). Follicular lymphoma (FL) is the largest category of indolent B-cell lymphomas (1). In 2000–2002, five-year relative survival (tool to approach net survival) in FL was estimated at 72.8% in Europe (4). An improvement in FL

survival has been observed for FL diagnosed in recent years in France: five-year net survival for FL improved from 70% (1995–97) to 77% (2001–04) (5).

In recent years, the treatment of NHL has improved, not only because of increased expertise in the disease and better care management, but also by the identification of new molecular therapeutic targets thanks to better knowledge of the cellular mechanisms involved in the occurrence and progression of most lymphomas. In particular, in 1993, the

therapeutic arsenal was enriched by interferon and in 1998 by the development of humanised monoclonal antibodies, especially rituximab, which targets CD20 (6). FLs were the first NHL to benefit from this new immunotherapy and became a model for the therapeutic management of all indolent B-cell lymphomas. Since then, several randomised clinical trials have shown the significant contribution of this new therapy to the improvement in overall survival, and interferon and rituximab were approved in France for the treatment of FL (7, 8). However, due to the many restrictive inclusion criteria, patients involved in clinical trials account for a small proportion of patients seen every day in medical practice, especially those under a wait-and-watch strategy.

The impact of changes in FL management on survival at a population level cannot be assessed by clinical trials and is still a topic of active research. Furthermore, the methods classically used to analyse survival in clinical research are based on the concept of overall survival (i.e. the overall mortality hazard rates) (9). These methods do not provide information on the excess mortality rate (EMR), that is, the mortality due (directly or not) to the disease. It thus seems appropriate to study the effect of therapeutic changes on population-based data and use an appropriate method to estimate excess mortality (10).

The method based on an excess mortality regression model is able to provide an answer to this issue (11–13). In this approach, the estimation of excess mortality is theoretically possible by removing from the observed (all-cause) mortality the other-cause-than-cancer mortality using the main assumption that the latter is correctly reflected by general population life tables (population with similar sex, age, area of residence, and year of death or last follow-up) (14). As the effects of some covariates may change over the follow-up time (i.e. exert a non-proportional effect over time) and the effects of continuous covariates may be nonlinear, Remontet *et al.* (13) proposed to model nonlinear and non-proportional effects of covariates simultaneously according to an additive excess hazard model. This makes it possible to explore the dynamics of the EMR over the follow-up period and the changes in these dynamics according to covariates of interest, such as age at diagnosis or year of diagnosis (15).

The main goal of this study was to estimate the change in EMR according to the year of diagnosis in a French population-based study on FL, and specifically by lymphoma stage, using an appropriate methodology.

Materials and methods

Study design and data collection

This retrospective population-based study used data provided by FRANCIM, the French network of cancer registries. The FRANCIM database collects all cancer cases diagnosed and

registered in 23 French Départements. The quality and completeness of the participating registries are assessed every four years via an audit by the Comité National des Registres. Quality controls are conducted at the registry and whole database levels. These controls use tools provided by the International Agency for Research on Cancer.

For this study, all FLs recorded in eight cancer registries and diagnosed since 1 January 1995 in patients aged >15 yr were included. The diagnosis of FL was defined according to the third edition of the International Classification of Diseases for Oncology (ICD-O-3), that is, using morphology code 9690/3 for follicular lymphoma NOS, 9691/3 for follicular lymphoma grade 2, 9695/3 for follicular lymphoma grade 1 and 9698/3 for follicular lymphoma grade 3. Information on vital status was updated via a request to the Répertoire National d'Identification des Personnes Physiques (RNIPP). The following information was available for each patient: sex, date of birth, date of diagnosis, morphology of FL, vital status, and date of death or last known vital status. For six cancer registries (General Cancer Registries of Bas-Rhin, Hérault, Isère, Somme, Tarn and the Registries of Haematological Malignancies of Côte d'Or), data for all stages together were available until 31 December 2004 with vital status information updated at 1 January 2008 (2% of the patients were lost to follow up). For two cancer registries, (the General Cancer Registry of Doubs and the Registry of Haematological malignancies of Côte d'Or), all FLs diagnosed until 31 December 2010 were available with vital status information updated at 1 January 2013 (1% of the patients were lost to follow up). In addition, the two latter registries provided the stage for 94% of the cases. First, changes in EMR were analysed using all of the data (eight registries – $n = 1477$ – period 1995–2004), all stages together. Secondly, a similar analysis was conducted taking into account the Ann Arbor stage (two registries – $n = 451$ – period 1995–2010). The characteristics of the patients were considered using numbers and relative frequencies for categorical variables and medians and interquartile ranges for continuous variables (Tables 1 and 2).

Statistical modelling of excess mortality

The EMR was modelled until five years after the diagnosis. The age at diagnosis and the year of diagnosis were analysed as continuous variables. To investigate simultaneously the nonlinear (effect may not be linear with the increase in units of the variable) and non-proportional effects (effect may not be constant over follow-up) of age at diagnosis and year of diagnosis on the EMR, we used the model of Remontet *et al.*, (13) which allows for flexibility through spline functions. Spline functions are smoothing functions based on different piecewise polynomial functions that join each other at the knots. Analysis was conducted in two steps.

Table 1 Characteristics of patients in the population-based study of 1477 cases of follicular lymphoma diagnosed between 1995 and 2004

Characteristic	Period of diagnosis			
	All periods n = 1477	[1995–1997] n = 366	[1998–2000] n = 421	[2001–2004] n = 690
Sex				
Men	686 (46) ¹	166 (45)	199 (47)	321 (47)
Women	791 (54)	200 (55)	222 (53)	369 (53)
Age at diagnosis (median, interquartile range)	62 (21)	62 (22)	62 (21)	62 (21)
Age groups				
[15–44]	192 (13)	44 (12)	59 (14)	89 (13)
[45–54]	291 (20)	68 (18)	87 (21)	136 (20)
[55–64]	348 (24)	90 (25)	93 (22)	165 (24)
[65–74]	358 (24)	85 (23)	104 (25)	169 (24)
[75–+]	288 (19)	79 (22)	78 (18)	131 (19)
Median follow-up (yr)	7	11	8	5
Deaths at 5 yr after diagnosis	435 (29)	134 (37)	123 (29)	178 (26)

¹Values are expressed as numbers of patients (percentages of column total).

In the analysis for all stages together, the EMR was modelled as a function of the duration of follow-up, the age at diagnosis and the year of diagnosis between 1995 and 2004. To model the baseline hazard in a continuous way, a cubic regression spline for time was used (with one interior knot at one year). As age has a known nonlinear, non-proportional effect on mortality (15), the nonlinear effect of age was modelled directly using a quadratic spline function (with one interior knot at the mean), and for the non-proportional effect, an interaction term between a cubic spline function of time and age was introduced into the model. The effect of the year of diagnosis on the EMR was explored using the following strategy. Four models were compared using the likelihood ratio test (LRT) (5% alpha error): a model assuming a non-proportional and nonlinear effect of the year of diagnosis, a model assuming a proportional and nonlinear effect of the year of diagnosis, a model assuming a proportional and linear effect of the year of diagnosis and a model with no effect of the year of diagnosis. To test the nonlinear effect, the year of diagnosis was modelled using a quadratic spline function (with one interior knot at 1998, the year of reference), and to test the non-proportional effect, an interaction term between a cubic spline function of time and year of diagnosis was introduced into the model (16). The effect of sex on the EMR was also checked with the LRT. Once the EMR had been estimated with the final model, the EMR

Table 2 Characteristics of patients in the population-based study of 451 cases of follicular lymphoma diagnosed between 1995 and 2010 with lymphoma stage available

Characteristic	Ann Arbor stage		
	All stage n = 451	Stage I–II n = 120	Stage III–IV n = 331
Sex			
Men	209 (46) ¹	63 (52)	146 (44)
Women	242 (54)	57 (48)	185 (56)
Age at diagnosis (median, interquartile range)	62 (20)	64 (22)	62 (19)
Age groups			
[15–44]	45 (10)	13 (11)	32 (10)
[45–54]	84 (19)	25 (21)	59 (18)
[55–64]	120 (27)	23 (19)	97 (29)
[65–74]	118 (26)	33 (27)	85 (26)
[75–+]	84 (19)	26 (22)	58 (18)
Transformation			
No	352 (78)	96 (80)	256 (78)
Yes	86 (19)	22 (18)	64 (19)
Unknown	13 (3)	2 (2)	11 (3)
Period of diagnosis			
[1995–1997]	62 (14)	11 (9)	51 (15)
[1998–2000]	76 (17)	28 (23)	48 (15)
[2001–2003]	93 (21)	21 (18)	72 (22)
[2004–2006]	103 (23)	30 (25)	73 (22)
[2007–2010]	117 (26)	30 (25)	87 (26)
Median follow-up (yr)	9	9	9
Deaths at 5 yr after diagnosis	117 (26)	25 (21)	92 (28)

¹Values are expressed as numbers of patients (percentages of column total).

ratio was calculated (i.e. the EMR estimated for any given year of diagnosis was divided by the EMR of 1998, the year of reference).

In the analysis taking into account the Ann Arbor stage, the EMR was modelled as a function of the duration of follow-up, the age at diagnosis, the year of diagnosis and the Ann Arbor stage group (two categories: I–II vs. III–IV). The effect of the year of diagnosis was explored using the previous strategy assuming a proportional effect of the stage on the EMR. The statistical modelling used functions previously developed by Remontet *et al.* (13) in R software.

Results

First population-based study for all stages together

The study analysed data on 1477 patients with FL all stages together. Table 1 shows their main characteristics by period of diagnosis (1995–1997) (1998–2000) and (2001–2004). There was a greater proportion of women than men whatever the period of diagnosis. FL affected all age groups with

a median age of 62 yr. The distribution across age groups was relatively stable over the periods of diagnosis. During the last period (2001–2004), the percentages were as follows: 13% for ages 15–44, 20% for ages 45–54, 24% for ages 55–64 and for ages 65–74 and 19% for ages 75 and older (Table 1). Over the whole study period, 435 patients (29% of the cohort) died during the first five years of follow-up.

The final model was based on a nonlinear and a non-proportional effect of age, and a linear and a proportional effect of the year of diagnosis on the EMR. The results for the comparison of models using the LRT are shown in the Supplemental data: Table S1.

Figure 1 shows the EMR ratio (with 1998 as the year of reference) according to the year of diagnosis as obtained with the final model. The excess mortality decreased progressively from 1995 to 2004 with a decrease in the EMR ratio equal to 0.75 in 2004 compared with 1998.

Figure 2 shows the EMRs for age at diagnosis (30, 50, 65 and 80) and the year of diagnosis (1995, 1998, 2000 and 2004). For ages 30 and 50, the EMR did not exceed 0.1 deaths per person-year during the first five years of follow-up over the study period. The EMR was very low during the first years of follow-up but increased after four years of follow-up (Fig. 2, panels A and B). For age 65, the EMR fluctuated more (Fig. 2, panel C). For age 80, the EMR was very high just after diagnosis but decreased thereafter to stabilise around 0.1 at five years of follow-up (Fig. 2, panels D). For all ages, the EMR decreased with more recent years of diagnosis.

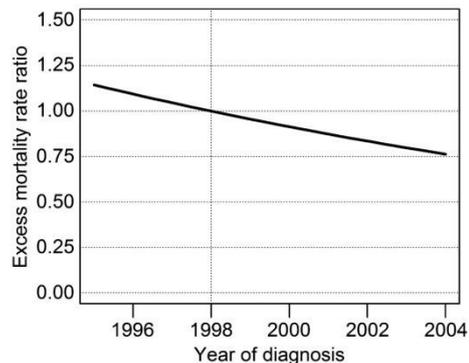


Figure 1 Excess mortality rate ratio according to the year of diagnosis (the year of reference is 1998) based on the population-based data for 1477 patients with follicular lymphoma diagnosed between 1995 and 2004, all stages together. Results based on final model that assumed a non-proportional and nonlinear effect of age at diagnosis and a proportional and linear effect of the year of diagnosis.

Second population-based study adjusted for lymphoma stage

A total of 451 patients with FL were enrolled in the second set of analyses (Table 2). With a greater proportion of women (54%) and a median age equal to 62 yr, the characteristics of this cohort were similar to those of the previous cohort. The overall median follow-up was 9 yr, and 26% of patients (117 patients) died during the first five years of follow-up. At diagnosis, 27% of patients had an Ann Arbor stage I–II and 73% had stage III–IV. There were more men in the I–II stage group than in III–IV stage group, but there was no age difference between stage groups. The distribution of stage at diagnosis was relatively stable over the study period. Among these 451 patients, 86 (19%) had a known transformation of their follicular lymphoma into a more aggressive lymphoma during follow-up. The proportion of patients with transformation was similar in both stage groups. Twenty-eight per cent of patients in the stage III–IV group had died at five years against 21% in the stage I–II group.

In these 451 cases of FL, the year of diagnosis had an almost significant ($P = 0.06$) linear and proportional effect on the EMR adjusted for the Ann Arbor stage group. This result is in accordance with outcome of the previous analysis. The results of the comparison of models using the LRT are shown in the Supplemental data, Table S2.

The overall EMR ratio resulting from the final model according to the year of diagnosis decreased since 1995 (Fig. 3). With 1998 as the year of reference, the decrease in the risk mirrored that in the previous cohort until 2004 and decreased by a factor of two between 1998 and 2010.

The EMR was estimated using the final model according to the duration of follow-up, for age at diagnosis (30, 50, 65 and 80) and the year of diagnosis (1995, 1998, 2000 and 2004) for stage I–II (Fig. 4) and for stage III–IV (Fig. 5). At 30 and 50 yr old, there was no difference in the EMR according to the stage group, except for a greater increase in the youngest patients at five years of follow-up in stage III–IV (Fig. 4, panels A and B and Fig. 5, panels A and B). However, in patients older than 65 yr at diagnosis, the EMR was higher in the advanced stages (Fig. 4, panels C and D and Fig. 5, panels C and D). The EMR decreased in more recent years of diagnosis in both stage groups during the first four years of follow-up with a more marked decrease for advanced stages, thus reducing the EMR gap between stage groups.

Discussion

The present study analyses data from eight exhaustive population-based FL registries and offers a strategy to model EMR over two periods of diagnosis (10 yr and 16 yr). To

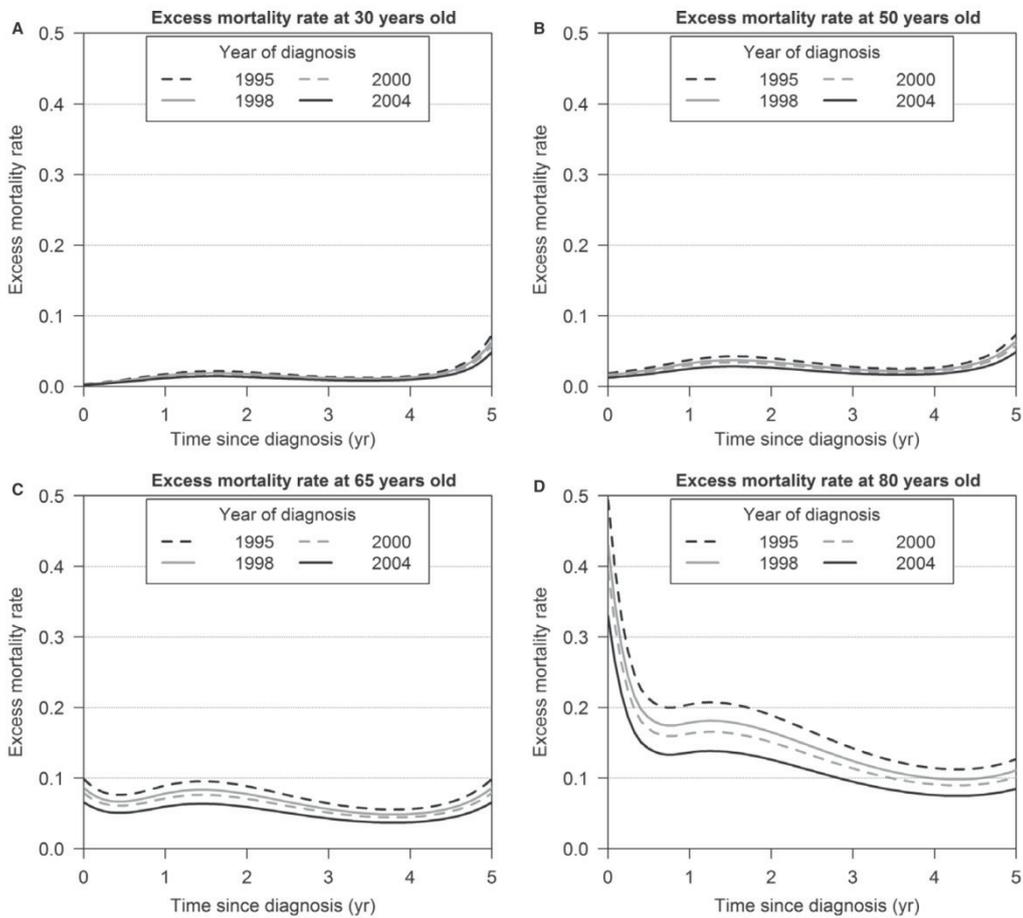


Figure 2 Dynamics of the excess mortality rates according to the year of diagnosis in patients aged 30, 50, 65 and 80 yr old at diagnosis based on the population-based data for 1477 patients with follicular lymphoma diagnosed between 1995 and 2004, all stages together. Results based on final model that assumed a non-proportional and nonlinear effect of age at diagnosis and a proportional and linear effect of the year of diagnosis.

our knowledge, these are the first estimations of EMR that simultaneously model the effect of the year of diagnosis and the age at diagnosis for FL.

The results revealed that FL mortality hazard rates decreased from 1995 to 2004. However, the prognosis remained poor in the oldest patient (> 65 yr), and we found no difference according to the sex. These results are in accordance with results for NHL in Europe and the USA (17, 18) and with a French study based on NHL subtypes (5).

Since the 1990s, the treatment of NHL has undergone profound changes with the use of immunotherapy (interferon, rituximab), leading to a significant improvement in prognosis

as shown in clinical trials (19). A clinical trial conducted by the German Low-Grade Lymphoma Study Group on 428 patients with advanced-stage FL diagnosed in 2000 showed significantly better two-year overall survival in patients treated with immunotherapy in association with a CHOP protocol than in patient with CHOP alone (7). Similar positive results were found by an American trial with long-term follow-up in patients with relapsing or newly diagnosed indolent NHL (20). However, the French FL2000 study, which evaluated the combination of rituximab with chemotherapy plus interferon (CHVP+ I vs. R-CHVP+ I) in the first-line treatment of FL with a high tumour burden, found no

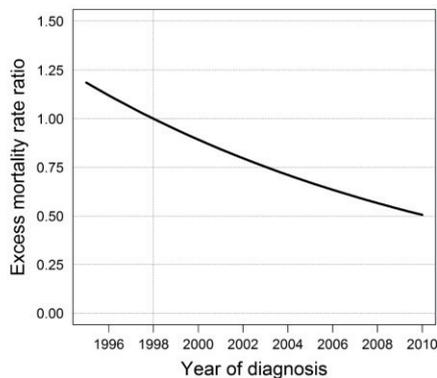


Figure 3 Excess mortality rate ratio according to the year of diagnosis (the year of reference is 1998) based on the population-based data for 451 patients with follicular lymphoma diagnosed between 1995 and 2010, taking into account the lymphoma stage. Results based on final model adjusted for lymphoma stage, which assumed a non-proportional and nonlinear effect of age at diagnosis and a proportional and linear effect of the year of diagnosis.

significant difference in overall survival (8). Full knowledge of the real impact of a new drug also requires a population-based evaluation on unsorted data (i.e. without specific patient inclusion criteria). To date, there have been few population-based cohort studies, and the event of interest is often the response to treatment with the duration of remission (21). In a recent population-based study involving 476 patients with FL diagnosed in 2003, Keegan *et al.* found no significant association between the use of rituximab and survival (22) while a German study involving 165 patients with FL recorded in a regional population-based cancer registry found a significant difference in five-year overall survival between patients receiving any first-line chemotherapy with rituximab and those receiving chemotherapy without rituximab (23). It is thus difficult to draw definite conclusions on the impact of immunotherapy at a population level because of the different strategies used to analyse the data and the contradictory results produced. In our study, we used the year of diagnosis as a proxy for the change over time in the management of FL. 1998 marked a turning point with the marketing authorisation (MA) for rituximab as a treatment in relapsing FL (24). Before 1998, the therapeutic management of FL included (i) surveillance without therapy for asymptomatic patients with a low tumour burden or (ii) an induction therapy phase, whose length varied according to the extension of the disease in symptomatic patients, followed by radiotherapy for symptomatic localised tumours, or poly-chemotherapy (essentially the CHOP protocol) combined with interferon for advanced disease, followed by a maintenance therapy phase using interferon. In 1998, the first clinical trials tested the efficacy of rituximab in patients with relapsing

low-grade FL (25). In 2002, the MA was obtained for rituximab as a first-line treatment for FL. Today, rituximab is an important part of FL management: two-thirds of patients receive rituximab in association with poly-chemotherapy as induction and maintenance treatments (26, 27). Even though there may be a delay between the MA and the real use of the drug at the population level, a large proportion of patients included in clinical trials could have benefited from immunotherapy before 1998 and between 1998 and 2002.

The major strength of the present study is the use of an appropriate methodology to estimate excess mortality due to cancer. Flexible regression models are now the reference methods for estimating net survival and are well suited to the analysis of survival trends (28). The use of a flexible model for EMR made it possible to (i) consider the impact of the year of diagnosis and the age at diagnosis on mortality with non-proportional and nonlinear effects; and (ii) model the baseline hazard as a continuous function. To reinforce confidence in modelling the EMR with regression splines, complementary statistical analyses were conducted: the baseline hazards were also modelled using a step function (10) and a supplementary regression analysis was performed using the year of diagnosis as a categorical variable. This analysis confirmed the good quality of the estimates obtained with the regression splines (results not shown) (13).

Because lymphoma stage was not available in all of the participating registries, this variable was not included in the first analyses. Nonetheless, this first analysis allowed an accurate quantification of the effect of the year of diagnosis on the EMR on a large data set. In the present results, the age at diagnosis had a non-proportional and nonlinear effect on the EMR. The high EMRs in the early follow-up of elderly people may have various explanations including longer delays to treatment or excessively aggressive treatments. In young patients, the progressive increase in EMR probably reflected transformations in more aggressive lymphomas, adverse drug effects or the occurrence of a second cancer, especially as these patients are more frequently included in clinical trials in which new molecules are used whatever the stage (29). The results adjusted for lymphoma stage showed no difference in EMR according to the stage in the youngest patients, but in the oldest patients, the EMRs were higher in patients with advanced stages (III–IV), probably due to the use of more aggressive treatments in younger patients with an advanced stage that would not be suitable in older patients due to the presence of comorbidities. In the second analysis, the delay between diagnosis and the latest follow-up or death was shorter than in the previous analysis (3 vs. 4 yr). Thus, the end of the curve should be interpreted with caution, especially regarding the oldest patients whose prognoses are generally poorer.

The transformation of a follicular lymphoma into a more aggressive lymphoma can affect the prognosis of a patient. In the second population-based cohort of FL, the information

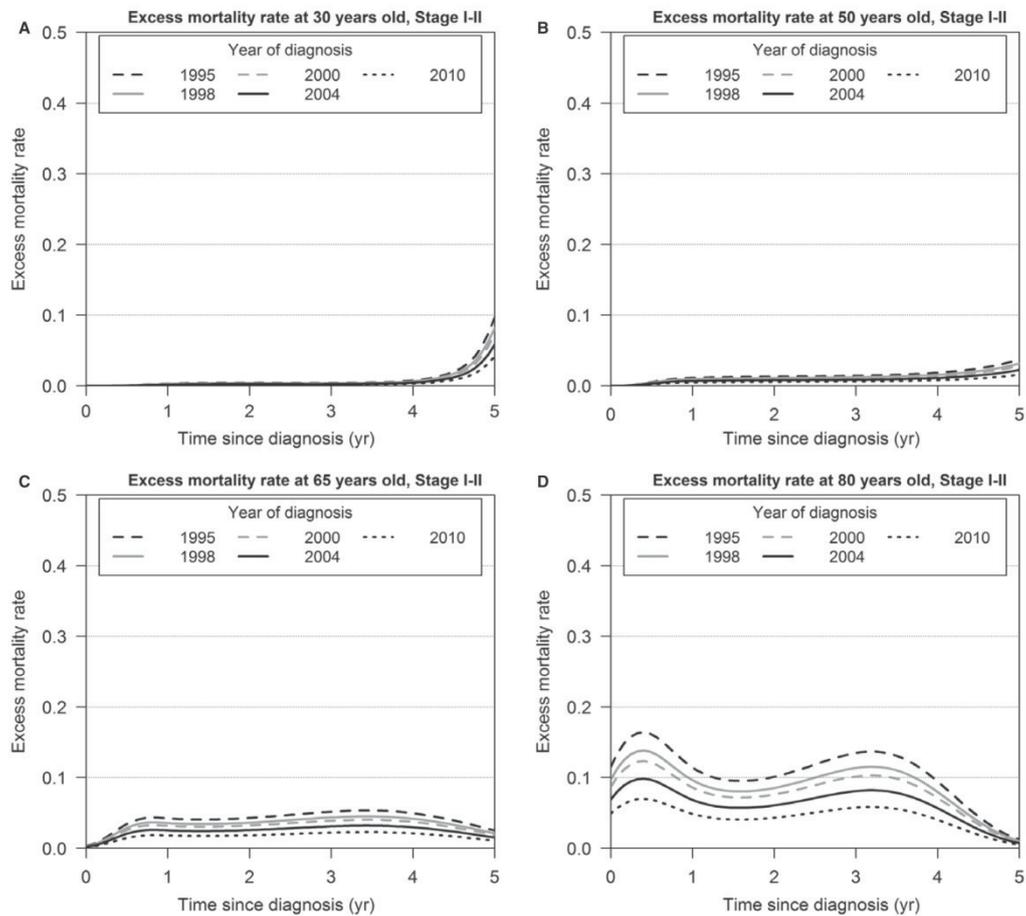


Figure 4 Dynamics of the excess mortality rates in Stage I-II according to the year of diagnosis in patients aged 30, 50, 65, and 80 yr old at diagnosis based on the population-based data for 451 patients with follicular lymphoma diagnosed between 1995 and 2010. Results based on final Model adjusted for lymphoma stage, which assumed a non-proportional and nonlinear effect of age at diagnosis and a proportional and linear effect of the year of diagnosis.

about transformation was available. A specific analysis performed on the 86 patients with known transformation, the trend of the EMR over time according to the year of diagnosis, was similar to that of the whole cohort in both stage groups, with an EMR slightly higher in the oldest patients with transformation than in the oldest patients without transformation (results not shown).

The year of diagnosis had a proportional and linear effect on the EMR, which decreased from 1995 to 2010. This decrease may be explained by better knowledge of the disease, changes in the management of FL, with the improvement in treatment (advent of immunotherapy with interferon since

1993 and rituximab since 1998) combined with earlier treatments [especially in the youngest patients, with the wider use of autologous stem cell transplantation in combination in cases of relapse (30)], a reduction in maximum doses in recent years of diagnosis, better surveillance of patients without therapy and earlier diagnosis.

The classification of neoplasms was updated in 2001 (ICD-O-2 to ICD-O-3) (31). For lymphoma, the new classification was largely based on the Revised European American Lymphoma Classification published in 1994 (32), and the incidence of FL was only slightly affected by changes in cancer classification. FL is commonly diagnosed incidentally

Mounier *et al.*

Trends in excess mortality in follicular lymphoma

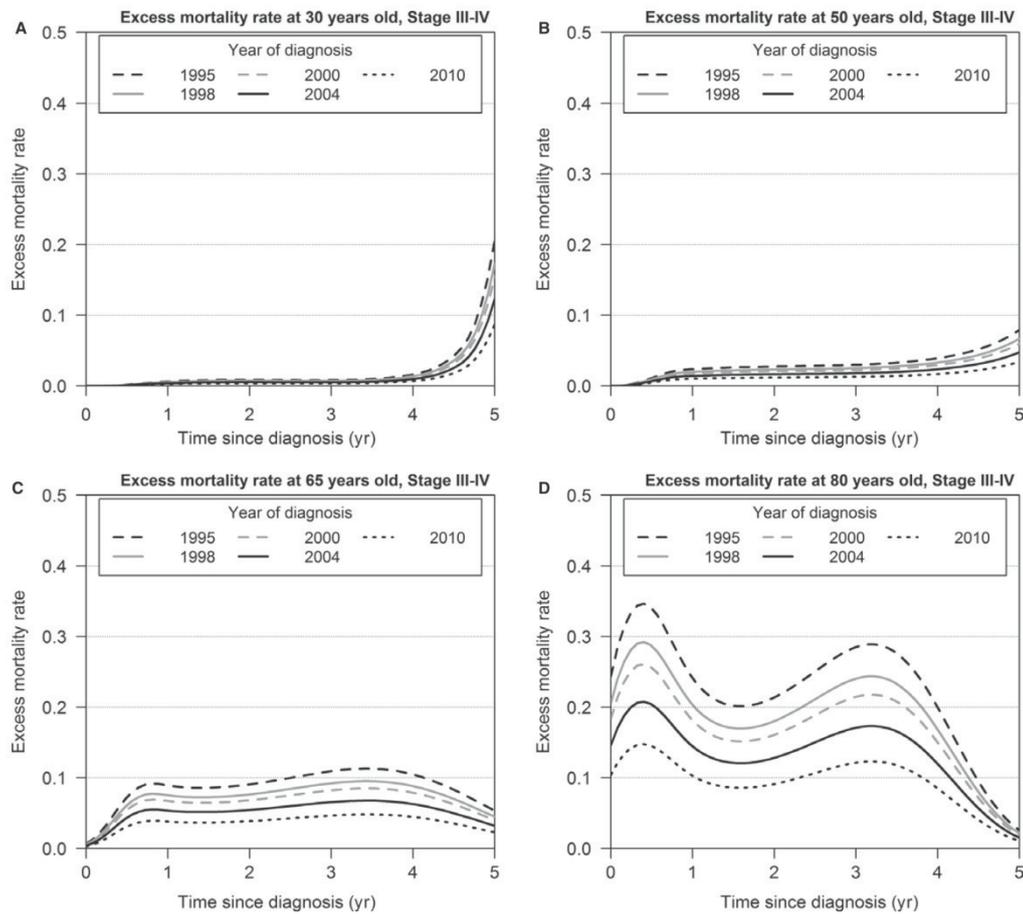


Figure 5 Dynamics of the excess mortality rates in stage III-IV according to the year of diagnosis in patients aged 30, 50, 65 and 80 yr old at diagnosis based on the second population-based data for 451 patients with follicular lymphoma diagnosed between 1995 and 2010. Results based on final model adjusted for lymphoma stage, which assumed a non-proportional and nonlinear effect of age at diagnosis and a proportional and linear effect of the year of diagnosis.

during unrelated medical diagnostic or therapeutic procedures. As such procedures became more common during the years of the study, there have been fewer cases of symptomatic disease and more incidental findings of FL in recent times. However, individual characteristics of the patients, such as age at diagnosis, and the distribution of lymphoma stages did not change during the periods of diagnosis studied.

In this era of immunotherapy, the Follicular Lymphoma International Prognostic Index 2 (FLIPI2) was published in 2008 (33), but the follow-up durations for our data did not allow us to evaluate the impact of this new index on practitioners' treatment decisions.

At a population level, advances in the management of FL have led to a decrease in excess mortality over time and have reduced differences between stage groups for mortality. However, they were not able to erase age-related differences in mortality.

Acknowledgements

The authors thank the following persons and cancer registries for their contribution: Marc Colonna (General Cancer Registry of Isère), Pascale Grosclaude (General Cancer Registry of Tam), Bénédicte Lapôtre-Ledoux (General Cancer

Registry of Somme), Brigitte Trétarre (General Cancer Registry of Hérault), Xavier Troussard (Registry of Hematological Malignancies of Calvados) and Michel Velten (General Cancer Registry of Bas-Rhin). The authors also thank Philip Bastable (Dijon School of Medicine, Burgundy University) for a revision of the final version of the manuscript. This work was presented at the meeting of the Groupe pour l'Epidémiologie et l'Enregistrement du cancer dans les Pays de Langue Latine (2013, Siracusa, Italy) and was awarded the Enrico Anglesio Prize (Fondo Anglesio Moroni per l'Oncologia, Torino, Italy).

Conflict of interest and source of funding

The authors have no conflict of interests to report. The present work is part of the regular tasks of the Registry of Hematological Malignancies of Côte d'Or, France. It was, however, partly supported by a grant from the French Agence Nationale de la Recherche (ANR, MESURE project, Grant Number ANR-09-BLAN-0357).

References

- 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**:3724–34.
- Belot A, Grosclaude P, Bossard N, *et al.* Cancer incidence and mortality in France over the period 1980–2005. *Rev Epidemiol Sante Publique* 2008;**56**:159–75.
- Sant M, Allemani C, De Angelis R, *et al.* Influence of morphology on survival for non-Hodgkin lymphoma in Europe and the United States. *Eur J Cancer* 2008;**44**:579–87.
- Marcos-Gragera R, Allemani C, Tereanu C, *et al.* Survival of European patients diagnosed with lymphoid neoplasms in 2000–2002: results of the HAEMACARE project. *Haematologica* 2011;**96**:720–8.
- Monnereau A, Troussard X, Belot A, Guizard A-V, Woronoff A-S, Bara S, Lapôtre-Ledoux B, Iwaz J, Trétarre B, Maynadié M. Unbiased estimates of long-term net survival of hematological malignancy patients detailed by major subtypes in France. *Int J Cancer* 2012. DOI: 10.1002/ijc.27889.
- Giné E, Gutiérrez-García G, López-Guillermo A. Current immunochemotherapy strategies in follicular lymphoma. *Adv Ther* 2010;**27**:704–13.
- Hiddemann W, Kneba M, Dreyling M, *et al.* Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. *Blood* 2005;**106**:3725–32.
- Salles G, Mounier N, de Guibert S, *et al.* Rituximab combined with chemotherapy and interferon in follicular lymphoma patients: results of the GELA-GOELAMS FL2000 study. *Blood* 2008;**112**:4824–31.
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;**53**:457–81.
- Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: elements for further discussion. *Stat Med* 1990;**9**:529–38.
- Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med* 2004;**23**:51–64.
- Giorgi R, Abrahamowicz M, Quantin C, *et al.* A relative survival regression model using B-spline functions to model non-proportional hazards. *Stat Med* 2003;**22**:2767–84.
- Remontet L, Bossard N, Belot A, Estève J. An overall strategy based on regression models to estimate relative survival and model the effects of prognostic factors in cancer survival studies. *Stat Med* 2007;**26**:2214–28.
- Danieli C, Remontet L, Bossard N, Roche L, Belot A. Estimating net survival: the importance of allowing for informative censoring. *Stat Med* 2012;**31**:775–86.
- Corn S, Roche L, Micol J-B, *et al.* Changes in the dynamics of the excess mortality rate in chronic phase-chronic myeloid leukemia over 1990–2007: a population study. *Blood* 2011;**118**:4331–7.
- Abrahamowicz M, MacKenzie TA. Joint estimation of time-dependent and non-linear effects of continuous covariates on survival. *Stat Med* 2007;**26**:392–408.
- van de Schans SAM, Gondos A, van Spronsen DJ, *et al.* Improving relative survival, but large remaining differences in survival for non-Hodgkin's lymphoma across Europe and the United States from 1990 to 2004. *JCO* 2011;**29**:192–9.
- De Angelis R, Sant M, Coleman MP, *et al.* Cancer survival in Europe 1999–2007 by country and age: results of EURO-CARE-5—a population-based study. *Lancet Oncol* 2014;**15**:23–34.
- Solal-Celigny P, Lepage E, Brousse N, *et al.* Recombinant interferon alfa-2b combined with a regimen containing doxorubicin in patients with advanced follicular lymphoma. Groupe d'Etude des Lymphomes de l'Adulte. *N Engl J Med* 1993;**329**:1608–14.
- Czuczman MS, Weaver R, Alkuzweny B, Berlefin J, Grillo-López AJ. Prolonged clinical and molecular remission in patients with low-grade or follicular non-Hodgkin's lymphoma treated with rituximab plus CHOP chemotherapy: 9-year follow-up. *J Clin Oncol* 2004;**22**:4711–6.
- Vidal L, Gafter-Gvili A, Salles G, *et al.* Rituximab maintenance for the treatment of patients with follicular lymphoma: an updated systematic review and meta-analysis of randomized trials. *J Natl Cancer Inst* 2011;**103**:1799–806.
- Keegan THM, Moy LM, Foran JM, *et al.* Rituximab use and survival after diffuse large B-cell or follicular lymphoma: a population-based study. *Leuk Lymphoma* 2013;**54**:743–51.
- Krause SW, Gerken M, Andreesen R, Hofstädter F, Klinkhammer-Schalke M. Treatment of B cell lymphoma with chemotherapy plus rituximab: a survival benefit can be

- demonstrated in the routine data of a regional cancer registry. *Ann Hematol* 2012;**91**:561–70.
24. Coiffier B, Haioun C, Ketterer N, *et al.* Rituximab (anti-CD20 monoclonal antibody) for the treatment of patients with relapsing or refractory aggressive lymphoma: a multicenter phase II study. *Blood* 1998;**92**:1927–32.
 25. Maloney DG, Grillo-López AJ, White CA, *et al.* IDEC-C2B8 (Rituximab) anti-CD20 monoclonal antibody therapy in patients with relapsed low-grade non-Hodgkin's lymphoma. *Blood* 1997;**90**: 2188–95.
 26. Freedman A. Follicular lymphoma: 2012 update on diagnosis and management. *Am J Hematol* 2012;**87**:988–95.
 27. Friedberg JW, Taylor MD, Cerhan JR, *et al.* Follicular lymphoma in the United States: first report of the national lymphocare study. *J Clin Oncol* 2009;**27**:1202–8.
 28. Roche L, Danieli C, Belot A, Grosclaude P, Bouvier A-M, Velten M, Iwaz J, Remontet L, Bossard N. Cancer net survival on registry data: use of the new unbiased Pohar-Perme estimator and magnitude of the bias with the classical methods. *Int J Cancer* 2012; DOI: 10.1002/ijc.27830.
 29. Morton LM, Curtis RE, Linet MS, *et al.* Second malignancy risks after non-Hodgkin's lymphoma and chronic lymphocytic leukemia: differences by lymphoma subtype. *J Clin Oncol* 2010;**28**:4935–44.
 30. Le Gouill S, De Guibert S, Planche L, *et al.* Impact of the use of autologous stem cell transplantation at first relapse both in naive and previously rituximab exposed follicular lymphoma patients treated in the GELA/GOELAMS FL2000 study. *Haematologica* 2011;**96**:1128–35.
 31. Turner JJ, Morton LM, Linet MS, *et al.* InterLymph hierarchical classification of lymphoid neoplasms for epidemiologic research based on the WHO classification (2008): update and future directions. *Blood* 2010;**116**:e90–8.
 32. Morton LM, Turner JJ, Cerhan JR, *et al.* Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). *Blood* 2007;**110**:695–708.
 33. Formica V, Norman AR, Cunningham D, Wotherspoon A, Oates J, Chong G. Utility of the follicular lymphoma international prognostic index and the international prognostic index in assessing prognosis and predicting first-line treatment efficacy in follicular lymphoma patients. *Acta Haematol* 2009;**122**:193–9.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Results of the Likelihood Ratio Test to select the final model on the first population-based data (1477 patients with diagnosis between 1995 to 2004).

Table S2. Results of the Likelihood Ratio Test to select the final model on the second population-based data (451 patients with diagnosis between 1995 to 2010).

SUPPLEMENTAL DATA

TABLE S1 – Results of the Likelihood Ratio Test to select the final model on the first population-based data (1477 patients with diagnosis between 1995 to 2004).

The non linear effect was modeling with a cubic regression spline for time with one interior knot at one year follow-up, a quadratic spline function for age with one interior knot at the mean and a quadratic spline for the year of diagnosis with one interior knot at 1998, the year of reference.

Model 1*	Model compared	Model 2*	Likelihood Ratio Test with a 5% alpha error	Degrees of freedom	p.value
Effect of the year on the EMR					
non-linear and non-proportional effect		non-linear and proportional effect		4	0,359
non-linear and proportional effect		linear and proportional effect		2	0,696
linear and proportional effect		no effect		1	0,047
Effect of the sex on the EMR					
proportional effect**		no effect**		1	0,57

* assuming a non linear effect of time of follow-up and a non-linear and non-proportional effect of age on the EMR

** assuming a non linear effect of time of follow-up, a linear and proportional effect of year and a non-linear and non-proportional effect of age on the EMR.

TABLE S2 – Results of the Likelihood Ratio Test to select the final model on the second population-based data (451 patients with diagnosis between 1995 to 2010).

The non linear effect was modeling with a cubic regression spline for time with one interior knot at one year follow-up, a quadratic spline function for age with one interior knot at the mean and a quadratic spline for the year of diagnosis with one interior knot at 1998, the year of reference.

Model 1*	Model compared	Model 2*	Likelihood Ratio Test with a 5% alpha error	Degrees of freedom	p.value
Effect of the year on the EMR					
non-linear and non-proportional effect		non-linear and proportional effect		4	0.479
non-linear and proportional effect		linear and proportional effect		2	0.355
linear and proportional effect		no effect		1	0.062
Effect of the sex on the EMR					
proportional effect**		no effect**		1	0.079

* assuming a non linear effect of time of follow-up, a non-linear and non-proportional effect of age and a proportional effect of the Ann Arbor stage on the EMR.

** assuming a non linear effect of time of follow-up, a linear and proportional effect of year, a non-linear and a non-proportional effect of age and a proportional effect of the Ann Arbor stage on the EMR.

III.5. RÉSULTATS COMPLEMENTAIRES

L'étude de ces deux bases de données de population a fait l'objet d'analyses complémentaires afin d'abord de tester l'effet d'autres covariables sur la mortalité en excès et ensuite de confirmer la validité de nos résultats.

Tout d'abord, nous nous sommes intéressés aux estimations de survie nette standardisées en fonction du sexe des patients.

Remarque 11. Dans le concept de survie nette, même si le taux de mortalité attendue est ajusté sur l'âge des patients, cela reste insuffisant pour tenir compte des éventuelles différences de structure d'âge entre les deux groupes. Aussi, si l'intérêt est de fournir une mesure de la survie nette à des fins de comparaisons entre plusieurs groupes, il est nécessaire de standardiser les probabilités de survie sur l'âge. Cela permet de «gommer» des différences de survie nette liées à des structures d'âge différentes. Cet indicateur est obtenu en estimant, dans chaque groupe, la probabilité pour chaque classe d'âge prédéfinie puis en effectuant une moyenne pondérée des probabilités obtenues. Les poids utilisés peuvent être définis en utilisant des poids externes (Corazziari *et al.* 2004) ou sinon à l'aide d'une pondération interne, en utilisant la structure d'âge de la cohorte d'origine. Au-delà de l'âge, la standardisation sur d'autres variables peut être également pertinente.

En se basant sur la cohorte 1995-2004, nous avons estimé la survie nette à l'aide de l'estimateur non paramétrique proposé par Pohar-Perme *et al.* (Perme *et al.* 2012) en utilisant le package « Relsurv » sous le logiciel R. La probabilité de survie nette globale (c'est-à-dire pour tous âges) à 5 ans a été standardisée sur l'âge en utilisant les poids de l'International Cancer Survival Standard (Corazziari *et al.* 2004).

Ces résultats sont présentés sur la Figure III.1 en fonction des périodes de diagnostic et par sexe.

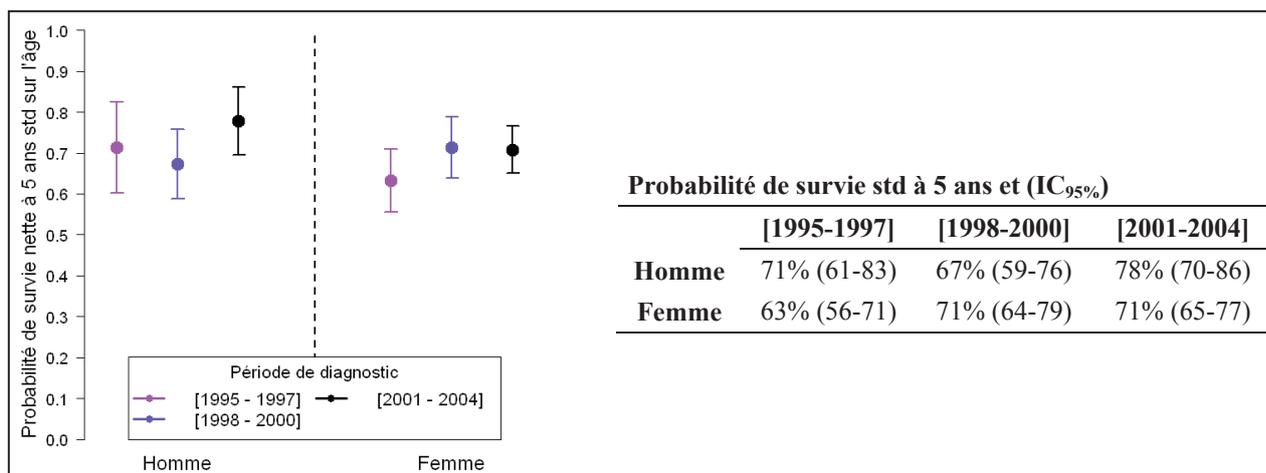


Figure III.1 - Probabilité de survie nette (et IC_{95%}) à 5 ans standardisée sur l'âge à la suite d'un diagnostic de LF, en fonction de la période de diagnostic et du sexe.

La probabilité de survie nette à 5 ans était légèrement plus élevée chez les hommes que chez les femmes avec respectivement 78% et 71% sur la dernière période de diagnostic [2001-2004] (Figure III.1). La survie s'améliorait entre la première et la dernière période de diagnostic dans les deux sexes.

L'effet du sexe sur la mortalité en excès a ensuite été testé en incluant cette variable avec un effet proportionnel et linéaire dans le modèle de Remontet *et al.*. L'effet ayant été retrouvé non significatif, nous n'avons alors pas ajusté notre modèle final sur cette variable.

Comme une transformation du LF en un lymphome d'évolutivité plus agressive peut être un facteur influençant la survie, nous avons mené une étude spécifique par sous-groupe sur la cohorte 1995-2010. Pour cela, l'effet de l'année de diagnostic a été modélisé d'une part sur les patients n'ayant pas connu de transformation et d'autre part sur ceux ayant connu une transformation. Ces résultats ont rapporté un effet de l'année de diagnostic toujours proportionnel et linéaire dans les deux groupes.

Une étude par classe d'âge a ensuite permis de s'assurer de l'effet de l'année de diagnostic au sein de chaque classe. Pour cela, la probabilité de survie nette à 5 ans a tout d'abord été estimée pour cinq classes d'âge en fonction des périodes de diagnostic à l'aide de l'estimateur non paramétrique de Pohar-Perme (Figure III.2).

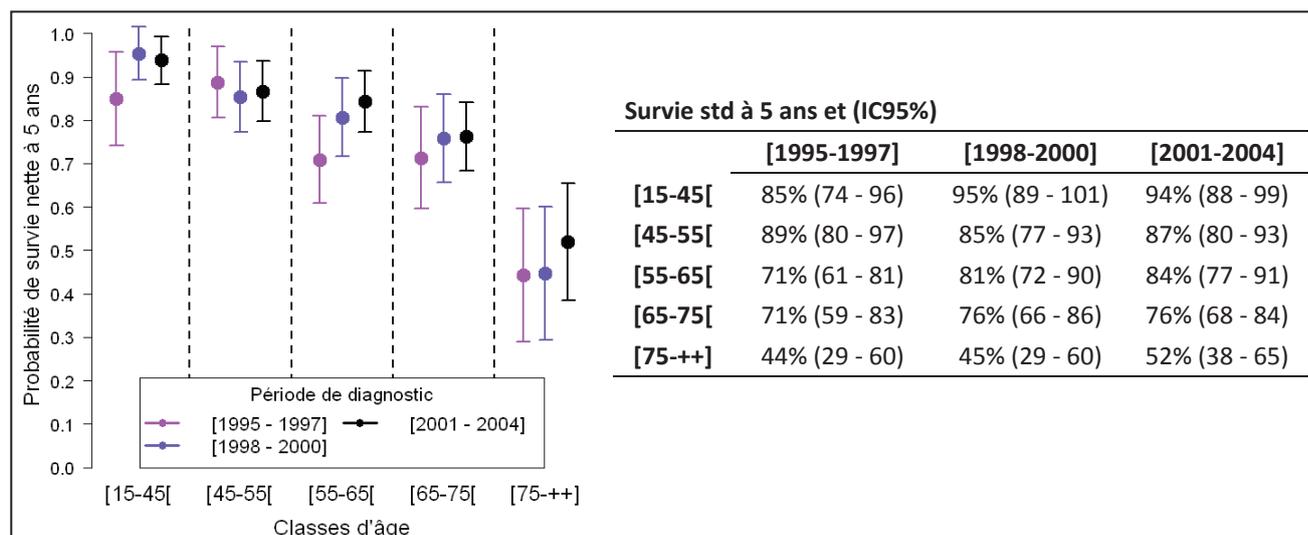


Figure III.2 - Probabilité de survie nette à 5 ans (et $IC_{95\%}$) par classe d'âge, à la suite d'un diagnostic de LF en fonction de la période de diagnostic.

Alors que la probabilité de survie nette à 5 ans est relativement bonne pour les trois premières classes d'âge (94%, 87% et 84% respectivement pour la période [2001-2004]), elle est plus basse dans les classes les plus âgées (76% pour les [65-75[et 52% pour les [75-++]) dans cette même période) (Figure III.2). À l'exception des [45-55[où la survie est stable, la survie nette s'améliorait avec les périodes de diagnostic récentes dans toutes les classes d'âge. L'amélioration la plus importante était observée chez les [55-65[où la survie à 5 ans passait de 71% sur la période [1995-1997] à 84% sur la période [2001-2004].

Une modélisation du taux de mortalité en excès spécifique par classe d'âge avec le modèle de Remontet *et al.* a rapporté un effet proportionnel et linéaire de l'année sur la mortalité dans chacune des classes. Du fait de ces résultats, nous n'avons pas modélisé d'interaction entre l'année de diagnostic et l'âge dans le modèle, en supposant ainsi un effet de l'année identique en fonction de l'âge du patient.

Afin de s'assurer du bon usage des fonctions flexibles dans la modélisation paramétrique, plusieurs analyses complémentaires ont été menées. La dynamique du taux de mortalité de base du modèle sans covariables a été comparée aux estimations obtenues par un modèle considérant un taux de base constant par intervalle de temps (Figure III.3). Ces résultats complémentaires ont permis de confirmer la bonne qualité de la modélisation basée sur l'utilisation de fonctions flexibles.

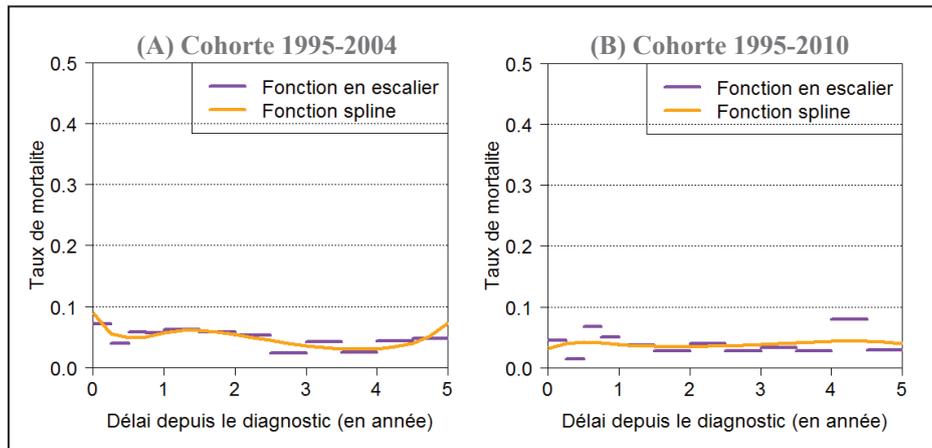


Figure III.3 - Taux de base de mortalité en excès (sans covariables) en fonction du temps de suivi dont l'effet est modélisé par une fonction en escalier ou par une fonction spline cubique (avec un nœud à 1 an de suivi). (A) Résultat de la cohorte 1995-2004, (B) Résultat de la cohorte 1995-2010.

Enfin, à l'issue de notre modèle final qui assumait un effet linéaire et proportionnel de l'année de diagnostic dans les deux cohortes étudiées, la modélisation en continu de l'effet de l'année a été vérifiée en comparant les ratios de mortalité obtenus avec ceux issus d'un modèle où l'effet de l'année était modélisé en variable catégorielle (Figure III.4).

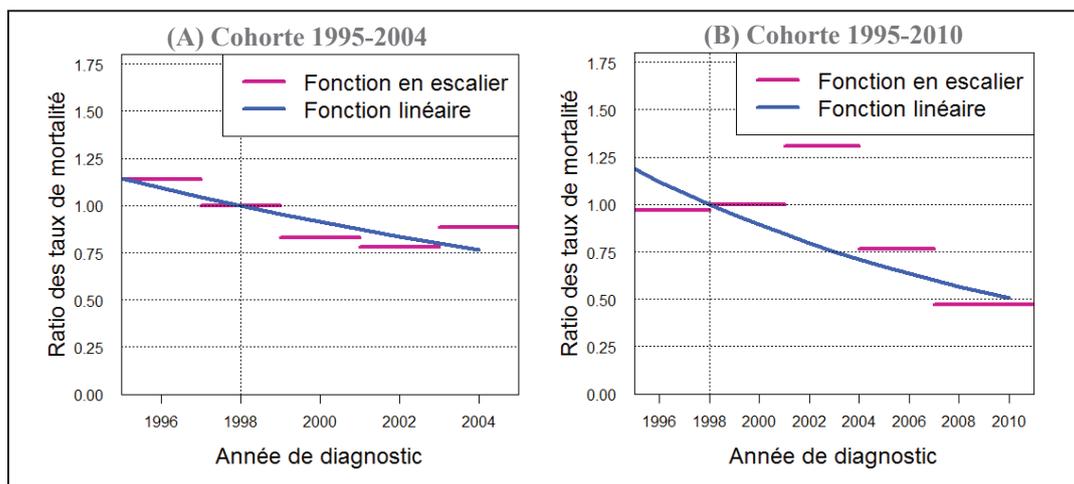


Figure III.4 - Ratio des taux de mortalité en fonction de l'année de diagnostic dont l'effet est modélisé par une fonction en escalier ou par une fonction linéaire (année de référence = 1998). (A) Résultat de la cohorte 1995-2004 (ajusté sur un effet NL-NP de l'âge, période de référence 1997-98 pour la fonction en escalier). (B) Résultat de la cohorte 1995-2010 (ajusté sur un effet NL-NP de l'âge et sur un effet PP du stade, période de référence 1998-2000 pour la fonction en escalier).

Les périodes de diagnostic n'étant pas larges, le nombre de patients dans chaque groupe est limité. Aussi, l'interprétation des ratios de risque issus du modèle où l'année est considérée comme une fonction en escalier doit être faite avec précaution, spécifiquement pour les résultats portant sur la cohorte 1995-2010.

III.6. DISCUSSION

Cette étude menée sur des données de population française a permis de mettre en évidence une diminution de la mortalité en excès à la suite d'un LF au cours des années de diagnostic depuis 1995 et jusqu'en 2010 en population générale. Un point fort de notre étude est la modélisation conjointe des effets de l'âge et de l'année de diagnostic sur le taux de mortalité en excès.

Dans les deux analyses, l'année de diagnostic avait un effet linéaire et proportionnel sur le taux de mortalité en excès, en ajustant sur l'âge au diagnostic des patients et sur le stade du lymphome. La première analyse basée sur un nombre conséquent de patients (1477 cas) a permis de quantifier de façon précise l'effet de l'année de diagnostic sur la mortalité en excès en population générale. La diminution de la mortalité en excès était linéaire au cours des années de diagnostic ce qui signifie que le pronostic des patients s'est amélioré de façon progressive entre 1995 et 2004. Aucun point de cassure dans la tendance de la mortalité en fonction des années de diagnostic n'ayant été retrouvé, ces résultats positifs pourraient être la résultante de plusieurs facteurs liés à la fois à : une détection plus précoce de la maladie, des outils diagnostiques plus performants, une meilleure surveillance des patients en abstention thérapeutique (Friedberg *et al.* 2009 ; Freedman 2012) et à l'utilisation de nouvelles molécules thérapeutiques telles que l'interféron et les anticorps monoclonaux anti-CD20. Il est à noter que l'utilisation de l'immunothérapie s'est faite de façon progressive au cours des années de diagnostic. D'une part, certains patients ont bénéficié de ce nouveau traitement avec l'inclusion dans des essais cliniques avant son AMM. D'autre part, il peut exister un décalage entre l'AMM du rituximab et son utilisation en pratique clinique. Cela peut expliquer pourquoi nous n'avons pas retrouvé de tendance particulière de la mortalité depuis 1998.

Malgré des diminutions de la mortalité en excès au cours des années de diagnostic, il persiste des différences de mortalité en fonction de l'âge au diagnostic des patients au détriment des sujets plus âgés, la dynamique du taux étant également différente en fonction de l'âge. Chez les sujets jeunes (30 ans), la mortalité est faible juste après le diagnostic et augmente à la fin du suivi, alors que chez les sujets âgés (80 ans), une mortalité précoce est observée juste après le diagnostic. Cette différence dans la dynamique du taux peut s'expliquer par des délais de prise en charge plus longs chez les sujets âgés à l'inverse des sujets plus jeunes qui ont plus rapidement accès aux nouvelles molécules en étant davantage inclus dans des essais thérapeutiques (McNamara *et al.* 2012).

Les résultats par stade issus de la deuxième analyse (portant sur 451 patients) ont mis en évidence qu'il n'existait pas de différence de taux de mortalité en fonction du stade chez les sujets jeunes (30 ans) alors que la mortalité en excès était plus élevée chez les sujets plus âgés (80 ans) pour les stades avancés (stade III-IV) par rapport au stade I-II. Une explication pourrait être l'utilisation de traitements plus agressifs chez les patients jeunes avec un stade avancé, non recommandée chez les sujets âgés de par la présence de comorbidités.

Notre étude connaît tout de même certaines limites. L'absence de données cliniques au diagnostic et sur le traitement des patients ne permet pas de mettre en évidence la part propre des divers changements dans la prise en charge des LF sur la diminution de la mortalité en excès. Lors de l'analyse spécifique basée sur le stade du lymphome, le nombre limité de cas ne nous a pas permis de modéliser spécifiquement l'effet de l'année de diagnostic par sous-groupe de stade. Nous avons alors considéré un effet de l'année similaire en fonction du stade du lymphome en modélisant le taux de mortalité sur l'ensemble de la cohorte.

Enfin, la stratégie de construction du modèle adoptée dans ce travail n'est pas la seule approche possible. D'autres approches permettant de modéliser les effets complexes du temps de suivi, de l'âge et de l'année de diagnostic peuvent être envisagées (Wynant et Abrahamowicz 2014).

Toujours dans la recherche d'une meilleure connaissance épidémiologique des LF, notre travail offre un indicateur supplémentaire en population générale en fournissant la tendance de la dynamique du taux de mortalité en excès au cours des années de diagnostic en France.

Chapitre IV



Application 2 : Tendances et comparaisons internationales

IV.1. RATIONNEL

Le cancer est un fléau mondial avec plus de 14 millions de nouveaux cas dans le monde pour la seule année 2012, selon le rapport du CIRC (International Agency for Research on Cancer 2014). Ce nombre devrait encore augmenter au cours des prochaines années en raison de l'allongement de l'espérance de vie dans les pays industrialisés. Afin de coordonner et d'associer les efforts dans la lutte contre le cancer au niveau international, des bases de données de population communes ont été créées, telles que la base EURO CARE ou la base CONCORD (Chapitre I.2.2.2.2). Ce regroupement de bases issues de différents registres des cancers permet la mise en place d'études comparatives d'indicateurs épidémiologiques entre zones géographiques et/ou par période de diagnostic afin d'une part de jouer un rôle d'observation du cancer et d'autre part d'évaluer la qualité du système de soins dans les différents pays.

Les premiers résultats de survie par sous-type de LMNH en Europe ont été publiés par le groupe HAEMACARE (Marcos-Gragera *et al.* 2011) et ont récemment été actualisés sur la période de diagnostic 1995-2008 (Sant *et al.* 2014). Sant *et al.* ont montré que la survie des LMNH s'est améliorée en population générale durant les dernières périodes de diagnostic pour toutes les zones européennes. Cette amélioration reflète les avancées majeures qu'ont connues les LMNH en matière de prise en charge, avec une meilleure connaissance de la maladie mais également avec l'utilisation de nouvelles thérapeutiques telles que la thérapie ciblée.

Néanmoins il persiste encore des disparités de survie entre zones géographiques et en fonction de l'âge au diagnostic des patients. Il apparaît alors pertinent de comprendre à quel moment du suivi sont observées ces divergences de mortalité entre différentes zones et différents âges. Sont-elles situées tout au long du suivi ou bien dans les premiers mois suivant le diagnostic ou encore seulement à partir d'un certain temps après le diagnostic?

La modélisation de la mortalité liée au cancer à des fins de comparaisons géographiques ou temporelles nécessite l'utilisation d'outils statistiques adaptés. La survie nette apparaît comme l'outil adéquat puisque ce concept permet de mettre en évidence des différences de mortalité liées au cancer et non liées à des différences de mortalité autres

causes entre les zones. Comme nous l'avons exprimé dans le *Chapitre III*, seul l'estimateur non paramétrique proposé par Pohar-Perme *et al.* (Perme *et al.* 2012) et les modèles régressifs ajustés sur l'âge sont des estimateurs non biaisés de la survie nette (Danieli *et al.* 2012).

Les LF et les LBDGC représentent les deux groupes de LMNH les plus fréquents en termes d'incidence et ont été les premiers à bénéficier des nouvelles thérapeutiques devenant ainsi la référence en matière de prise en charge pour les autres LMNH.

L'objectif de ce travail était de comparer les tendances de la survie nette et de la dynamique du taux de mortalité en excès entre différentes zones européennes, pour les patients atteints de LF ou d'un LBDGC. L'année de diagnostic a été utilisée comme une mesure indirecte de l'évolution de la prise en charge des patients.

IV.2. DONNÉES DE SURVIE EN POPULATION EUROPÉENNE

Les données proviennent de la base de population EUROCORE-5 en se focalisant sur les LF et les LBDGC diagnostiqués entre 1996 à 2004. Les codes morphologiques ICD-O-3 retenus pour les LF étaient les codes 9690/3, 9695/3, 9691/3 et 9698/3 respectivement pour les LF sans autres informations, de grade 1, de grade 2 et de grade 3. Le code 9675/3 qui était utilisé dans la classification précédente avant 2000, a également été pris en compte¹¹. Pour les LBDGC, étaient considérés les codes morphologiques 9678/3, 9679/3, 9680/3 et 9684/3 respectivement pour les lymphomes B des séreuses, médiastinaux à grandes cellules, médiastinaux diffus et sans autres informations.

Des critères de sélection ont dans un premier temps porté sur l'inclusion des registres de cancer dans l'étude. Afin d'assurer un recueil exhaustif des cas, seuls les registres permettant la distinction par sous-type de lymphome B¹² et disposant d'un enregistrement continu des cas (en termes d'année d'enregistrement) sur toute la période d'étude ont été inclus. Un critère de la qualité des enregistrements a consisté en l'étude du nombre de cas de

¹¹ Peu de cas était codés en 9675/3, les diagnostics ayant été repris dans la plupart des registres de cancer lors du changement de classification.

¹² Ce critère a exclu un seul registre des cancers (national).

lymphomes codés sans autres informations (SAI) sur la période d'étude. Les codes ICD-O-3 associés étaient le code 9590/3 pour les lymphomes-SAI et le code 9591/3 pour les LMNH-SAI. Les registres qui comptaient plus de 30% de cas enregistrés en lymphome-SAI par rapport au nombre total de cas de LMNH sur la période 1996-2004 ont été exclus de l'analyse.¹³

Des critères de sélection ont dans un second temps porté sur l'inclusion des cas dans l'étude. Tous les patients diagnostiqués d'un LF ou d'un LBDGC entre le 1^{er} janvier 1996 et le 31 décembre 2004, âgés de 15 ans ou plus ont été inclus dans l'analyse. Les cas diagnostiqués sur une autopsie ou sur un certificat de décès seul ont été exclus de l'analyse. Les patients ayant connu des tumeurs multiples (c'est-à-dire lorsque le diagnostic du lymphome précédait ou faisait suite à une autre tumeur maligne) ont été retenus dans l'analyse (Rosso *et al.* 2009). Ces critères d'inclusion sont similaires à ceux définis lors des précédentes études du groupe EURO CARE.

Au total, l'analyse a porté sur 13 988 diagnostics de LF et 25 320 diagnostics de LBDGC provenant de 43 registres de cancers (15 pays représentés) sur la période 1996 à 2004. La date de point était au 31 décembre 2008 pour tous les registres, à l'exception des registres français pour lesquels le statut vital des patients a été actualisé au 31 décembre 2007. Les tables des taux de mortalité attendue étaient disponibles pour chaque registre par sexe, âge (en année) et par année calendaire. Cinq zones européennes ont été définies en se basant sur les regroupements habituels des précédentes études EURO CARE : Europe du Nord, Ecosse/Pays de Galles, Europe centrale, Europe du Sud et Europe de l'Est.

IV.3. MÉTHODOLOGIE

Estimation ponctuelle de la survie nette en fonction des périodes de diagnostic.

L'année de diagnostic a été considérée comme une variable catégorielle et découpée selon 3 périodes de diagnostic : 1996-1998, 1999 -2001 et 2002-2004. Pour chaque sous-type de lymphome et pour chaque zone européenne, la survie nette à 5 ans a été estimée au sein de

¹³ Onze registres de cancer ont été exclus de par un pourcentage de lymphome-SAI dépassant 30% par rapport au nombre total de LMNH sur la période d'étude.

chaque classe d'âge en fonction des périodes de diagnostic avec l'estimateur de Pohar-Perme (Perme *et al.* 2012), puis standardisée en utilisant les poids ICSS (Corazziari *et al.* 2004).

Modélisation du taux de mortalité en excès en fonction de l'année de diagnostic.

Afin de mettre en évidence les effets de l'année de diagnostic sur la mortalité en excès, une approche basée sur la modélisation du taux en excès a été utilisée (Remontet *et al.* 2007). Dans cette modélisation, l'année a été conservée en variable continue. Pour chaque zone européenne et pour chaque sous-type de lymphome, le taux de mortalité en excès ainsi que les effets des covariables sur cette mortalité ont été modélisés jusqu'à 5 ans après le diagnostic du lymphome. Comme des précédentes études de survie ont rapporté un effet souvent complexe de l'âge au diagnostic sur le taux de mortalité, chacun des modèles a ici été ajusté sur un effet non linéaire et non proportionnel de l'âge. Une stratégie de construction de modèle proposée par Wynant et Abrahamowicz (Wynant et Abrahamowicz 2014) a été retenue afin de tester les effets de l'année de diagnostic sur la mortalité tout en ajustant sur les effets complexes des autres covariables du modèle (ici l'âge au diagnostic). Cinq modèles ont été comparés en considérant :

- un effet non linéaire et non proportionnel de l'année (NL / NP) :

Modèle 1 : $\ln[\lambda_c(t, \hat{age}, année)] = f(t) + g(\hat{age}) + h(t) * \hat{age} + i(année) + j(t) * année$

- un effet linéaire et non proportionnel de l'année (LL / NP) :

Modèle 2 : $\ln[\lambda_c(t, \hat{age}, année)] = f(t) + g(\hat{age}) + h(t) * \hat{age} + année + j(t) * année$

- un effet non linéaire et proportionnel de l'année (NL / PP) :

Modèle 3 : $\ln[\lambda_c(t, \hat{age}, année)] = f(t) + g(\hat{age}) + h(t) * \hat{age} + i(année)$

- un effet linéaire et proportionnel de l'année (LL / PP) :

Modèle 4 : $\ln[\lambda_c(t, \hat{age}, année)] = f(t) + g(\hat{age}) + h(t) * \hat{age} + année$

- pas d'effet de l'année :

Modèle 5 : $\ln[\lambda_c(t, \hat{age}, année)] = f(t) + g(\hat{age}) + h(t) * \hat{age}$

λ_c correspond au taux instantané de mortalité en excès due au cancer, t au temps depuis le diagnostic et les fonctions f, g, h, i et j représentent:

- f : une spline cubique avec un nœud à 1 an de suivi pour l'effet du temps depuis le diagnostic.
- g : une spline quadratique avec un nœud à l'âge moyen pour l'effet non linéaire de l'âge.
- h : une spline cubique avec un nœud à 1 an de suivi pour l'effet non proportionnel de l'âge.
- i : une spline quadratique avec un nœud à l'année 1998 pour les LF et à l'année 2002 pour les LBDGC, pour l'effet non linéaire de l'année.
- j : une spline cubique avec un nœud à 1 an de suivi pour l'effet non proportionnel de l'année.

Les années de référence 1998 et 2002 marquent l'arrivée de la thérapie ciblée dans la prise en charge thérapeutique des LF et des LBDGC.

Notre stratégie a consisté en une élimination pas à pas descendante des effets non significatifs de l'année de diagnostic en comparant les modèles emboîtés deux à deux à l'aide d'un test de rapport de vraisemblance (au seuil de 0,05). Deux bras de comparaison ont été effectués en partant du modèle le plus complexe (Figure II.6). Comme l'année peut ne pas avoir un effet identique en fonction des zones européennes, une stratégie de construction spécifique à chaque zone et pour chaque sous-type de lymphome, a été effectuée.

A partir de chaque modèle retenu, deux indicateurs ont été estimés :

- la dynamique du taux de mortalité en excès en fonction du temps de suivi depuis le diagnostic, estimée pour trois années fixées (1996, 2000 et 2004), pour tous âges et à âge fixé (à 30, 55 et 75 ans, ces âges étant représentatifs d'une éventuelle prise en charge différente du LMNH)
- la probabilité de survie nette à 5 ans en fonction de l'année de diagnostic, pour tous âges et à âge fixé (à 30, 55 et 75 ans).

L'ensemble des analyses a été réalisé sur le logiciel R avec le package « Relsurv » développé par Pohar-Perme et avec les fonctions développées par Remontet *et al.*.

IV.4. ARTICLE

Les résultats de ce travail ont été acceptés dans la revue *The Lancet Haematology* sous forme d'un article scientifique et sont en cours de publication.

Changes in the dynamics of the excess mortality rates after diagnosis of follicular or diffuse large B-cell lymphoma: comparisons between European population-based data

Morgane Mounier (MSc)¹⁻⁶, Nadine Bossard (MD)^{2-4,7}, Laurent Remontet (MSc)^{2-4,7}, Aurélien Belot (PhD)⁸, Pamela Minicozzi (PhD)⁹, Roberta De Angelis (MSc)¹⁰, Riccardo Capocaccia (MSc)¹⁰, Jean Iwaz (PhD)^{2-4,7}, Alain Monnereau (MD)^{11,12}, Xavier Troussard (MD)¹³, Milena Sant (MD)⁹, Marc Maynadié (MD)^{1,14}, Roch Giorgi (MD)^{5,6}, the EURO CARE-5 Working Group¹⁵ and CENSUR Working survival Group.

¹ Registre des Hémopathies Malignes de Côte d'Or, EA 4184, Université de Bourgogne Franche-Comté, Dijon, France;

² Université de Lyon, Lyon, France;

³ Université Lyon 1, Villeurbanne, France;

⁴ CNRS UMR5558, Laboratoire de Biométrie et Biologie Evolutive, Equipe Biostatistique Santé, Villeurbanne, France;

⁵ INSERM, UMR_S 912 "Economics and Social Sciences Applied to Health & Analysis of Medical Information" (SESSTIM), Marseille, France;

⁶ Aix Marseille Université, UMR_S 912, IRD, Marseille, France;

⁷ Service de Biostatistique, Hospices Civils de Lyon, Lyon, France;

⁸ Cancer Research UK Cancer Survival Group, Department of Non-Communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom;

⁹ Analytical Epidemiology and Health Impact Unit, Department of Preventive and Predictive Medicine, Fondazione IRCCS, Istituto Nazionale dei Tumori, Milan, Italy;

¹⁰ Centro Nazionale di Epidemiologia, Sorveglianza e Promozione della Salute (CNESPS), Istituto Superiore di Sanità, Roma, Italy;

¹¹ Registre des hémopathies Malignes de la Gironde, Institut Bergonié, Bordeaux, France;

¹² Centre Inserm U897, CIC 1401, Bordeaux, France;

¹³ Registre régional des hémopathies Malignes de la Basse Normandie, Centre Hospitalier Universitaire de Caen, Caen, France;

¹⁴ Service d'Hématologie Biologique, Centre Hospitalier Universitaire de Dijon, Dijon, France;

¹⁵ Members listed in Appendix.

SUMMARY

Background: Since 2001, the World Health Organisation classification of lymphoid neoplasms and the International Classification of Diseases for Oncology, 3rd Edition has improved data collection of lymphoma subtypes in most European cancer registries and allowed reporting on the major non-Hodgkin lymphoma subtypes. Besides, the treatment of non-Hodgkin lymphoma has undergone profound changes that were first beneficial to patients with follicular lymphoma or diffuse large B-cell lymphoma (DLBCL). Using data on survival from the largest number of collaborative European population-based cancer registries (EUROCARE), the aim was to compare trends in dynamics of the cancer mortality after follicular lymphoma or DLBCL between five large European areas.

Methods: 13988 follicular lymphoma cases and 25320 DLBCL cases in patients over 13 years, diagnosed between 01-01-1996 and 12-31-2004 and recorded in 43 cancer registries, were considered over five areas: Scotland/Wales plus Northern, Central, Eastern, and Southern Europe. In most cases, the vital status could be updated on 12-31-2008. Changes in net survival and excess mortality rate (EMR) were obtained using, respectively, the Pohar-Perme estimator and the flexible parametric model of Remontet *et al.* according to age and the year of diagnosis.

Findings: In the most recent years of diagnosis (past 2002), improvements of the five-year net survival were observed after both cancers, especially follicular lymphoma. An area disparity was seen over the 2002-2004 period: the age-standardized five-year net survival probabilities for follicular lymphoma and DLBCL were, respectively, 61% [95% CI: 54-69] and 46% [43-50] in Eastern Europe vs. 75% [69-80] and 58% [4-62] in Northern Europe. There was a significant effect of the year of diagnosis on the EMR at all ages in all areas, except for DLBCL in Eastern Europe. The EMR was not constant over the follow-up period: a high rate was observed early for both lymphomas except for follicular lymphoma in Northern Europe.

Interpretation: Despite the recent progress in the management of lymphoma, area-related low chances of survival persist, especially in the youngest and oldest patients.

KEY WORDS: EURO CARE-5 population based, excess mortality, net survival, follicular lymphoma, diffuse large B-cell lymphoma, joint modelling of complex effects of covariates.

Funding: The study was supported by Compagnia di San Paolo, Fondazione Cariplo Italy, the Italian Ministry of Health (Ricerca Finalizzata 2009, RF-2009-1529710), the European Commission (European Action Against Cancer, EPAAC, Joint Action No20102202), the Registre des Hémopathies Malignes de Côte d'Or, and the French Agence Nationale de la Recherche (ANR CENSUR project, grant number ANR-12-BSV1-0028).

INTRODUCTION

Non-Hodgkin lymphoma is a large heterogeneous group of lymphoid neoplasms with distinct biological, clinical, and prognostic features. This highlights the importance of reporting epidemiological results per subtype.¹ The 2001 World Health Organisation (WHO) classification of lymphoid neoplasms² and the International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3)³ have improved data collection on lymphoma subtypes in most European cancer registries and allowed reporting on the major non-Hodgkin lymphoma subtypes. For lymphoma, the new classification was largely based on the Revised European American Lymphoma Classification published in 1994.

Meanwhile, two major advances occurred: better disease knowledge and the development of the anti-CD20 chimeric monoclonal antibody; i.e., rituximab. Follicular lymphoma and diffuse large B-cell Lymphoma (DLBCL) patients were the first to benefit from the latter immunotherapy which improved overall survival in many clinical trials or population-based studies.^{4,5} However, disparities in drug availability between European regions created survival differences over periods of diagnosis or ages at diagnosis.⁶

The HAEMACARE project improved the coding rules of lymphoid neoplasm in Europe and published the first relative survival probabilities after major Non-Hodgkin lymphoma subtypes between 2000 and 2002.⁷ In a recent update, the changes in 5-year relative survivals after major haematological malignancy subtypes between 1997-99 and 2006-2008 showed persistent survival differences across Europe.⁸

Interpreting differences in all-cause survival after cancer between areas is difficult because differences in observed mortalities may be due to differences in expected mortality and not differences in cancer-specific mortalities. An alternative is comparing net survivals;

i.e., the probabilities of survival the disease being the only cause of death.⁹ The net survival is obtained from excess mortality rate (EMR) which is estimated by removing from the observed mortality rate of the cancer patients the expected mortality rate from other causes. The latter rate is considered equal to the all-cause mortality in the general population assuming the disease-specific mortality is negligible vs. other-cause mortality. Briefly, the EMR estimates how higher is the mortality rate of cancer patients with respect to the mortality rate expected in the general population.

Obtaining unbiased estimates of net survival requires allowing for informative censoring (i.e., due to covariates with simultaneous effects on disease-specific and other-cause mortalities, such as age).^{10,11} This is ensured by two methods: the Pohar-Perme non-parametric estimator¹² and the additive excess-mortality model^{13,14} adjusted on demographic covariates. Because the effects of some covariates (age or year of diagnosis) on the excess hazard of death may change over time elapsed since diagnosis and because the effects of some continuous covariates may be non-linear, a flexible excess hazard model was developed¹⁴ to model these effects simultaneously. This allows exploring the dynamics of the EMR over the follow-up period and the change of these dynamics according to the year of diagnosis.

The present work uses data on survival collected by the largest number of collaborative European population-based cancer registries.¹⁵ It compares trends in dynamics of the follicular lymphoma and DLBCL EMR and in net survival between European countries using the most recent net survival methodology.

MATERIALS AND METHODS

Study design and data collection

The current EURO CARE-5 project includes more than 21 million cancers recorded in 107 population-based registries from 30 European countries.¹⁵ The following is available on each patient: sex, date of birth, date of cancer diagnosis, cancer morphology and topography, basis of diagnosis, vital status, and date of death or last known vital status.

The neoplasm morphology was coded according to the ICD-O-3.³ A standardised quality control¹⁵ searched for inconsistencies such as mismatches between sex, age, topography, and morphology. Dubious records were checked at the registries and confirmed or corrected.⁶ The study considered eligible all follicular lymphomas and DLBCLs diagnosed Jan 1, 1996 to Dec 31, 2004 in persons aged ≥ 15 years. As usual in EURO CARE and in all cancer survival studies, cases incidentally diagnosed at autopsy or known from death certificates only were excluded.

The lymphoma groups were based on the WHO classification.² For follicular lymphoma, we considered ICD-O-3 codes 9675/3 (Follicular lymphoma, used before 2000 but not thereafter for mixed small and large cell, diffuse malignant lymphoma), 9690/3 (Follicular lymphoma, not otherwise specified, NOS), as well as codes 9695/3, 9691/3, and 9698/3 (Follicular lymphomas grades 1, 2, and 3, respectively). For DLBCL, we considered codes 9678/3 (Primary effusion lymphoma), 9679/3 (Mediastinal large B-cell lymphoma), 9680/3 (Diffuse large B-cell lymphoma), and 9684/3 (Immunoblastic large B-cell diffuse lymphoma).

The vital status was followed until Dec 31, 2008 in all registries but the French ones (Dec 31, 2007).

After receiving the agreements of the cancer registries to participate, we kept only cancer registries with complete incidence data registration. The indicators of exhaustiveness were: the percentage of unspecified cases and the percentage of loss to follow-up. Consistently with previous EURO CARE studies, we excluded cancer registries that reported >30% lymphoid neoplasms NOS out of all Non-Hodgkin lymphomas diagnosed 1996-2004 (i.e., ICD-O-3 codes 9590/3 for Lymphoma NOS and 9591/3 for Non-Hodgkin lymphoma NOS).

Finally, the study included 43 population-based cancer registries grouped into five areas: Northern Europe (Iceland and Norway), Scotland/Wales (Scotland and Wales), Central Europe (France, Germany, Switzerland, and The Netherlands), Southern Europe (Italy, Malta, Slovenia, and Spain), and Eastern Europe (Estonia, Lithuania, and Slovakia). Table 1 shows the quality of follicular lymphoma and DLBCL records in each included cancer registry.

Estimation of net survival

Net survival probabilities (with 95% confidence intervals, CIs) were estimated by the Pohar-Perme estimator.¹² Five-year net survival was estimated per subtype, area, age at diagnosis (in four classes), period of diagnosis (in three classes) and sex.

The general population life-tables were obtained from registry-specific official mortality data stratified by sex, age (in one-year classes), and calendar year. Net survival estimates were age-standardized according to the International Cancer Survival Standards.¹⁶

These survival analyses used package “*relsurv*” version 1.6.4 (www.R-project.org).

Statistical modelling of the EMR

We used a flexible version of the excess-mortality models¹⁴ to quantify the effects of the year of diagnosis on mortality. This approach uses regression splines and allows for non-proportional (i.e., time-dependent) and/or non-linear covariate effects. For each area and subtype, the EMR was modelled up to five years after diagnosis, the baseline hazard being modelled as a continuous function.

The covariates considered were the age at diagnosis and the year of diagnosis. The effect of age is known to be non-linear and time-dependent¹⁷; thus, both characteristics were systematically specified in the considered models. Regarding the year of diagnosis, the effect(s) to include (non-linear and/or time-dependent) were selected with a specific model-building strategy¹⁸ applied separately for each area and each subtype (Appendix p.7-9). The EMR dynamics and the trends of the 5-year net survival at specific ages and/or specific years of diagnosis were obtained from the final model: the curves (for EMR or net survival) are then continuous and possibly non-linear according to the retained model.

For this, we used functions previously developed in R software.¹⁴

Roles of the funding sources

The funding source had no role in study design, data collection, data analysis, result interpretation or writing of the report.

RESULTS

Data quality controls

The data-quality-control results are shown in Table 1. The loss to follow-up in each area was $\leq 1.9\%$. Only 0.5% of all follicular lymphoma and DLBCL cases were excluded from the analyses because diagnosed at autopsy or registered from death certificates.

Characteristics of the study cohort

The study analysed 13988 follicular lymphoma and 25320 DLBCL cases whose characteristics are shown in Appendix (Tables S1 and S2, p.2-3).

Women constituted the majority of follicular lymphoma cases: 53% of all cases but 58% of Eastern European cases (Appendix Table S1, p.2). Follicular lymphoma affected all age classes. The median age at diagnosis was 62 years (13% in class 15-44, 44% in class 45-64, 24% in class 65-74, and 19% in people aged 75 and older). In the whole cohort, 35% of follicular lymphoma patients died within five years after diagnosis: 32% in Northern and Southern Europe vs. 45% in Eastern Europe. The median follow-up over all areas was 8 years.

Men constituted the majority of DLBCL cases: 52% of all cases, 46% of Eastern European cases but 53% of Northern European cases (Appendix Table S2, p.3). DLBCL affected rather elderly patients. The median age at diagnosis was 68 years (13% in class 15-44, 29% in class 45-64, 26% in class 65-74, and 32% in people aged 75 and older). In the whole cohort, 57% of DLBCL patients died within five years after diagnosis. The median follow-up over all areas was 7 years.

Temporal trends of the 5-years net survival by period of diagnosis

Age-specific and age-standardized net survival probabilities at five-year were estimated over three-year periods of diagnosis for each area and subtype (Figure 1). The number at risk in each group and the number of deaths at five-years is shown in the Appendix (Tables S3 and S4, p.4-5).

In follicular lymphoma patients, the net survival increased in the last years (2002-04) in all areas, except in Scotland/Wales where it remained stable after 2001. Whatever the age class, the largest net survival disparity occurred between Eastern and Northern Europe with in period 2002-04 the five-year age-standardized net survival was equal to 61% [54-69] vs. 75% [69-80] for follicular lymphoma and 46% [43-50] vs. 58% [54-62] for DLBCL.

Whatever the subtype, the net survival was the lowest in the oldest patients. The largest difference in net survival according to age in period 2002-04 occurred in Southern Europe for follicular lymphoma and Northern Europe for DLBCL: the absolute differences the youngest and the oldest patients were 40% and 49%, respectively.

As regards sex, the 5-year net survival in follicular lymphoma was slightly higher in women than in men in both European areas; this difference was less marked in DLBCL (Appendix Figure S1, p.6).

Dynamics of the excess mortality rate according to the year of diagnosis

Figures 2 and 3 report the excess mortality rate according to time since diagnoses in follicular lymphoma - and DLBCL- in each European area as estimated at three fixed years of diagnosis (1996, 2000, and 2004) and three ages at diagnosis (30, 55, and 75 years).

In follicular lymphoma patients of all ages, the EMR dynamic differed widely between areas (Figure 2): in Northern Europe, the EMR remained stable over five years after diagnosis but in the other areas, especially in Eastern Europe, the EMR was high just after diagnosis. After one year, the EMRs were comparable between areas and did not exceed 0.2 deaths per person-year. In patients aged 30 years, a peculiar EMR dynamic was seen in the Scotland/Wales and Eastern Europe. In patients aged ≥ 75 years, high short-term mortality rates were seen in all areas but Northern Europe. The EMR decreased during the last years of diagnosis in all age classes and all areas. In 1996, in the Scotland/Wales, Southern Europe, and Eastern Europe, the EMR was higher than in the two other areas but it decreased over time to reach similar values in 2004 (Figure 2).

In DLBCL patients of all ages, the EMR dynamic was nearly the same in all areas: the high rates seen just after diagnosis decreased thereafter (Figure 3). In patients aged 30 years, the mortality rates just after diagnosis were higher in Central, Southern, and Eastern Europe than in Northern Europe and the Scotland/Wales. In patients aged 55 years, during the first year, the EMR was higher in Eastern Europe than in the other areas; this was less marked in patients ≥ 75 years.

The EMRs did not differ much according to the year of diagnosis: no significant effect was seen in Eastern Europe but a slight effect was seen in the last years of diagnosis, especially at the start of follow-up, in Northern, Central, and Southern Europe (Figure 3). A peculiar trend was observed in the Scotland/Wales: during the first year, the EMR was higher in 2004 than in 1996 whatever the age.

The trends of the 5-year net survival according to the year of diagnosis as continuous variable were estimated by area- and subtype-specific models at ages 30, 55, and 75 years

(Figure 4). Figure 4 shows how the changes with year of diagnosis in the dynamics of EMR finally impact the 5 year net survival.

In follicular lymphoma, the net survival improved progressively in Northern, Central, and Eastern Europe with a significant effect (Appendix p.9). In the Scotland/Wales and Southern Europe, the year of diagnosis had a significant non-linear effect on the EMR. In the Scotland/Wales, the 5-year net survival increased until 1998 and stabilised thereafter. In Southern Europe, the 5-year net survival started to increase after 1998 (Figure 4). Despite these increases, the area disparity observed in Figure 1 was also seen in Figure 4: in patients 30 years old, the net survival was shorter in the Scotland/Wales and Eastern Europe than in Northern, Central, and Southern Europe. In patient aged 55 or 75 years old, the net survival in Eastern Europe was shorter than in the other areas (Figure 4).

In DLBCL, the net survival increased slightly over the last years of diagnosis except in Eastern Europe. The slight increase seen in Figure 4 for "all ages" would reflect a change in age structure over the years of diagnosis. In Central and Southern Europe, the year of diagnosis had a significant linear effect on the EMR (Appendix p.9) but it had a non-linear effect on the EMR in the Scotland/Wales and Northern Europe where the net survival improved significantly after 2002 for all ages (Figure 4). Survival disparities were more noticeable in patients aged 30 years; the prognosis was better in Northern Europe than in the other areas (Figure 4).

DISCUSSION

This retrospective EUROCARE-5 population-based study assessed the dynamics of the EMR and the net survival of the two most frequent Non-Hodgkin lymphoma subtypes (follicular lymphoma and DLBCL) diagnosed in several European areas over nine years.

Regarding follicular lymphoma, we showed higher net survival probabilities during the last years of diagnosis (2002-04) in all areas and age groups and a constant effect of the year of diagnostic on the EMR. Regarding DLBCL, the net survival increased slightly over the study period, except in Eastern Europe. Despite different statistical approaches and cohorts, these results agree with the previous EUROCARE-5 report on haematological malignancies⁸ and with other population-based studies^{19,20} confirming the analyses robustness and the European dataset reliability. Unlike previous EUROCARE studies,^{6,8} Northern Europe data here relied essentially on Norway cancer registry; however, Non-Hodgkin lymphoma survival seems homogenous in Nordic countries.²¹ Similarly, the UK-England data here relied on Wales and Scotland and did not include England, Ireland, and Northern Ireland; this calls for caution though survival differences between these regions seem negligible.²²

Several facts may explain the improvement of follicular lymphoma and DLBCL survival during the last study years in all areas. First, the progresses in diagnosis and staging tools (i.e., molecular markers, immunophenotyping, imaging techniques) led to better diagnoses, classifications, disease extension evaluations, and therapeutic decisions.²³ Second, the use of new drug (the anti-CD20 chimeric monoclonal antibody rituximab) for follicular lymphoma and DLBCL in combination with the conventional chemotherapies,^{4,5} and the wider recourse to autologous stem-cell transplantation in the youngest patients with relapse²⁴ increased survival. Third, the guidelines for daily practice^{25,26} issued in most European

countries (i.e., rituximab plus polychemotherapy for induction and maintenance) or “watchful waiting” for follicular lymphoma with high tumour burden or DLBCL were also beneficial.

Nevertheless, age-related differences in prognosis still exist between European areas. Various studies have already reported poorer survival in elderly patients.^{1,6,8} Here, we show more marked differences in Southern Europe for follicular lymphoma and Northern Europe for DLBCL. Two main explanations could be: i) difference in care provision or clinical practices; i.e., more aggressive treatments in young patients or sub-optimal in the elderly;²⁷ ii) more inclusions of young patients in clinical trials with shorter delays to treatment.²⁷

Besides, area-related differences in survival were expected and may have several major explanations: i) different paces in the use of new drugs; ii) lack of access to modern treatments in Eastern Europe;⁸ iii) uneven guideline implementation or accuracy (e.g., radiotherapy doses for follicular lymphoma stages I and II); iv) optimal chemotherapy and costly and short lead-time treatments in the youngest patients (which reduces early mortality).²¹ Other minor explanations may be: different case-mixes, suboptimal “wait and watch” procedures, work in multidisciplinary teams,²⁸ different socio-economic conditions and life habits, different health insurance schemes (that may delay diagnosis), or different HIV prevalence rates.²⁹ As in every survival study, another potential cause of differences in survival could be a non-comparability of follicular lymphoma and DLBCL incidence rates over the study period due to different sensitivities of the medical systems between European areas, different qualities of recording by cancer registries, and different percentages of unspecified cases (Table 1). In fact, a low follicular lymphoma or DLBCL incidence rate may be explained by non-detection of the most serious cases; the probability of survival would be then biased toward an overestimation.

This study shows mainly the EMR dynamics over time of follow-up. In Northern Europe, the EMR was low and constant; however, EMRs were observed during the first year in the other areas for follicular lymphoma and in all areas for DLBCL. Possible explanations could be longer delays to treatment or access to aggressive treatments.

All these assumptions require confirmation and detailed database, since not routinely collected by the cancer registries, on: i) lymphoma stage: an unequal repartition of stage over the study period would explain survival differences between European areas; ii) cancer management: the use of new drugs at various moments in several countries could be an explanation for these survival differences. The lack of data on treatment in EURO CARE-5 database did not allow investigating the differences in therapeutic management in the present study. We used the year of diagnosis as an indirect indicator of change in follicular lymphoma and DLBCL therapeutic management over time in each European area; and, iii) better recording of co-morbidities and socio-economic facts. Additionally, the period of diagnosis of our data did not allow us to evaluate recent advances in the management of lymphoma (mainly, the use of new drugs) or the impact of new prognostic indexes on practitioners' treatment decisions.

One asset and originality of our study is its estimation methodology. First, Pohar-Perme's estimator provided unbiased punctual estimations of net survival and age-standardisation allowed comparisons with other European studies. Actually, the relative 5-year survival estimated by Ederer II method were close to those estimated by the Pohar-Perme estimator;³⁰ hence, the present results are close to previous EURO CARE-5.⁸ Second, the flexible regression was essential to describe the EMR dynamics over the years of diagnosis. We focused our analysis on the dynamics of the EMR according to the year of diagnosis and age in both European areas but did not compare survivals and excess mortality rates between

areas. Third, modelling with the year of diagnosis as continuous variable allowed us to display curves changes, which were not obvious on Figure 1. Our model-building strategy¹⁸ allowed considering simultaneously the impacts of the year of diagnosis and that of the age at diagnosis on the EMR, with possible non-linear and/or non-proportional effects for both covariates. A special attention was paid to the trend in the expected mortality according to the age at diagnosis in each cancer registries. The study of these trends by sex showed that the variability was low according to the year of diagnosis and between areas (results not shown).

We focused our analysis on the effect of the year of diagnosis and of the age at diagnosis on mortality rather than on the effect of sex which is less influential on patients' prognosis. Modelling the additional effect of sex would require a far more complex modelling strategy and the need of a specific study. Results were thus shown for all patients together.

This EURO CARE population-based study showed that despite international guidelines^{25,26} and progress of healthcare, age-related differences in survival after follicular lymphoma or DLBCL still exist between European areas. However, the decreasing EMRs should foster the search for more and better means of improving the access to modern care, especially for the youngest patients in Eastern Europe.

RESEARCH IN CONTEXT

Previous evidence: European collaborations allowed building the EURO CARE cancer database. Simultaneously, the HAEMACARE project improved the availability and homogeneity of haematological cancer data in Europe. Both enabled more reliable comparisons between countries. A previous EURO CARE-5 study has shown improvements in survival after non-Hodgkin lymphomas in Europe, especially follicular lymphoma and diffuse large B-cell lymphoma (DLBCL). The present study assessed the change in dynamics of the EMR and the trend of net survival after follicular lymphoma or DLBCL and compared survival between five large European areas. It considered follicular lymphoma and DLBCL cases diagnosed Jan 1, 1996 - Dec 31, 2004 in persons aged ≥ 15 years and tracked the vital status until Dec 31, 2008. It included 43 registries.

Added value: This study reported comparison using the probability of net survival (the most appropriate indicator for comparisons between areas) and its associated excess mortality rate (EMR, an interesting indicator of the dynamics of death that was never or rarely used previously), along the five years of follow-up. In addition, the study offers an original flexible parametric modelling of the EMR in follicular lymphoma and DLBCL, allowing the effects of age and year of diagnosis being linear/non-linear, and proportional/non-proportional. Results showed significant decreases in follicular lymphoma - and DLBCL-related EMRs according to the year of diagnosis in all areas, except for patient with DLBCL in Eastern Europe. More or less high EMRs were observed early after follicular lymphoma or DLBCL diagnosis according to area or age. Disparities between areas were seen past 2002; the largest EMR differences being observed between Eastern and Northern Europe, especially regarding the youngest and oldest patients.

Implications: The study of the dynamics of EMRs and of the corresponding net survival values help visualizing variations during follow-up and between area disparities. The latest encouraging survival results should foster the search for more and better means of improving the access to adequate care, especially for the youngest patients in Eastern Europe.

AUTHORS' CONTRIBUTIONS

R. Giorgi, N. Bossard, M. Maynadié, L. Remontet, A. Belot and M. Mounier designed the study. EURO CARE provided the access to its European database. P. Minicozzi, R. De Angelis, R. Capocaccia and M. Sant extracted, pooled, and tabulated//organized the data. M. Mounier, N. Bossard, R. Giorgi, L. Remontet, A. Belot, M. Maynadié, P. Minicozzi, R. De Angelis, R. Capocaccia analyzed and interpreted the data. M. Mounier, R. Giorgi, N. Bossard, A. Belot, L. Remontet, M. Maynadié, P. Minicozzi, R. De Angelis, R. Capocaccia, M. Sant, A. Monnereau, X. Troussard and J. Iwaz wrote the successive versions of the manuscript. All authors read and approved the final version of the manuscript.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

- 1 Sant M, Allemani C, De Angelis R, et al. Influence of morphology on survival for non-Hodgkin lymphoma in Europe and the United States. *Eur J Cancer* 2008; 44: 579–87.
- 2 Jaffe ES, Harris NL, Stein H. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. IARC. Press: Lyon: Vardiman J.W., 2001.
- 3 Fritz AG. International Classification of Diseases for Oncology: ICD-O. World Health Organization, 2000.
- 4 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.
- 5 Salles G, Mounier N, de Guibert S, et al. Rituximab combined with chemotherapy and interferon in follicular lymphoma patients: results of the GELA-GOELAMS FL2000 study. *Blood* 2008; 112: 4824–31.
- 6 De Angelis R, Sant M, Coleman MP, et al. Cancer survival in Europe 1999-2007 by country and age: results of EUROCORE--5-a population-based study. *Lancet Oncol* 2014; 15: 23–34.
- 7 Marcos-Gragera R, Allemani C, Tereanu C, et al. Survival of European patients diagnosed with lymphoid neoplasms in 2000-2002: results of the HAEMACARE project. *Haematologica* 2011; 96: 720–8.

- 8 Sant M, Minicozzi P, Mounier M, et al. Survival for haematological malignancies in Europe between 1997 and 2008 by region and age: results of EURO CARE-5, a population-based study. *Lancet Oncol* 2014; 15: 931–42.
- 9 Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: elements for further discussion. *Stat Med* 1990; 9: 529–38.
- 10 Danieli C, Remontet L, Bossard N, Roche L, Belot A. Estimating net survival: the importance of allowing for informative censoring. *Stat Med* 2012; 31: 775–86.
- 11 Roche L, Danieli C, Belot A, et al. Cancer net survival on registry data: use of the new unbiased Pohar-Perme estimator and magnitude of the bias with the classical methods. *Int J Cancer* 2013; 132: 2359–69.
- 12 Perme MP, Stare J, Estève J. On estimation in relative survival. *Biometrics* 2012; 68: 113–20.
- 13 Giorgi R, Abrahamowicz M, Quantin C, et al. A relative survival regression model using B-spline functions to model non-proportional hazards. *Stat Med* 2003; 22: 2767–84.
- 14 Remontet L, Bossard N, Belot A, Estève J, French network of cancer registries FRANCIM. An overall strategy based on regression models to estimate relative survival and model the effects of prognostic factors in cancer survival studies. *Stat Med* 2007; 26: 2214–28.
- 15 EURO CARE Survival of cancer patients in Europe. <http://www.eurocare.it> (accessed Aug 11, 2014).
- 16 Corazziari I, Quinn M, Capocaccia R. Standard cancer patient population for age standardising survival ratios. *Eur J Cancer* 2004; 40: 2307–16.

- 17 Bossard N, Velten M, Remontet L, et al. Survival of cancer patients in France: a population-based study from The Association of the French Cancer Registries (FRANCIM). *Eur J Cancer* 2007; 43: 149–60.
- 18 Wynant W, Abrahamowicz M. Impact of the model-building strategy on inference about nonlinear and time-dependent covariate effects in survival analysis. *Stat Med* 2014; 33: 3318–37.
- 19 Bosetti C, Bertuccio P, Malvezzi M, et al. Cancer mortality in Europe, 2005–2009, and an overview of trends since 1980. *Ann Oncol* 2013; 24: 2657–71.
- 20 Monnereau A, Troussard X, Belot A, et al. Unbiased estimates of long-term net survival of hematological malignancy patients detailed by major subtypes in France. *Int J Cancer* 2013; 132: 2378–87.
- 21 Storm HH, Klint Å, Tryggvadóttir L, et al. Trends in the survival of patients diagnosed with malignant neoplasms of lymphoid, haematopoietic, and related tissue in the Nordic countries 1964–2003 followed up to the end of 2006. *Acta Oncol* 2010; 49: 694–712.
- 22 Quaresma M, Coleman MP, Rachet B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971-2011: a population-based study. *The Lancet* 2015; 385: 1206–18.
- 23 Al-Hamadani M, Go RS. Stage redistribution and survival among Medicare beneficiaries before and after the approval of positron emission tomography scan for non-Hodgkin lymphoma. *Leuk Lymphoma* 2014; nov-5: 1–3.(epub ahead of print)
- 24 Le Gouill S, De Guibert S, Planche L, et al. Impact of the use of autologous stem cell transplantation at first relapse both in naive and previously rituximab exposed follicular

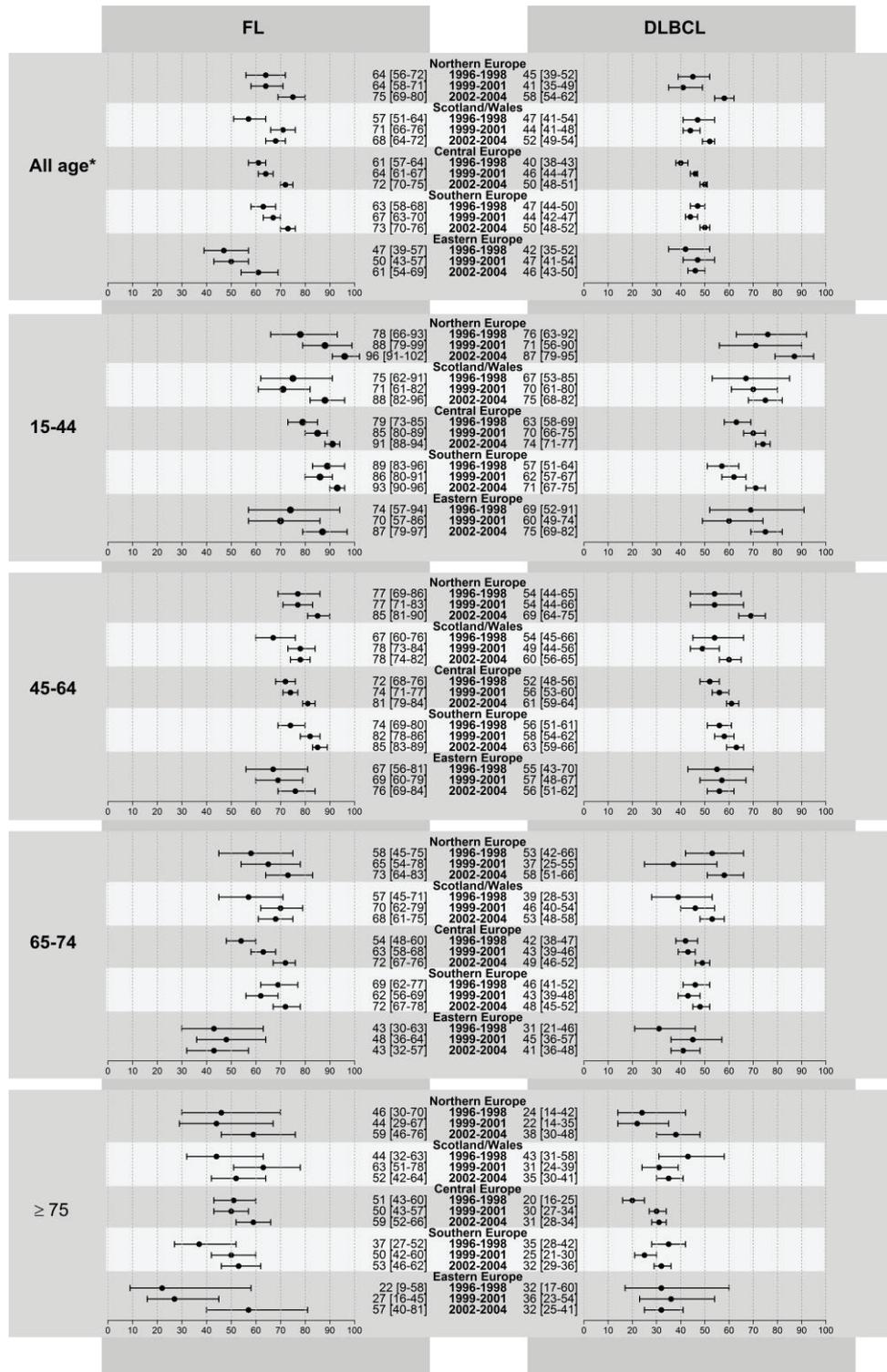
- lymphoma patients treated in the GELA/GOELAMS FL2000 study. *Haematologica* 2011; 96: 1128–35.
- 25 Hiddemann W, European Society for Medical Oncology. ESMO Minimum Clinical Recommendations for diagnosis, treatment and follow-up of newly diagnosed follicular lymphoma. *Ann Oncol* 2003; 14: 1163–4.
- 26 ESMO minimum clinical recommendations for diagnosis, treatment and follow-up of newly diagnosed large cell non-Hodgkin's lymphoma. *Ann Oncol* 2001; 12: 1209–10.
- 27 Fields PA, Linch DC. Treatment of the elderly patient with diffuse large B cell lymphoma. *Br J Haematol* 2012; 157: 159–70.
- 28 Munro AJ. Leukaemia and lymphoma: why has survival improved? *Lancet Oncol* 2014; 15: 906–7.
- 29 Tao L, Foran JM, Clarke CA, Gomez SL, Keegan THM. Socioeconomic disparities in mortality after diffuse large B-cell lymphoma in the modern treatment era. *Blood* 2014; 123: 3553–62.
- 30 Seppä K, Hakulinen T, Pokhrel A. Choosing the net survival method for cancer survival estimation. *Eur J Cancer* 2013; published online Oct 29. DOI:10.1016/j.ejca.2013.09.019

Area, country, region	Cases 1996-2004				Quality indicators for DLBCL and FL	Excluded for DLBCL and FL	Included	
	All NHLs N	DLBCL N (%)	FL N (%)	Unspecified N (%)	Loss to follow-up N (%)	(DCO or autopsy) N (%)	DLBCL N (%)	FL N (%)
<i>Northern Europe</i>	8581	1468 (17)	1235 (14)	1863 (22)	3 (0.1)	4 (0.2)	1465 (99)	1234 (99)
Iceland	451	129 (29)	71 (16)	17 (4)	0 (0.0)	3 (1.5)	126 (98)	71 (100)
Norway	8130	1339 (16)	1164 (14)	1846 (23)	3 (0.1)	1 (0.1)	1339 (100)	1163 (99)
<i>Scotland/Wales</i>	16491	2972 (18)	1985 (12)	4408 (27)	1 (0.2)	10 (0.2)	2964 (99)	1983 (99)
Scotland	10338	2012 (19)	1327 (13)	2564 (25)	1 (0.3)	8 (0.2)	2006 (99)	1325 (99)
Wales	6153	960 (16)	658 (11)	1844 (29)	0 (0.0)	2 (0.1)	958 (99)	658 (100)
<i>Central Europe</i>	44550	11923 (27)	6507 (15)	4452 (10)	227 (1.2)	93 (0.5)	11840 (99)	6495 (99)
France								
Bas Rhin	1763	410 (23)	186 (11)	226 (13)	16 (2.7)	0 (0.0)	410 (100)	186 (100)
Côte d'Or	1135	258 (23)	141 (12)	85 (7)	3 (0.8)	0 (0.0)	258 (100)	141 (100)
Doubs	959	218 (23)	128 (13)	35 (4)	3 (0.9)	0 (0.0)	218 (100)	128 (100)
Hérault	1720	463 (27)	265 (15)	78 (5)	39 (5.4)	0 (0.0)	463 (100)	265 (100)
Isère	1999	525 (26)	271 (14)	141 (7)	17 (2.1)	0 (0.0)	525 (100)	271 (100)
Somme	942	277 (29)	106 (11)	69 (7)	5 (1.3)	0 (0.0)	277 (100)	106 (100)
Tarn	708	206 (29)	107 (15)	32 (5)	4 (1.3)	0 (0.0)	206 (100)	107 (100)
Germany								
Hambourg	3438	516 (15)	452 (13)	890 (26)	26 (2.7)	3 (0.3)	513 (99)	452 (100)
Saarland	2125	493 (23)	314 (15)	369 (17)	0 (0.0)	1 (0.1)	492 (99)	314 (100)
Switzerland								
Basel	825	275 (33)	101 (12)	69 (8)	40 (10.6)	14 (3.7)	262 (95)	100 (99)
Grisons	403	102 (25)	44 (11)	68 (17)	1 (0.7)	0 (0.0)	102 (100)	44 (100)
St. Gallen	920	242 (26)	134 (15)	84 (9)	2 (0.5)	3 (0.8)	239 (99)	134 (100)
Valais	492	112 (23)	61 (12)	47 (10)	3 (1.7)	0 (0.0)	112 (100)	61 (100)
The Netherlands	27121	7826 (29)	4197 (15)	2259 (8)	68 (0.6)	72 (0.6)	7763(99) ¹	4186 (99) ¹
<i>Southern Europe</i>	31123	7340 (24)	3483 (11)	4718 (15)	205 (1.9)	50 (0.5)	7296 (99)	3477 (99)
Italy								
Alto Adige	975	234 (24)	92 (9)	45 (5)	0 (0.0)	0 (0.0)	234 (100)	92 (100)
Biella	556	128 (23)	90 (16)	31 (6)	0 (0.0)	0 (0.0)	128 (100)	90 (100)
Ferrara	1097	278 (25)	90 (8)	132 (12)	4 (1.1)	6 (1.6)	272 (98)	90 (100)
Genova	2748	654 (24)	404 (15)	265 (10)	0 (0.0)	1 (0.1)	653 (99)	404 (100)
Latina	721	185 (26)	78 (11)	200 (28)	10 (3.8)	0 (0.0)	185 (100)	78 (100)
Modena	1848	488 (26)	188 (10)	66 (4)	9 (1.3)	0 (0.0)	488 (100)	188 (100)
Napoli	750	172 (23)	114 (15)	208 (28)	6 (2.1)	0 (0.0)	172 (100)	114 (100)
Parma	1235	339 (27)	91 (7)	344 (28)	12 (2.8)	0 (0.0)	339 (100)	91 (100)
Ragusa	467	86 (18)	53 (11)	89 (19)	0 (0.0)	0 (0.0)	86 (100)	53 (100)
Reggio Emilia	1261	269 (21)	105 (8)	321 (25)	3 (0.8)	0 (0.0)	269 (100)	105 (100)
Romagna	3042	772 (25)	243 (8)	408 (13)	0 (0.0)	1 (0.1)	772 (100)	242 (99)
Sassari	934	243 (26)	117 (13)	106 (11)	0 (0.0)	0 (0.0)	243 (100)	117 (100)
Trentino	974	176 (18)	131 (13)	264 (27)	0 (0.0)	0 (0.0)	176 (100)	131 (100)
Varese	2148	432 (20)	245 (11)	466 (22)	13 (1.9)	2 (0.3)	432 (100)	243 (99)
Malta	535	93 (17)	71 (13)	134 (25)	1 (0.6)	0 (0.0)	93 (100)	71 (100)
Slovenia	2810	660 (23)	177 (6)	450 (16)	0 (0.0)	19 (2.3)	641 (97)	177 (100)
Spain								
Basque Country	3471	854 (25)	522 (15)	554 (16)	40 (2.9)	12 (0.9)	844 (99)	520 (99)
Girona	1080	270 (25)	155 (14)	68 (6)	10 (2.4)	1 (0.2)	269 (99)	155 (100)
Granada	992	257 (26)	127 (13)	75 (8)	3 (0.8)	0 (0.0)	257 (100)	127 (100)
Murcia	1399	253 (18)	183 (13)	216 (15)	94 (21.6)	1 (0.2)	252 (99)	183 (100)
Navarra	966	249 (26)	99 (10)	117 (12)	0 (0.0)	3 (0.9)	246 (99)	99 (100)
Tarragona	1114	248 (22)	108 (10)	159 (14)	0 (0.0)	4 (1.1)	245 (99)	107 (99)
<i>Eastern Europe</i>	11072	1782 (16)	823 (7)	2096 (19)	20 (0.8)	51 (2.0)	1755 (98)	799 (97)
Estonia	1772	303 (17)	185 (10)	251 (14)	0 (0.0)	2 (0.4)	302 (99)	184 (99)
Lithuania	3775	469 (12)	103 (3)	896 (24)	20 (3.5)	5 (0.9)	465 (99)	102 (99)
Slovakia	5525	1010 (18)	535 (10)	949 (17)	0 (0.0)	44 (2.9)	988 (98)	513 (96)
EUROCARE Pool	111817	25485 (23)	14033 (13)	17537 (16)	456 (1.1)	208 (0.5)	25320 (99)	13988 (99)

NHL: Non-Hodgkin lymphoma - FL: Follicular Lymphoma - DLBCL: Diffuse large B-cell lymphoma – DCO: Death certificate only.

¹ One patient excluded due to sex unspecified.

Table 1 - Indicators of data quality in the contributing cancer registries regarding follicular lymphomas and diffuse large B-cell lymphomas recorded between 1996 and 2004.



* Age-standardized net survivals using the International Cancer Survival Standards.

Figure 1 - Age-standardized net survival and age-classes net survival at five years with 95% confidence intervals (95% CI), after follicular lymphoma (FL) or diffuse large B-cell lymphoma (DLBCL) according to European area and period of diagnosis.

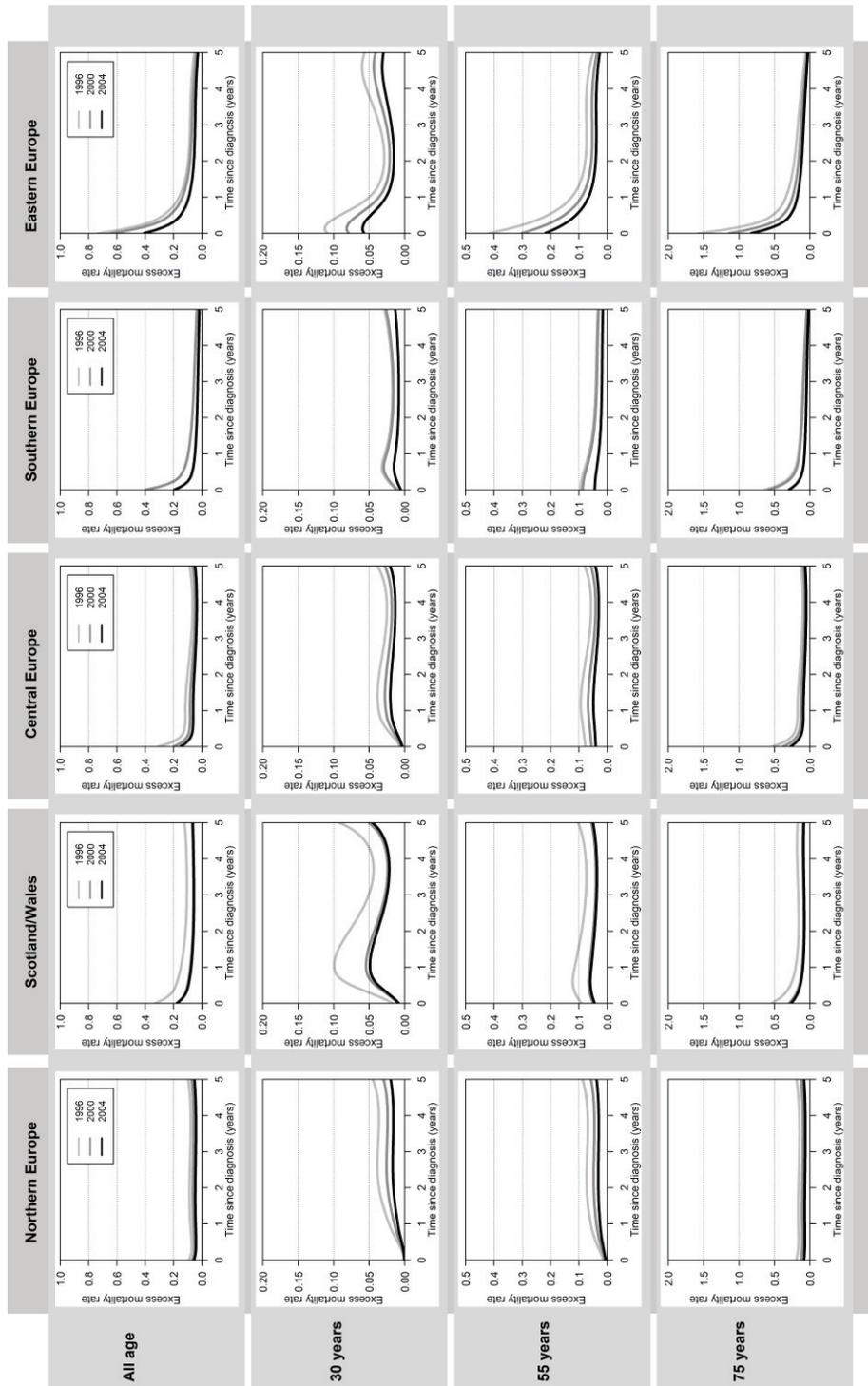


Figure 2 - Dynamics of the excess mortality rates in follicular lymphoma according to the European area, in all patients and in patients aged 30, 55, and 75 years old at diagnosis and for the three years of diagnosis 1996, 2000, and 2004. Results of the finally retained model for each European area.

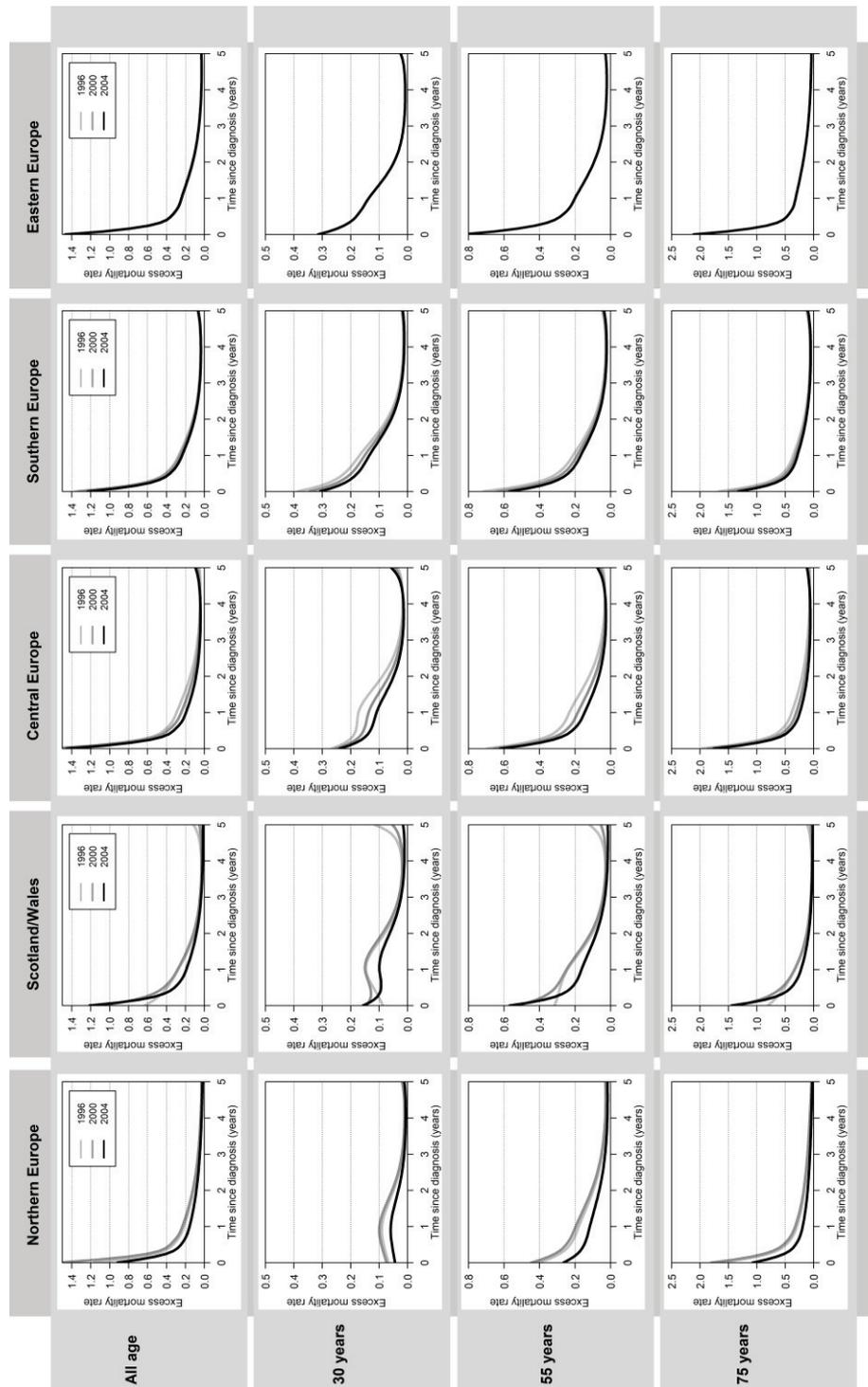


Figure 3 - Dynamics of the excess mortality rates in diffuse large B-cell lymphoma according to the European area in all patients and in patients aged 30, 55, and 75 years old at diagnosis and for the three years of diagnosis at 1996, 2000, and 2004. Results of the finally retained model for each European area.

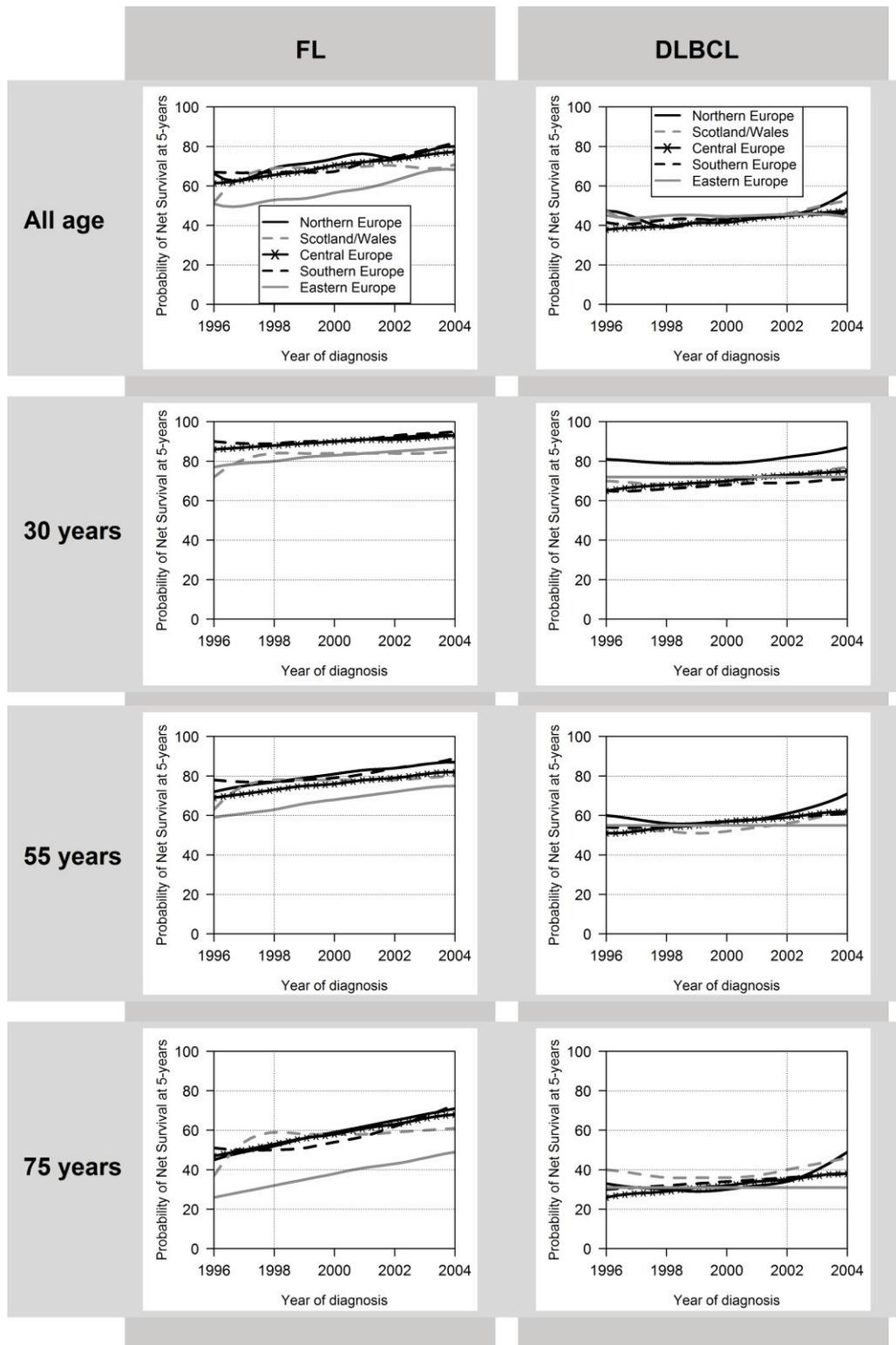


Figure 4 - Five-year net survival probabilities according to the year of diagnosis in follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL) according to the European area in all patients and in patients aged 30, 55, and 75 years old at diagnosis. Results of the finally retained model for each European area.

APPENDIX

EUROCARE-5 Working group

Estonia: M. Mägi (*Estonian CR*); T. Aareleid (*National Institute for Health Development, Tallinn*); **France:** M. Velten (*Bas Rhin CR*); J. Faivre* (*Côte d'Or Digestive Tract Registry, Bourgogne*); M. Maynadié (*Côte d'Or Haematological Malignancies CR*); A. S. Woronoff (*Doubs CR*); B. Tretarre (*Hérault CR*); N. Bossard* (*Hospices Civils de Lyon*); A. Belot (*Hospices Civils de Lyon, FRANCIM*); M. Colonna* (*Isère CR*); B. Lapôtre-Ledoux (*Somme CR*); P. Grosclaude (*Tarn CR*); **Germany:** H. Brenner (*German Cancer Research Center*); A. Nennecke (*Hamburg CR*); B. Holleczeck (*Saarland CR*); A. Katalinic* (*Schleswig-Holstein CR*); **Iceland:** J. G. Jónasson, L. Tryggvadóttir (*Icelandic CR*); **Italy:** G. Mazzoleni, A. Bulatko (*Alto Adige CR*); C. Buzzoni (*Associazione Italiana Registri Tumore*); A. Giacomini (*Biella CR*); S. Ferretti (*Ferrara CR*); G. Gatta*, M. Sant*, H. Amash, C. Amati, P. Baili*, F. Berrino*, S. Bonfarnuzzo, L. Botta, F. Di Salvo, R. Foschi, C. Margutti, E. Meneghini, P. Minicozzi, A. Trama (*Fondazione IRCCS Istituto Nazionale dei Tumori, Milan*); R. De Angelis*, M. Caldora, R. Capocaccia*, E. Carrani, S. Francisci*, S. Mallone, D. Pierannunzio, P. Roazzi, S. Rossi, M. Santaquilani, A. Tavilla (*Istituto Superiore di Sanità, Roma*); F. Pannozzo, M. Natali (*Latina CR*); L. Bonelli, M. Vercelli (*Liguria CR, IRCCS AOU SM-IST*); C. Marchesi, C. Cirilli (*Modena CR*); M. Fusco, M. F. Vitale (*Napoli 3 South CR*); M. Michiara (*Parma CR*); R. Tumino (*Ragusa CR*); P. Giorgi Rossi, M. Vicentini (*Reggio Emilia CR*); F. Falcini (*Romagna CR*); O. Sechi, R. Cesaraccio (*Sassari CR*); S. Piffer (*Servizio Epidemiologia Clinica e Valutativa, Trento*); G. Tagliabue, P. Contiero (*Varese Province CR, Fondazione IRCCS Istituto Nazionale dei Tumori*); **Lithuania:** G. Smalyte (*Lithuanian CR*); **Malta:** N. Calleja, R. Micallef (*Malta National Cancer Registry, Health Information and Research*); **Norway:** T. B. Johannesen (*Norwegian CR*); **Poland:** M. Bielska-Lasota* (*National Institute of Public Health-NIH, Warszawa*); **Slovakia:** C. Safaei Diba (*Slovakian National CR*); **Slovenia:** M. Primic-Zakelj* (*Cancer Registry of Republic of Slovenia*); **Spain:** M. Errezola, N. Larrañaga (*Basque Country CR*); R. Marcos-Gragera*, L. Vilardell (*Girona CR*); M. J. Sanchez, E. Molina (*Granada CR, CIBERESP, ibs.Granada*); C. Navarro, M. D. Chirlaque (*Murcia CR, CIBERESP, IMB-Arrixaca*); C. Moreno-Iribas, E. Ardanaz (*Navarra CR, CIBERESP*); J. Galceran, M. Carulla (*Tarragona CR*); **Switzerland:** M. Mousavi (*Basel CR*); S. M. Ess, H. Frick (*Grisons-Glarus CR*); M. Lorez (*NICER*); S. M. Ess, C. Herrmann (*St. Gallen CR*); I. Konzelmann (*Valais CR*); **The Netherlands:** O. Visser, V. Lemmens (*The Netherlands CR*); R. Otter*; **UK-England:** M. Coleman*, C. Allemani, B. Rachet (*London School of Hygiene and Tropical Medicine*); J. Verne*, N. Easey, G. Lawrence, T. Moran, J. Rashbass, M. Roche, J. Wilkinson (*Public Health England*); **UK-Scotland:** D. H. Brewster (*Scottish CR*); **UK-Wales:** D. W. Huws, C. White (*Welsh Cancer Intelligence and Surveillance Unit*).

*EUROCARE Steering Committee

Area, country, region	Sex		Age classes					Period of diagnosis					Deaths at 5 y N (%)	Median FU (y)		
	Cases	Men N (%)	Women N (%)	15-44		45-64		65-74		≥ 75		1996-98 N (%)			1999-2001 N (%)	2002-04 N (%)
				Age at diagnosis Median (SD)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)					
<i>Northern Europe</i>	1234	598 (48)	636 (52)	125 (10)	630 (51)	267 (22)	212 (17)	251 (20)	395 (32)	388 (48)	389 (32)	389 (32)	389 (32)	389 (32)	8	
Iceland	71	38 (54)	33 (46)	12 (17)	28 (39)	18 (25)	15 (18)	8 (11)	23 (32)	40 (56)	23 (32)	40 (56)	23 (32)	40 (56)	7	
Norway	1163	560 (48)	603 (52)	113 (10)	602 (52)	249 (21)	199 (17)	249 (21)	372 (32)	548 (47)	548 (47)	548 (47)	548 (47)	548 (47)	8	
<i>Scotland/Wales</i>	1983	885 (45)	1098 (55)	198 (10)	885 (45)	415 (21)	355 (18)	355 (18)	653 (33)	975 (49)	975 (49)	975 (49)	975 (49)	975 (49)	8	
Scotland	1325	593 (45)	732 (55)	135 (10)	560 (42)	356 (27)	274 (21)	274 (21)	447 (34)	633 (48)	633 (48)	633 (48)	633 (48)	633 (48)	8	
Wales	658	292 (44)	366 (56)	63 (10)	297 (45)	157 (24)	141 (21)	110 (17)	206 (31)	342 (52)	342 (52)	342 (52)	342 (52)	342 (52)	8	
<i>Central Europe</i>	6495	3147 (48)	3348 (52)	835 (13)	2948 (45)	1480 (23)	1232 (19)	1485 (23)	2109 (32)	2901 (45)	2901 (45)	2901 (45)	2901 (45)	2901 (45)	8	
France	186	95 (51)	91 (49)	29 (16)	71 (38)	49 (26)	37 (20)	50 (27)	49 (26)	87 (47)	87 (47)	87 (47)	87 (47)	87 (47)	6	
Bas Rhin	141	61 (43)	80 (57)	15 (11)	62 (44)	32 (23)	32 (23)	32 (23)	42 (30)	67 (48)	67 (48)	67 (48)	67 (48)	67 (48)	7	
Côte d'Or	128	55 (43)	73 (57)	18 (14)	52 (41)	36 (28)	22 (17)	24 (19)	46 (36)	58 (45)	58 (45)	58 (45)	58 (45)	58 (45)	7	
Doubs	265	121 (46)	144 (54)	28 (11)	113 (43)	68 (26)	54 (20)	54 (20)	94 (35)	117 (44)	117 (44)	117 (44)	117 (44)	117 (44)	7	
Hérault	271	120 (44)	151 (56)	60 (14)	29 (11)	133 (49)	64 (24)	45 (17)	86 (32)	126 (46)	126 (46)	126 (46)	126 (46)	126 (46)	7	
Iserre	106	60 (57)	46 (43)	64 (13)	10 (9)	47 (44)	36 (34)	13 (12)	23 (22)	30 (28)	30 (28)	30 (28)	30 (28)	30 (28)	7	
Somme	107	53 (50)	54 (50)	10 (9)	42 (39)	29 (27)	26 (24)	24 (22)	27 (25)	56 (52)	56 (52)	56 (52)	56 (52)	56 (52)	7	

Table S1 - Characteristics of the 13 988 follicular lymphoma cases recorded between 1996 and 2004 in various European areas, countries, and regions.

Area, country, region	Cases	Sex		Age at diagnosis					Age classes					Period of diagnosis				Median FU (y)
		Men N (%)	Women N (%)	Median (SD)	15-44		45-64		65-74		≥75		1996-98 N (%)	1999-2001 N (%)	2002-04 N (%)	Deaths at 5y		
					N (%)	N (%)	N (%)	N (%)										
<i>Northern Europe</i>	1465	782 (53)	683 (47)	69 (15)	143 (10)	454 (31)	354 (24)	514 (35)	284 (19)	270 (18)	911 (62)	790 (54)	6					
Iceland	126	70 (56)	56 (44)	71 (17)	14 (11)	33 (26)	30 (24)	49 (39)	28 (22)	46 (37)	52 (41)	64 (51)	8					
Norway	1339	712 (53)	627 (47)	68 (15)	129 (10)	421 (31)	324 (24)	465 (35)	256 (19)	224 (17)	859 (64)	726 (54)	6					
<i>Scotland/Wales</i>	2964	1491 (50)	1473 (50)	69 (15)	286 (10)	881 (30)	853 (29)	944 (32)	313 (11)	895 (30)	1756 (59)	1705 (58)	7					
Scotland	2006	970 (48)	1036 (52)	69 (15)	196 (10)	596 (30)	574 (29)	640 (32)	255 (13)	639 (32)	1112 (55)	1160 (58)	7					
Wales	958	521 (54)	437 (46)	69 (15)	90 (9)	285 (30)	279 (29)	304 (32)	58 (6)	256 (27)	644 (67)	545 (57)	6					
<i>Central Europe</i>	11840	6247 (53)	5593 (47)	68 (16)	1491 (13)	3444 (29)	2963 (25)	3942 (33)	2212 (19)	3675 (31)	5953 (50)	6843 (58)	7					
France	410	202 (49)	208 (51)	68 (17)	55 (13)	122 (30)	93 (23)	140 (34)	48 (12)	123 (30)	239 (58)	225 (55)	6					
Bas Rhin	258	142 (55)	116 (45)	70 (17)	63 (24)	73 (28)	63 (24)	87 (34)	54 (21)	79 (31)	125 (48)	143 (55)	7					
Côte d'Or	218	114 (52)	104 (48)	67 (17)	29 (13)	58 (27)	59 (27)	72 (33)	40 (18)	76 (35)	102 (47)	117 (54)	7					
Doubs	463	251 (54)	212 (46)	69 (17)	71 (15)	110 (24)	111 (24)	171 (37)	90 (19)	165 (36)	208 (45)	267 (58)	7					
Hérault	525	280 (53)	245 (47)	67 (17)	72 (14)	164 (31)	121 (23)	168 (32)	93 (18)	178 (34)	254 (48)	274 (52)	7					
Isère	277	145 (52)	132 (48)	69 (16)	34 (12)	78 (28)	84 (30)	84 (30)	48 (17)	89 (32)	140 (51)	149 (54)	7					
Somme	206	108 (52)	98 (48)	72 (13)	8 (4)	53 (26)	60 (29)	85 (41)	36 (17)	73 (35)	97 (47)	116 (56)	6					
Tarn	513	247 (48)	266 (52)	67 (16)	68 (13)	167 (33)	117 (23)	161 (31)	134 (26)	158 (31)	221 (43)	253 (49)	8					
Hambourg	492	236 (48)	256 (52)	69 (15)	54 (11)	142 (29)	128 (26)	168 (34)	52 (11)	170 (35)	270 (55)	255 (52)	8					
<i>Switzerland</i>	262	132 (50)	130 (50)	68 (16)	29 (11)	75 (29)	66 (25)	92 (35)	74 (28)	80 (31)	108 (41)	156 (60)	7					
Basel	102	49 (48)	53 (52)	72 (16)	3 (8)	28 (27)	20 (20)	46 (45)	25 (25)	24 (24)	53 (52)	65 (64)	7					
Grisons	239	125 (52)	114 (48)	69 (18)	33 (14)	65 (26)	55 (22)	90 (38)	50 (21)	78 (33)	111 (46)	137 (57)	8					
St. Gallen	112	67 (60)	45 (40)	69 (16)	16 (14)	29 (26)	22 (20)	39 (35)	22 (20)	39 (35)	51 (46)	56 (50)	8					
Valais	763	4149 (53)	3614 (47)	68 (16)	979 (13)	2282 (29)	1963 (25)	2539 (33)	1446 (19)	2343 (30)	3974 (51)	4630 (60)	8					
<i>The Netherlands</i>	7296	3759 (52)	3537 (48)	68 (17)	1142 (16)	1900 (26)	1983 (27)	2271 (31)	1404 (19)	2319 (32)	3573 (49)	4140 (57)	8					
<i>Italy</i>	234	113 (48)	121 (52)	70 (17)	36 (15)	62 (26)	51 (22)	85 (36)	44 (19)	71 (30)	119 (51)	126 (54)	7					
Alto Adige	128	50 (39)	78 (61)	71 (15)	15 (12)	31 (24)	36 (28)	46 (36)	27 (21)	46 (36)	55 (43)	73 (57)	8					
Biella	272	162 (60)	110 (40)	70 (16)	55 (13)	58 (21)	85 (31)	94 (35)	57 (21)	82 (30)	133 (49)	173 (64)	8					
Ferrara	653	331 (51)	322 (49)	70 (16)	88 (13)	165 (25)	174 (27)	228 (35)	132 (20)	218 (33)	303 (46)	393 (60)	8					
Genova	185	96 (52)	89 (48)	63 (17)	37 (20)	63 (34)	52 (28)	33 (18)	61 (33)	31 (17)	93 (50)	83 (45)	9					
Liguria	488	256 (52)	232 (48)	69 (16)	52 (11)	130 (27)	133 (27)	173 (35)	102 (21)	158 (32)	228 (47)	292 (60)	8					
Modena	172	91 (53)	81 (47)	62 (16)	35 (20)	57 (33)	49 (28)	31 (18)	21 (12)	54 (31)	97 (56)	92 (53)	7					
Napoli	339	162 (48)	177 (52)	72 (16)	41 (12)	67 (20)	95 (28)	136 (40)	63 (19)	101 (30)	175 (52)	196 (58)	7					
Parma	86	38 (44)	48 (56)	68 (17)	12 (14)	21 (24)	24 (28)	29 (34)	9 (10)	30 (35)	47 (55)	47 (55)	8					
Ragusa	269	134 (50)	135 (50)	70 (17)	38 (14)	58 (22)	70 (26)	103 (38)	44 (16)	84 (31)	141 (52)	169 (63)	7					
Reggio Emilia	772	421 (55)	351 (45)	70 (16)	102 (13)	178 (23)	220 (28)	272 (35)	183 (24)	258 (33)	331 (43)	426 (55)	8					
Romagna	243	130 (53)	113 (47)	64 (17)	46 (19)	81 (33)	57 (23)	59 (24)	51 (21)	86 (35)	106 (44)	142 (58)	9					
Sassari	176	87 (49)	89 (51)	72 (15)	13 (7)	38 (22)	54 (31)	71 (40)	39 (22)	82 (47)	55 (31)	88 (50)	10					
Trentino	432	232 (54)	200 (46)	67 (17)	77 (18)	111 (26)	132 (31)	112 (26)	83 (19)	122 (28)	227 (53)	240 (56)	7					
Varese	93	53 (57)	40 (43)	65 (17)	19 (20)	26 (28)	23 (25)	25 (27)	21 (23)	38 (41)	34 (37)	55 (59)	10					
Malta	641	291 (45)	350 (55)	68 (16)	81 (13)	195 (30)	192 (30)	173 (27)	109 (17)	182 (28)	330 (52)	334 (52)	8					
Slovenia	844	452 (54)	392 (46)	66 (18)	184 (22)	216 (26)	205 (24)	239 (28)	135 (16)	205 (24)	444 (53)	485 (57)	7					
<i>Spain</i>	269	143 (53)	126 (47)	64 (17)	47 (17)	89 (33)	65 (24)	68 (25)	53 (20)	78 (29)	138 (51)	166 (62)	7					
Basque Country	257	141 (55)	116 (45)	66 (17)	49 (19)	71 (28)	65 (25)	43 (17)	78 (30)	136 (53)	157 (61)	173 (66)	7					
Girona	252	130 (52)	122 (48)	60 (24)	60 (24)	64 (25)	59 (23)	69 (27)	59 (23)	80 (30)	104 (41)	146 (58)	7					
Granada	246	109 (44)	137 (56)	69 (17)	37 (15)	56 (23)	71 (29)	82 (33)	40 (16)	91 (37)	115 (47)	122 (50)	7					
Murcia	245	137 (56)	108 (44)	68 (17)	38 (16)	65 (27)	64 (26)	75 (31)	28 (11)	75 (31)	142 (58)	135 (55)	8					
Navarra	1755	802 (46)	953 (54)	65 (16)	266 (15)	582 (33)	516 (29)	391 (22)	188 (11)	388 (22)	1179 (67)	1006 (57)	6					
Tarragona	302	142 (47)	160 (53)	68 (15)	40 (13)	85 (28)	103 (34)	74 (25)	37 (12)	75 (25)	190 (63)	185 (61)	6					
<i>Eastern Europe</i>	465	207 (45)	258 (55)	65 (16)	75 (16)	149 (32)	147 (32)	94 (20)	6 (1)	90 (13)	399 (86)	248 (53)	6					
Estonia	988	453 (46)	535 (54)	64 (16)	151 (15)	348 (35)	266 (27)	273 (23)	145 (15)	252 (26)	590 (60)	573 (58)	7					
Lithuania	25320	13081 (52)	12239 (48)	68 (16)	3328 (13)	7261 (29)	6669 (26)	8062 (32)	4401 (17)	7547 (30)	13372 (53)	14484 (57)	7					
Slovakia	EURO-CARE Pool																	

Table S2 - Characteristics of the 25 320 diffuse large B-cell lymphoma cases recorded between 1996 and 2004 in various European areas, countries, and regions.

	FL														
	All age			15-44			45-64			65-74			≥75		
	Cases	Deaths at 5 y N (%)		Cases	Deaths at 5 y N (%)		Cases	Deaths at 5 y N (%)		Cases	Deaths at 5 y N (%)		Cases	Deaths at 5 y N (%)	
Northern Europe	1996-1998 251	96 (38)		36	8 (22)		114	29 (25)		55	27 (49)		46	32 (70)	
	1999-2001 395	135 (34)		41	5 (12)		214	55 (26)		84	36 (43)		56	39 (70)	
	2002-2004 588	158 (27)		48	2 (4)		302	50 (17)		128	43 (34)		110	63 (57)	
Scotland/Wales	1996-1998 355	165 (46)		35	9 (26)		152	54 (36)		85	44 (52)		83	58 (70)	
	1999-2001 653	237 (36)		70	21 (30)		277	69 (25)		175	70 (40)		131	77 (59)	
	2002-2004 975	351 (36)		93	11 (12)		428	106 (25)		253	102 (40)		201	132 (66)	
Central Europe	1996-1998 1485	620 (42)		201	44 (22)		644	196 (30)		335	177 (53)		305	203 (67)	
	1999-2001 2109	800 (38)		277	44 (16)		947	273 (29)		475	210 (44)		410	273 (67)	
	2002-2004 2901	851 (29)		357	32 (9)		1357	276 (20)		670	231 (34)		517	312 (60)	
Southern Europe	1996-1998 669	246 (37)		97	11 (11)		267	75 (28)		187	73 (39)		118	87 (74)	
	1999-2001 1186	414 (35)		177	26 (15)		448	92 (21)		319	141 (44)		242	155 (64)	
	2002-2004 1622	451 (28)		264	20 (8)		649	106 (16)		378	128 (34)		331	197 (60)	
Eastern Europe	1996-1998 155	77 (50)		22	6 (27)		66	24 (36)		47	30 (64)		20	17 (85)	
	1999-2001 275	135 (49)		39	12 (31)		113	39 (35)		66	39 (59)		57	45 (79)	
	2002-2004 369	146 (40)		58	8 (14)		166	46 (28)		85	54 (64)		60	38 (63)	

Table S3 - The number of follicular lymphoma cases and the number of deaths at 5 years per period of diagnosis and age-class in various European areas.

		DLBCL											
		All age		15-44		45-64		65-74		≥75			
		Cases	Deaths at 5 y N (%)	Cases	Deaths at 5 y N (%)	Cases	Deaths at 5 y N (%)	Cases	Deaths at 5 y N (%)	Cases	Deaths at 5 y N (%)		
Northern Europe	1996-1998	284	162 (57)	33	8 (24)	92	44 (48)	80	43 (54)	79	67 (85)		
	1999-2001	270	176 (65)	27	8 (30)	84	40 (48)	48	32 (67)	111	96 (86)		
	2002-2004	911	452 (50)	83	11 (13)	278	90 (32)	226	108 (48)	324	243 (75)		
Scotland/Wales	1996-1998	313	189 (60)	33	11 (33)	89	43 (48)	80	54 (68)	111	81 (73)		
	1999-2001	895	546 (61)	97	30 (31)	273	144 (53)	250	152 (61)	275	220 (80)		
	2002-2004	1756	970 (55)	156	40 (26)	519	219 (42)	523	284 (54)	558	427 (77)		
Central Europe	1996-1998	2212	1394 (63)	305	114 (37)	664	334 (50)	570	360 (63)	673	586 (87)		
	1999-2001	3675	2187 (60)	469	141 (30)	1020	465 (46)	926	580 (63)	1260	1001 (79)		
	2002-2004	5953	3262 (55)	717	185 (26)	1760	705 (40)	1467	809 (55)	2009	1563 (78)		
Southern Europe	1996-1998	1404	824 (59)	248	106 (43)	373	172 (46)	384	227 (59)	399	319 (80)		
	1999-2001	2319	1370 (59)	376	142 (38)	606	266 (44)	626	380 (61)	711	582 (82)		
	2002-2004	3573	1946 (54)	518	153 (30)	921	358 (39)	973	546 (56)	1161	889 (77)		
Eastern Europe	1996-1998	188	116 (62)	22	7 (32)	65	32 (49)	65	48 (74)	36	29 (81)		
	1999-2001	388	223 (57)	61	25 (41)	123	58 (47)	122	76 (62)	82	64 (78)		
	2002-2004	1179	667 (57)	183	47 (26)	394	186 (47)	329	214 (65)	273	220 (81)		

Table S4 - The number of diffuse large B-cell lymphoma cases and the number of deaths at 5 years per period of diagnosis and age-class in various European areas.

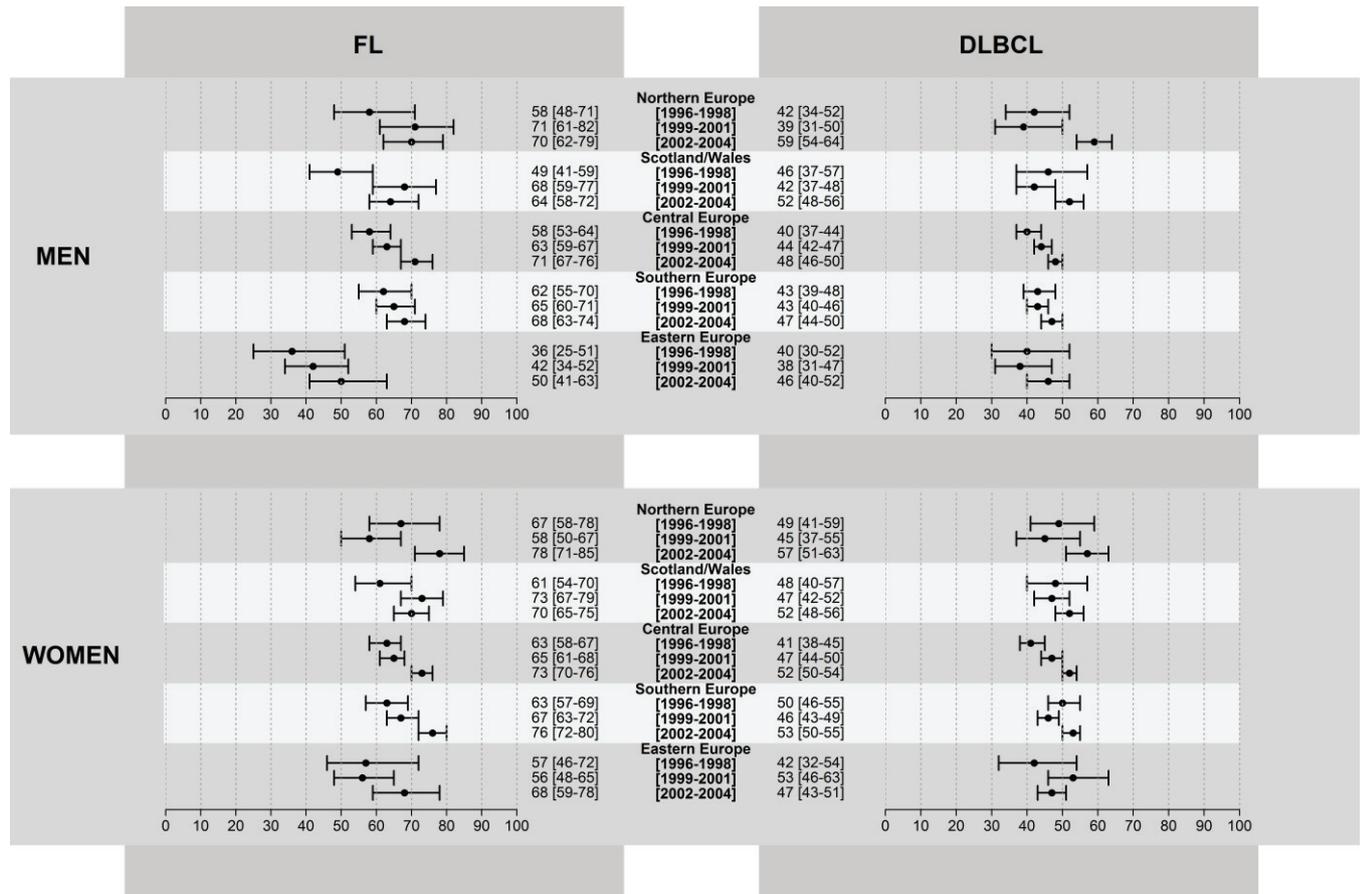


Figure S1 - Age-standardized net survival at five years with 95% confidence intervals (95% CI) by sex, after follicular lymphoma (FL) or diffuse large B-cell lymphoma (DLBCL) according to European area and period of diagnosis.

Estimating the effect of the year of diagnosis on the excess mortality rate after follicular lymphoma and diffuse large B-cell lymphoma in each European area: model-building strategy and results

We explain here the strategy used to model the effect of the year of diagnosis on the excess mortality rate (EMR) and show its results. This effect was explored using a model-building strategy on inference about non-linear and time dependent covariate effects within the context of survival analysis as proposed by Wynant and Abrahamowicz.¹⁸ This effect is studied independently for each lymphoma subtype and each European area.

In the following, LL stands for "log linear", NL for "non-linear", TD for "time-dependent", and PH for "proportional hazard".

Time since diagnosis, age at diagnosis, and the year of diagnosis were considered as continuous variables.

First, we modelled the baseline hazard as a continuous function (i. e. $f(\text{time})$ below) and we assumed systematically a NL and TD effect of age at diagnosis on the EMR (respectively $g(\text{age})$ and $h(\text{time}) \times \text{age}$). Then, five models with different year of diagnosis effect were considered:

Model 1: $\log[\lambda_c(\text{time}, \text{age}, \text{year})] = f(\text{time}) + g(\text{age}) + h(\text{time}) \times \text{age} + i(\text{year}) + j(\text{time}) \times \text{year}$

Model 2: $\log[\lambda_c(\text{time}, \text{age}, \text{year})] = f(\text{time}) + g(\text{age}) + h(\text{time}) \times \text{age} + \text{year} + j(\text{time}) \times \text{year}$

Model 3: $\log[\lambda_c(\text{time}, \text{age}, \text{year})] = f(\text{time}) + g(\text{age}) + h(\text{time}) \times \text{age} + i(\text{year})$

Model 4: $\log[\lambda_c(\text{time}, \text{age}, \text{year})] = f(\text{time}) + g(\text{age}) + h(\text{time}) \times \text{age} + \text{year}$

Model 5: $\log[\lambda_c(\text{time}, \text{age}, \text{year})] = f(\text{time}) + g(\text{age}) + h(\text{time}) \times \text{age}$

In these equations, λ_c is the EMR and functions $f, g, h, i,$ and j are as follows:

Effect	Function	Kind of function	Location of the knot
Time since diagnosis	f	Cubic spline	At one year
Linearity of age	g	Quadratic spline	At mean age
Proportionality of age	h	Cubic spline	At one year
Linearity of year of diagnosis	i	Quadratic spline	1998 for FL* 2002 for DLBCL*
Proportionality of year of diagnosis	j	Cubic spline	At one year

* We used year 1998 for FL and 2002 for DLBCL to take into account the offset due to the introduction of anti-CD20 chimeric monoclonal antibody in the therapeutic management of these non-Hodgkin lymphoma types.

Concerning the effect of the year of diagnosis:

- Model 1 corresponds to a NL-TD effect
- Model 2 corresponds to a LL-TD effect
- Model 3 corresponds to a NL-PH effect
- Model 4 corresponds to a LL-PH effect
- Model 5 corresponds to a model with no effect of year of diagnosis

To determine which of these five models should be retained, we used a backward elimination procedure (as recommended by Wynant and Abrahamowicz)¹⁸ using the likelihood ratio test: we started with the full flexible Model 1 and eliminated the effects which were found not statistically significant ($p > 0.05$). This backward selection can be performed in two arms (Cf. figure below, where $F = f(t) + g(\text{age}) + h(t) \times \text{age}$, t = time since diagnosis, df = degree of freedom). The first arm tested whether we could eliminate: i) the NL effect (test 1); ii) the TD effect (test 3); and, iii) the whole effect (test 5) of the year of diagnosis. The second arm tested whether we could eliminate: i) the TD effect (test 2); ii) the NL effect (test 4); and, iii) the whole effect (test 5) of the year of diagnosis.

The results of all tests and the model finally retained for each European area and each lymphoma subtype are shown in the figure below. Black frames correspond to the models retained for FL and the grey frames to the models retained for DLBCL. Note that model which is retained is always the same regardless of the two arms, (except in DLBCL for Scotland/Wales in which the result of the test of the LL was not similar if we assumed the TD or the PH of the effect of the year).

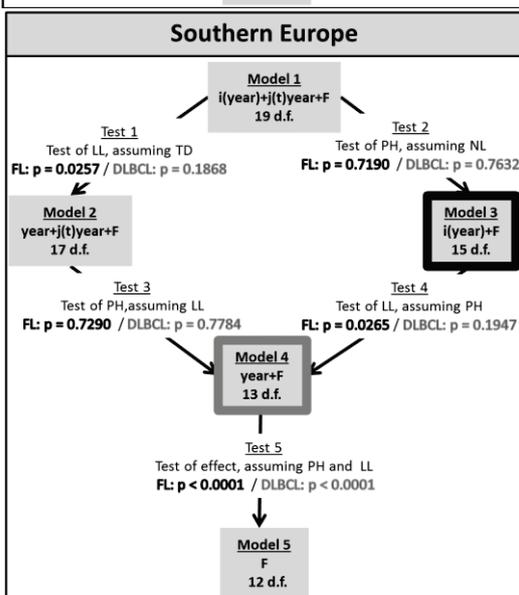
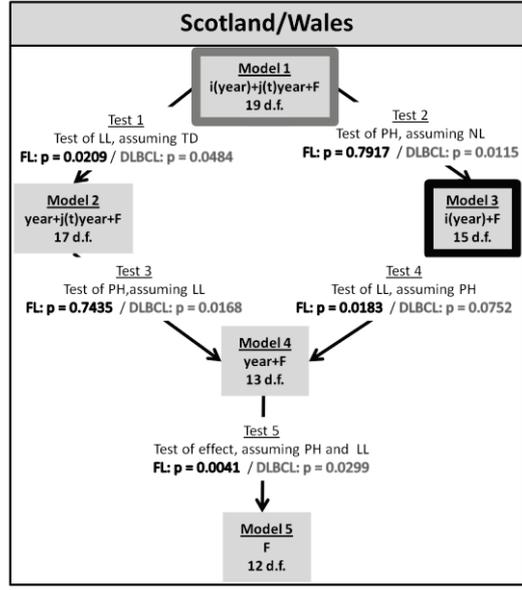
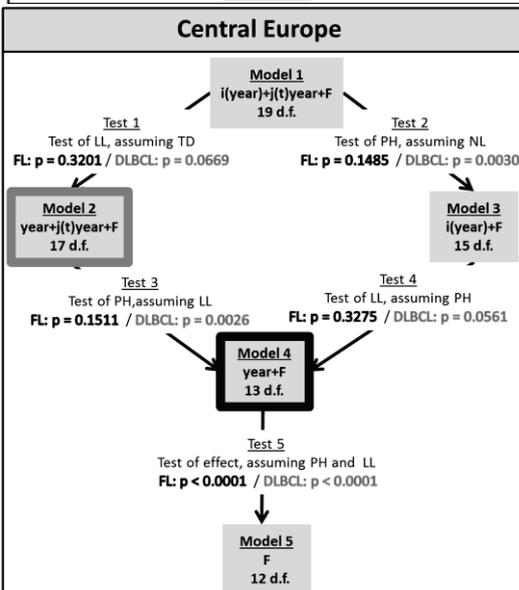
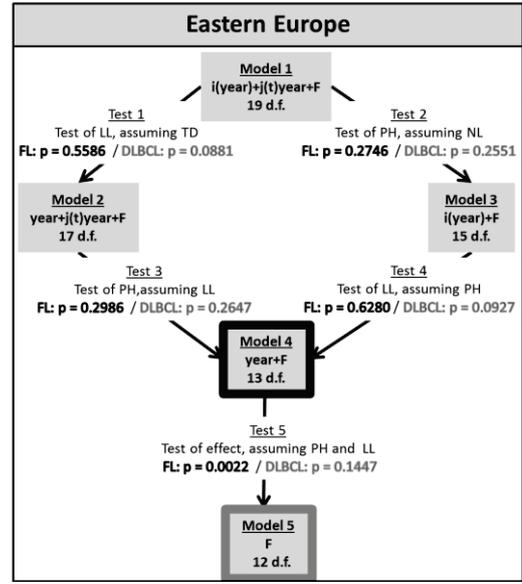
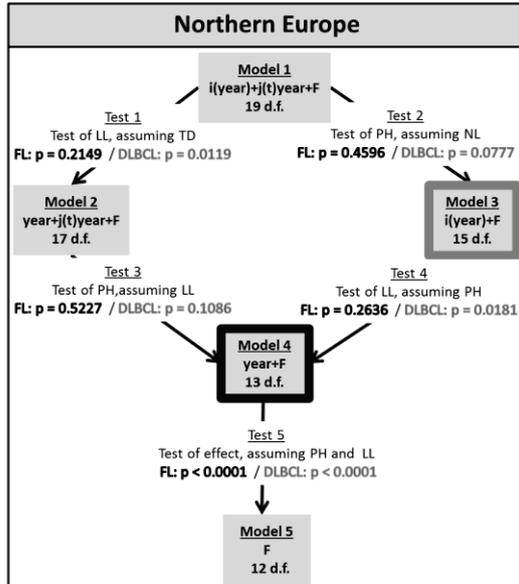
For each European area, the EMR and the five-year net survival at specific ages and/or specific years of diagnosis were estimated from the retained model.

The EMR for all ages at time t , $\lambda_{all\ ages}(t)$, that is the marginal EMR for the whole population, in Figure 2 and 3 is given separately for each year of diagnosis by:

$$\lambda_{all\ ages}(t) = \sum_{i=1}^n \frac{S_i(t)}{\sum_{i=1}^n S_i(t)} \lambda_i(t)$$

S_i and λ_i being respectively the net survival and the EMR for individual i (this quantities being predicted by the model from the age of diagnosis and the year of diagnosis of this individual).

The prediction of five-year survival for all ages in Figure 4 was obtained with the average of the five-year survivals for all patients, separately for each year of diagnosis.



Year of diagnosis with a :

	FL
Northern Europe	LL & PH effect
Scotland/Wales	NL & PH effect
Central Europe	LL & PH effect
Southern Europe	NL & PH effect
Eastern Europe	LL & PH effect
	DLBCL
Northern Europe	NL & PH effect
Scotland/Wales	NL & TD effect
Central Europe	LL & TD effect
Southern Europe	LL & PH effect
Eastern Europe	No effect

IV.5. RÉSULTATS COMPLÉMENTAIRES

Des analyses complémentaires ont été menées sur ces données européennes afin d'une part de tester l'effet d'autres covariables sur la mortalité, comme le sexe, et d'autre part de s'assurer du bon usage des outils statistiques utilisés.

L'effet du sexe sur la mortalité a tout d'abord été étudié, en estimant la survie nette de façon ponctuelle chez les hommes puis chez les femmes.

La figure S1 (Appendix p.6 de l'article) présente les probabilités de survie nette à 5 ans standardisées sur l'âge par sexe, période de diagnostic et par zone européenne, obtenues à l'aide de l'estimateur Pohar-Perme.

Dans les cinq zones européennes, la survie nette à 5 ans était plus élevée chez les femmes que chez les hommes pour les LF, la différence d'effet se réduisant avec les périodes de diagnostic. Pour les LBDGC et sur la dernière période 2002-04, il n'existait pas de différences notables de survie à 5 ans en fonction du sexe dans chacune des zones.

Il aurait été souhaitable de pouvoir modéliser l'effet du sexe sur la mortalité en excès, notamment pour les LF pour lesquels des différences de survie nette sont observées. Cependant dans notre travail, nous avons privilégié la modélisation de l'année et de l'âge au diagnostic, ces variables ayant un effet plus important sur le pronostic des LNH par rapport à l'effet du sexe. L'ajout de la modélisation de l'effet du sexe avec notamment les interactions entre sexe et année de diagnostic, doit se baser sur une stratégie d'analyse complexe ce qui nécessite une étude spécifique.

Afin de visualiser graphiquement les tendances du risque au cours des années de diagnostic, les ratios des taux de mortalité attachés à l'année de diagnostic ont été estimés à l'aide du modèle retenu dans chaque zone européenne et pour chaque sous-type de lymphomes (Figure IV.1). Le ratio correspond au taux de mortalité pour une année donnée divisé par le taux de l'année de référence, fixée ici à 1998 pour les LF et à 2002 pour les LBDGC.

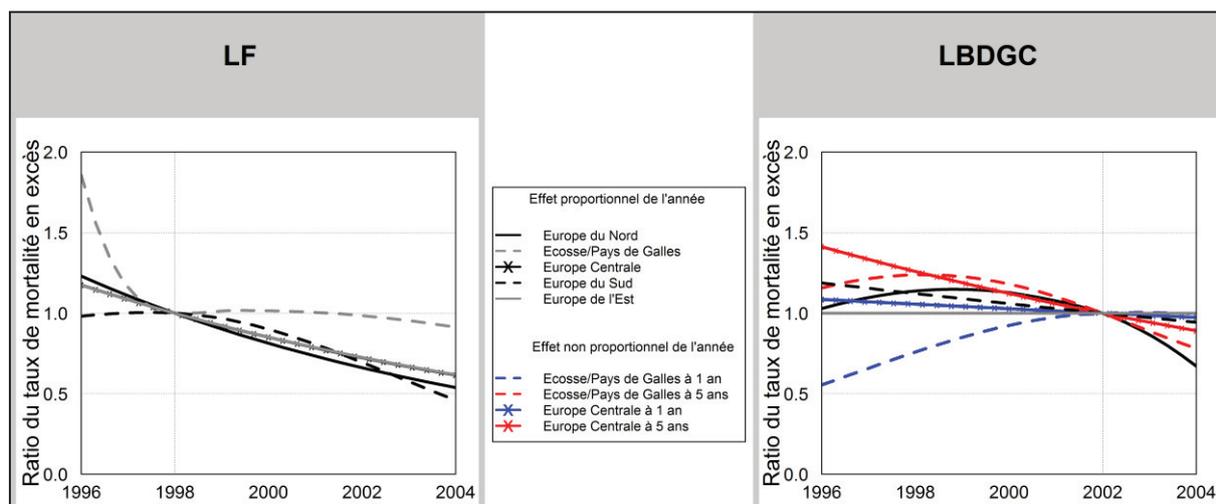


Figure IV.1 - Ratio des taux de mortalité en excès en fonction de l'année de diagnostic, issu du modèle retenu par zone et par sous-type de lymphome. (Pour les LBDGC, le modèle retenu supposait un effet NP de l'année en Ecosse/Pays de Galles et en Europe Centrale. Aussi pour ces deux zones, le ratio est présenté à 1 an et 5 ans après le diagnostic)

Pour les LF, l'effet de l'année a été retrouvé proportionnel dans les cinq zones européennes. Alors que l'évolution est similaire en Europe du Nord, Centrale et de l'Est, l'évolution est plus atypique pour l'Europe du Sud avec un ratio diminuant seulement après 1998 et pour l'Ecosse/Pays de Galles avec un ratio se stabilisant après 1998.

En ce qui concerne les LBDGC, alors qu'en Europe de l'Est, l'année de diagnostic n'avait pas d'effet sur la mortalité avec un ratio constant entre 1996 et 2004, l'effet de l'année a été retrouvé proportionnel en Europe du Nord et du Sud et non proportionnel en Ecosse/Pays de Galles et en Europe Centrale. En Ecosse/Pays de Galles, le ratio de risque à 1 an augmente jusqu'en 2002 alors qu'il diminue en fonction de l'année de diagnostic à 5 ans. Ce résultat assez original soulève des questions sur l'évolution de la prise en charge de ces patients au cours de la période d'étude. En Europe Centrale, le risque à court terme (à 1 an) reste inchangé avec les années de diagnostic alors que le risque à plus long terme (à 5 ans) connaît une diminution depuis 1996. Cette figure IV.1 permet de visualiser des différences dans les tendances du risque de mortalité en fonction de l'année de diagnostic entre les zones. Néanmoins, les disparités de niveau des taux de mortalité entre zones ne sont pas observables avec les ratios. Nous avons donc privilégié la présentation des taux de mortalité en fonction de l'année de diagnostic dans chacune des zones plutôt que les ratios.

D'autres analyses supplémentaires ont permis de s'assurer d'une bonne utilisation des outils méthodologiques.

L'estimateur proposé par Pohar-Perme *et al.* peut être instable en présence de taux de mortalité attendus trop fluctuants, ce qui est souvent observé chez les sujets âgés. Comme les tables de mortalité des registres n'étaient pas lissées, un travail plus approfondi a porté sur la tendance des taux attendus en fonction de l'âge pour chaque registre de cancer. L'étude des taux attendus a montré que la variabilité entre les années de sortie était acceptable pour tous les registres inclus dans notre étude avec des taux attendus chez les sujets âgés restant dans une fourchette similaire entre les registres.

Tout comme pour la précédente application, une étude de l'effet de l'année de diagnostic sur la mortalité dans chaque zone européenne spécifiquement par classe d'âge a permis de mettre en évidence que l'effet de l'année était le même dans chacune des classes, évitant ainsi l'ajout d'une interaction entre l'âge et l'année de diagnostic dans chacun des modèles (résultats non présentés).

IV.6. DISCUSSION

Sur une période de diagnostic de 9 ans, cette étude a permis d'évaluer les tendances de survie nette et de la dynamique du taux de mortalité en Europe à l'aide d'outils méthodologiques adaptés. Les classifications en hématologie ayant évolué au cours du temps, nous avons considéré que le codage des lymphomes par sous-type était homogène durant la période d'étude dans les différentes zones, puisque après 1996, les classifications épidémiologiques qui se sont succédé étaient largement basées sur la classification REAL (Harris *et al.* 1994).

Les résultats montrent, pour les LF, une amélioration de la survie nette à 5 ans avec les années de diagnostic récentes dans toutes les zones européennes et pour tous les âges. Pour les LBDGC, cette amélioration en fonction de l'année de diagnostic est moins prononcée voire inexistante dans le cas de l'Europe de l'Est. Ces évolutions positives de survie reflètent les progrès dans la prise en charge des lymphomes au cours de la période d'étude. En effet les

LMNH ont bénéficié de l'amélioration des outils diagnostiques devenus de plus en plus efficaces permettant de définir de façon plus précise le stade du lymphome et ainsi d'optimiser le traitement des patients (Al-Hamadani et Go 2014). La prise en charge thérapeutique des LMNH s'est également améliorée grâce d'une part à l'utilisation de nouvelles thérapeutiques avec le développement de la thérapie ciblée (anticorps monoclonaux anti-CD20) (Coiffier 2006) et d'autre part avec la mise en place de guides de recommandation de prise en charge des lymphomes dans les différents pays (McNamara *et al.* 2012 ; Tilly *et al.* 2012 ; Ghielmini *et al.* 2013).

Malgré ces résultats positifs, il persiste des différences de pronostic en fonction de l'âge au diagnostic, au détriment des patients plus âgés. Les résultats montrent également des différences de survie en fonction des zones européennes, avec un écart assez remarquable pour l'Europe de l'Est par rapport aux autres zones européennes. Ces différences peuvent s'expliquer par des systèmes de soins à plusieurs vitesses entre pays (Sant *et al.* 2014).

L'apport de ce travail réside dans la présentation de la dynamique du taux de mortalité en excès qui est variable en fonction des zones européennes. Pour les LF, la mortalité en excès est stable durant les cinq premières années de suivi en Europe du Nord, à l'inverse des autres zones européennes où une mortalité précoce est observée juste après le diagnostic. Cela pourrait s'expliquer par une prise en charge des patients différente en Europe du Nord (avec des délais de prise en charge thérapeutique réduits ou par l'utilisation de traitements moins agressifs, suppositions qui nécessitent d'être vérifiées). Pour les LBDGC, la dynamique du taux est relativement identique entre les zones européennes avec une mortalité très importante juste après le diagnostic qui diminue ensuite. Au regard de l'année de diagnostic, le taux de mortalité diminue légèrement avec les années récentes.

L'étude de la survie nette à 5 ans en fonction de l'année de diagnostic modélisée de façon continue a permis de mettre en évidence des éventuels points de cassure non visibles lors de l'étude par période de diagnostic.

Pour les LF, alors que la survie s'améliore de façon linéaire à partir de 1996 en Europe du Nord, Centrale et de l'Est, la survie évolue de façon non linéaire en Europe du Sud avec une augmentation observée à partir de 2000 et en Ecosse/Pays de Galles avec une amélioration de la survie jusqu'en 1998 et qui se stabilise ensuite. Une moins bonne survie à 5 ans est observée en Ecosse/Pays de Galles et en Europe de l'Est chez les sujets jeunes et seule

l'Europe de l'Est connaît une survie en dessous de celles des autres zones pour les sujets plus âgés (à 55 et 75 ans).

Pour les LBDGC, aucun effet de l'année de diagnostic sur la survie n'est retrouvé en Europe de l'Est. La survie à 5 ans évolue de façon linéaire en Europe Centrale et du Sud, et de façon non linéaire en Europe du Nord et en Ecosse/Pays de Galles avec une survie qui s'améliore de façon plus prononcée après 2002. La survie des sujets jeunes (30 ans) en Europe du Nord est meilleure que dans les autres zones européennes alors que pour les sujets de 55 et 75 ans, les différences de survie en fonction des zones sont davantage perceptibles après 2002 avec une meilleure survie observée en Europe du Nord et en Ecosse/Pays de Galles.

Notre étude connaît néanmoins certaines limites. Au niveau méthodologique, nous avons considéré un effet de l'année de diagnostic non dépendant de l'âge, en n'incluant pas d'interaction entre ces deux termes dans la modélisation. Même si l'étude spécifique par classe d'âge a montré que l'année avait un effet similaire dans chaque classe au sein de chaque zone européenne, il aurait tout de même été souhaitable de modéliser la relation âge/année.

Au niveau clinique, l'absence de données cliniques au diagnostic et sur la prise en charge thérapeutique des patients dans la base EUROCARE-5 limite l'interprétation des résultats.

De plus, la période de diagnostic de l'étude ne permet pas de mettre en évidence l'impact des améliorations récentes dans la prise en charge des lymphomes telles que l'utilisation de nouvelles molécules (Hernandez-Ilizaliturri *et al.* 2004 ; Younes *et al.* 2014) ou encore l'impact de la mise en place de l'index pronostic international FLIPI2 (Formica *et al.* 2009) sur la décision de prise en charge des praticiens pour les LF.

Enfin, le taux de mortalité en excès n'est modélisé que jusque 5 ans après le diagnostic. Alors que l'évolution naturelle de ces lymphomes est marquée par des rechutes régulières, il aurait été intéressant de pouvoir modéliser ce taux ainsi que les effets de l'année sur la mortalité à plus long terme.



Conclusion & Perspectives

V.1. CONCLUSION

Ce travail de thèse offre une application des méthodes statistiques les plus avancées pour l'étude du pronostic des patients atteints d'un LF ou d'un LBDGC, dans le but de fournir des indicateurs épidémiologiques de qualité en population générale.

Les besoins en production d'indicateurs plus fins pour permettre de mesurer l'impact grandissant du cancer nécessitent des méthodes statistiques adaptées. Nous avons vu que deux méthodes sont capables de fournir des estimations non biaisées de la survie nette : l'estimateur non paramétrique proposé par Pohar-Perme *et al.* (Perme *et al.* 2012) et la modélisation paramétrique du taux de mortalité en excès prenant en compte les covariables de la table de mortalité, essentiellement l'âge.

Dans ce travail nous sommes restés dans le cadre d'une relation additive entre la mortalité attendue et la mortalité induite par la maladie, cette relation étant plus réaliste cliniquement qu'une relation multiplicative (Buckley 1984).

L'avantage d'une approche de survie non paramétrique est l'absence du choix de la forme de la fonction de risque de base mais elle ne permet pas d'étudier simultanément l'effet de plusieurs covariables sur l'excès de mortalité. Aussi lorsque l'intérêt réside dans la modélisation des effets des variables sur la mortalité, l'approche paramétrique est à privilégier. Afin de coller au mieux à la réalité clinique, les modèles régressifs se sont complexifiés pour devenir plus flexibles, permettant désormais de modéliser de façon continue le taux de mortalité de base et de modéliser des effets non proportionnels et non linéaires de variables continues sur l'excès de mortalité. C'est le cas de la variable âge au diagnostic dont l'effet complexe sur la mortalité en excès a été démontré dans plusieurs études (Monnereau *et al.* 2013b ; Jooste *et al.* 2013) et qui ne devrait plus faire l'objet d'une catégorisation dans les études de survie du cancer. Par ailleurs, l'effet non proportionnel des variables de quelque nature qu'elles soient, devrait toujours être exploré lors de la construction d'un modèle.

Dès lors que l'on souhaite explorer les effets non linéaires et non proportionnels d'une ou plusieurs covariables, ainsi que leurs éventuelles interactions, se pose la question de la

stratégie de construction de modèle à adopter. Comme l'ont suggéré certains auteurs (Abrahamowicz et MacKenzie 2007 ; Wynant et Abrahamowicz 2014), l'effet non proportionnel et non linéaire d'une variable continue doit être testé simultanément afin d'éviter que l'étude de la proportionnalité ne compromette celle de la linéarité et inversement. Wynant et Abrahamowicz ont proposé plusieurs stratégies dans cette situation et les ont évaluées. Leurs conclusions ont guidé notre choix. La question de la « meilleure » stratégie de construction de modèle reste toutefois un sujet difficile. Nous n'avons pas réalisé de recherche bibliographique exhaustive à ce sujet. Cette question fait cependant l'objet d'une tâche de travail méthodologique dans le projet CENSUR.

Pour chacune des deux applications présentées aux *Chapitres III et IV* de ce travail, nous nous sommes heurtés à quelques difficultés dans la stratégie de modélisation alors même que le nombre de variables restait somme toute limité. Plusieurs choix ont dû être faits.

Pour la partie de notre travail appliquée sur les données françaises (*Chapitre III*), des résultats préliminaires ont porté sur la modélisation de l'effet de l'année de diagnostic spécifiquement par classe d'âge, rapportant un effet de l'année similaire dans chacune des classes. Dans la modélisation de l'effet de l'année portant sur l'ensemble de la population (tous âges), nous avons donc fait le choix de ne pas inclure d'interaction entre l'année et l'âge au diagnostic dans nos modèles.

Dans cette même application sur les données françaises, lors de la modélisation de l'effet de l'année sur la mortalité prenant en compte le stade du lymphome, nous n'avons pas tenu compte de l'interaction entre ces deux variables dans le modèle. Nous étions ici dans une situation d'effectifs limités et il n'était pas raisonnable de complexifier le modèle dans ces conditions.

Parallèlement, dans l'application sur les données européennes (*Chapitre IV*), l'intérêt était de mettre en évidence des éventuelles tendances de survie nette différentes entre zones européennes. Nous avons privilégié une modélisation de l'effet de l'année propre à chaque zone européenne. Néanmoins, ce choix ne nous a pas permis de tester des effets statistiquement différents en fonction des zones. Une possibilité aurait pu être l'utilisation d'un seul modèle sur la totalité de la cohorte européenne en incluant une variable zone européenne dans le modèle et en ajoutant une interaction entre cette variable et l'année de diagnostic. Si tel avait été le cas, il aurait fallu également s'assurer de l'étude d'un éventuel

effet non proportionnel de la zone européenne sur la mortalité, ce qui aurait complexifié encore la modélisation.

Nous voyons bien là les difficultés engendrées par la prise en compte simultanée des effets complexes de variables sur la mortalité. A notre connaissance, il s'agit des premiers travaux conciliant à la fois la modélisation des effets complexes de l'âge et de l'année de diagnostic sur le taux de mortalité en excès à la suite d'un cancer.

Alors que la modélisation permet d'estimer directement le taux instantané de mortalité en excès, cet indicateur n'est que peu présenté dans les publications. Pourtant, l'étude de sa dynamique au cours du temps de suivi apparaît comme un indicateur clinique utile. Nous avons pu voir que le taux de mortalité instantané n'était quasiment jamais constant durant les cinq premières années de suivi et que cette dynamique était différente en fonction de certaines caractéristiques, telles que l'âge ou le stade du lymphome au diagnostic. L'étude de la dynamique du taux de mortalité à la suite d'un LF en Europe a mis en évidence un taux de mortalité en excès stable en Europe du Nord à l'inverse des autres zones dans lesquelles il existait une mortalité précoce. Ces dynamiques de mortalité différentes n'étaient pas rendues visibles sur les résultats de survie à 5 ans. Pour les LBDGC, la dynamique du taux était relativement identique entre zones européennes.

L'étude de l'effet de l'année de diagnostic sur la mortalité en excès à la suite d'un LF rapporte un effet proportionnel retrouvé à la fois sur les données françaises et sur les données européennes. Ce résultat est très intéressant puisqu'il permet de conclure que les progrès de prise en charge des LF ont permis de réduire à la fois la mortalité en excès à court terme et à long terme sur notre période d'étude. L'effet de l'année est retrouvé linéaire sur données françaises et en Europe du Nord, Centrale et de l'Est à l'inverse de l'Europe du Sud où l'amélioration est observée à partir de 2000 et en Ecosse/Pays de Galles où la survie s'est améliorée jusqu'en 1998 pour se stabiliser ensuite.

En ce qui concerne l'effet de l'année sur la mortalité en excès à la suite d'un LBDGC, nous avons retrouvé un effet non proportionnel de l'année en Ecosse/Pays de Galles et en Europe Centrale avec une amélioration de la mortalité à long terme. Un effet non linéaire de l'année est également retrouvé en Europe du Nord et du Sud où la survie s'améliore davantage après 2002, marquant à priori un tournant dans la prise en charge des patients dans ces deux zones à partir de cette date. En Europe de l'Est, la prise en charge de LBDGC ne

semble pas avoir évolué avec une stabilisation de la mortalité au cours des années de diagnostic.

Ce travail conforte également l'idée que les sous-types de LMNH n'ont pas les mêmes caractéristiques épidémiologiques avec des différences retrouvées sur la probabilité de survie, sur la dynamique des taux et sur les effets de variables telles que l'année. Il est donc indispensable de dissocier les sous-types de LMNH dans les travaux futurs.

Il sera nécessaire par la suite d'étudier la dynamique du taux de mortalité en excès au-delà de 5 ans après le diagnostic pour ces deux types de LMNH.

Les résultats issus de bases de population permettent de fournir des indicateurs d'observation de la survie en population générale. En l'absence d'informations cliniques, il apparaît en revanche délicat de définir avec exactitude la part des diverses composantes ayant joué un rôle sur cette diminution de la mortalité, à savoir la part des progrès dans la prévention, la détection, le diagnostic ou dans la prise en charge thérapeutique des patients. Aussi, les résultats de nos travaux permettent d'apprécier l'efficacité globale du système de soins des LF et LBDGC.

Les bases collaboratives de population permettent d'obtenir des résultats assez robustes en se basant sur un nombre conséquent de cas diagnostiqués et de fournir des comparaisons géographiques pour mieux appréhender la prise en charge du cancer. Ces études collaboratives nécessitent tout de même l'assurance de l'homogénéité et de l'exhaustivité des cas. Dans l'étude des LMNH, ces lymphomes ont connu plusieurs évolutions successives dans les classifications internationales. Aussi, des études comparatives temporelles ou géographiques par sous-type de LMNH ne peuvent être réalisables qu'à partir d'une certaine période de diagnostic. Dans nos travaux, nous avons considéré une homogénéité acceptable du codage par sous-type de LMNH en France et ailleurs à partir de l'année 1995, date à laquelle les classifications épidémiologiques se sont largement basées sur la classification REAL utilisée en hématologie et en se basant également sur des études de tendance temporelle de l'incidence.

Ce travail a permis de coupler les avancées méthodologiques récentes dans le domaine de la survie nette avec l'étude du pronostic des patients atteints d'un LF ou d'un LBDGC, en répondant à des besoins cliniques réels. Les méthodes discutées dans ce rapport peuvent

également être utilisées pour l'analyse de survie des maladies autres que le cancer. Nous espérons que ce travail amènera le lecteur à prendre pleinement conscience de l'importance d'utiliser une méthodologie adaptée dans l'étude du pronostic des patients à la suite d'un diagnostic de cancer.

V.2. PERSPECTIVES

A la suite de ce travail basé sur l'étude de la survie nette en population générale des patients atteints d'un LMNH, plusieurs perspectives apparaissent intéressantes : des perspectives portant sur l'amélioration des outils statistiques et d'autres pour approfondir l'étude du pronostic des patients à la suite d'un diagnostic de LMNH.

V.2.1. ÉVOLUTION DU CADRE MÉTHODOLOGIQUE DE LA SURVIE NETTE

Ces dernières années, le cadre méthodologique de la survie nette a connu une avancée majeure avec de nouveaux outils disponibles. Actuellement, plusieurs axes méthodologiques sont en cours de réflexion :

- *Stratification des tables de mortalité sur d'autres facteurs pronostiques :*

Les travaux de Graffeo *et al.* (Grafféo *et al.* 2012) ont mis en évidence que l'absence de stratification pour un facteur pronostique que l'on souhaite étudier, biaise les estimations des effets de cette variable sur la mortalité en excès mais aussi les estimations des effets des autres variables. Par exemple, l'étude de l'impact des inégalités socio-économiques sur la mortalité en excès nécessite de disposer en France de tables de mortalité ajustées sur un indicateur socio-économique.

- Comparaison de distribution de survie nette :

Le test du Log-rank permet de comparer les fonctions de survie « toutes causes » estimées à partir d'un estimateur non paramétrique. L'équivalent d'un tel test serait appréciable dans le cadre de la survie nette. Grafféo *et al.* ont proposé la construction d'un test permettant la comparaison de distributions de survie nette estimées à partir de l'estimateur Pohar-Perme (Grafféo 2014). L'utilisation de ce test permettra par la suite d'associer une significativité aux différences de probabilités de survie nette estimées de façon ponctuelle. Ce test est en cours d'implémentation sous le logiciel R.

- Vérification des différentes hypothèses d'un modèle paramétrique du taux en excès :

La mise en place d'outils permettant de tester les principales hypothèses d'un modèle du taux en excès concernant la fonction de lien, la forme fonctionnelle d'une covariable et le caractère proportionnel ou non de son effet, serait un atout majeur. Ce travail est en cours dans le cadre du projet CENSUR .

- Mise à disposition de ces nouveaux outils dans les logiciels de statistiques :

L'estimateur Pohar-Perme est disponible sous le logiciel R avec le package « Relsurv » (Pohar et Stare 2006) et sous le logiciel STATA avec la commande « stns » (Clerc-Urmes *et al.* 2014). L'implémentation du modèle régressif de Remontet *et al.* (Remontet *et al.* 2007) est en cours de réalisation sur ces deux logiciels afin de le rendre accessible à tout utilisateur.

Ces diverses évolutions méthodologiques représentent des axes de recherche du projet CENSUR, auxquels nous pouvons également ajouter des développements propres : aux outils non paramétriques (avec l'étude de la variance de l'estimateur Pohar-Perme (Dickman *et al.* 2013)), aux outils paramétriques (avec la comparaison de différentes stratégies de construction de modèle) ou encore aux outils de projection de survie.

V.2.2. ÉTUDE DU PRONOSTIC DES LMNH

A la suite de ce travail, des perspectives spécifiques à la conduite des études de survie des patients atteints d'un LMNH sont envisagées.

- Nécessité de disposer du stade du lymphome de façon systématique:

L'application des méthodes de survie nette sur les données françaises présentée en Chapitre III a rapporté une différence de mortalité en excès en fonction du stade du lymphome au diagnostic. Il aurait été souhaitable de pouvoir disposer du stade du lymphome également sur les données européennes. Le stade étant connu comme un facteur pronostique important, il est nécessaire d'aller vers un recueil systématique de cette variable. Les LF et LBDGC font actuellement l'objet d'une étude de faisabilité dans le cadre d'EUROCARE, afin de pouvoir disposer de facteurs cliniques au diagnostic et de la prise en charge thérapeutique des patients à plus grande échelle.

- Etude des phases évolutives à la suite d'un diagnostic de LMNH :

Malgré une amélioration de la survie nette des patients, les LF connaissent une évolution lente mais souvent caractérisée par des rechutes récurrentes ou par une évolution clinique accélérée par la survenue d'une transformation histologique en un lymphome d'évolutivité plus agressive (Ducassou *et al.* 2011). De moins bon pronostic, l'évolution d'un LBDGC peut également être marquée par des rechutes. Plusieurs études rapportent également un risque augmenté de développer une seconde tumeur à la suite d'un diagnostic de LMNH. La toxicité des traitements des lymphomes en est souvent la cause principale (Mudie *et al.* 2006),(Sacchi *et al.* 2008). Pour ces deux sous-types de LMNH qui connaissent différentes phases dans leur évolution, le décès n'apparaît alors pas comme le seul événement d'intérêt.

Plusieurs essais cliniques ont rapporté des résultats en termes de survie sans progression s'intéressant à la fois à la rechute et au décès du patient (Karlin et Coiffier 2013). Cette méthodologie repose sur des outils de survie toutes causes. Néanmoins dans cette approche, les événements transitoires tels que la rechute ne sont plus observables après le décès du patient et sont donc censurés par cet événement terminal induisant ainsi une censure non indépendante. Par ailleurs, ces méthodes ne se basent que sur l'étude du premier événement observé et des effets des covariables sur cet événement. Un intérêt majeur réside dans l'étude simultanée des diverses possibilités d'évolution des LMNH en tenant compte des

événements transitoires, tels que l'apparition de rechutes, de la transformation en un lymphome plus agressif, ou de seconds cancers pour amener à l'événement terminal qui est le décès du patient. Afin de visualiser l'impact réel de la maladie sur la mortalité, il apparaît plus pertinent de ne s'intéresser qu'à la mortalité induite par le lymphome, ce qui amènerait à se positionner dans le cadre de la survie nette.

Dans le but de répondre à ces besoins cliniques, Belot *et al.* (Belot *et al.* 2011) ont proposé un modèle de survie nette en présence d'événements concurrents dans le concept des risques compétitifs s'intéressant ainsi aux effets de divers covariables sur l'apparition du premier événement. Dans le cadre des modèles multi-états, Huszti *et al.* (Huszti *et al.* 2012) ont proposé un modèle de Markov non-homogène en dissociant les décès dus au cancer des autres causes. Extensions des modèles de risques concurrents, ces modèles multi-états représentent un outil adapté pour étudier les différents états successifs d'un patient à la suite d'un diagnostic de cancer, permettant d'étendre l'analyse au-delà du premier événement et d'étudier les forces de passage entre ces différents états (Putter *et al.* 2007). Le modèle proposé par Huszti *et al.* repose sur l'hypothèse que la probabilité de transiter dans un état dépend du temps depuis l'état initial, l'état initial étant en général choisi comme le diagnostic du cancer. Cependant d'un point de vue clinique, cette hypothèse n'est pas toujours vérifiée, les états dépendants davantage du temps passé dans l'état précédent que du temps depuis le diagnostic. Aussi Gillaizeau *et al.* (Gillaizeau *et al.* 2015) ont proposé tout récemment un modèle semi-Markovien qui considère que l'intensité de transition entre deux états dépend seulement du temps passé dans l'état précédent.

Une autre possibilité statistique que nous souhaiterions également explorer est l'utilisation d'une modélisation conjointe d'événements récurrents avec un événement terminal (Mazroui *et al.* 2013). Le choix d'un modèle conjoint à fragilité permettrait de tenir compte des rechutes successives que peuvent connaître les patients au cours de leur suivi, ce qui est difficilement réalisable à l'aide d'un modèle multi-états. Récemment, un modèle conjoint associant la modélisation d'un événement récurrent avec un événement terminal dans le concept de la survie nette a été publié (Belot *et al.* 2014).

Ce travail aura pour objectif d'estimer les effets de diverses variables cliniques sur les événements évolutifs du patient (rechutes, transformation, second cancer, décès dû au cancer) à la suite d'un diagnostic de LF ou d'un LBDGC à l'aide d'une méthodologie adaptée et en se basant sur des données de population. Les données proviendront de la base de données ISO-

Lymph, projet financé par l'INCA et piloté par le Registre des Hémopathies de Gironde (coordinateur : Sandra Leguyader-Peyrou) qui s'appuie sur une étude rétrospective et exhaustive de tous les cas de LF et LBDGC diagnostiqués sur la période de 2002 à 2008 issus des bases de trois registres français spécialisés en hématologie. Ce travail se basant sur la modélisation conjointe de trois événements (dont un récurrent) et d'un événement terminal représenté ici par les décès dus au lymphome, nécessiterait une extension des outils actuels.



Annexes

Survival for haematological malignancies in Europe between 1997 and 2008 by region and age: results of EUROCARE-5, a population-based study



Milena Sant, Pamela Minicozzi, Morgane Mounier, Lesley A Anderson, Hermann Brenner, Bernd Holleczek, Rafael Marcos-Gragera, Marc Maynadié, Alain Monnereau, Gemma Osca-Gelis, Otto Visser, Roberta De Angelis, the EUROCARE-5 Working Group*

Summary

Background More effective treatments have become available for haematological malignancies from the early 2000s, but few large-scale population-based studies have investigated their effect on survival. Using EUROCARE data, and HAEMACARE morphological groupings, we aimed to estimate time trends in population-based survival for 11 lymphoid and myeloid malignancies in 20 European countries, by region and age.

Methods In this retrospective observational study, we included patients (aged 15 years and older) diagnosed with haematological malignancies, diagnosed up to Dec 31, 2007, and followed up to Dec 31, 2008. We used data from the 30 cancer registries (across 20 countries) that provided continuous incidence and good quality data from 1992 to 2007. We used a hybrid approach to estimate age-standardised and age-specific 5-year relative survival, for each malignancy, overall and for five regions (UK, and northern, central, southern, and eastern Europe), and for four 3-year periods (1997–99, 2000–02, 2003–05, 2006–08). For each malignancy, we also estimated the relative excess risk of death during the 5 years after diagnosis, by period, age, and region.

Findings We analysed 560 444 cases. From 1997–99 to 2006–08 survival increased for most malignancies: the largest increases were for diffuse large B-cell lymphoma (42.0% [95% CI 40.7–43.4] to 55.4% [54.6–56.2], $p < 0.0001$), follicular lymphoma (58.9% [57.3–60.6] to 74.3% [72.9–75.5], $p < 0.0001$), chronic myeloid leukaemia (32.3% [30.6–33.9] to 54.4% [52.5–56.2], $p < 0.0001$), and acute promyelocytic leukaemia (50.1% [43.7–56.2] to 61.9% [57.0–66.4], $p = 0.0038$, estimate not age-standardised). Other survival increases were seen for Hodgkin's lymphoma (75.1% [74.1–76.0] to 79.3% [78.4–80.1], $p < 0.0001$), chronic lymphocytic leukaemia/small lymphocytic lymphoma (66.1% [65.1–67.1] to 69.0% [68.1–69.8], $p < 0.0001$), multiple myeloma/plasmacytoma (29.8% [29.0–30.6] to 39.6% [38.8–40.3], $p < 0.0001$), precursor lymphoblastic leukaemia/lymphoma (29.8% [27.7–32.0] to 41.1% [39.0–43.1], $p < 0.0001$), acute myeloid leukaemia (excluding acute promyelocytic leukaemia, 12.6% [11.9–13.3] to 14.8% [14.2–15.4], $p < 0.0001$), and other myeloproliferative neoplasms (excluding chronic myeloid leukaemia, 70.3% [68.7–71.8] to 74.9% [73.8–75.9], $p < 0.0001$). Survival increased slightly in southern Europe, more in the UK, and conspicuously in northern, central, and eastern Europe. However, eastern European survival was lower than that for other regions. Survival decreased with advancing age, and increased with time only slightly in patients aged 75 years or older, although a 10% increase in survival occurred in elderly patients with follicular lymphoma, diffuse large B-cell lymphoma, and chronic myeloid leukaemia.

Interpretation These trends are encouraging. Widespread use of new and more effective treatment probably explains much of the increased survival. However, the persistent differences in survival across Europe suggest variations in the quality of care and availability of the new treatments. High-resolution studies that collect data about stage at diagnosis and treatments for representative samples of cases could provide further evidence of treatment effectiveness and explain geographic variations in survival.

Funding Compagnia di San Paolo, Fondazione Cariplo, European Commission, and Italian Ministry of Health.

Introduction

New targeted drugs for treating haematological malignancies became available from the early 2000s, including anti-CD20 monoclonal antibodies for B-cell lymphomas^{1,2} and acute lymphoblastic leukaemia,³ imatinib for chronic myeloid leukaemia,⁴ and proteasome inhibitors for myeloma.⁵ These treatments have greatly improved prognosis for chronic myeloid leukaemia, many lymphoid neoplasms and, to a lesser extent, multiple myeloma.^{2,6–8} However these improvements

were mainly documented in hospital-based series or controlled trials. Few population-based studies by cancer registries have investigated survival for haematological malignancies.⁸ Cancer registries include all incident cases in a registry area and thus reflect clinical practice and survival everywhere, not just in centres of excellence or selected groups of patients. Past EUROCARE studies⁹ have shown substantial differences in population-based survival for haematological malignancies across Europe; however, because of changes in disease classification and

Lancet Oncol 2014; 15: 931–42

Published Online

July 14, 2014

[http://dx.doi.org/10.1016/S1470-2045\(14\)70282-7](http://dx.doi.org/10.1016/S1470-2045(14)70282-7)

This online publication has been corrected.

The corrected version first appeared at thelancet.com/oncology on September 1, 2014

See [Comment](#) page 906

*Members listed in appendix

Analytical Epidemiology and Health Impact Unit, Department of Preventive and Predictive Medicine,

Fondazione IRCCS, Istituto Nazionale dei Tumori, Milan, Italy (M Sant MD,

P Minicozzi PhD); Registre des hémopathies malignes de Côte d'Or, EA 4184, University of Burgundy and University Hospital of Dijon, Dijon, France (M Mounier MSc,

Prof M Maynadié MD); Cancer Epidemiology and Health Services Research Group,

Centre for Public Health, Queen's University Belfast, Belfast, Northern Ireland

(LA Anderson PhD); Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DFKZ),

Heidelberg, Germany

(Prof H Brenner MD); German

Cancer Consortium (DKTK),

Heidelberg, Germany

(Prof H Brenner); Saarland

Cancer Registry, Saarbrücken,

Germany (B Holleczek MSc);

Epidemiology Unit and Girona

Cancer Registry, Oncology

Coordination Plan, Department

of Health, Autonomous

Government of Catalonia,

Catalan Institute of Oncology,

Girona Biomedical Research

Institute, Girona, Spain

(R Marcos-Gragera PhD,

G Osca-Gelis MSc); Registre des

hémopathies malignes de la

Gironde, Institut Bergonié,

Bordeaux, France

(A Monnereau MD); Centre

INSERM U897, Centre

Articles

d'Investigation Clinique, Bordeaux, France (A Monnerieu); Comprehensive Cancer Centre the Netherlands, Department of Registration and Research, Utrecht, Netherlands (O Visser PhD); and Centro Nazionale di Epidemiologia, Sorveglianza e Promozione della Salute (CNEPS), Istituto Superiore di Sanità, Rome, Italy (R De Angelis MSc)

Correspondence to: Dr Milena Sant, Analytical Epidemiology and Health Impact Unit, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Via Venezian 1, 20133 Milan, Italy milena.sant@istitutotumori.mi.it

See Online for appendix

definitions of malignancy over recent decades, investigating changes in haematological malignancy survival over time was not possible. The HAEMACARE project¹⁰ promoted the standardisation of data registration for haematological malignancies by EURO CARE cancer registries, and encouraged adoption of the updated International Classification of Disease for Oncology.^{11,12} As a result, survival time trends can now be estimated in different European populations based on comparable data.

The primary aim of the present study was to investigate changes in 5-year relative survival for adult European patients diagnosed with haematological malignancies by morphology, age at diagnosis, and European region; we were particularly interested in morphologies for which new, more effective treatments became available from the early 2000s. We aimed to corroborate our results with multivariate models to estimate relative excess risks of death in these patients by calendar period, age at diagnosis, and European region, in the 5 years after diagnosis.

Methods

Participants

We included adult (age ≥ 15 years) patients archived in the EURO CARE-5 database,¹³ diagnosed with haematological malignancies up to Dec 31, 2007 and followed up to Dec 31, 2008. We used standardised procedures to check data quality and data completeness, as described elsewhere.¹³ Cases were coded according to the International Classification of Disease for Oncology 3rd edition,¹¹ grouped according to WHO¹² and HAEMACARE^{10,14} criteria, and divided into four categories according to age at diagnosis: 15–54 years, 55–64 years, 65–74 years, and ≥ 75 years.

We estimated survival time trends for six specific lymphoid morphologies: Hodgkin's lymphoma; chronic lymphocytic leukaemia/small lymphocytic lymphoma; follicular lymphoma; diffuse large B-cell lymphoma; multiple myeloma/plasmacytoma; and precursor lymphoblastic leukaemia/lymphoma (see appendix for International Classification of Disease for Oncology codes). We also estimated survival time trends for five specific myeloid morphologies: acute promyelocytic leukaemia; acute myeloid leukaemia (excluding acute proteomyelocytic leukaemia); chronic myeloid leukaemia; other myeloproliferative neoplasms (excluding chronic myeloid leukaemia); and myelodysplastic syndromes (appendix p 2).

For more uncommon cancer morphologies, we could not estimate meaningful survival trends because of the small numbers of cases. We completely excluded mast-cell tumours because they were particularly rare (about 300 cases). We estimated time trends for 5-year relative survival¹⁵ using a modified period approach (hybrid approach)¹⁶ for patients under observation in four 3-year periods (1997–99, 2000–02, 2003–05, and 2006–08). This

approach enables 5-year survival for patients diagnosed in the most recent period (without 5 years of follow-up) to be estimated from the survival of patients diagnosed in preceding years (appendix p 3). For estimates for earlier periods we had to include survival data from patients diagnosed before 1996. Therefore, of the 102 EURO CARE-5 cancer registries archiving adult cases, only the 35 able to provide continuous incidence data from 1992 to 2007 were eligible for inclusion. Of these, five were excluded as having more than 14% of cases known by death certificate only or diagnosed at autopsy; more than 30% of cases with unknown histological or cytological verification; or more than 30% of cases with not otherwise specified morphology (including lymphoid, non-Hodgkin lymphoma, and myeloid not otherwise specified; appendix p 2).

Thus, 30 cancer registries contributed to the study (table 1), grouped into five European regions: northern Europe (Denmark, Iceland, and Norway), the UK (England, Northern Ireland, Scotland, and Wales), central Europe (Austria, Côte d'Or haematological, Saarland, Basel, Geneva, Grisons, St Gallen, Valais, and Netherlands), eastern Europe (Bulgaria, Estonia, Lithuania, Kielce, and Slovakia), and southern Europe (Ferrara, Modena, Parma, Ragusa, Romagna, Sassari, Torino, Malta, and Slovenia).

The Danish cancer registry provided information in broad morphology categories but without specific International Classification of Disease for Oncology morphology codes. Analyses of data quality showed that only for some categories (Hodgkin's lymphoma, chronic lymphocytic leukaemia/small lymphocytic lymphoma, multiple myeloma/plasmacytoma, precursor lymphoblastic leukaemia/lymphoma, acute myeloid leukaemia, and chronic myeloid leukaemia) were proportions of total cases compatible with proportions in other national cancer registries geographically close to Denmark. We included only these categories in the survival analyses, whereas we included all Danish cases in the descriptive analyses.

Statistical analysis

We estimated 5-year relative survival, with 95% CIs, by European region and age at diagnosis using the hybrid approach.¹⁶ Relative survival¹⁵ is the ratio of observed survival in the patient group to expected survival in a comparable group of the general population assumed to be free of the cancer of interest. Expected survival was estimated by the Ederer 2 method¹⁷ using registry-specific lifetables stratified by age, sex, and year, smoothed¹⁸ and checked against official published mortality data. Because the age structure of European populations varies, relative survival estimates were age-standardised with international cancer survival standards.¹⁹ For acute promyelocytic leukaemia, the small number of cases in each age class made it expedient to estimate non-age-standardised survival. For myelodysplastic syndromes, we present

time trends in 3-year survival from 2003–2005 to 2006–2008, because of large between-region variations in case distribution before 2001. We excluded cases recorded on death certificate or diagnosed at autopsy from the survival analyses.

For each haematological malignancy, we also estimated the relative excess risk of death (with 95% CIs), during the 5 years after diagnosis, by calendar period, age at diagnosis, and European region, using generalised linear

models that assumed a Poisson distribution for the observed number of deaths.¹⁵ The models estimated the extent to which excess risk of death (ie, excess risk of death because of cancer compared with the risk of death in the general population) in the category under consideration differed from the excess risk of death in the reference category (calendar period 1997–99, age 15–54 years, UK region, or first year of follow-up after diagnosis) after controlling for the covariates (calendar

Cancer registry	All cases	Recorded only on death certificate	Diagnosed only at autopsy	Eligible for survival analyses					
				All cases	HL	NHL	MM/P	Myeloid malignancies	NOS
Northern Europe	44 402	308 (0.7%)	130 (0.3%)	43 964	2754 (6.3%)	24 187 (55.0%)	7369 (16.8%)	7333 (16.7%)	2321 (5.2%)
Denmark Denmark	22 416	0 (0.0%)	98 (0.4%)	22 318	1449 (6.5%)	14 146 (63.4%)	3408 (15.3%)	3271 (14.7%)	44 (0.1%)
Iceland Iceland	1113	5 (0.4%)	9 (0.8%)	1099	79 (7.2%)	584 (53.1%)	184 (16.8%)	225 (20.5%)	27 (2.4%)
Norway Norway	20 873	303 (1.5%)	23 (0.1%)	20 547	1226 (6.0%)	9457 (46.0%)	3777 (18.5%)	3837 (18.7%)	2250 (10.8%)
UK	315 608	11342 (3.6%)	65 (0.0%)	304 201	17 067 (5.6%)	106 579 (35.1%)	45 121 (14.8%)	84 667 (27.8%)	50 767 (16.7%)
England England	269 578	10695 (4.0%)	0 (0.0%)	258 883	14 270 (5.5%)	87 002 (33.5%)	37 251 (14.4%)	76 858 (29.7%)	43 502 (16.9%)
Northern Ireland	6454	64 (1.0%)	23 (0.4%)	6367	421 (6.6%)	2105 (33.0%)	1326 (20.8%)	862 (13.5%)	1653 (26.1%)
Scotland Scotland	24 606	99 (0.4%)	42 (0.2%)	24 465	1547 (6.3%)	11 092 (45.3%)	3970 (16.2%)	4584 (18.7%)	3272 (13.5%)
Wales Wales	14 970	484 (3.2%)	0 (0.0%)	14 486	829 (5.7%)	6380 (44.0%)	2574 (17.8%)	2363 (16.3%)	2340 (16.2%)
Central Europe	114 377	507 (0.4%)	407 (0.4%)	113 463	7309 (6.4%)	53 987 (47.6%)	16 818 (14.8%)	24 580 (21.7%)	10 769 (9.5%)
Austria Austria	28 287	0 (0.0%)	0 (0.0%)	28 287	1875 (6.6%)	10 091 (35.6%)	3948 (14.0%)	5753 (20.3%)	6620 (23.5%)
France Côte d'Or haematological	3175	0 (0.0%)	0 (0.0%)	3175	161 (5.1%)	1548 (48.8%)	406 (12.8%)	934 (29.4%)	126 (3.9%)
Germany Saarland	4905	412 (8.4%)	0 (0.0%)	4493	289 (6.4%)	2280 (50.7%)	670 (14.9%)	748 (16.6%)	506 (11.4%)
Switzerland Basel	2217	64 (2.9%)	71 (3.2%)	2082	140 (6.7%)	1048 (50.3%)	303 (14.6%)	504 (24.2%)	87 (4.2%)
Switzerland Geneva	2294	10 (0.4%)	13 (0.6%)	2271	162 (7.1%)	950 (41.8%)	264 (11.6%)	695 (30.6%)	200 (8.9%)
Switzerland Grisons	1178	0 (0.0%)	2 (0.2%)	1176	66 (5.6%)	529 (45.0%)	163 (13.9%)	332 (28.2%)	86 (7.3%)
Switzerland St Gallen	2512	6 (0.2%)	27 (1.1%)	2479	170 (6.9%)	1233 (49.7%)	367 (14.8%)	585 (23.6%)	124 (5.0%)
Switzerland Valais	1390	15 (1.1%)	1 (0.1%)	1374	105 (7.6%)	641 (46.7%)	195 (14.1%)	351 (25.5%)	82 (6.1%)
Netherlands Netherlands	68 419	0 (0.0%)	293 (0.4%)	68 126	4341 (6.4%)	35 667 (52.4%)	10 502 (15.4%)	14 678 (21.5%)	2938 (4.3%)
Eastern Europe	48 570	2103 (4.3%)	330 (0.7%)	46 137	4762 (10.3%)	17 200 (37.3%)	6342 (13.7%)	10 076 (21.8%)	7757 (16.9%)
Bulgaria Bulgaria	15 783	1 (0.0%)	0 (0.0%)	15 782	1973 (12.5%)	4482 (28.4%)	1800 (11.4%)	3246 (20.6%)	4281 (27.1%)
Estonia Estonia	4523	16 (0.4%)	41 (0.9%)	4466	364 (8.2%)	2073 (46.5%)	605 (13.5%)	1029 (23.0%)	395 (8.8%)
Lithuania Lithuania	10 044	233 (2.3%)	0 (0.0%)	9811	924 (9.4%)	3832 (39.1%)	1372 (14.0%)	2275 (23.2%)	1408 (14.3%)
Poland Kielce	3451	55 (1.6%)	4 (0.1%)	3392	354 (10.4%)	1180 (34.8%)	588 (17.3%)	543 (16.0%)	727 (21.5%)
Slovakia Slovakia	14 769	1798 (12.2%)	285 (1.8%)	12 686	1147 (9.0%)	5633 (44.4%)	1977 (15.6%)	2983 (23.5%)	946 (7.5%)
Southern Europe	37 487	380 (1.0%)	125 (0.3%)	36 982	2191 (5.9%)	16 835 (45.5%)	5912 (16.1%)	8758 (23.7%)	3286 (8.8%)
Italy Ferrara	2793	29 (1.0%)	11 (0.4%)	2753	170 (6.2%)	1279 (46.5%)	426 (15.4%)	644 (23.4%)	234 (8.5%)
Italy Modena	4759	10 (0.2%)	0 (0.0%)	4749	244 (5.1%)	2440 (51.3%)	710 (15.0%)	1267 (26.7%)	88 (1.9%)
Italy Parma	3258	2 (0.1%)	17 (0.5%)	3239	174 (5.4%)	1259 (38.9%)	564 (17.4%)	744 (23.0%)	498 (15.3%)
Italy Ragusa	1486	24 (1.6%)	0 (0.0%)	1462	92 (6.3%)	490 (33.6%)	217 (14.8%)	518 (35.4%)	145 (9.9%)
Italy Romagna	8026	182 (2.3%)	9 (0.1%)	7835	397 (5.1%)	3641 (46.5%)	1487 (19.0%)	1609 (20.5%)	701 (8.9%)
Italy Sassari	2191	19 (0.9%)	0 (0.0%)	2172	176 (8.1%)	1048 (48.3%)	313 (14.4%)	456 (21.0%)	179 (8.2%)
Italy Torino	6159	68 (1.1%)	1 (0.0%)	6090	325 (5.3%)	2733 (44.9%)	979 (16.0%)	1385 (22.7%)	668 (11.1%)
Malta Malta	1429	46 (3.2%)	2 (0.1%)	1381	119 (8.6%)	558 (40.4%)	150 (10.8%)	350 (25.3%)	204 (14.9%)
Slovenia Slovenia	7386	0 (0.0%)	85 (1.2%)	7301	494 (6.8%)	3387 (46.4%)	1066 (14.6%)	1785 (24.4%)	569 (7.8%)
All cases	560 444	14 640 (2.6%)	1057 (0.2%)	544 747	34 083 (6.3%)	218 788 (40.1%)	81 562 (15.0%)	135 414 (24.9%)	74 900 (13.7%)

Data are n (%). HL=Hodgkin's lymphoma. NHL=non-Hodgkin lymphoma. MM/P=multiple myeloma/plasmacytoma. NOS=not otherwise specified.

Table 1: Haematological malignancies diagnosed in 1996–2007, by European region and country

Articles

period, age at diagnosis, region, and year of follow-up, except that under study). Relative excess risk of death for the reference categories was set to 1; thus, a relative excess risk of death of greater than 1 implies greater risk of death than the reference category.

We also used generalised linear models, adjusted for age at diagnosis where relevant, and year of follow-up, to investigate the significance of linear trends in survival by European region and age category. Wald test *p* values less than 0.05 (when calendar period was included as covariate) indicated a significant change in survival over time. We did the analyses with STATA (version 12).

Role of the funding sources

The funders had no role in study design, collection, analysis or interpretation of data, or writing of the report. The corresponding author had full access to all the data and had final responsibility to submit for publication.

Results

560 444 haematological malignancies were diagnosed between 1996 and 2007 in the regions covered: 7% from southern Europe, 8% from northern Europe, 9% from eastern Europe, 20% from central Europe, and 56% from the UK (table 1). Only 2.8% of cases overall were recorded on death certificate only or diagnosed at autopsy, with high proportions in Wales (3.2%), Malta (3.3%), England (4.0%), Basel (6.1%), Saarland (8.4%), and Slovakia (14.0%; table 1). Excluding these cases, and grouping the remaining cases into five categories (appendix p 2), non-Hodgkin lymphoma was the most common haematological malignancy (40.1%), followed by myeloid malignancies (24.9%), multiple myeloma/plasmacytoma (15.0%), not otherwise specified (13.7%), and Hodgkin's lymphoma (15.0%; table 1). Northern Europe had the highest proportion of non-Hodgkin lymphoma (55.0%) and lowest proportion of malignancies that were not otherwise specified (5.2%). The UK and eastern Europe had the lowest proportions of non-Hodgkin lymphoma (35.1% and 37.3%, respectively) and highest proportions of not otherwise specified (16.7% and 16.9%, respectively). However, the proportion of not otherwise specified malignancies varied greatly within regions in central Europe (4–24%) and southern Europe (2–15%; table 1).

Patients with Hodgkin's lymphoma were generally younger (70% were aged 15–54 years) than were those with the other haematological malignancies (figure 1). By contrast, about 60% of patients with myelodysplastic syndromes and 40% of those with chronic lymphocytic leukaemia/small lymphocytic lymphoma, and multiple myeloma/plasmacytoma were aged 75 years or older at diagnosis (appendix pp 4–7). There was little regional variation in age distribution for each malignancy, although eastern European patients tended to be slightly younger. 54% of all patients were men, with little variation in sex ratio between region within broad haematological malignancy groupings (figure 1) or for specific lymphoid or myeloid malignancies (appendix pp 4–7). Of the lymphoid malignancies, multiple myeloma/plasmacytoma was the most common, followed by chronic lymphocytic leukaemia/small lymphocytic lymphoma, diffuse large B-cell lymphoma, Hodgkin's lymphoma, follicular lymphoma, and precursor lymphoblastic leukaemia/lymphoma (table 2). Acute myeloid leukaemia was the most common myeloid malignancy, ranging from 29.7% in the UK to 57.5% in northern Europe, followed by other myeloproliferative neoplasms (excluding chronic myeloid leukaemia), myelodysplastic syndromes, chronic myeloid leukaemia, and acute promyelocytic leukaemia (table 3).

5-year relative survival increased significantly over time for all lymphoid malignancies (figure 2, table 4), with the greatest increases for follicular lymphoma and diffuse large B-cell lymphoma, with more modest improvements for chronic lymphocytic leukaemia/small lymphocytic

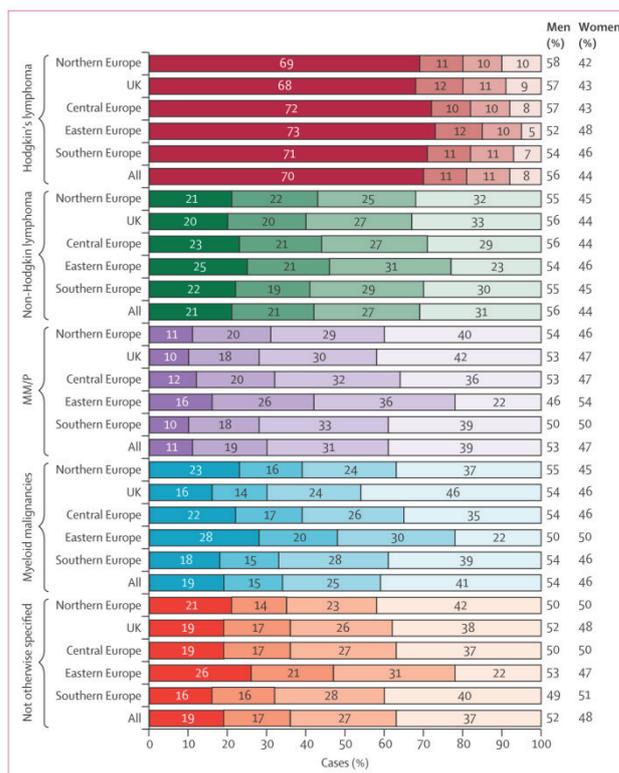


Figure 1: Age and sex distributions of patients diagnosed 1996–2007, according to region
Age categories (15–54, 55–64, 65–74, ≥75 years) are represented by four shades (dark to light). MM/P=multiple myeloma/plasmacytoma.

	Northern Europe	UK	Central Europe	Eastern Europe	Southern Europe	All cases
Hodgkin's lymphoma	2754 (7.6%)	17 067 (7.9%)	7309 (8.3%)	4762 (13.8%)	2191 (8.0%)	34 083 (8.4%)
NHL	24 187 (66.5%)	106 579 (49.3%)	53 987 (61.3%)	17 200 (50.2%)	16 835 (60.7%)	218 788 (54.3%)
CLL/SLL	6686 (18.4%)	35 675 (16.5%)	14 625 (16.6%)	8017 (23.4%)	4737 (17.1%)	69 740 (17.3%)
Follicular lymphoma	1794 (4.9%)	18 572 (8.6%)	8387 (9.5%)	1418 (4.1%)	1939 (7.0%)	32 110 (8.0%)
Diffuse large B-cell lymphoma	2439 (6.7%)	28 410 (13.1%)	15 212 (17.3%)	3570 (10.4%)	4730 (17.1%)	54 361 (13.5%)
LBL/L	699 (1.9%)	4190 (1.9%)	1866 (2.1%)	1405 (4.1%)	656 (2.4%)	8816 (2.2%)
Other NHL*	12 569 (34.6%)	19 732 (9.1%)	13 897 (15.8%)	2790 (8.1%)	4773 (17.2%)	53 761 (13.4%)
Multiple myeloma/plasmacytoma	7369 (20.3%)	45 121 (20.9%)	16 818 (19.1%)	6342 (18.5%)	5912 (21.3%)	81 562 (20.3%)
Not otherwise specified	2046 (5.6%)	47 457 (21.9%)	9964 (11.3%)	5982 (17.5%)	2777 (10.0%)	68 226 (17.0%)
NHL not otherwise specified	1854 (5.1%)	27 773 (12.8%)	6194 (7.0%)	3594 (10.5%)	2309 (8.3%)	41 724 (10.4%)
Lymphoid not otherwise specified	192 (0.5%)	19 684 (9.1%)	3770 (4.3%)	2388 (7.0%)	468 (1.7%)	26 502 (6.6%)

Data are n (%). CLL/SLL=chronic lymphocytic leukaemia/small lymphocytic lymphoma. LBL/L=precursor lymphoblastic leukaemia/lymphoma. NHL=non-Hodgkin lymphoma.
*Excluded from survival analyses because of uncommon morphology or (Danish cases only) with morphologies whose incidence proportions were not compatible with those in other national cancer registries geographically close to Denmark (see methods).

Table 2: Lymphoid malignancies diagnosed 1996–2007, by European region

	Northern Europe	UK	Central Europe	Eastern Europe	Southern Europe	All cases
Myeloid malignancies	7333 (96.3%)	84 667 (96.2%)	24 580 (96.8%)	10 076 (85.0%)	8758 (94.5%)	135 414 (95.3%)
APL	72 (0.9%)	841 (1.0%)	393 (1.5%)	115 (1.0%)	127 (1.4%)	1548 (1.1%)
AML	4374 (57.5%)	26 155 (29.7%)	10 793 (42.5%)	3830 (32.3%)	2934 (31.7%)	48 086 (33.8%)
CML	1242 (16.3%)	7586 (8.6%)	3385 (13.3%)	2671 (22.5%)	1253 (13.5%)	16 137 (11.4%)
MPN	732 (9.6%)	23 480 (26.7%)	4103 (16.2%)	2360 (19.9%)	2348 (25.3%)	33 023 (23.2%)
MDS	665 (8.7%)	23 851 (27.1%)	4716 (18.6%)	969 (8.2%)	1882 (20.3%)	32 083 (22.6%)
Other myeloid malignancies*	248 (3.3%)	2754 (3.1%)	1190 (4.7%)	131 (1.1%)	214 (2.3%)	4537 (3.2%)
Not otherwise specified	275 (3.6%)	3310 (3.8%)	805 (3.2%)	1775 (15.0%)	509 (5.5%)	6674 (4.7%)

Data are n (%). APL=acute promyelocytic leukaemia. AML=acute myeloid leukaemia excluding APL. CML=chronic myeloid leukaemia. MPN=other chronic myeloproliferative neoplasms, excluding CML. MDS=myelodysplastic syndromes. *Excluded from survival analyses because of uncommon morphology or (Danish cases only) with morphologies whose incidence proportions were not compatible with those in other national cancer registries geographically close to Denmark (see methods).

Table 3: Myeloid malignancies diagnosed 1996–2007, by European region

lymphoma and Hodgkin's lymphoma. For myeloid malignancies (figure 2, table 4), 5-year relative survival increased most for chronic myeloid leukaemia, but significant increases were also seen for acute promyelocytic leukaemia, acute myeloid leukaemia, and other myeloproliferative neoplasms excluding chronic myeloid leukaemia (table 4). 3-year survival for myelodysplastic syndromes did not change significantly in Europe from 2003–05 to 2006–08 ($p=0.5010$; table 4).

For all European regions, the survival increases for follicular lymphoma, diffuse large B-cell lymphoma, precursor lymphoblastic leukaemia/lymphoma, and chronic myeloid leukaemia were greater after 2000 (figure 2). Survival increased slightly in southern Europe, more in the UK, and conspicuously in northern, central, and eastern Europe (figure 2). Survival improvements in eastern Europe were greatest for diffuse large B-cell lymphoma, follicular lymphoma, chronic myeloid leukaemia, Hodgkin's lymphoma, and other myeloproliferative neoplasms even though survival remained lower in eastern Europe than elsewhere (appendix pp 8–10). 3-year survival for myelodysplastic syndromes

decreased slightly (but not significantly) in northern and southern Europe.

For each haematological malignancy (figure 3), younger patients had better 5-year relative survival than did older patients, and the survival improvement over time was generally greater in younger than in older patients. The increase in survival for Hodgkin's lymphoma was large in 55–64 year olds: from 60.8% (95% CI 57.1–64.3) in 1997–99 to 70.7% (67.6–73.6) in 2006–08 ($p<0.0001$; appendix p 11); whereas for older patients with Hodgkin's lymphoma, most of the survival increase occurred between 1997–99 and 2000–02, and hardly changed thereafter (figure 3). For follicular lymphoma, diffuse large B-cell lymphoma, and multiple myeloma/plasmacytoma, survival increased steadily for all ages, although survival increased only slightly in those aged 75 years or older. For chronic lymphocytic leukaemia/small lymphocytic lymphoma, and precursor lymphoblastic leukaemia/lymphoma, the survival improvement was also mainly confined to patients younger than 75 years (figure 3).

Survival was poor for patients with acute myeloid leukaemia, varying from close to 0% in the oldest patients,

Articles

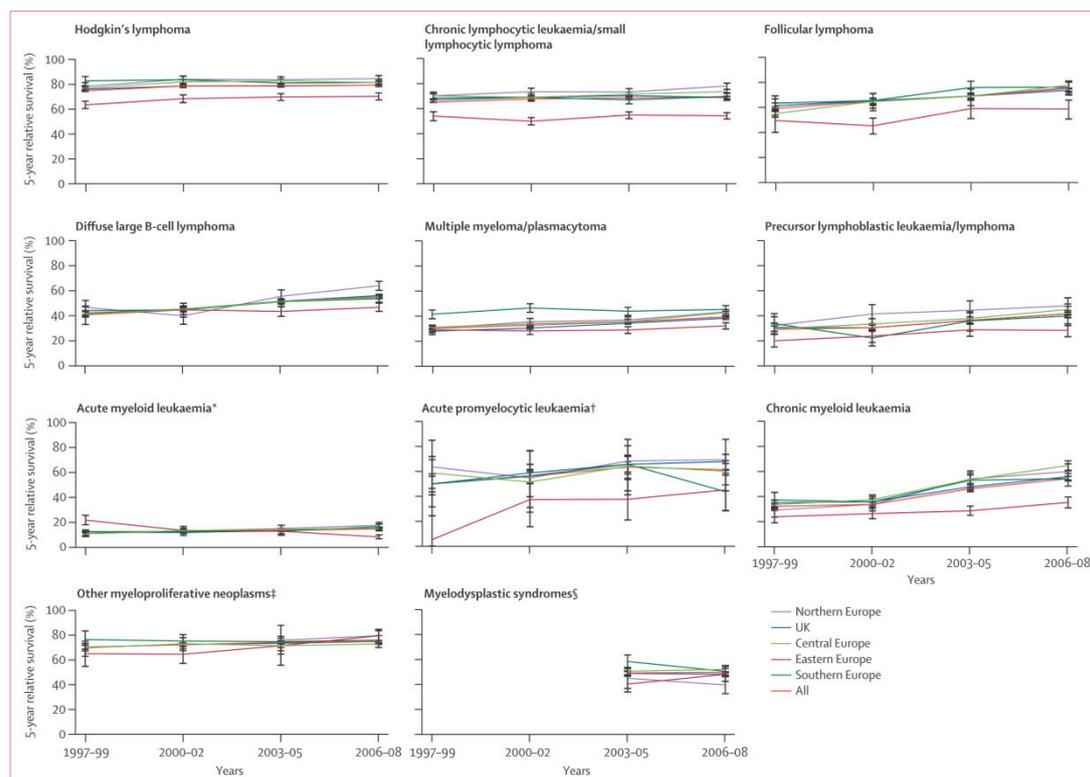


Figure 2: Changes in age-standardised 5-year relative survival, according to region

*Excluding acute promyelocytic leukaemia. †Non-age-standardised estimates. ‡Excluding chronic myeloid leukaemia; not enough cases in northern Europe 1997–2002 to estimate survival (<10 cases). §3-year relative survival.

	1997-99	2000-02	2003-05	2006-08	p value for trend
HL	75.1% (74.1-76.0)	78.5% (77.5-79.4)	78.9% (78.0-79.7)	79.3% (78.4-80.1)	<0.0001
CLL/SLL	66.1% (65.1-67.1)	67.0% (66.0-67.9)	68.7% (67.8-69.5)	69.0% (68.1-69.8)	<0.0001
FL	58.9% (57.3-60.6)	64.1% (62.5-65.5)	69.0% (67.5-70.3)	74.3% (72.9-75.5)	<0.0001
DLBCL	42.0% (40.7-43.4)	44.8% (43.7-46.0)	50.9% (49.9-51.8)	55.4% (54.6-56.2)	<0.0001
MM/P	29.8% (29.0-30.6)	32.7% (31.9-33.5)	35.2% (34.4-35.9)	39.6% (38.8-40.3)	<0.0001
LBL/L	29.8% (27.7-32.0)	30.7% (28.7-32.7)	36.1% (34.0-38.2)	41.1% (39.0-43.1)	<0.0001
AML	12.6% (11.9-13.3)	12.8% (12.2-13.4)	13.7% (13.1-14.3)	14.8% (14.2-15.4)	<0.0001
APL*	50.1% (43.7-56.2)	55.8% (50.2-61.1)	63.1% (57.9-67.8)	61.9% (57.0-66.4)	0.0038
CML	32.3% (30.6-33.9)	33.8% (32.2-35.4)	45.7% (43.9-47.5)	54.4% (52.5-56.2)	<0.0001
MPN	70.3% (68.7-71.8)	72.4% (70.9-73.8)	74.2% (73.0-75.5)	74.9% (73.8-75.9)	<0.0001
MDS†	48.9% (47.5-50.3)	48.8% (47.5-50.1)	0.5010

Data are relative survival (95% CI) unless stated otherwise. HL=Hodgkin's lymphoma. CLL/SLL=chronic lymphocytic leukaemia/small lymphocytic lymphoma. FL=follicular lymphoma. DLBCL=diffuse large B-cell lymphoma. MM/P=multiple myeloma/plasmacytoma. LBL/L=precursor lymphoblastic leukaemia/lymphoma. APL=acute promyelocytic leukaemia. AML=acute myeloid leukaemia (excluding APL). CML=chronic myeloid leukaemia. MPN=other myeloproliferative neoplasms (excluding CML). MDS=myelodysplastic syndromes. *Not age-standardised. †3-year relative survival.

Table 4: Time trends in age-standardised 5-year relative survival

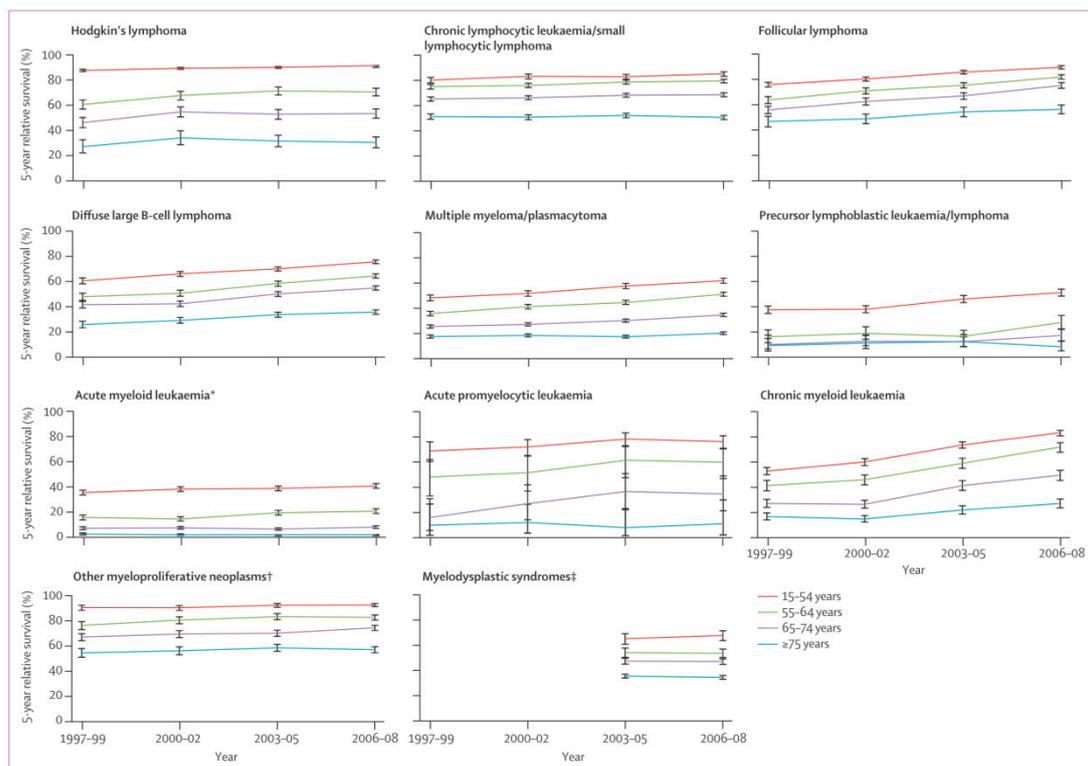


Figure 3: Changes in 5-year relative survival, according to age at diagnosis
 *Excluding acute promyelocytic leukaemia. †Excluding chronic myeloid leukaemia. ‡3-year relative survival.

to around 40% in the youngest. Survival improved slightly in patients aged 15–64 years over the 11 years assessed, but not in older patients. Survival for acute promyelocytic leukaemia was quite good, except for patients younger than 75 years; for younger patients survival increased steadily up to 2003–05 and stabilised afterwards (figure 3). The most notable survival increase across all ages was for chronic myeloid leukaemia ($p < 0.0001$; appendix p 12). This was evident from 2000–02, and was particularly notable for patients younger than age 65 years, although a 10% increase from 1997–99 to 2006–08 also occurred in patients 75 years or older. Survival for other myeloproliferative neoplasms was higher than for chronic myeloid leukaemia for each age group, and improved only slightly (by 2–7 percentage points) over the study period. For myelodysplastic syndromes, 3-year survival was stable for all ages from 2003 to 2008 (figure 3).

Table 5 shows relative excess risks of death during the 5 years (3 years for myelodysplastic syndrome) after

diagnosis, by period, age at diagnosis, and region. After adjustment for age, region, and year of follow-up, the relative excess risks of death for follicular lymphoma and chronic myeloid leukaemia in 2006–08 were 0.50 (95% CI 0.46–0.53) and 0.55 (0.51–0.59), compared with 1997–99 (table 5). Over the same period, the relative excess risk of death also fell for chronic lymphocytic leukaemia/small lymphocytic lymphoma, acute myeloid leukaemia, and other malignancies, except myelodysplastic syndromes, for which no changes in excess risk of death occurred from 2003 to 2008 (table 5).

After adjustment for period, region, and year of follow-up, the excess risk of death increased significantly with age for all morphologies. Particularly, for patients aged 75 years or older, the relative excess risk of death when compared to patients aged between 15 and 54 years old was 13.4 (95% CI 12.5–14.3) for Hodgkin's lymphoma, 6.8 (6.1–7.6) for other myeloproliferative neoplasms, 6.3 (5.0–7.9) for acute

Articles

	Period (reference: 1997-99)			Age at diagnosis (reference: 15-54 years)			European region (reference: UK)			
	2000-02	2003-05	2006-08	55-64 years	65-74 years	≥75 years	Northern Europe	Central Europe	Eastern Europe	Southern Europe
HL	0.83 (0.77-0.89)	0.82 (0.77-0.89)	0.83 (0.77-0.89)	3.69 (3.43-3.98)	6.55 (6.12-7.01)	13.35 (12.47-14.30)	0.70 (0.63-0.79)	0.85 (0.79-0.91)	1.73 (1.61-1.85)	0.75 (0.67-0.85)
CLL/SLL	0.97 (0.92-1.01)	0.90 (0.86-0.94)	0.90 (0.86-0.94)	1.37 (1.27-1.48)	2.14 (2.00-2.29)	4.14 (3.88-4.43)	0.72 (0.68-0.77)	0.89 (0.85-0.93)	1.70 (1.62-1.77)	1.00 (0.94-1.07)
FL	0.81 (0.76-0.87)	0.66 (0.61-0.70)	0.50 (0.46-0.53)	1.69 (1.57-1.82)	2.37 (2.21-2.54)	4.10 (3.82-4.40)	0.90 (0.80-1.01)	1.02 (0.97-1.08)	2.01 (1.83-2.22)	0.89 (0.80-1.00)
DLBCL	0.92 (0.88-0.96)	0.81 (0.78-0.84)	0.73 (0.70-0.76)	1.51 (1.44-1.58)	1.99 (1.91-2.08)	3.50 (3.36-3.64)	0.85 (0.80-0.91)	1.02 (0.99-1.05)	1.22 (1.13-1.26)	1.01 (0.96-1.05)
MM/P	0.93 (0.91-0.96)	0.89 (0.86-0.91)	0.78 (0.76-0.81)	1.39 (1.34-1.45)	2.05 (1.97-2.12)	3.17 (3.06-3.29)	0.82 (0.79-0.85)	0.87 (0.85-0.89)	1.15 (1.11-1.19)	0.72 (0.69-0.75)
LBL/L	0.93 (0.87-1.00)	0.85 (0.79-0.91)	0.73 (0.68-0.79)	2.05 (1.90-2.21)	2.71 (2.52-2.92)	3.75 (3.48-4.04)	0.83 (0.74-0.92)	0.94 (0.88-1.01)	1.17 (1.08-1.26)	1.15 (1.04-1.27)
AML	0.99 (0.96-1.02)	0.97 (0.94-1.00)	0.92 (0.89-0.94)	1.78 (1.72-1.84)	2.72 (2.64-2.81)	4.66 (4.52-4.81)	0.88 (0.85-0.91)	0.97 (0.95-1.00)	1.12 (1.08-1.16)	0.99 (0.95-1.03)
APL	0.88 (0.69-1.11)	0.73 (0.57-0.93)	0.73 (0.58-0.93)	2.01 (1.59-2.54)	3.78 (3.03-4.71)	6.27 (4.98-7.91)	0.68 (0.44-1.06)	0.93 (0.76-1.13)	1.97 (1.51-2.58)	1.01 (0.75-1.36)
CML	0.93 (0.88-0.99)	0.70 (0.65-0.74)	0.55 (0.51-0.59)	1.50 (1.39-1.62)	2.59 (2.42-2.78)	4.61 (4.33-4.91)	0.92 (0.83-1.01)	0.88 (0.82-0.93)	1.56 (1.47-1.66)	0.95 (0.87-1.04)
MPN	0.90 (0.83-0.97)	0.83 (0.77-0.90)	0.79 (0.74-0.86)	2.41 (2.13-2.72)	3.93 (3.52-4.40)	6.81 (6.11-7.59)	0.71 (0.55-0.92)	0.97 (0.89-1.06)	1.04 (0.92-1.17)	0.87 (0.77-0.98)
MDS*	1.03 (0.99-1.07)	1.65 (1.46-1.86)	2.00 (1.80-2.22)	2.86 (2.58-3.16)	1.04 (0.93-1.16)	0.85 (0.80-0.89)	1.13 (1.03-1.25)	0.87 (0.81-0.95)

Data are relative excess risk (95% CI). Models were adjusted for each variable included in the table and also for years of follow-up. HL=Hodgkin's lymphoma. CLL/SLL=chronic lymphocytic leukaemia/small lymphocytic lymphoma. FL=follicular lymphoma. DLBCL=diffuse large B-cell lymphoma. MM/P=multiple myeloma/plasmacytoma. LBL/L=precursor lymphoblastic leukaemia/lymphoma. APL=acute promyelocytic leukaemia. AML=acute myeloid leukaemia (excluding APL). CML=chronic myeloid leukaemia. MPN=other myeloproliferative neoplasms (excluding CML). MDS=myelodysplastic syndromes.
*Relative excess risk of death during the 3 years after diagnosis.

Table 5: Relative excess risks of death in the 5 years after diagnosis

promyelocytic leukaemia, and ranged from 3 to 5 for other malignancies (table 5).

Compared with the UK, the excess risk of death was significantly lower in northern, central, and southern Europe for most morphologies. The excess risk of death in eastern Europe was significantly higher for all morphologies, except other myeloproliferative neoplasms (table 5). To try to understand the higher excess risk of death in eastern Europe, we hypothesised that later provision or availability of treatments to patients in eastern Europe might have affected survival. To further investigate this possibility, in the absence of information for the actual treatments given to patients, we used multivariate modelling to compare survival in the EURO CARE population with that in controlled clinical trials. After adjustment for age and region, relative excess risk of death in the 2 years after diagnosis of follicular lymphoma in 2006-08 (vs 1997-99) was 0.6 (95% CI 0.54-0.67). This result is similar to the hazard ratio for mortality (0.6, 95% CI 0.51-0.79) for rituximab-treated patients compared with those treated in the pre-rituximab era, reported by a Cochrane meta-analysis⁶ of randomised trials involving 1480 patients with follicular lymphoma followed up for a median of 24 months (range 18-39). The similar reduction in the risk of death in the population-based and clinical series suggests that the improvement reported in clinical trials was transferred to the entire population.

Discussion

Our data suggest that survival improved for most haematological malignancies over the study period. Multivariate analysis showed a significant decrease in excess risk of death for all malignancies except myelodysplastic syndromes. However, despite this encouraging change, improvements in survival were not uniform across Europe, which could be a result of persisting inequalities in the provision of care. Even after adjustment for age, period, and year of follow-up, survival for most haematological malignancies was significantly lower in eastern Europe, and significantly higher in northern, central, and southern Europe in comparison with the UK (reference region). Additionally, after adjustment for calendar period and region, the excess risk of death increased significantly from the youngest to the oldest ages for all malignancies. Geographical variations in survival across Europe and poor survival in elderly patients have been repeatedly documented for solid tumours, and also for non-Hodgkin lymphoma by a EURO CARE study.¹³ Poor survival in elderly patients is generally attributed to the inability to give potentially curative treatments because patients are frail or have comorbidities. Late diagnosis and under-evaluation of disease symptoms could also play a part; for haematological malignancies treated by stem-cell transplantation, older patients might not be able to

support myeloablation. Finally, most clinical trials do not include older patients or those with low performance status, so treatment protocols are not optimised for the elderly.

Controlled clinical trials^{26,7} have shown that targeted treatments can improve prognoses for chronic myeloid leukaemia, many lymphoid neoplasms, and multiple myeloma. Some population-based studies also show significant improvements in survival for patients with follicular lymphoma²⁰ and diffuse large B-cell lymphoma²¹ when treated with rituximab compared with patients treated by chemotherapy alone. The use of radiotherapy in association with systemic treatments can also improve outcomes.²²

We were unable to find complete information about when drugs were released or when regional or national health services or insurance agencies started providing drug reimbursement. We know that rituximab, imatinib, thalidomide, and bortezomib were first made available for general use in Europe in 1997, 2001, 1998, and 2003, respectively.²³ The years following general release of these drugs coincided with large increases in survival for chronic myeloid leukaemia, diffuse large B-cell lymphoma, and follicular lymphoma; with a smaller but still significant survival increase for multiple myeloma/plasmacytoma. However, the dissemination of these drugs was not uniform across Europe during the years covered by our study. For example, approval of the new drugs occurred earlier and uptake was greater in France than in the UK;²³ survival for all haematological malignancies included in the present study was higher in France than in the UK. Moreover, market uptake of imatinib, rituximab, and bortezomib was lower in eastern Europe than elsewhere.²³ These differences might help explain the consistently lower survival in eastern Europe revealed by our study. Despite our post-hoc analysis showing a similar reduction in the risk of death for follicular lymphoma in population-based and clinical series, suggesting that the improvement reported in clinical trials was transferred to the entire population, results of another study²⁴ show that survival at the population-level for patients with chronic myeloid leukaemia was lower than that for patients enrolled in clinical trials. This difference is not surprising because trial patients generally have more favourable clinical characteristics than the general patient population.

The modest improvement for Hodgkin's lymphoma can be explained by the fact that effective treatments have been available since the early 1970s,²⁵ so in most regions baseline survival was high at the beginning of the study period. The survival increase in older patients with Hodgkin's lymphoma suggests less toxic chemotherapy regimens that can be given to older patients are being adopted; most elderly patients do not receive full-dose treatments because of contraindications or comorbidities. There is also evidence²⁶ that new drug

combinations have reduced the risk of developing iatrogenic myelodysplastic syndromes and acute myeloid leukaemia.

The large between-region variations in survival increase for follicular lymphoma and diffuse large B-cell leukaemia are probably a result of different adoption of rituximab, which is safe and effective in older as well as younger patients.¹² Improved supportive care with better control of comorbidities, especially for older patients, might also have contributed to improved survival for lymphoid malignancies in general.²⁷

Chronic lymphocytic leukaemia/small lymphocytic lymphoma mainly affect older patients and do not always require specific anti-cancer treatment. When treatment is indicated, rituximab or alemtuzumab are combined with chemotherapy. Improved supportive treatment for vascular conditions and infections²⁸ also probably contributed to the survival improvement. Survival in 2003–05 for chronic lymphocytic leukaemia/small lymphocytic lymphoma was lower than that reported for US patients (68.7% vs 75.2%), and only northern Europe had survival similar to the USA.²⁸ Lower survival in Europe could be a result of a lower quality of treatment or restricted availability of new treatments, but variations in diagnostic intensity (eg, in the frequency of blood examination to detect abnormalities, especially in elderly patients) and disease registration practices might also have played a part. Under-reporting of chronic lymphocytic leukaemia/small lymphocytic lymphoma might have resulted in survival underestimation as a result of selective inclusion of cases requiring treatment rather than medical surveillance.²⁹ Survival for multiple myeloma/plasmacytoma was poor (29.8%) in 1997–99 but had increased to 39.6% by 2006–08. This improvement, similar to that reported in the USA,³⁰ is probably caused by the widespread introduction of thalidomide, lenalidomide, and proteasome inhibitors, as adjuncts to chemotherapy and bone-marrow transplantation.⁵

Although the survival improvement over time in precursor lymphoblastic leukaemia/lymphoma was significant, it occurred mainly in 15–54 year olds, with practically no increase in the oldest patients; so for all cases combined, survival was a modest 41% in 2006–08. The disease is rare in adults (2% of lymphoid malignancies in the EURO CARE adult database) and typically occurs in children.³¹ Lack of optimised treatments for what might be an atypical disease in elderly patients could have contributed to poor survival in adults. Furthermore, younger patients might receive higher doses of effective treatments than older patients because they have fewer comorbidities and better health; younger patients might also respond better to cytotoxic treatments than would older patients.

The poor survival for acute myeloid leukaemia (about 15% in 2006–08) and good survival for acute

Panel: Research in context**Systematic review**

Results of EURO CARE population-based studies^{13,31} show that survival for adult Europeans with non-Hodgkin lymphoma improved between 1999 and 2007, and that survival for several haematological malignancies in European children improved over the same period. Other studies^{24,28,30,34} analysed population-based SEER data about US patients and showed improved survival for several haematological malignancies. However, interpretation of survival time trends was made difficult by changes in disease classification. New targeted treatments became available for many haematological malignancies in the early 2000s and clinical trials of selected patients showed that the new treatments were effective.^{1,7} To our knowledge, only two population-based studies^{20,21} specifically linked improved outcomes to the introduction of targeted treatment (rituximab). We therefore did a European-wide population-based study, using the latest EURO CARE data to assess the extent of survival improvements and also using the more detailed updated HAEMACARE criteria for classification of haematological malignancies. To interpret our findings and assess to what extent survival improvements correlated with the introduction and use of new treatments, we examined the results of hospital-based and controlled clinical trials that reported improvements in survival in relation to application of these new treatments.

Interpretation

As far as we are aware, ours is the largest population-based study of survival of European adults with specific haematological malignancies. We found substantial improvements in survival with time for most diseases, especially follicular lymphoma, diffuse large B-cell lymphoma, chronic myeloid leukaemia, and acute promyelocytic leukaemia. These improvements were conspicuous in northern, central, and eastern Europe, less so in the UK and southern Europe, and are plausibly linked to the widespread introduction of targeted and other new treatments. However, persisting geographical differences in risks of death suggest inequalities in the provision of care and availability of the new treatments. Since EURO CARE does not have information about treatments actually given, smaller high-resolution studies, which collect data for stage at diagnosis and treatments for representative samples of cancer registry cases, might suggest reasons for geographical variations in survival, and provide evidence of the effectiveness of new treatments.

promyelocytic leukaemia (about 62% in 2006–08) are consistent with our previous findings for 2000–02.³² Survival for acute promyelocytic leukaemia improved in all regions except southern Europe, where the apparent fall from 2003–05 to 2006–08 is probably an artifact caused by low numbers of cases of the disease. Acute promyelocytic leukaemia occurs mainly in children, and

is unusual in older patients (those aged 65 years and older), which could have contributed to poor survival.

The most striking increase in survival occurred for chronic myeloid leukaemia, mainly starting from 2000–02. The improvement can be attributed to the adoption of imatinib as first-line treatment for the disease.⁴ 5-year relative survival for chronic myeloid leukaemia was higher in the US Surveillance, Epidemiology, and End Results (SEER) database population (around 53% in 1999–2005)³⁴ than for European patients (46% in 2003–05). Better survival in the USA could be caused by a wider availability of effective treatments than in Europe. However, direct comparisons between EURO CARE and SEER are difficult because the data were collected with different protocols; furthermore, information about treatment is not available in either the EURO CARE or the SEER databases. In the USA, imatinib was approved by the US Food and Drug Administration in 2001.³³

To our knowledge this study is the largest population-based cohort of European haematological malignancies analysed (panel), making it possible to assess and compare survival across Europe. We included EURO CARE registries with complete registration from 1992 to 2007 and few cases with not otherwise specified morphology. However, in some countries registration coverage is not national; thus, our findings might not be representative of those countries as a whole.¹³

Another strength of our study is our use of the modified period (hybrid) approach to estimate relative survival, which captures survival improvement earlier than does the traditional cohort approach. However, even hybrid estimates tend to underestimate the survival of recently diagnosed patients when long-term survival is improving.³⁴ Thus, our survival estimates are likely to be somewhat pessimistic for malignancies for which survival is continuing to improve.

The main limitation of our study is the absence of information about treatments, stage at diagnosis, comorbidities, and disease transformations (eg, follicular lymphoma into the more aggressive diffuse large B-cell lymphoma). Thus, although our estimates are robust, evidence linking them to improvements in treatments in general, and the introduction of targeted treatments in particular, is only indirect. Information for treatments, stage at diagnosis, and comorbidities is not available to EURO CARE because it is not uniformly collected by cancer registries. High resolution studies that use clinical records to collect detailed clinical information for representative samples of cancer registry cases can more directly link treatments and clinical characteristics to survival, but are too expensive to do across the whole of Europe.

Other potential causes of improved survival include earlier diagnosis (lead-time bias) as a result of improved diagnostic tests, and changes in disease classification. For example, myelodysplastic syndrome was upgraded

from borderline to malignant in 2000. Adoption of this change could have artificially inflated incidence, and variable timing of adoption by cancer registries would have further complicated the matter. It is for this reason that we only estimated survival trends for myelodysplastic syndrome from 2003–05 to 2006–08.

For follicular lymphoma and diffuse large B-cell lymphoma, we detected an association between survival and incidence (data not shown), suggesting that incidence increased over the study period for these diseases. Specific studies²⁹ report increased incidence for follicular lymphoma and diffuse large B-cell lymphoma but only in elderly patients. It is difficult to establish whether this increase is a real change in disease incidence, or an artificial inflation caused by the inclusion of (formerly excluded) benign lesions.

Overall, we found that survival for almost all haematological malignancies improved over the 11-year study period. Most of these improvements are plausibly linked to the widespread use of new and more effective treatments. Notwithstanding this encouraging and pervasive positive trend, survival differences persisted across Europe, just as they did for solid cancers. The improvements we note have probably continued in the years since the study period ended, as newer target treatments have been introduced into clinical practice. The next priority is to determine what has driven these improvements in survival of haematological malignancies. Cancer registries will remain essential for monitoring improvements in cancer survival, while smaller high-resolution studies, which collect data about stage at diagnosis and treatments for representative samples of cancer registry cases, can provide further evidence as to whether the new treatments are mainly responsible for the improved survival.

Contributors

MS drafted the study protocol, designed the study, supervised the entire work, and drafted most of the initial version of the report. She also advised on the analyses and presentation of the results. PM designed the study and drafted the study protocol, did all statistical analyses, and wrote the relevant parts of the report. MS, PM, and RDA prepared the data and did the quality controls. The EURO CARE-5 Working Group revised the study protocol, collected, prepared, and transmitted raw data to the study database, corrected data after quality controls, and checked the results of the analyses. MS, PM, MMo, LAA, HB, BH, RM-G, AM, MMA, GO-G, OV, RDA, and the EURO CARE-5 Working Group revised the paper. All authors contributed to data interpretation and rewriting the paper, and reviewed and approved the final version.

Declaration of interests

MMA has received personal fees from Roche, outside the submitted work. The other authors declare no competing interests.

Acknowledgments

We thank Chiara Margutti and Simone Bonfarnuzzo for secretarial assistance, and Don Ward for help with the English.

References

- 1 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.
- 2 Plosker GL, Figgitt DP. Rituximab: a review of its use in non-Hodgkin's lymphoma and chronic lymphocytic leukaemia. *Drugs* 2003; **63**: 803–43.
- 3 Thomas DA, O'Brien S, Kantarjian HM. Monoclonal antibody therapy with rituximab for acute lymphoblastic leukemia. *Hematol Oncol Clin North Am* 2009; **23**: 949–71.
- 4 O'Brien SG, Guilhot F, Larson RA, et al. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med* 2003; **348**: 994–1004.
- 5 Armand JP, Burnett AK, Drach J, Harousseau JL, Löwenberg B, San Miguel J. The emerging role of targeted therapy for hematologic malignancies: update on bortezomib and tipifarnib. *Oncologist* 2007; **12**: 281–90.
- 6 Schulz H, Bohlius J, Skoetz N, et al. Chemotherapy plus rituximab versus chemotherapy alone for B-cell non-Hodgkin's lymphoma. *Cochrane Database Syst Rev* 2007; **4**: CD003805.
- 7 Druker BJ, Guilhot F, O'Brien SG, et al. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *N Engl J Med* 2006; **355**: 2408–17.
- 8 Monnereau A, Troussard X, Belot A, et al. Unbiased estimates of long-term net survival of hematological malignancy patients detailed by major subtypes in France. *Int J Cancer* 2013; **132**: 2378–87.
- 9 EURO CARE Survival of cancer patients in Europe. <http://www.eurocare.it> (accessed May 5, 2014).
- 10 HAEMACARE Working Group. Manual for coding and reporting Haematological Malignancies. *Tumori* 2010; **96**: i-A32.
- 11 Fritz A, Percy C, Jack A, et al. International classification of disease for oncology (ICD-O), 3rd edn. Geneva: World Health Organization, 2000.
- 12 Swerdlow SH, Campo E, Harris NL, et al. WHO classification of tumours of haematopoietic and lymphoid tissues, 4th edn (WHO Classification of Tumours, Volume 2). Geneva: World Health Organization, 2008.
- 13 De Angelis R, Sant M, Coleman MP, et al. Cancer survival in Europe 1999–2007 by country and age: results of EURO CARE-5—a population-based study. *Lancet Oncol* 2014; **15**: 23–34.
- 14 Sant M, Allemani C, Tereanu C, et al. Incidence of hematologic malignancies in Europe by morphologic sub-type: results of the HAEMACARE Project. *Blood* 2010; **116**: 3724–34.
- 15 Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med* 2004; **23**: 51–64.
- 16 Brenner H, Rachet B. Hybrid analysis for up-to-date long-term survival rates in cancer registries with delayed recording of incident cases. *Eur J Cancer* 2004; **40**: 2494–501.
- 17 Ederer F, Heise H. Instructions to IBM 650 Programmers in processing survival computations. Methodological note No. 10. End Results Evaluation Section. National Cancer Institute, Bethesda, MD; 1959.
- 18 Baili P, Micheli A, Montanari A, Capocaccia R. Comparison of four methods for estimating complete life tables from bridge life tables using mortality data supplied to EURO CARE-3. *Math Popul Stud* 2005; **12**: 183–98.
- 19 Corazzari I, Quinn MJ, Capocaccia R. Standard cancer patient population for age standardizing survival ratios. *Eur J Cancer* 2004; **40**: 2307–16.
- 20 Krause SW, Gerken M, Andreesen R, Hofstädter F, Klinkhammer-Schalke M. Treatment of B cell lymphoma with chemotherapy plus rituximab: a survival benefit can be demonstrated in the routine data of a regional cancer registry. *Ann Hematol* 2012; **91**: 561–70.
- 21 Sehn LH, Donaldson J, Chhanabhai M, et al. Introduction of combined CHOP plus rituximab therapy dramatically improved outcome of diffuse large B-cell lymphoma in British Columbia. *J Clin Oncol* 2005; **23**: 5027–33.
- 22 Phan J, Mazloom A, Medeiros LJ, et al. Benefit of consolidative radiation therapy in patients with diffuse large B-cell lymphoma treated with R-CHOP chemotherapy. *J Clin Oncol* 2010; **28**: 4170–76.
- 23 Wilking N, Jönsson B, Högberg D. Comparator report on patient access to cancer drugs in Europe. http://www.lmi.no/dm_documents/comparator_report_on_patient_access_to_cancer_drugs_in_europe_jan_15_09_wbki.pdf (accessed May 19, 2014).
- 24 Pulte D, Gondos A, Redaniel MT, Brenner H. Survival of patients with chronic myelocytic leukemia: comparisons of estimates from clinical trial settings and population based cancer registries. *Oncologist* 2011; **16**: 663–71.

 Articles

- 25 De Vita Jr, Simon RM, Hubbard SM, et al. Curability of advanced Hodgkin's disease with chemotherapy. Long-term follow-up of MOPP-treated patients at the National Cancer Institute. *Ann Intern Med* 1980; **92**: 587-95.
- 26 Koontz MZ, Horning SJ, Balise R, et al. Risk of therapy-related secondary leukemia in Hodgkin lymphoma: the Stanford University experience over three generations of clinical trials. *J Clin Oncol* 2013; **31**: 592-98.
- 27 Azoulay E, Mokart D, Pène F, et al. Outcomes of critically ill patients with hematologic malignancies: prospective multicenter data from France and Belgium—a groupe de recherche respiratoire en réanimation onco-hématologique study. *Clin Oncol* 2013; **31**: 2810-18.
- 28 Brenner H, Gondos A, Pulte D. Trends in long-term survival of patients with chronic lymphocytic leukemia from the 1980s to the early 21st century. *Blood* 2008; **111**: 4916-21.
- 29 Morton LM, Wang SS, Devesa SS, Hartge P, Weisenburger DD, Linet MS. Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001. *Blood* 2006; **107**: 265-76.
- 30 Pulte D, Gondos A, Brenner H. Improvement in survival of older adults with multiple myeloma: results of an updated period analysis of SEER data. *Oncologist* 2011; **16**: 1600-03.
- 31 Gatta G, Botta L, Rossi S, et al. Childhood cancer survival in Europe 1999-2007: results of EUROCARE-5—a population-based study. *Lancet Oncol* 2014; **15**: 35-47.
- 32 Maynadié M, De Angelis R, Marcos-Gragera R, et al. Survival of European patients diagnosed with myeloid malignancies: a HAEMACARE study. *Haematologica* 2013; **98**: 230-38.
- 33 Cohen MH, Moses ML, Pazdur R. Gleevec[®] for the treatment of chronic myelogenous leukemia: U.S. Food and Drug Administration regulatory mechanisms, accelerated approval, and orphan drug status. *Oncologist* 2002; **7**: 390-92.
- 34 Brenner H, Hakulinen T. Up-to-date long-term survival curves of patients with cancer by period analysis. *J Clin Oncol* 2002; **20**: 826-32.

Références bibliographiques

- Aalen O. Nonparametric Inference for a Family of Counting Processes. *Ann Stat.* 1978;6(4):701-26.
- Abrahamowicz M, MacKenzie TA. Joint estimation of time-dependent and non-linear effects of continuous covariates on survival. *Stat Med.* 2007;26(2):392-408.
- Abramson JS, Zelenetz AD. Recent advances in the treatment of non-Hodgkin's lymphomas. *J Natl Compr Cancer Netw.* 2013;11(5 Suppl):671-5.
- Agopian J, Navarro J-M, Gac A-C, Lecluse Y, Briand M, Grenot P, *et al.* Agricultural pesticide exposure and the molecular connection to lymphomagenesis. *J Exp Med.* 2009; 206(7):1473-83.
- Al-Hamadani M, Go RS. Stage redistribution and survival among Medicare beneficiaries before and after the approval of positron emission tomography scan for non-Hodgkin lymphoma. *Leuk Lymphoma.* 2014;1-3. [Epub ahead of print].
- Allemani C, Weir HK, Carreira H, Harewood R, Spika D, Wang X-S, *et al.* Global surveillance of cancer survival 1995-2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). *Lancet.* 2015;385(9972):977-1010.
- Andersen PK, Vaeth M. Simple parametric and nonparametric models for excess and relative mortality. *Biometrics.* 1989;523-35.
- Armitage JO. Staging non-Hodgkin lymphoma. *CA Cancer J Clin.* 2005;55(6):368-76.
- Bailar JC, Gornik HL. Cancer undefeated. *N Engl J Med.* 1997;336(22):1569-74.
- Becker N, Fortuny J, Alvaro T, Nieters A, Maynadié M, Foretova L, *et al.* Medical history and risk of lymphoma: results of a European case-control study (EPILYMPH). *J Cancer Res Clin Oncol.* 2009;135(8):1099-107.

-
- Belot A, Grosclaude P, Bossard N, Jouglu E, Benhamou E, Delafosse P, *et al.* Cancer incidence and mortality in France over the period 1980-2005. *Rev Dépidémiologie Santé Publique.* 2008;56(3):159-75.
- Belot A, Remontet L, Launoy G, Jooste V, Giorgi R. Competing risk models to estimate the excess mortality and the first recurrent-event hazards. *BMC Med Res Methodol.* 2011;11:78.
- Belot A, Rondeau V, Remontet L, Giorgi R, CENSUR working survival group. A joint frailty model to estimate the recurrence process and the disease-specific mortality process without needing the cause of death. *Stat Med.* 2014;33(18):3147-66.
- Berkson J. The Calculation of Survival Rates, in Carcinoma and Other Malignant Lesions of the Stomach, In Walters W, Gray H.K., and Priestley J.T. (editors) *Carcinoma and aother malignt lesions of the stomach*, pages 467-484. Sanders. 1942.
- Berkson J, Gage RP. Calculation of survival rates for cancer. *Proc Staff Meet Mayo Clin.* 1950;25(11):270-86.
- Binder-Foucard F, Bossard N, Delafosse P, Belot A, Woronoff A-S, Remontet L, *et al.* Cancer incidence and mortality in France over the 1980-2012 period: solid tumors. *Rev Dépidémiologie Santé Publique.* 2014;62(2):95-108.
- Böhmer PE. *Theorie de unabhingige Wahrscheinlichkeit. Rapports, Memoires et Proces-verbaux de Septieme Congres International d'Actuaires.* 1912.
- Bolard P, Quantin C, Esteve J, Faivre J, Abrahamowicz M. Modelling time-dependent hazard ratios in relative survival: application to colon cancer. *J Clin Epidemiol.* 2001;54(10):986-96.
- Breslow NE. Analysis of Survival Data under the Proportional Hazards Model. *Int Stat Rev Rev Int Stat.* 1975;43(1):45-57.
- Buckley JD. Additive and multiplicative models for relative survival rates. *Biometrics.* 1984;40(1):51-62.
- Campo E, Swerdlow SH, Harris NL, Pileri S, Stein H, Jaffe ES. The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. *Blood.* 2011;117(19):5019-32.

-
- Cancer ECTAM of PA. Measurement of Progress Against Cancer. *J Natl Cancer Inst.* 1990;82(10):825-35.
- Casey R, Brennan P, Becker N, Boffetta P, Cocco P, Domingo-Domenech E, *et al.* Influence of familial cancer history on lymphoid neoplasms risk validated in the large European case-control study epilymph. *Eur J Cancer.* 2006;42(15):2570-6.
- Clerc-Urmes I, Grzebyk M, Hedelin G. Net survival estimation with stns. *Stata J.* 2014;14(1):87-102.
- Cocco P, Satta G, Dubois S, Pili C, Pilleri M, Zucca M, *et al.* Lymphoma risk and occupational exposure to pesticides: results of the Epilymph study. *Occup Environ Med.* 2013;70(2):91-8.
- Coiffier B, Haioun C, Ketterer N, Engert A, Tilly H, Ma D, *et al.* Rituximab (Anti-CD20 Monoclonal Antibody) for the Treatment of Patients With Relapsing or Refractory Aggressive Lymphoma: A Multicenter Phase II Study. *Blood.* 1998;92(6):1927-32.
- Coiffier B, Lepage E, Briere J, Herbrecht R, Tilly H, Bouabdallah R, *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(4):235-42.
- Corazziari I, Quinn M, Capocaccia R. Standard cancer patient population for age standardising survival ratios. *Eur J Cancer.* 2004;40(15):2307-16.
- Coiffier B. Treatment of non-Hodgkin's lymphoma: a look over the past decade. *Clin Lymphoma Myeloma.* 2006;7 Suppl 1:S7-13.
- Cox DR. Regression Models and Life-Tables. *J R Stat Soc Ser B Methodol.* 1972;34(2):187-220.
- Dal-Maso L, Franceschi S. Epidemiology of non-Hodgkin lymphomas and other haemolymphopietic neoplasms in people with AIDS. *Lancet Oncol.* 2003;4(2):110-9.
- Danieli C, Remontet L, Bossard N, Roche L, Belot A. Estimating net survival: the importance of allowing for informative censoring. *Stat Med.* 2012;31(8):775-86.

-
- Dickman PW, Gibberd RW, Hakulinen T. Estimating potential savings in cancer deaths by eliminating regional and social class variation in cancer survival in the Nordic countries. *J Epidemiol Community Health*. 1997;51(3):289-98.
- Dickman PW, Auvinen A, Voutilainen ET, Hakulinen T. Measuring social class differences in cancer patient survival: is it necessary to control for social class differences in general population mortality? A Finnish population-based study. *J Epidemiol Community Health*. 1998;52(11):727-34.
- Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med*. 2004;23(1):51-64.
- Dickman PW, Adami H-O. Interpreting trends in cancer patient survival. *J Intern Med*. 2006;260(2):103-17.
- Dickman PW, Lambert PC, Coviello E, Rutherford MJ. Estimating net survival in population-based cancer studies. *Int J Cancer*. 2013;133(2):519-21.
- Ducassou S, Ferlay C, Bergeron C, Girard S, Laureys G, Pacquement H, *et al*. Clinical presentation, evolution, and prognosis of precursor B-cell lymphoblastic lymphoma in trials LMT96, EORTC 58881, and EORTC 58951. *Br J Haematol*. 2011;152(4):441-51.
- Ederer F. and Heise H. Instructions to IBM 650 programmers in processing survival computations, methodological note no.10, end results evaluation section. Technical report, National Cancer Institute, Bethesda MD. 1959.
- Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Natl Cancer Inst Monogr*. 1961;6:101-21.
- Edwards BK, Noone A-M, Mariotto AB, Simard EP, Boscoe FP, Henley SJ, *et al*. Annual Report to the Nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. *Cancer*. 2014;120(9):1290-314.
- Elie C, Landais P, De Rycke Y. A model combining excess and relative mortality for population-based studies. *Stat Med*. 2014;33(2):275-88.

-
- Engholm G, Ferlay J, Christensen N, Bray F, Gjerstorff ML, Klint A, *et al.* NORDCAN--a Nordic tool for cancer information, planning, quality control and research. *Acta Oncol Stockh Swed.* 2010;49(5):725-36.
- Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: elements for further discussion. *Stat Med.* 1990;9(5):529-38.
- EUROCARE Survival of cancer patients in Europe. <http://www.eurocare.it> (accessed Aug 11, 2014).
- Faivre J, Legoux JL, Martin F, Michiels R, Cabanne F, Klepping C. Incidence of large bowel cancer in Côte-d'Or (Burgundy). *Digestion.* 1979;19(4):221-7.
- Federico M, Bellei M, Marcheselli L, Luminari S, Lopez-Guillermo A, Vitolo U, *et al.* Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. *J Clin Oncol.* 2009;27(27):4555-62.
- Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer J Int Cancer.* 2015;136(5):E359-86.
- Flinn IW, van der Jagt R, Kahl BS, Wood P, Hawkins TE, Macdonald D, *et al.* Randomized trial of bendamustine-rituximab or R-CHOP/R-CVP in first-line treatment of indolent NHL or MCL: the BRIGHT study. *Blood.* 2014;123(19):2944-52.
- Forman D, Bray F, Brewster D, Gombe Mbalawa C, Kohler BA, Pineros M, *et al.* *Cancer Incidence in five continents volume X.* 2013.
- Formica V, Norman AR, Cunningham D, Wotherspoon A, Oates J, Chong G. Utility of the Follicular Lymphoma International Prognostic Index and the International Prognostic Index in assessing prognosis and predicting first-line treatment efficacy in follicular lymphoma patients. *Acta Haematol.* 2009;122(4):193-9.
- Freedman A. Follicular lymphoma: 2012 update on diagnosis and management. *Am J Hematol.* 2012;87(10):988-95.

-
- Friedberg JW, Mauch P, Rimza L. Lymphomas: non-Hodgkin's lymphomas. *Cancer Princ Pract Oncol*. 8th Édition. Philadelphia: Devita, V. T., Jr., Lawrence, T. S., & Rosenberg, S. A.; 2008. p. Chapter 51. Section 2: pp. 2098-143.
- Friedberg JW, Taylor MD, Cerhan JR, Flowers CR, Dillon H, Farber CM, *et al*. Follicular Lymphoma in the United States: First Report of the National LymphoCare Study. *J Clin Oncol*. 2009;27(8):1202-8.
- Fritz AG. *International Classification of Diseases for Oncology*. World Health Organ. Third Edition. Geneva; 2001.
- Ghielmini M, Vitolo U, Kimby E, Montoto S, Walewski J, Pfreundschuh M, *et al*. ESMO Guidelines consensus conference on malignant lymphoma 2011 part 1: diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL) and chronic lymphocytic leukemia (CLL). *Ann Oncol Off J Eur Soc Med Oncol ESMO*. 2013;24(3):561-76.
- Gillaizeau F, Dantan E, Giral M, Foucher Y. A multistate additive relative survival semi-Markov model. *Stat Med*. 2015. pii: 0962280215586456. [Epub ahead of print].
- Giorgi R, Abrahamowicz M, Quantin C, Bolard P, Esteve J, Gouvernet J, *et al*. A relative survival regression model using B-spline functions to model non-proportional hazards. *Stat Med*. 2003;22(17):2767-84.
- Grafféo N, Jooste V, Giorgi R. The impact of additional life-table variables on excess mortality estimates. *Stat Med*. 2012;31(30):4219-30.
- Grafféo N. *Méthodes d'analyse de la survie nette: utilisation des tables de mortalité, test de comparaison et détection d'agrégats spatiaux*. Thèse de doctorat d'Université. Aix-Marseille Université. 2014.
- Grech G, Zhan X, Yoo BC, Bubnov R, Hagan S, Danesi R, *et al*. EPMA position paper in cancer: current overview and future perspectives. *EPMA J*. 2015;6(1):9.
- Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics*. 1982;38(4):933-42.
- Hakulinen T, Tenkanen L. Regression Analysis of Relative Survival Rates. *J R Stat Soc Ser C Appl Stat*. 1987;36(3):309-17.

-
- Harris NL, Jaffe ES, Stein H, Banks PM, Chan JK, Cleary ML, *et al.* A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood*. 1994;84(5):1361-92.
- Hernandez-Ilizaliturri FJ, Gowda A, Czuczman MS. Development of targeted therapies for B-cell non-Hodgkin lymphoma and multiple myeloma. *Clin Adv Hematol Oncol HO*. 2004;2(9):606-18.
- Hiddemann W, Kneba M, Dreyling M, Schmitz N, Lengfelder E, Schmits R, *et al.* Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. *Blood*. 2005;106(12):3725-32.
- Hill C, Laplanche A, Rezvani A. Comparison of the mortality of a cohort with the mortality of a reference population in a prognostic study. *Stat Med*. 1985;4(3):295-302.
- Hodgkin T. On some morbid appearances of the absorbent glands and spleen. *Med Chir Soc Tr*. 1832.
- Huszti E, Abrahamowicz M, Alioum A, Binquet C, Quantin C. Relative survival multistate Markov model. *Stat Med*. 2012;31(3):269-86.
- ICMJE. <http://www.icmje.org/> . (accessed Jun 5, 2015).
- International Agency for Research on Cancer. World Cancer Report 2014. Stewart B.W et Wild C.P; 2014.
- Jaffe ES, Harris NL, Stein H. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. IARC. Press: Lyon: Vardiman J.W., 2001
- Jensen RE, Arora NK, Bellizzi KM, Rowland JH, Hamilton AS, Aziz NM, *et al.* Health-related quality of life among survivors of aggressive non-Hodgkin lymphoma. *Cancer*. 2013;119(3):672-80.

-
- Jooste V, Grosclaude P, Remontet L, Launoy G, Baldi I, Molinié F, *et al.* Unbiased estimates of long-term net survival of solid cancers in France. *Int J Cancer J Int Cancer.* 2013;132(10):2370-7.
- Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data.* John Wiley & Sons; 2011.
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 1958;53:457-81.
- Karlin L, Coiffier B. Improving survival and preventing recurrence of diffuse large B-cell lymphoma in younger patients: current strategies and future directions. *OncoTargets Ther.* 2013;6:289-96.
- Le Gouill S, De Guibert S, Planche L, Brice P, Dupuis J, Cartron G, *et al.* Impact of the use of autologous stem cell transplantation at first relapse both in naïve and previously rituximab exposed follicular lymphoma patients treated in the GELA/GOELAMS FL2000 study. *Haematologica.* 2011;96(8):1128-35.
- Lennert K, Mohri N, Stein H, Kaiserling E. *The Histopathology of Malignant Lymphoma.* *Br J Haematol.* 1975;31:193-203.
- Lukes RJ, Collins RD. Immunologic characterization of human malignant lymphomas. *Cancer.* 1974;34(S8):1488-503.
- Mahboubi A, Abrahamowicz M, Giorgi R, Biquet C, Bonithon-Kopp C, Quantin C. Flexible modeling of the effects of continuous prognostic factors in relative survival. *Stat Med.* 2011;30(12):1351-65.
- Marcos-Gragera R, Allemani C, Tereanu C, De Angelis R, Capocaccia R, Maynadie M, *et al.* Survival of European patients diagnosed with lymphoid neoplasms in 2000-2002: results of the HAEMACARE project. *Haematologica.* 2011;96(5):720-8.
- Matsuda T, Saika K. Worldwide burden of cancer incidence in 2002 extrapolated from cancer incidence in five continents Vol. IX. *Jpn J Clin Oncol.* 2012;42(11):1111-2.
- Mazroui Y, Mathoulin-Pélissier S, Macgrogan G, Brouste V, Rondeau V. Multivariate frailty models for two types of recurrent events with a dependent terminal event: application to breast cancer data. *Biom J Biom Z.* 2013;55(6):866-84.

-
- McNamara C, Davies J, Dyer M, Hoskin P, Illidge T, Lyttelton M, *et al.* Guidelines on the investigation and management of follicular lymphoma. *Br J Haematol.* 2012;156(4):446-67.
- Monnereau A, Remontet L, Maynadié M, Binder-Foucard F, Belot A, Troussard X, *et al.* Estimation nationale de l'incidence des cancers en France entre 1980 et 2012. Partie2 - Hémopathies malignes. Saint-Maurice (Fra): Institut de veille sanitaire. 2013a.
- Monnereau A, Troussard X, Belot A, Guizard A-V, Woronoff A-S, Bara S, *et al.* Unbiased estimates of long-term net survival of hematological malignancy patients detailed by major subtypes in France. *Int J Cancer.* 2013b;132(10):2378-87.
- Morton LM, Turner JJ, Cerhan JR, Linet MS, Treseler PA, Clarke CA, *et al.* Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). *Blood.* 2007;110(2):695-708.
- Morton LM, Hartge P, Holford TR, Holly EA, Chiu BCH, Vineis P, *et al.* Cigarette smoking and risk of non-Hodgkin lymphoma: a pooled analysis from the International Lymphoma Epidemiology Consortium (interlymph). *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol.* 2005;14(4):925-33.
- Morton LM, Slager SL, Cerhan JR, Wang SS, Vajdic CM, Skibola CF, *et al.* Etiologic heterogeneity among non-Hodgkin lymphoma subtypes: the InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr.* 2014;2014(48):130-44.
- Mudie NY, Swerdlow AJ, Higgins CD, Smith P, Qiao Z, Hancock BW, *et al.* Risk of Second Malignancy After Non-Hodgkin's Lymphoma: A British Cohort Study. *J Clin Oncol.* 2006;24(10):1568-74.
- Nelson W. Theory and Applications of Hazard Plotting for Censored Failure Data. *Technometrics.* 1972;42(1):12-25.
- Nelson CP, Lambert PC, Squire IB, Jones DR. Flexible parametric models for relative survival, with application in coronary heart disease. *Stat Med.* 2007;26(30):5486-98.
- Parkin DM. The evolution of the population-based cancer registry. *Nat Rev Cancer.* 2006;6(8):603-12.

-
- Percy C, Stanek E, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. *Am J Public Health*. 1981;71(3):242-50.
- Percy C, Van Holten V, Muir C. *International Classification of Diseases for Oncology*. World Health Organ. Second Edition. Geneva; 1990.
- Perme MP, Stare J, Estève J. On estimation in relative survival. *Biometrics*. 2012;68(1):113-20.
- Peveling-Oberhag J, Arcaini L, Hansmann M-L, Zeuzem S. Hepatitis C-associated B-cell non-Hodgkin lymphomas. Epidemiology, molecular signature and clinical management. *J Hepatol*. 2013;59(1):169-77.
- Plan Cancer. <http://www.e-cancer.fr/le-plan-cancer>. (accessed Jan 29, 2015).
- Pohar M, Stare J. Relative survival analysis in R. *Comput Methods Programs Biomed*. 2006;81(3):272-8.
- Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med*. 2007;26(11):2389-430.
- Quantin C, Abrahamowicz M, Moreau T, Bartlett G, MacKenzie T, Tazi MA, *et al*. Variation over time of the effects of prognostic factors in a population-based study of colon cancer: comparison of statistical models. *Am J Epidemiol*. 1999;150(11):1188-200.
- Rappaport, H. Tumors of the hematopoietic system. 1966;97-161.
- Remontet L, Bossard N, Belot A, Estève J, French network of cancer registries FRANCIM. An overall strategy based on regression models to estimate relative survival and model the effects of prognostic factors in cancer survival studies. *Stat Med*. 2007;26(10):2214-28.
- Robb-Smith AHT. U.S. National Cancer Institute Working formulation of Non-Hodgkin's Lymphomas for clinical use. *The Lancet*. 1982;320(8295):432-4.
- Robins JM. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. Alexandria, Virginia, U.S.; 1993. p. 24-33.

-
- Roche L, Danieli C, Belot A, Grosclaude P, Bouvier A-M, Velten M, *et al.* Cancer net survival on registry data: use of the new unbiased Pohar-Perme estimator and magnitude of the bias with the classical methods. *Int J Cancer J Int Cancer.* 2013;132(10):2359-69.
- Rosso S, De Angelis R, Ciccolallo L, Carrani E, Soerjomataram I, Grande E, *et al.* Multiple tumours in survival estimates. *Eur J Cancer.* 2009;45(6):1080-94.
- Royston P, Parmar MKB. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med.* 2002;21(15):2175-97.
- Sacchi S, Marcheselli L, Bari A, Marcheselli R, Pozzi S, Luminari S, *et al.* Secondary malignancies after treatment for indolent non-Hodgkin's lymphoma: a 16-year follow-up study. *Haematologica.* 2008;93(3):398-404.
- Salles G, Mounier N, de Guibert S, Morschhauser F, Doyen C, Rossi J-F, *et al.* Rituximab combined with chemotherapy and interferon in follicular lymphoma patients: results of the GELA-GOELAMS FL2000 study. *Blood.* 2008;112(13):4824-31.
- Salles G, Seymour JF, Offner F, López-Guillermo A, Belada D, Xerri L, *et 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 Lond Engl.* 1 janv 2011;377(9759):42-51.
- Sant M, Allemani C, De Angelis R, Carbone A, de Sanjosè S, Gianni AM, *et al.* Influence of morphology on survival for non-Hodgkin lymphoma in Europe and the United States. *Eur J Cancer.* 2008;44(4):579-87.
- Sant M, Allemani C, Santaquilani M, Knijn A, Marchesi F, Capocaccia R. EURO CARE-4. Survival of cancer patients diagnosed in 1995-1999. Results and commentary. *Eur J Cancer.* 2009;45(6):931-91.
- Sant M, Allemani C, Tereanu C, De Angelis R, Capocaccia R, Visser O, *et al.* Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood.* 2010;116(19):3724-34.

-
- Sant M, Minicozzi P, Mounier M, Anderson LA, Brenner H, Holleczeck B, *et al.* Survival for haematological malignancies in Europe between 1997 and 2008 by region and age: results of EURO CARE-5, a population-based study. *Lancet Oncol.* 2014;15(9):931-42.
- Sattar T, Griffeth LK, Latifi HR, Glass J, Munker R, Lilien DL. PET imaging today: contribution to the initial staging and prognosis of patients with non-Hodgkin's lymphomas. *J La State Med Soc Off Organ La State Med Soc.* 2006;158(4):193-201.
- Satten GA, Datta S, Robins J. Estimating the marginal survival function in the presence of time dependent covariates. *Stat Probab Lett.* 2001;54(4):397-403.
- Schaffar R, Rapiti E, Rachet B, Woods L. Accuracy of cause of death data routinely recorded in a population-based cancer registry: impact on cause-specific survival and validation using the Geneva Cancer Registry. *BMC Cancer.* 2013;13(1):609.
- Schans SAM van de, Gondos A, Spronsen DJ van, Rachtan J, Holleczeck B, Zanetti R, *et al.* Improving Relative Survival, But Large Remaining Differences in Survival for Non-Hodgkin's Lymphoma Across Europe and the United States From 1990 to 2004. *J Clin Oncol.* 2011;29(2):192-9.
- Seppä K, Hakulinen T, Pokhrel A. Choosing the net survival method for cancer survival estimation. *Eur J Cancer.* 2013. doi: 10.1016/j.ejca.2013.09.019. Epub 2013 Oct 29.
- Smith SM. New drugs for the treatment of non-Hodgkin lymphomas. *Chin Clin Oncol.* 2015;4(1):14.
- Solal-Celigny P, Lepage E, Brousse N, Reyes F, Haioun C, Lepage M, *et al.* Recombinant interferon alfa-2b combined with a regimen containing doxorubicin in patients with advanced follicular lymphoma. Groupe d'Etude des Lymphomes de l'Adulte. *N Engl J Med.* 1993;329(22):1608-14.
- Stare J, Henderson R, Pohar M. An individual measure of relative survival. *J R Stat Soc Ser C.* 2005;54(1):115-26.
- Talback M, Dickman PW. Estimating expected survival probabilities for relative survival analysis--exploring the impact of including cancer patient mortality in the calculations. *Eur J Cancer.* 2011;47(17):2626-32.

-
- The International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N Engl J Med.* 1993;329(14):987-94.
- Tilly H, Vitolo U, Walewski J, da Silva MG, Shpilberg O, André M, *et al.* Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol Off J Eur Soc Med Oncol ESMO.* 2012;23 Suppl 7:vii78-82.
- Wagner G. Cancer registration: historical aspects. *IARC Sci Publ.* 1985;(66):3-12.
- Wagner G. History of cancer registration. *IARC Sci Publ.* 1991;(95):3-6.
- Wang SS, Slager SL, Brennan P, Holly EA, De Sanjose S, Bernstein L, *et al.* Family history of hematopoietic malignancies and risk of non-Hodgkin lymphoma (NHL): a pooled analysis of 10 211 cases and 11 905 controls from the International Lymphoma Epidemiology Consortium (InterLymph). *Blood.* 2007;109(8):3479-88.
- Wilks S. Enlargement of the lymphatic glands and spleen (or, Hodgkin's disease) with remarks. *Guys Hosp Rep.* 1856.
- Willett EV, Morton LM, Hartge P, Becker N, Bernstein L, Boffetta P, *et al.* Non-Hodgkin lymphoma and obesity: a pooled analysis from the InterLymph Consortium. *Int J Cancer J Int Cancer.* 2008;122(9):2062-70.
- World Health Organization. International Classification of Diseases for Oncology. First Edition. Geneva; 1976.
- Wynant W, Abrahamowicz M. Impact of the model-building strategy on inference about nonlinear and time-dependent covariate effects in survival analysis. *Stat Med.* 2014;33(19):3318-37.
- Younes A, Thieblemont C, Morschhauser F, Flinn I, Friedberg JW, Amorim S, *et al.* Combination of ibrutinib with rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) for treatment-naïve patients with CD20-positive B-cell non-Hodgkin lymphoma: a non-randomised, phase 1b study. *Lancet Oncol.* 2014;15(9):1019-26.
- Zelenetz AD, Gordon LI, Wierda WG, Abramson JS, Advani RH, Andreadis CB, *et al.* Non-Hodgkin's lymphomas, version 2.2014. *J Natl Compr Cancer Netw.* 2014;12(6):916-46.