Méthodologie statistique pour la prédiction du risque et la construction de score pronostique en transplantation rénale et en oncologie : une pierre angulaire de la médecine de précision

Dewi Vernerey

▶ **To cite this version:**

Dewi Vernerey. Méthodologie statistique pour la prédiction du risque et la construction de score pronostique en transplantation rénale et en oncologie : une pierre angulaire de la médecine de précision. Cancer. Université de Franche-Comté, 2016. Français. NNT : 2016BESA3004 . tel-01659921

HAL Id: tel-01659921
https://theses.hal.science/tel-01659921

Submitted on 9 Dec 2017

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.

UNIVERSITE DE FRANCHE-COMTE Ecole Doctorale n° 554 « Environnements-Santé »

DOCTORAT DE SANTÉ PUBLIQUE

ÉPIDÉMIOLOGIE ET RECHERCHE CLINIQUE

Méthodologie statistique pour la prédiction du risque et la construction de score pronostique en oncologie et en transplantation rénale: une pierre angulaire de la médecine de précision

---

Dewi Vernerey

Thèse codirigée par le Professeur Franck Bonnetain et le Docteur Alexandre Loupy

Présentée et soutenue publiquement le 8 Décembre 2016

Composition du Jury :

Directeur : Pr Franck Bonnetain (Université de Bourgogne Franche-Comté)

Co-directeur: Dr Alexandre Loupy (Université Paris Descartes; Paris V)

Président : Pr Aimery De Gramont (Université Pierre et Marie Curie; Paris VI)

Rapporteur : Pr Simone Mathoulin-Pélissier (Université de Bordeaux)

Rapporteur : Pr Olivier Thaunat (Université Claude Bernard Lyon I)

Examinateur: Pr Xavier Jouven (Université Paris Descartes; Paris V)

Examinateur: Pr Didier Ducloux (Université de Bourgogne Franche-Comté)

Examinateur: Pr Pascal Hammel (Université Paris Diderot; Paris VII)

« La prédiction est un exercice très compliqué, spécialement quand elle concerne le futur. »

Niels Bohr, 1885 - 1962

Physicien danois

Prix Nobel de physique (1922)

Je remercie sincèrement chacun des membres du jury d'avoir accepté de juger cette thèse.
Le Professeur Aimery de Gramont d'avoir accepté de présider le jury, c'est un réel honneur.
Les Professeurs Simone Mathoulin-Pélissier et Olivier Thaunat pour avoir accepté d'examiner et de discuter de cette thèse.
Je suis ravi que les Professeurs Pascal Hammel et Didier Ducloux aient aussi accepté d'évaluer ce travail.
Au Professeur Xavier Jouven, c'est un plaisir de te compter parmi mon jury. Merci pour ta confiance et la grande liberté que tu m'as données lorsque j'étais quotidiennement dans ton unité.
Au Professeur Franck Bonnetain, pour m'avoir fait confiance dans le cadre de la construction de son équipe de recherche à Besançon et aussi pour m'avoir encouragé à réaliser cette thèse. Entre Bison et Bisontin il y a un socle commun ! Merci de me transmettre un peu de ton expertise.
Au Docteur Alexandre Loupy, merci de m'avoir guidé dans ce domaine passionnant de la transplantation et enseigné les fondamentaux de la spécialité, ce qui m'a permis de m'exprimer au mieux. Carmen, je ne peux pas te dissocier de ces remerciements. Merci pour votre confiance, votre enthousiasme et les bons moments passés ensemble. Votre vision de la recherche et de la médecine m'a beaucoup apportée. Longue vie au Paris Transplant Group !
A Olivier Aubert, pour toutes ces soirées de travail acharné, comme tu le dis de Paris à San Francisco en passant par Besançon et même… Pouilley-les-Vignes ! C'est toujours un plaisir de travailler avec toi.
Je tiens aussi à remercier mes collègues de l'Unité de Méthodologie et de Qualité de Vie en Cancérologie du CHRU de Besançon. Mes remerciements s'adressent en particulier à ma camarade Sophie pour sa bienveillance, ainsi qu'à Julie pour sa bonne humeur quotidienne. Merci à Amélie de m'avoir évité de sombrer dans la fragilité dans la dernière ligne droite. A mes supers étudiants Angélique, Marie-Justine et Edouard pour leur enthousiasme, sympathie et pour m'avoir transmis leurs connaissances médicales, c'était un vrai échange scientifique, clinique et humain ça fait du bien.

J’exprime ma reconnaissance aux membres de l’équipe d’oncologie médicale du CHRU de Besançon, avec une attention particulière à Marine pour ta gentillesse et tes compétences ; à Christophe, c’est toujours très agréable de travailler avec quelqu’un d’aussi impliqué et performant dans sa spécialité, ainsi qu’à Olivier Adotevi pour son soutien, ses nombreuses visites et son optimisme légendaire !
Merci aussi à toute l’équipe de l’unité INSERM U970 du PARCC ainsi qu’à tous les membres des unités de recherche dans lesquelles je me suis formé à l’épidémiologie et ouvert l’esprit à la recherche lors des 10 années de mon aventure parisienne.

A ma famille et amis :
A mes grands-parents disparus trop tôt mais qui m’ont transmis leurs valeurs humaines et de travail.
A mon oncle Gérard pour sa disponibilité et son hospitalité ; rien ne vaut un bon moment tous ensemble à Grandfontaine sur Creuse ; à très vite donc !
A mes parents, merci de m’avoir toujours encouragé et soutenu. Votre parcours a toujours été un exemple pour moi et le sera toujours.
A mes fidèles amis, qui m’accompagnent depuis le début, Anthony, Franck, Fredo, Thomas, Jojo, Goss, Vilette, Jacques et tous les autres, merci pour les soirées risk (décidemment un sujet pas choisi au hasard), poker, les tennis, fous rires et beaucoup d’autres choses que je ne peux pas écrire sur une thèse !

J’exprime ma profonde gratitude à ma femme Carine et ma fille Olivia qui ont souvent dû composer sans moi. Je vous aime et vous dédie ce travail.

Je remercie enfin toutes les personnes que je n’ai pas citées mais qui m’ont soutenues et qui comptent tout autant.

Liste des publications relatives à la thèse

Oncologie

Publications principales:

Vernerey D, Huguet F, Vienot A, Goldstein D, Paget-Bailly S, Van Laethem J-L, Glimelius B, Artru P, Moore MJ, André T, Mineur L, Chibaudel B, Benetkiewicz M, Louvet C, Hammel P, Bonnetain F
Prognostic nomogram and score to predict overall survival in locally advanced untreated pancreatic cancer (PROLAP). **British Journal of Cancer** 115: 281–289, doi:10.1038/bjc.2016.212. (2016)
Impact factor: 5.6

Jary M, Lecomte T, Bouché O, Kim S, Dobi E, Queiroz L, Ghiringhelli F, Etienne H, Léger J, Godet Y, Balland J, Lakkis Z, Adotevi O, Bonnetain F, Borg C, Vernerey D
Prognostic value of baseline seric Syndecan-1 in initially unresectable metastatic colorectal cancer patients: a simple biological score. **International Journal of Cancer** doi:10.1002/ijc.30367. (2016)
Impact factor: 5.5

Vienot A, Beinse G, Louvet C, de Mestier L, Meurisse A, Fein F, Heyd B, Cleau D, d' Engremont C, Dupont-Gossart A-C, Lakkis Z, Tournigand C, Bouché O, Rousseau B, Neuzillet C, Bonnetain F, Borg C**,** Vernerey D
Development and validation of a prognostic nomogram and score to predict overall survival and usefulness of second-line chemotherapy in advanced pancreatic adenocarcinoma. **Journal of the National Cancer Institute in review**.

Publications annexes:

Fiteni F*, Vernerey D*, Bonnetain F, Vaylet F, Sennélart H, Trédaniel J, Moro-Sibilot D, Herman D, Laizé H, Masson P, Derollez M, Clément-Duchêne C, Milleron B, Morin F, Zalcman G, Quoix E, Westeel V (2016) Prognostic value of health-related quality of life for overall survival in elderly non-small-cell lung cancer patients. **European Journal Cancer** Oxf Engl 1990 52: 120–128, doi:10.1016/j.ejca.2015.10.004 (2016)
*Co first authorship
Impact factor: 6.2

Transplantation rénale

Publication principale:

Aubert O*, Kamar N*, Vernerey D*, Viglietti D, Martinez F, Duong-Van-Huyen J-P, Eladari D, Empana J-P, Rabant M, Verine J, Rostaing L, Congy N, Guilbeau-Frugier C, Mourad G, Garrigue V, Morelon E, Giral M, Kessler M, Ladrière M, Delahousse M, Glotz D, Legendre C, Jouven X, Lefaucheur C, Loupy A
Long term outcomes of transplantation using kidneys from expanded criteria donors: prospective, population based cohort study. **British Medical Journal** 351: h3557. (2015)
*Co first authorship
Impact factor: 20.0

Publication annexe:

Loupy A*, Vernerey D*, Viglietti D*, Aubert O, Duong Van Huyen J-P, Empana J-P, Bruneval P, Glotz D, Legendre C, Jouven X, Lefaucheur C (2015) Determinants and Outcomes of Accelerated Arteriosclerosis: Major Impact of Circulating Antibodies. **Circulation Research** 117: 470–482, doi:10.1161/CIRCRESAHA.117.306340. (2015)
*Co first authorship
Impact factor: 11.6

Publications complémentaires à la thèse servant de support pour la partie méthode:

Lefaucheur C, Loupy A, Vernerey D, Duong-Van-Huyen J-P, Suberbielle C, Anglicheau D, Vérine J, Beuscart T, Nochy D, Bruneval P, Charron D, Delahousse M, Empana J-P, Hill GS, Glotz D, Legendre C, Jouven X
Antibody-mediated vascular rejection of kidney allografts: a population-based study. **Lancet** 381: 313–319, doi:10.1016/S0140-6736(12)61265-3. (2013)
Impact factor: 44.0

Loupy A, Lefaucheur C, Vernerey D, Prugger C, Duong van Huyen J-P, Mooney N, Suberbielle C, Frémeaux-Bacchi V, Méjean A, Desgrandchamps F, Anglicheau D, Nochy D, Charron D, Empana J-P, Delahousse M, Legendre C, Glotz D, Hill GS, Zeevi A, Jouven X
Complement-binding anti-HLA antibodies and kidney-allograft survival. **New England Journal of Medicine** 369: 1215–1226, doi:10.1056/NEJMoa1302506. (2013)
Impact factor: 54.4

Loupy A, Lefaucheur C, Vernerey D, Chang J, Hidalgo LG, Beuscart T, Verine J, Aubert O, Dubleumortier S, Duong van Huyen J-P, Jouven X, Glotz D, Legendre C, Halloran PF
Molecular Microscope Strategy to Improve Risk Stratification in Early Antibody-Mediated Kidney Allograft Rejection. **Journal of the American Society of Nephrology** doi:10.1681/ASN.2013111149. (2014)
Impact factor: 9.3

Principales valorisations scientifiques

Communications orales internationales majeures

Vernerey D, Huguet F, Vienot A, Goldstein D, Paget-Bailly S, Van Laethem J-L, Glimelius B, Artru P, Moore MJ, André T, Mineur L, Chibaudel B, Benetkiewicz M, Louvet C, Hammel P, Bonnetain F
Prognostic nomogram and score to predict overall survival in locally advanced untreated pancreatic cancer (PROLAP).

- **Oral Communication** at the Annual meeting of the French Pancreatic Club in 2014, Paris, France.

- **Oral Communication** at the Annual meeting of the European Pancreatic Club in 2015, Toledo, Spain.

- **Oral Communication** at the American Society of Clinical Oncology Gastrointestinal Cancers Symposium in 2015, San Francisco, United States of America

- **Oral Communication** at the European Society for Medical Oncology World Congress on Gastrointestinal Cancer in 2014, Barcelona, Spain.

- **Poster discussion** at the American Society of Clinical Oncology Annual meeting in 2014, Chicago, United States of America**.**

- **Selection for poster of distinction** at the American Pancreatic Association in 2015, San Diego, United States of America.

Récompenses

Vernerey D, Huguet F, Vienot A, Goldstein D, Paget-Bailly S, Van Laethem J-L, Glimelius B, Artru P, Moore MJ, André T, Mineur L, Chibaudel B, Benetkiewicz M, Louvet C, Hammel P, Bonnetain F
Prognostic nomogram and score to predict overall survival in locally advanced untreated pancreatic cancer (PROLAP).

- **Best Abstract in Clinical Science** at the Annual meeting of the French Pancreatic Club in 2014, Paris, France

- **Best Abstract in Clinical Science** at the Annual meeting of the European Pancreatic Club in 2015, Toledo, Spain.

Jary M, Lecomte T, Bouché O, Kim S, Dobi E, Queiroz L, Ghiringhelli F, Etienne H, Léger J, Godet Y, Balland J, Lakkis Z, Adotevi O, Bonnetain F, Borg C, Vernerey D
Prognostic value of baseline seric Syndecan-1 in initially unresectable metastatic colorectal cancer patients: a simple biological score

- **Best Abstract in Colorectal Cancer** at the JFHOD annual congress in 2016, Paris, France

Sollicitations

Symposium CELGENE aux Journées Francophone d'Hépato-gastroentérologie et d'Oncologie Digestive, 2015, Paris, France
Le cancer du pancréas métastatique : quelle prise en charge demain ?
Revue de la littérature et des preuves scientifiques disponibles dans la prise en charge de la 1ère ligne du cancer du pancréas métastatique : le point du vue du clinicien / le point de vue du statisticien. P Artru et D Vernerey.

Demande de rédaction d'un article à destination des professionnels spécialisés en Oncologie dans la revue ONKO+ : « Estimation de la survie globale chez des patients atteints d'un cancer pancréatique localement avancé ».

Résumé

Le pronostic est depuis longtemps un concept de base de la médecine. Hippocrate envisageait déjà le pronostic des maladies par l'étude des circonstances passées, l'établissement des faits présents, et enfin la prédiction des phénomènes à venir. Pour lui, tout l'art du pronostic était de savoir interpréter intelligemment ces informations, et ainsi moduler le pronostic en fonction de leur valeur relative. Une recherche à visée pronostique consiste toujours actuellement en l'examen des relations entre un état de santé connu au moment de l'investigation et un évènement futur. L'augmentation de l'espérance de vie implique que de plus en plus de personnes vivent avec une ou plusieurs maladies ou problèmes altérant leur santé. Dans ce contexte, l'étude du pronostic n'a jamais été aussi importante. Cependant, contrairement au domaine des essais cliniques randomisés dans lequel les recommandations CONSORT sont appliquées depuis plus de 20 ans et garantissent une recherche de qualité, la recherche pronostique commence seulement à se doter d'initiatives similaires. En effet, des recommandations TRIPOD ont été élaborées en 2015 et un groupe de travail, PROGRESS, s'est constitué en 2013 au Royaume-Uni et a fait le constat que les recherches a visée pronostique sont réalisées de façon très hétérogènes et malheureusement ne respectent pas toujours des standards de qualité nécessaires pour supporter leurs conclusions et garantir la reproductibilité des résultats.

Ce travail vise à participer à l'amélioration et la standardisation de la recherche pronostique. Le premier objectif a été d'identifier les étapes clefs des recherches à visée pronostique et de proposer une approche méthodologique adéquate ainsi que des développements pour répondre aux problématiques. Le second objectif de cette thèse est de répondre à des problématiques pronostiques en oncologie digestive et en transplantation rénale avec i) le développement et la validation d'outils de prédiction individuelle et populationnelle au diagnostic pour la survie globale chez des patients atteints d'un cancer du pancréas localement avancé et aussi ii) avant l'initiation du traitement de seconde ligne pour les patients atteints d'un cancer du pancréas avancé, iii) l'étude du rôle pronostique du CD138 dans sa forme soluble sur la survie globale chez les patients atteints d'un cancer colorectal métastatique, iv) l'identification des facteurs pronostiques de la survie des reins à critères élargis et isoler des groupes de patients transplantés dont le pronostic des greffons est distinct afin d'optimiser leur attribution.

Dans cette thèse quatre étapes ont été identifiées. Pour chacune d'entre elles une approche méthodologique et une stratégie d'analyse adéquate et rigoureuse ont été proposées. Très schématiquement, la première étape consiste en l'identification et la caractérisation du phénotype de la pathologie à étudier. Dans un second temps, s'en suit la recherche de facteurs pronostiques de l'évènement d'intérêt dans la population initialement considérée. Ensuite, la troisième étape a pour objectif de déterminer la valeur additionnelle apportée par un nouveau facteur pronostique identifié aux facteurs conventionnels et fonctionnels dans la pratique clinique. Enfin, la

quatrième étape identifiée consiste à considérer le spectre de paramètres pronostiques de l'occurrence de l'évènement étudié et développer des outils de prédiction adéquats permettant une stratification du risque d'évènement, à l'échelle individuelle et populationnelle.

Dans le champ de l'oncologie digestive, deux travaux ont mis en évidence une hétérogénéité de survie considérable chez les patients atteints d'un cancer pancréatique, d'une part au stade localement avancé au moment du diagnostic, et d'autre part au stade avancé à l'initiation de la seconde ligne de chimiothérapie. Dans ces deux situations, des outils de prédiction individuelle (nomogramme) et populationnelle (score) ont été proposés (http://www.umqvc.org/en/tool/prolap.html; http://www.umqvc.org/en/tool/proscap.html). Ces travaux vont permettre de guider la construction des futurs essais cliniques en améliorant la sélection des patients ou encore la procédure de randomisation. Les résultats obtenus dans ces études ont aussi des retombées directes en pratique clinique notamment pour les patients atteints d'un cancer pancréatique avancé. Aujourd'hui, dans cette population, il n'existe pas de consensus sur la sélection des patients qui tirent bénéfice de l'administration d'une seconde ligne de chimiothérapie. L'outil pronostique développé dans cette thèse arbore un excellent pouvoir prédictif du bénéfice en survie globale de l'administration de la seconde ligne de chimiothérapie et a ainsi permis de répondre à une question clinique importante. La troisième étude proposée a permis de démontrer la valeur pronostique pour la survie globale du Syndecan-1 au diagnostic chez des patients atteints d'un cancer colorectal métastatique initialement non résécable. Sur la base du score pronostique biologique proposé intégrant le dosage sérique du Syndecan-1 et des lactates déshydrogénases, nous avons pu isoler une population conséquente de patients de très mauvais pronostic pour laquelle une intensification de traitement doit être envisagée. Un essai thérapeutique va débuter très prochainement pour évaluer le bien-fondé de cette hypothèse.

Le travail de ce manuscrit de thèse, qui relève du champ de la transplantation rénale, démontre l'impact délétère de la présence des anticorps dirigés contre le donneur circulants, le jour de la transplantation, sur la survie des greffons, dans un contexte de donneurs à critères élargis, et a ainsi permis de valider *a posteriori* la politique de greffe actuelle en France, en encourageant l'utilisation de reins à critères élargis. Cela constitue une solution acceptable en réponse à la pénurie d'organes à condition que celle-ci ne s'adresse pas à des patients immunisés le jour de la transplantation.

En conclusion, une approche méthodologique rigoureuse dans le cadre d'analyses à visée pronostique de cohortes exhaustives et bien caractérisées avec des données de qualité a permis d'apporter des réponses à des questions cliniques pertinentes susceptibles de modifier la prise en charge des patients.

Mots-clés : pronostic, prédiction, oncologie digestive, cancer pancréatique, cancer colorectal, transplantation rénale, anticorps anti-HLA, stratification du risque.

Abstract

Prognosis is historically a basic concept of medicine. Hippocrates already considered the prognosis of disease as the study of the past circumstances, the establishment of the present state of health and finally the prediction of future events. He presented the prognosis as the ability to interpret these elements and to adapt the prognosis regarding their relative values. Currently, the prognostic research is still based on the examination of the relationship between a well-established health condition at the time of the investigation and the occurrence of an event. The increase in life expectancy implies that more and more people are living with one or more diseases or with problems that can impair their health status. In this context, the study of the prognosis has never been more important. However, in comparison with the field of randomized clinical trials in which the CONSORT statement recommendations are implemented for more than 20 years in order to guarantee quality research, the prognostic research only begins to develop similar initiatives. Indeed, in 2015 the TRIPOD statement recommendations were provided and in 2013 a working group called PROGRESS was constituted in the United Kingdom and its members made the observation that prognostic researches are developed with considerable heterogeneity in the methodology used and unfortunately do not always meet the quality standards required to support their conclusions and their reproducibility.

The work provided in this thesis is then in line with the need of improvement and standardization of the prognostic research. Its first objective is to identify the key stages implied in a prognostic research and to review, propose and develop adequate statistical methodology to answer the issues raised in each of them. The second objective of this thesis is to address important prognostic issues in the context of two medical specialities that are digestive oncology and renal transplantation with i) the development and validation of prediction tools for overall survival at diagnosis in patients with locally advanced pancreatic cancer and also ii) at the initiation of second line chemotherapy for patients with an advanced pancreatic cancer, iii) the assessment and validation of the independent prognostic value of baseline serum CD138 for overall survival in patients with metastatic colorectal cancer, iv) the identification of prognostic factors for kidney allograft survival in expanded criteria donor transplant patients and the determination of groups of patients with distinct kidney survival to optimize their attribution.

In this thesis four steps were identified. For each of them a methodological approach and an adequate and rigorous analysis strategy have been proposed. Very briefly, the first step consists in the identification and characterisation of the phenotype of the pathology to be studied. The second step is naturally the search of prognostic factors for the event of interest in the population initially considered. Then, the third step is to determine the additional value of a new prognostic factor identified to a set of conventional factors already known and functional in clinical practice. Finally, the fourth step identified consists in the consideration of the whole spectrum of prognostic

parameters for the occurrence of the event and then in the development of adequate prediction tools based on a final multivariate model allowing stratification of the risk of events, at an individual and population level.

In the gastrointestinal oncology field, two studies have demonstrated a considerable survival heterogeneity, firstly in patients with locally advanced pancreatic cancer at diagnosis and secondly in advanced disease at the second line chemotherapy initiation. In these two settings, individual (nomogram) and population-level (score) prediction tools have been proposed (http://www.umqvc.org/en/tool/prolap.html;http://www.umqvc.org/en/tool/proscap.html). This work will improve the design of future clinical trials in guiding the selection of patients or by enhancing the randomization procedure. The results obtained in these studies have also an impact in clinical practice particularly in the context of the second line chemotherapy since, to date, there is no consensual and validated strategy to select patients with advanced pancreatic cancer eligible for chemotherapy second line. Indeed, the prognosis tool developed in this work features an excellent predictive ability to estimate survival benefit of chemotherapy administration allowing to answer to an important clinical question. The third study proposed demonstrated the additional prognostic value for overall survival of Syndecan-1 at diagnosis in patients with metastatic colon cancer initially unresectable. Based on the suggested biological prognostic score involving Syndecan-1 and lactate dehydrogenase, we highlighted a substantial population with very poor survival prognosis for which treatment intensification should be considered. A randomized therapeutic trial will begin soon to assess the validity of this assumption.

Concerning kidney transplantation this thesis provides a validation *a posteriori* of the actual transplant policy in France with the encouragement of the use of expanded criteria kidneys as an acceptable solution in response to the shortage of organs. Nevertheless, the recognition of the deleterious impact of the presence the day of transplantation of anti-HLA antibodies against the donor circulating in patients transplanted from donors with expended criteria suggests to not address these grafts into patients immunized the day of transplantation.

In conclusion, a rigorous methodological approach in the context of prognostic analyses in exhaustive and well-phenotyped cohorts provides answers to relevant clinical questions and may modify the management of the patients.

Key words: prognosis, prediction, digestive oncology, pancreatic cancer, colorectal cancer, kidney transplantation, anti-HLA antibodies, and risk-stratification.

Laboratoire d’accueil

Unité de recherche EA 3181

Carcinogenèse épithéliale : facteurs prédictifs et pronostiques

Directrice : Pr Christiane Mougin

19 rue Ambroise Paré

UFR SMP

Les Hauts du Chazal

F-25041 Besançon cedex

Liste des abréviations

SORT : Standardized Reporting of Trials

ConSORT : CONsolidated Standards Of Reporting Trials

TRIPOD : Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

PROGRESS : PROGnosis RESearch Strategy Partnership

HLA : Human Leucocyte Antigen

FDA : Food and Drug Administration

PC : Cancer Pancréatique

CRC : Cancer ColoRectal

LAPC : Cancer Pancréatique Localement Avancé

aPC : Cancer Pancréatique avancé

mCRC : Cancer ColoRectal métastatique

ACP : Analyse en Composantes Principales

DSA : Anticorps anti-HLA Spécifiques dirigés contre le Donneur

TCMR : Rejet Cellulaire aigu Médié par les lymphocytes T

ABMR : Rejet humoral aigu Médié par les Anticorps

SG : Survie Globale

OMS : Organisation Mondiale de la Santé

DMCI : Différence Minimale Cliniquement Importante

RECIST : Response Evaluation Criteria In Solid Tumours

DATECAN : Definition for the Assessment of Time-to-event Endpoints in CANcer trials

DIVAT : Données Informatiques Validées en Transplantation

AUC : Area Under the Curve

ROC : Receiver operating characteristic

C-index : Concordance index

NRI : Net Reclassification Improvement

IDI : Integrated Discrimination Index

DCA : Decision Curve Analysis

SSP : Survie Sans Progression

PS : Performance Status

SDC-1 / CD138 : Syndecan-1

ELISA : Enzyme-Linked Immunosorbent Assays

L2 : seconde ligne de chimiothérapie

L1: première ligne de chimiothérapie

ECD : Donneur à Critères Elargis

SCD : Donneurs à Critères Standards

PRA : Panel Reactive Antibody

WHOQOL : World Health Organization Quality of Life

IC : Intervalle de Confiance

QdV : Qualité de Vie relative à la santé

PROs : Patient-Reported Outcomes

Table des matières simplifiée

I Introduction **23**
I.1 La recherche pronostique *23*
I.2 L'oncologie digestive *28*
I.2.1 Le cancer pancréatique 28
I.2.2 Le cancer colorectal 30
I.3 La transplantation rénale *34*

II Objectifs du travail **37**

III Méthodologie **39**
III.1 Caractérisation et identification du phénotype de la pathologie à étudier *41*
III.1.1 Caractérisation : Choix selon des caractéristiques déjà établies 42
III.1.2 Identification par des techniques de statistique descriptive multidimensionnelle 45
III.2 Recherche de facteurs pronostiques de l'évènement d'intérêt *58*
III.2.1 Choix de l'évènement d'intérêt et de la variable à expliquer 59
III.2.2 L'analyse des données de survie 68
III.3 Détermination de la valeur additionnelle apportée par un nouveau facteur pronostique identifié aux facteurs conventionnels déjà connus et fonctionnels dans la pratique clinique *93*
III.3.1 L'indice de discrimination ou de concordance 95
III.3.2 Les indices de reclassification 104
III.4 Construction d'outils à visée pronostique sur la base de modèles intégratifs *118*
III.4.1 Le modèle de Cox multivarié 119
III.4.2 Outils pronostiques exploitant l'information issue d'un modèle de Cox 136

IV Résultats **157**
IV.1 Axe oncologie *158*
IV.1.1 Développement d'un nomogramme et d'un score pronostique pour prédire la survie globale chez des patients atteints d'un cancer pancréatique localement avancé au diagnostic avant l'initiation du traitement d'induction. 158
IV.1.2 Etude de la valeur pronostique pour la survie globale du Syndecan-1 (CD138) au diagnostic chez des patients atteints d'un cancer colorectal métastatique initialement non résécable. 194
IV.1.3 Développement et validation d'un nomogramme et d'un score pronostique pour prédire la survie globale sous seconde ligne de chimiothérapie chez des patients atteints d'un cancer pancréatique avancé. 223
IV.2 Axe transplantation *282*
IV.2.1 Pronostic des transplantations rénales réalisées à partir de donneurs à critères élargis. 283

V Discussion **305**
V.1 Discussion des travaux de la thèse *305*
V.1.1 Objectif méthodologique 305
V.1.2 Objectif clinique 309
V.1.3 Mise en perspective de l'ensemble des travaux cliniques dans le processus méthodologique proposé 321

V.2 Discussion générale *323*

VI Perspectives du travail de recherche **328**

VI.1 Evaluation et enrichissement de la méthodologie *328*

VI.2 Score pronostique dans la première année post-transplantation rénale de la survie du greffon à long terme *329*

VI.3 Score pronostique le jour de la transplantation rénale de la survie du greffon à long terme *333*

VI.4 Travaux en cours en oncologie digestives *334*

VII Références de la thèse **335**

VIII Annexes **348**

VIII.1 Annexe 1: Publications complémentaires à la thèse servant de support pour la partie méthode : *348*

VIII.1.1 Identification d'une nouvelle forme de rejet aigu en transplantation rénale 348

VIII.1.2 Etude du rôle pronostique des anticorps anti-HLA fixant le complément sur le risque de perte du greffon 367

VIII.1.3 Démonstration de l'intérêt pronostique du microscope moléculaire dans la stratification du risque de perte du greffon 390

VIII.2 Annexe 2: *413*

VIII.3 Annexe 3: *425*

Table des matières détaillée

I Introduction 23

I.1 La recherche pronostique 23

I.2 L'oncologie digestive 28

I.2.1 Le cancer pancréatique 28

I.2.2 Le cancer colorectal 30

I.3 La transplantation rénale 34

II Objectifs du travail 37

III Méthodologie 39

III.1 Caractérisation et identification du phénotype de la pathologie à étudier 41

III.1.1 Caractérisation : Choix selon des caractéristiques déjà établies 42

III.1.2 Identification par des techniques de statistique descriptive multidimensionnelle 45

III.1.2.1 Classification hiérarchique par cluster (dendogramme) et illustration par une « Heat map » 45

III.1.2.2 Analyse en Composantes Principales 50

III.2 Recherche de facteurs pronostiques de l'évènement d'intérêt 58

III.2.1 Choix de l'évènement d'intérêt et de la variable à expliquer 59

III.2.1.1 Evènements centrés sur le patient 59

III.2.1.1.1 Le décès 59

III.2.1.1.2 Les évènements reportés par la patient (Patient Reported Outcome) et en particulier la qualité de vie relative à la santé 60

III.2.1.2 Evènements centrés sur la maladie 61

III.2.1.2.1 La perte du greffon 61

III.2.1.2.2 La progression tumorale 61

III.2.1.3 Critère de jugement du type « temps jusqu'à événement » ou censuré à droite 62

III.2.1.4 Définition des variables du type « temps jusqu'à évènement» utilisées dans cette thèse 66

III.2.1.4.1 En transplantation 66

III.2.1.4.1.1 La survie globale du greffon 66

III.2.1.4.1.2 La survie globale du patient 66

III.2.1.4.2 En oncologie 67

III.2.1.4.2.1 La survie sans progression 67

III.2.1.4.2.2 La survie globale 67

III.2.2 L'analyse des données de survie 68

III.2.2.1 La distribution de survie 68

III.2.2.1.1 Estimation de la distribution de survie 69

III.2.2.1.1.1 Estimation paramétrique 69

III.2.2.1.1.2 Estimation non paramétrique de Kaplan-Meier 69

III.2.2.1.1.3 Comparaison de la distribution de survie de deux ou plusieurs groupes estimés par la méthode non paramétrique de Kaplan-Meier : le test du log-rank 72

III.2.2.1.2 Modélisation de la distribution de survie pour la prise en compte de plusieurs facteurs de risque 74

III.2.2.1.2.1 Modélisation paramétrique 74

III.2.2.1.2.2 Modélisation semi-paramétrique : le modèle de Cox 75

III.2.2.1.2.2.1 Estimation et test des paramètres du modèle de Cox 77
III.2.2.1.2.2.2 Adéquation de l'hypothèse des risques proportionnels pour un modèle de Cox 79
III.2.2.1.2.2.3 Codage des variables à inclure dans le modèle de Cox 80
III.2.2.1.2.2.3.1 Variables catégorielles 81
III.2.2.1.2.2.3.2 Variables quantitatives 81
III.2.2.1.2.2.3.2.1 Approche non paramétrique : méthode des splines cubiques restreintes 84
III.2.2.1.2.2.3.2.2 Approche paramétrique : méthode des polynômes fractionnaires 86
III.2.2.1.2.2.3.3 Considération des valeurs manquantes 88
III.2.2.1.2.2.3.4 Considération de la colinéarité des variables explicatives 90
III.2.2.1.2.2.3.5 Considération du terme d'interaction 92
III.3 Détermination de la valeur additionnelle apportée par un nouveau facteur pronostique identifié aux facteurs conventionnels déjà connus et fonctionnels dans la pratique clinique 93
III.3.1 L'indice de discrimination ou de concordance 95
III.3.1.1 Evaluation de la valeur additionnelle du nouveau facteur pronostique en termes de discrimination. 98
III.3.1.2 Evaluation de la significativité statistique de la valeur additionnelle par procédure de ré-échantillonnage 99
III.3.2 Les indices de reclassification 104
III.3.2.1 Net Reclassification Improvement (NRI) 105
III.3.2.1.1 Le NRI catégoriel 105
III.3.2.1.2 Le NRI continu 111
III.3.2.2 Integrated discrimination improvement (IDI) 114
III.4 Construction d'outils à visée pronostique sur la base de modèles intégratifs 118
III.4.1 Le modèle de Cox multivarié 119
III.4.1.1 Performance du modèle 119
III.4.1.1.1 La discrimination 119
III.4.1.1.2 La calibration 120
III.4.1.1.2.1 Le test de Hosmer-Lemeshow 123
III.4.1.1.2.2 Le graphique de calibration issu du test de Hosmer-Lemeshow 125
III.4.1.1.2.3 La courbe de calibration lissée 127
III.4.1.2 Validation du modèle : 128
III.4.1.2.1 Validation interne du modèle 129
III.4.1.2.2 Validation interne dans un échantillon provenant d'un fractionnement des données 129
III.4.1.2.3 Validation externe du modèle 130
III.4.1.3 Evaluation de l'intérêt clinique du modèle multivarié : l'analyse par courbe de décision 131
III.4.2 Outils pronostiques exploitant l'information issue d'un modèle de Cox 136
III.4.2.1 Approche individuelle : le nomogramme 136
III.4.2.1.1 Estimation de la probabilité de survie sans évènement à un temps donné 142
III.4.2.1.2 Estimation de la médiane de survie pour le patient j 143
III.4.2.1.3 Intervalle de confiance pour les estimations de survie 144
III.4.2.1.4 Points du Nomogramme 146
III.4.2.2 Approche populationnelle : Construction d'un score pronostique 147
III.4.2.2.1 Identification de populations de pronostics différents 147
III.4.2.2.2 Choix du nombre de groupes à risque à identifier 148

III.4.2.2.3 Choix des seuils du score permettant de définir les populations 149
III.4.2.2.4 Possibilité d'évaluation de l'interaction entre les prédictions de survie et le traitement 155

IV Résultats 157
IV.1 Axe oncologie 158
IV.1.1 Développement d'un nomogramme et d'un score pronostique pour prédire la survie globale chez des patients atteints d'un cancer pancréatique localement avancé au diagnostic avant l'initiation du traitement d'induction. 158
IV.1.1.1 Positionnement du problème 158
IV.1.1.2 Population de l'étude 160
IV.1.1.3 Approche pronostique méthodologique 161
IV.1.1.4 Résultats 163
IV.1.1.5 Conclusion de l'étude 175
IV.1.1.6 Manuscrit scientifique 175
IV.1.1.7 Manuscrit à destination des professionnels spécialisés en Oncologie dans la revue ONKO+ 185
IV.1.2 Etude de la valeur pronostique pour la survie globale du Syndecan-1 (CD138) au diagnostic chez des patients atteints d'un cancer colorectal métastatique initialement non résécable. 194
IV.1.2.1 Positionnement du problème 194
IV.1.2.2 Population de l'étude 196
IV.1.2.3 Approche pronostique méthodologique 197
IV.1.2.4 Résultats 199
IV.1.2.5 Conclusion de l'étude 210
IV.1.2.6 Manuscrit 210
IV.1.3 Développement et validation d'un nomogramme et d'un score pronostique pour prédire la survie globale sous seconde ligne de chimiothérapie chez des patients atteints d'un cancer pancréatique avancé. 223
IV.1.3.1 Positionnement du problème 223
IV.1.3.2 Population de l'étude 225
IV.1.3.3 Approche pronostique méthodologique 225
IV.1.3.4 Résultats 228
IV.1.3.5 Conclusion de l'étude 249
IV.1.3.6 Manuscrit 250
IV.2 Axe transplantation 282
IV.2.1 Pronostic des transplantations rénales réalisées à partir de donneurs à critères élargis. 283
IV.2.1.1 Positionnement du problème 283
IV.2.1.2 Population de l'étude 285
IV.2.1.3 Données 285
IV.2.1.4 Approche méthodologique 286
IV.2.1.5 Résultats 287
IV.2.1.6 Conclusions 295
IV.2.1.7 Manuscrit 295

V Discussion 305
V.1 Discussion des travaux de la thèse 305
V.1.1 Objectif méthodologique 305
V.1.2 Objectif clinique 309
V.1.2.1 Travaux d'intérêt de la thèse en oncologie digestive 309

V.1.2.2 Travaux complémentaires à la thèse en transplantation rénale 316
V.1.2.1 Travail d'intérêt de la thèse en transplantation rénale 319
V.1.2.2 Travaux de la thèse présentés en annexes 319
V.1.3 Mise en perspective de l'ensemble des travaux cliniques dans le processus méthodologique proposé 321
V.2 Discussion générale *323*

VI Perspectives du travail de recherche **328**
VI.1 Evaluation et enrichissement de la méthodologie *328*
VI.2 Score pronostique dans la première année post-transplantation rénale de la survie du greffon à long terme *329*
VI.3 Score pronostique le jour de la transplantation rénale de la survie du greffon à long terme *333*
VI.4 Travaux en cours en oncologie digestives *334*

VII Références de la thèse **335**

VIII Annexes **348**
VIII.1 Annexe 1: Publications complémentaires à la thèse servant de support pour la partie méthode : *348*
VIII.1.1 Identification d'une nouvelle forme de rejet aigu en transplantation rénale 348
VIII.1.1.1 Contexte 348
VIII.1.1.2 Population de l'étude 349
VIII.1.1.3 Approche méthodologique 349
VIII.1.1.4 Résultats 350
VIII.1.1.5 Conclusion de l'étude 357
VIII.1.1.6 Manuscrit 358
VIII.1.2 Etude du rôle pronostique des anticorps anti-HLA fixant le complément sur le risque de perte du greffon 367
VIII.1.2.1 Positionnement du problème 367
VIII.1.2.2 Population d'étude 368
VIII.1.2.3 Approche méthodologique 369
VIII.1.2.4 Résultats 370
VIII.1.2.5 Conclusion de l'étude 377
VIII.1.2.6 Manuscrit 377
VIII.1.3 Démonstration de l'intérêt pronostique du microscope moléculaire dans la stratification du risque de perte du greffon 390
VIII.1.3.1 Positionnement du problème 390
VIII.1.3.2 Population de l'étude 391
VIII.1.3.3 Données cliniques 392
VIII.1.3.4 Extraction de RNA et procédure d'analyse en microarray 393
VIII.1.3.5 Approche méthodologique 393
VIII.1.3.6 Résultats 394
VIII.1.3.7 Conclusion de l'étude 401
VIII.1.3.8 Manuscrit 401
VIII.2 Annexe 2: *413*
VIII.3 Annexe 3: *425*

I Introduction

I.1 La recherche pronostique

Hippocrate a inclus le pronostic comme un concept de base de la médecine. Son travail a beaucoup contribué à la description des symptômes, des signes physiques, du traitement chirurgical et du pronostic. En effet, l'école hippocratique était une médecine fondée sur l'observation de l'évolution naturelle des maladies. Hippocrate, dans son ouvrage *Pronostic* (IV^{ème} siècle avant J-C), envisage le pronostic des maladies par l'étude des circonstances passées, l'établissement des faits présents, et enfin la prédiction des phénomènes à venir. Tout l'art du pronostic est de savoir interpréter intelligemment ces informations rassemblées, et ainsi moduler le pronostic en fonction de leur valeur relative. Cette médecine basée sur l'utilisation approfondie systématique de l'interrogatoire et de l'examen clinique parvint à traiter efficacement les maladies et a ainsi permis un grand développement de la pratique clinique (Hippocrate (0460 - 0377 av. J.-C.)).

En médecine, dans un contexte de pratique clinique, l'étude du pronostic consiste en l'estimation, à un temps donné, d'une probabilité pour un individu, avec une maladie diagnostiquée ou non, de développer un état particulier de santé (évènement) en fonction de ses caractéristiques cliniques et non cliniques (Figure). Les évènements sont souvent des évènements spécifiques centrés sur le patient tels que le décès, l'évolution de la douleur ou encore des mesures de qualité de vie relative à la santé (QdV), mais ils peuvent aussi être centrés sur la maladie comme par exemple la mesure de la progression ou la récidive de la maladie.

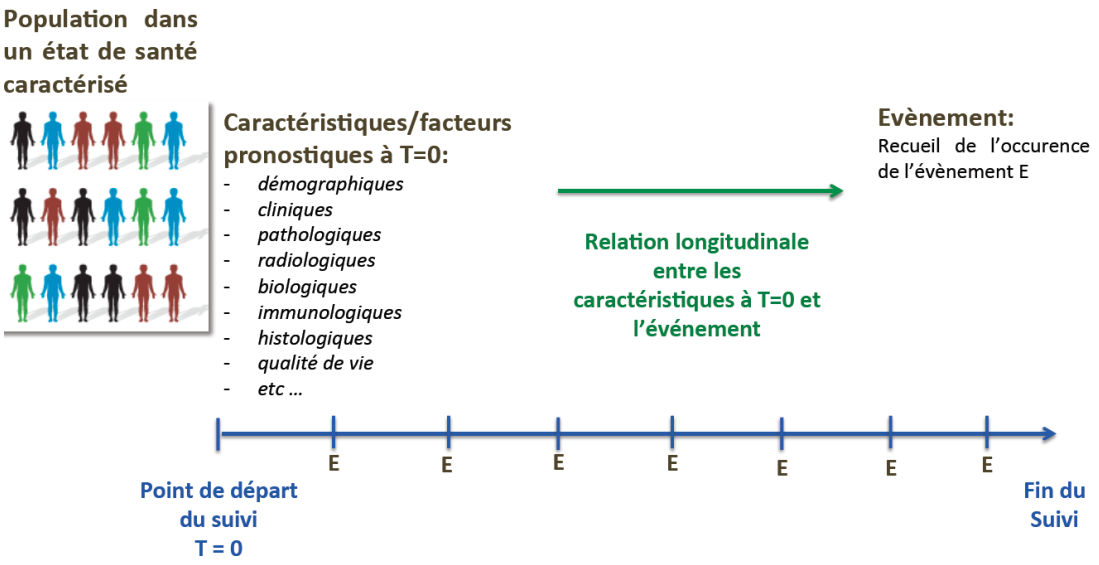


Figure : Représentation du design d'une étude pronostique

Dans la littérature médicale, le pronostic fait fréquemment référence à l'évolution de la santé des malades. Cependant, il est important de noter que le terme pronostic n'est pas limité aux personnes ayant un diagnostic de maladie établi mais peut être utilisé pour des individus en bonne santé. Par exemple, le score de Framingham mesure le risque coronarien à 10 ans chez des sujets sains au moment de l'évaluation.

Aujourd'hui, l'augmentation de l'espérance de vie implique que de plus en plus de personnes vivent avec une ou plusieurs maladies ou problèmes altérant leur santé. Dans ce contexte, l'étude du pronostic n'a jamais été aussi importante. En effet, la recherche pronostique peut fournir des informations fondamentales pour comprendre, expliquer et prédire des issues cliniques précises et ainsi apporter des éléments clefs pour guider les décisions cliniques et les stratégies de recherches.

Schématiquement, la recherche pronostique consiste donc en l'investigation des relations entre un évènement futur et un état de santé connu au moment de l'investigation (randomisation, diagnostic, début de traitement, etc.).

Dans le domaine des essais cliniques randomisés, dès 1993, un groupe d'experts, formé de chercheurs, de méthodologistes et d'éditeurs, s'est réuni pour élaborer les recommandations SORT (Standardized Reporting of Trials) puis en 1996 ConSORT (CONsolidated Standards Of Reporting Trials). Il en résulte une liste de recommandations ayant pour objectif l'amélioration et l'homogénéisation de la publication des essais cliniques, permettant ainsi aux lecteurs de mieux comprendre la conception (Begg *et al*, 1996) (Moher *et al*, 2001) (Moher *et al*, 2010), la conduite, l'analyse statistique et ainsi d'évaluer la validité des résultats. La recherche pronostique quant à elle commence seulement à se doter d'initiatives similaires, d'une part en 2015 avec la sortie des recommandations TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) (Collins *et al*, 2015), et d'autre part en 2013 avec la constitution du groupe de travail appelé PROGRESS (PROGnosis RESearch Strategy Partnership) au Royaume Uni.

Les membres du groupe PROGRESS ont fait dans un premier temps le constat que les recherches a visée pronostique sont réalisées de façon très hétérogène et malheureusement ne respectent pas toujours des standards de qualité nécessaires : design peu rigoureux, biais de sélection de population, analyses statistiques incomplètes et restitution des résultats biaisés…(Hemingway *et al*, 2013)

Aujourd'hui, une recherche pronostique avec des standards de qualité précis est requise pour exporter les résultats des études dans la pratique clinique, afin d'améliorer la prise en charge des patients et/ou ouvrir des pistes de recherche.

Dans ce contexte et dans un second temps, le groupe PROGRESS a identifié quatre axes clefs délimitant le champ d'application de la recherche pronostique.

Le premier domaine d'application appelé *« recherche pronostique fondamentale »* vise à fournir, dans un contexte clinique spécifique, le pronostic moyen des patients, aussi appelé risque de base, offrant ainsi la possibilité de quantifier l'impact et/ou la qualité des soins actuels, et donnant

aussi l'éventualité d'identifier et de hiérarchiser des priorités en termes de stratégie de recherche.

Le second axe est celui de la « *recherche de facteurs pronostiques* » c'est-à-dire la recherche d'une mesure, chez des personnes avec un état de santé donné, qui est associée avec un résultat clinique ultérieur. L'identification de facteurs pronostiques permet d'aider à la caractérisation d'une maladie à un moment T et d'être aussi informatif pour les décisions cliniques et thérapeutiques. Les facteurs pronostiques sont aussi essentiels pour améliorer la conception des essais cliniques et permettent d'identifier des cibles pour de nouvelles interventions et donc pour de nouvelles voies de recherche. Cependant, aujourd'hui il existe des limites majeures dans les recherches visant à identifier de nouveaux facteurs pronostiques qui par ailleurs ne sont pas toujours répliquées par la suite.

S'en suit naturellement la « *recherche de modèle pronostique* », le troisième axe clef de la recherche pronostique. Compte tenu de la variabilité entre les patients et de l'étiologie des maladies, un simple prédicteur procure rarement une estimation adéquate et précise du pronostic des patients. Un modèle pronostique est, en effet, la considération conjointe de plusieurs facteurs pronostiques pour prédire précisément et individuellement un futur évènement clinique chez les patients. Cela permet ainsi de définir le spectre des caractéristiques associées à l'évènement et donc potentiellement d'améliorer l'information et la prise en charge des patients. Sur la base d'un modèle pronostique multivarié, des outils peuvent être proposés pour estimer des probabilités de faire l'évènement associé à différentes combinaisons de prédicteurs. Ces outils sont généralement appelés modèle pronostique, modèle prédictif, score pronostique ou encore nomogramme. L'objectif de ces outils est double. Le premier est de fournir aux cliniciens des informations pronostiques précises sur la base de multiples prédicteurs et le second est de minimiser les biais liés à l'hétérogénéité de pronostic des patients atteints d'une même pathologie dans les stratégies de recherches futures (sélection de patients, similarité de pronostic entre les patients dans les différents bras d'un essai clinique, etc.). L'approche multivariée donne aussi la possibilité de confronter les facteurs pronostiques entre eux. Ainsi, il est possible d'évaluer la valeur prédictive additionnelle d'un facteur pronostique nouvellement

identifié à ceux déjà classiquement utilisés dans la pathologie. Cette démarche est essentielle pour quantifier la part attribuable à l'explication du pronostic des patients au regard des autres facteurs considérés et aussi pour comprendre l'information qu'il véhicule (meilleure identification des patients de bon ou de mauvais pronostic ou les deux). Ceci est d'autant plus nécessaire lorsque le facteur pronostique nouvellement identifié est par exemple plus invasif ou plus coûteux par rapport aux autres paramètres d'intérêt. Un aspect plus technique, consiste en l'élaboration du modèle pronostique qui repose sur quatre étapes clefs : sa construction, son évaluation, sa validation externe et l'évaluation de son impact clinique.

Le dernier axe de la recherche pronostique s'inscrit dans le cadre de la médecine de précision et s'intitule « *la recherche en médecine stratifiée* ». L'idée est de pouvoir identifier des biomarqueurs pouvant prédire la réponse d'un individu à un traitement et donc de permettre d'identifier des sous-populations pour lesquelles le traitement serait le plus efficace. Les cliniciens parlent souvent de la valeur prédictive d'un biomarqueur qu'ils opposent régulièrement de manière trop rapide au terme pronostic. En effet, l'évaluation de la réponse d'un patient à un traitement passe par la prédiction d'un évènement (centré sur le patient ou la tumeur) et donc par l'élaboration d'un modèle pronostique de l'évènement. L'évaluation de la réponse à un traitement selon l'expression d'un biomarqueur se fait donc à l'aide d'une modélisation pronostique de l'évènement d'intérêt. Il est d'ailleurs intéressant de noter que parfois la réponse à un traitement peut être associée, non pas à l'expression d'un biomarqueur, mais à un gradient de pronostic de survie des patients. Par exemple, les thérapies cytotoxiques agressives sont connues pour être en général plus efficaces et sont privilégiées pour les patients en relativement bon état général et ayant donc à priori un pronostic de survie favorable.

Le développement intensif des connaissances en sciences médicales ainsi que la facilité d'accès à la technologie de pointe confrontent les médecins à un tsunami d'informations de différentes natures complexifiant l'appréhension du pronostic des patients. Dans ce contexte, il est primordial, via le développement de la méthodologie statistique, de rendre l'ensemble de ces informations exploitables et donc de résumer, hiérarchiser et simplifier

l'information afin de guider le management des patients, dans le but de faciliter la prise de décision.

A ce jour, il n'existe pas de travail identifiant et synthétisant clairement les étapes clefs des recherches à visée pronostique avec une proposition d'outils méthodologiques adéquats pour répondre aux problématiques posées dans chacune d'entre elles.

L'oncologie et la transplantation rénale sont deux domaines dans lesquels des recherches pronostiques, respectivement au moment du diagnostic (ou à l'initiation d'un traitement) et de la mise en œuvre de la transplantation (ou après la première année post-transplantation), doivent être réalisées ou remises à jour avec les connaissances actuelles.

I.2 L'oncologie digestive

Les cancers digestifs regroupent les cancers de l'œsophage, de l'estomac, du côlon, du rectum, du canal anal, du pancréas, les hépato-carcinomes, les cholangio-carcinomes et les tumeurs endocrines digestives. L'incidence des cancers digestifs est très variable selon les localisations, mais globalement progresse principalement en raison de l'augmentation de l'incidence du cancer colorectal et de celui du pancréas. De par leur fréquence et leur gravité, ces deux pathologies représentent aujourd'hui un enjeu important pour la santé publique.

I.2.1 Le cancer pancréatique

Le cancer pancréatique (PC) est la quatrième cause de décès par cancer dans les pays développés (Siegel *et al*, 2016) et devrait devenir la deuxième cause de décès par cancer en 2030 (Rahib *et al*, 2014), en raison d'une incidence en constante augmentation et de l'absence de changement significatif du taux de mortalité entre 1997 et 2013 (Cancer Statistics Review, 1975-2013 - SEER Statistics). En effet, le cancer pancréatique est

notamment caractérisé par le pronostic de survie le plus sombre des cancers digestifs avec un taux de survie à 5 ans de l'ordre de 5-7%. Cette évolution défavorable et la sévérité de la maladie soulignent l'importance de donner la priorité à la recherche dans la prévention et la prise en charge thérapeutique de ce cancer. En guise d'illustration, en 2014 dans l'Union Européenne le nombre de décès par cancer pancréatique était estimé à 82 300. Par analogie, cette prédiction était similaire à celle des décès de femmes atteintes d'un cancer du sein (n=89 300) (Malvezzi *et al*, 2014).

La résection chirurgicale dans un contexte de maladie localisée est le seul traitement qui peut fournir une survie globale (SG) prolongée. Cependant, en l'absence de symptôme spécifique au début de la maladie, le diagnostic est fait à un stade avancé dans la grande majorité des cas (> 80%). La médiane de SG est d'environ 9-15 mois chez les patients avec un cancer pancréatique localement avancé, et de 6-9 mois au stade de maladie métastatique (Ryan *et al*, 2014; Neuzillet *et al*, 2015)

Lorsque le cancer pancréatique est non résécable (localement avancé ou métastatique), il demeure une maladie non curable dont la prise en charge doit être améliorée. En effet pour ces patients, une chimiothérapie par gemcitabine est restée le traitement de référence de première ligne pendant plus d'une décennie, jusqu'en 2010. Cependant, au cours des cinq dernières années, des progrès significatifs ont été réalisés pour les patients en situation de maladie métastatique et en bon état général, avec l'approbation successive de deux combinaisons cytotoxiques actives : FOLFIRINOX (5-fluorouracil [5-FU], irinotecan, et oxaliplatine) et l'association gemcitabine plus *nab*-paclitaxel. Ces deux nouveaux schémas se sont avérés supérieurs à la gemcitabine seule permettant d'obtenir respectivement des médianes de survie globale de 11,1 et 8,5 mois, (Conroy *et al*, 2011; Von Hoff *et al*, 2013).

À l'heure actuelle, les recommandations définissants les pratiques cliniques pour le traitement en première ligne des patients atteints d'un cancer pancréatique sont généralement bien établies (Ducreux *et al*, 2015).

Malgré les récents progrès mentionnés ci-dessus les stratégies thérapeutiques pour traiter le cancer pancréatique restent limitées. De plus, des incertitudes demeurent ; par exemple, dans l'identification des patients atteints d'un cancer pancréatique localement avancé qui pourraient bénéficier de chimioradiothérapie, ou dans l'identification des patients avec une maladie avancée pour lesquels une seconde ligne de chimiothérapie pourrait ne pas être bénéfique après échec de la première.
L'absence actuelle de consensus, pour la définition de profils de risque spécifiques pour la survie globale chez les patients localement avancés, peut induire une confusion dans le développement de stratégies de recherche et aussi dans le management des patients. Il est donc important de proposer et de valider un nouveau modèle pronostique pour la survie globale pour ces patients, en étudiant un spectre important de paramètres disponibles au diagnostic avant l'initiation du traitement d'induction et de développer sur la base de ce nouveau modèle des outils de prédiction individuelle et populationnelle. Le même besoin est nécessaire chez les patients avec une maladie avancée dans un contexte de seconde ligne de chimiothérapie.

In fine, une recherche pronostique permettant d'améliorer les connaissances sur le cancer pancréatique peut apporter des opportunités d'optimisation des standards actuels de prise en charge et de clarifier les stratégies de recherche, comme la conception des futurs essais cliniques.

I.2.2 Le cancer colorectal

Le cancer colorectal (CRC) est, quant à lui, le deuxième cancer le plus couramment diagnostiqué en Europe et une des principales causes de décès attribuable au cancer à la fois en Europe et dans le monde entier (Ferlay *et al*, 2013, 2015). En 2012, il y eu 447 000 nouveaux cas diagnostiqués et 215 000 décès par cancer colorectal en Europe.

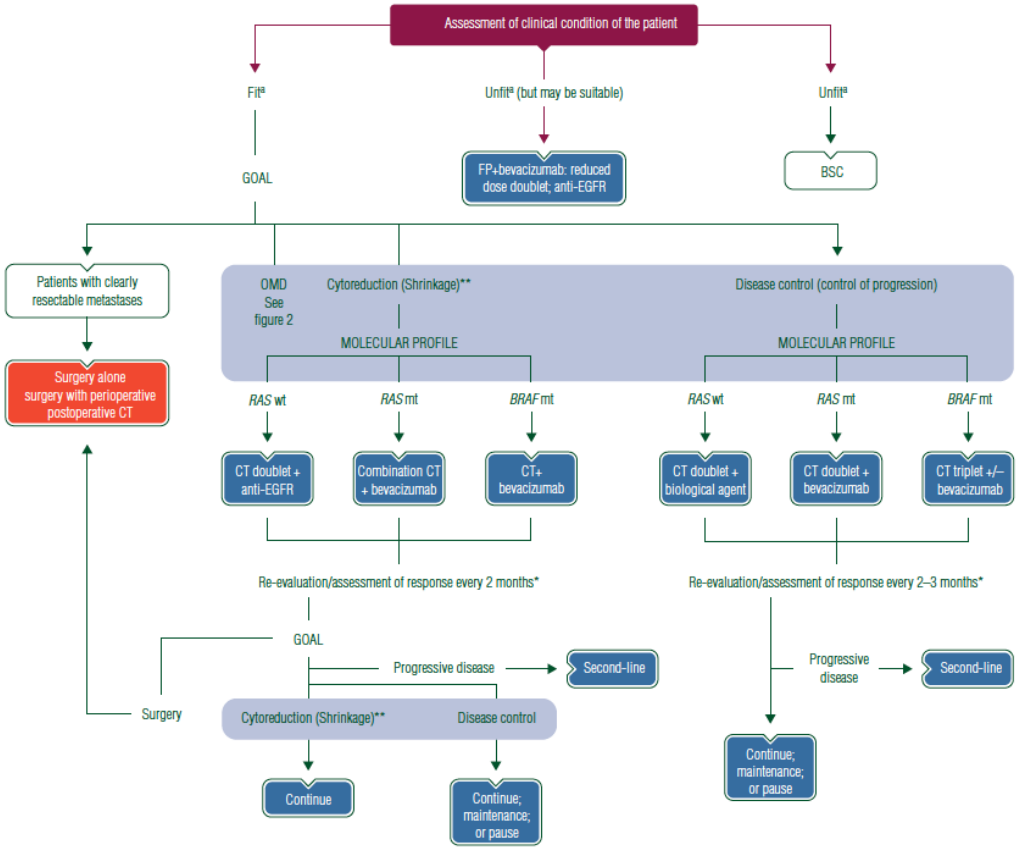
Le cancer colorectal est un cancer de bon pronostic lorsqu'il est diagnostiqué à un stade précoce : la survie relative à 5 ans est de 91% pour les stades localisés, de 70% pour les stades avec envahissement locorégional. En revanche, la survie à 5 ans est d'environ 11% dans les situations métastatiques, ce qui représente environ 25% des patients au moment du diagnostic. En effet, en raison de son évolution souvent asymptomatique dans un premier temps, le cancer colorectal est diagnostiqué parfois tardivement et nécessite alors des traitements lourds.

Néanmoins, au cours de la dernière décennie, la prise en charge clinique pour les patients atteints d'un cancer colorectal métastatique (mCRC) s'est considérablement améliorée. En effet, la médiane de survie globale pour ces patients a doublée en 20 ans pour être aujourd'hui de l'ordre de 30 mois.

Cette tendance favorable est liée d'une part à une augmentation du nombre de patients ayant une résection chirurgicale de leur maladie métastatique, et d'autre part à une approche plus stratégique concernant l'utilisation des diverses options thérapeutiques disponibles.

En effet de nos jours, plusieurs interventions thérapeutiques sont disponibles dans un contexte de mCRC (Fakih, 2015; Van Cutsem *et al*, 2016)

Figure: Recommandations consensuelles sur l'utilisation des produits cytotoxiques et biologiques dans le traitement de première ligne et au-delà pour des patients atteints de mCRC (Van Cutsem *et al*, 2016).



Zurich treatment algorithm. BSC, best supportive care; CT, chemotherapy; EGFR, epidermal growth factor receptor; FP, fluoropyrmidine; mt, mutant, NED, no evidence of disease; OMD, oligometastatic disease; wt, wild-type. [a]Patients assessed as fit or unfit according to medical condition not due to malignant disease. *After two re-evaluations, consider maintenance. **(A) Includes two subgroups: (1) those for whom intensive treatment is appropriate with the goal of cytoreduction (tumour shrinkage) and conversion to resectable disease; (2) those who need an intensive treatment, although they will never make it to resection or LAT, since they need a rapid reduction of tumour burden because of impending clinical threat, impending organ dysfunction, severe symptoms.

Dans ce contexte, il y a un besoin urgent d'outils pronostiques afin d'optimiser la sélection des patients éligibles pour les différentes options de traitement afin de proposer une prise en charge plus personnalisée. Ceci repose sur une meilleure caractérisation du pronostic des patients et ainsi sur un effort important de recherche pronostique.

L'identification de biomarqueurs pronostiques au diagnostic permettant une meilleure stratification du décès reste un défi à relever chez les patients atteints de mCRC.

Les deux scores pronostiques de la SG actuellement les plus utilisés pour les patients mCRC sont le score de Kohne et celui du Groupe Coopérateur Multidisciplinaire en Oncologie GERCOR (Köhne *et al*, 2002; Chibaudel *et al*, 2011). Le score de Khone repose sur l'évaluation du « performance status » (PS), du nombre de globules blancs, du nombre de sites métastatiques, et du taux de phosphatases alcalines. Cependant, une étude a montré que dans un

contexte de thérapie ciblée seulement 2 de ces paramètres (le PS et le nombre de globules blancs) demeurent associés à la SG (Desot *et al*, 2013). Le score du GERCOR quant à lui s'articule aussi sur le PS auquel il associe l'évaluation des lactates déshydrogénases (LDH), un paramètre biologique. Le score du GERCOR tend à être aujourd'hui le score de référence puisqu'il est plus discriminant pour le SG que celui de Khone et est de plus bien adapté au contexte des thérapies ciblées.

Ces deux scores sont donc basés sur le PS qui est un paramètre ayant pour objet d'évaluer l'état fonctionnel d'un patient. A cet effet, il est établi sur la base d'un interrogatoire qui peut véhiculer la subjectivité du clinicien et parfois ne pas être reproductible. Le PS peut être influencé par de multiples facteurs liés aux patients (anorexie, état psychologique) et/ou liés à la maladie tumorale (inflammation, taille tumorale, anémie, etc.).

La substitution du PS par un paramètre biologique plus reproductible et aussi plus spécifique et représentatif de la tumeur ou de l'état métabolique du patient pourrait être intéressant pour mieux individualiser les différentes options thérapeutiques au regard de leur mode d'action. Si nous considérons le profil moléculaire de la tumeur, le statut mutationnel BRAF peut constituer un paramètre pronostique clef mais ne caractérise que 8 à 10% des patients mCRC (Popovici *et al*, 2012; Venderbosch *et al*, 2014; Schirripa *et al*, 2015; Sorbye *et al*, 2015).

Dans cette pathologie, un enjeu majeur réside donc dans l'identification et la validation de biomarqueurs spécifiques de la maladie tumorale dans le but d'améliorer la stratification du risque de décès. L'invasion, l'angiogenèse et le contrôle du microenvironnement tumoral sont des caractéristiques communes de la progression du cancer dans lesquels peu de biomarqueurs sont disponibles au diagnostic et offrent donc un terrain pour améliorer les connaissances du mCRC et la prise en charge des patients.

Nous avons récemment montré que l'angiopoïétine-2, qui est impliquée dans l'angiogenèse, est un paramètre pronostique intéressant pour la SG chez des patients mCRC (Jary *et al*, 2015). Le Syndecan-1 (SDC-1, aussi appelé CD138) est un protéoglycane transmembranaire de type héparane sulfate impliqué dans l'angiogenèse, la progression tumorale et aussi dans la transition épithélio-mesenchymateuse (Maeda *et al*, 2006; Yang *et al*, 2011; Mitselou *et al*, 2012). Le CD138 constitue un biomarqueur d'intérêt puisqu'il a déjà été décrit comme jouant un rôle important dans la séquence adénome-adénocarcinome(Leppä *et al*, 1996; Day *et al*, 1999) et son rôle pronostique a déjà été reconnu comme intéressant dans plusieurs cancers solides et hématologiques (Seidel *et al*, 2000; Akl *et al*, 2015).

Dans les cancers colorectaux, l'expression de CD138 a été décrite dans le stroma des tumeurs primitives comme étant corrélée à des caractéristiques histopathologiques de mauvais pronostique lorsqu'il existe une perte d'expression membranaire et une augmentation de l'expression stromale en immunohistochimie. Ex vivo, les cellules cancéreuses issues de patients présentant des taux sériques élevés de CD138 présentent une moins bonne sensibilité à la chimiothérapie (Hashimoto *et al*, 2008; Wang *et al*, 2014). Une hypothèse est donc l'augmentation de la chimiorésistance en cas de CD138 sérique élevé, mais celle ci reste à démontrer. Cependant, à ce jour, le rôle pronostique du CD138 dans sa forme soluble sur la SG n'a jamais été investigué chez les patients mCRC et mérite donc d'être évalué.

I.3 La transplantation rénale

Depuis la première greffe de rein réalisée en 1952 à l'hôpital Necker (Paris, France), suivie d'un taux initialement élevé de perte de greffon, la transplantation rénale est devenue aujourd'hui le traitement de référence pour les patients atteints d'une insuffisance rénale chronique en phase terminale. Cette pathologie est en constante augmentation dans les pays

industrialisés. Désormais, la transplantation rénale permet à la fois d'augmenter la survie et la qualité de vie de ces patients (Evans *et al*, 1985). La prévalence mondiale des receveurs d'organe est de plus de 1.5 million avec 150 000 transplantations chaque année, dont 70% sont des greffes rénales (ONT-WHO).

Actuellement, 33 298 patients transplantés rénaux sont recensés en France. La liste d'attente des candidats à la greffe est en constante augmentation puisqu'elle a doublée dans la dernière décennie. En raison de la pénurie d'organes et de la forte incidence des maladies rénales chroniques, la transplantation demeure un enjeu majeur de santé publique. En effet, le déséquilibre entre l'offre et la demande de greffons est responsable de nombreux décès. Par conséquence, optimiser la survie des greffons est dans ce contexte une nécessité pour la communauté médicale (Garcia *et al*, 2012).

Cependant, malgré des progrès considérables, le rejet du greffon demeure toujours une menace importante pour les greffons fonctionnels. Chaque année le rejet d'organes greffés provoque l'échec de milliers de transplantations dans le monde avec des conséquences immédiates en termes de mortalité, de morbidité et de coût pour la société. Un changement majeur dans l'histoire naturelle du rejet en transplantation rénale est survenu suite à la diminution considérable de l'incidence du rejet cellulaire (rejet médié par les lymphocytes T) avec des conséquences limitées pour la survie à long terme du greffon. C'est pourquoi, l'attention s'est désormais tournée vers le rejet humoral, qui est médié par les anticorps anti-HLA circulants dirigés contre le donneur. Au cours de la dernière décennie, des études convergentes ont indiqué que ce rejet humoral représente actuellement la principale cause d'échec de la transplantation rénale (Nankivell & Alexander, 2010; Sellarés *et al*, 2012). Cependant, aujourd'hui peu de choses sont connues sur la physiopathologie du rejet médié par les anticorps et par conséquent il est difficile de déterminer l'importance des lésions, l'activité ou le stade de la maladie. Des progrès doivent être aussi faits sur la prévention et le traitement du rejet humoral. Dans la quête d'une médecine de précision (Jameson & Longo, 2015) et compte tenu de la complexité du rejet médié par les anticorps et les limites inhérentes aux systèmes existants qui

reposent sur des outils histologiques conventionnels, le rejet humoral a besoin de plus de soutien, comme l'illustre ce récent « appel à agir » lancé par la FDA (Food and Drug Administration) à Arlington en septembre.

Aujourd'hui, la transplantation rénale doit améliorer ses standards concernant le diagnostic de rejet, la mesure de l'activité et du stade de la maladie, la stratification du risque, et aussi optimiser ses standards pour la mesure de la réponse aux traitements. Les principales limites pour atteindre ces objectifs dans la pratique clinique actuelle sont le faible niveau de phénotypage, le sous-diagnostic fourni par les outils actuellement de référence (biopsie du greffon), l'absence d'outils modernes pour la surveillance des patients et le manque de thérapie personnalisée. La rareté des plateformes de diagnostic est une faiblesse dans le domaine de la transplantation, ce qui conduit à des erreurs de diagnostic et une classification incorrecte de la maladie. Dans ce contexte, une recherche pronostique de qualité peut constituer une arme pour apporter des réponses concrètes afin d'améliorer les standards actuels de soins en transplantation rénale. En particulier, l'étude du pronostic du greffon, dans le contexte de la réponse allo-immune induite par la transplantation à partir d'un donneur génétiquement différent qui représente aujourd'hui un obstacle majeur au succès de la greffe, est nécessaire. Ceci a fait l'objet de trois travaux qui sont présentés comme des publications complémentaires à la thèse permettant d'illustrer la partie méthode (Lefaucheur *et al*, 2013; Loupy *et al*, 2013, 2014)

Aujourd'hui, la politique de greffe, en réponse à la pénurie d'organes, a consisté ces dernières années en une augmentation de l'âge des donneurs. Ceci a conduit à la définition de donneurs à critères dits élargis. Ces greffons rénaux représentent en France environ la moitié des reins prélevés chez des donneurs décédés et donc une étude approfondie de leur pronostic de survie présente un intérêt majeur.

II Objectifs du travail

Ce travail s'inscrit dans une volonté d'amélioration et de standardisation de la recherche pronostique en sciences médicales et a l'ambition de répondre à un objectif double méthodologique et clinique.

1- Le premier objectif a été a) d'identifier les étapes clefs des recherches à visée pronostique et b) de proposer une approche méthodologique adéquate et rigoureuse ainsi que des développements pour répondre aux problématiques posées. Je propose ainsi avec ce travail une stratégie analytique globale et aussi des développements et applications dans le cadre de la recherche d'outils de prédiction et de la stratification du risque.

2- Le second objectif de cette thèse est de répondre à des problématiques pronostiques importantes en oncologie digestive et en transplantation rénale en intègrant l'ensemble des progrès thérapeutiques, diagnostiques et techniques et en appliquant la stratégie analytique préalablement proposée.

a- en oncologie digestive

i. proposer et valider un nouveau modèle pronostique pour la survie globale pour les patients atteints d'un cancer du pancréas localement avancé au diagnostic avant l'initiation du traitement d'induction et développer sur la base de ce nouveau modèle des outils de prédiction individuelle et populationnelle.

ii. proposer et valider un nouveau modèle pronostique pour la survie globale pour les patients atteints d'un cancer du pancréas avancé avant l'initiation du traitement de seconde ligne de chimiothérapie et développer sur la base de ce nouveau modèle des outils de prédiction individuelle et populationnelle.

iii. étudier le rôle pronostique du CD138 dans sa forme soluble sur la survie globale chez patients atteints d'un mCRC; déterminer sa valeur discriminante additionnelle aux paramètres déjà connus pour stratifier le risque de décès des patients et proposer un score pronostique de la SG basé sur des paramètres biologiques robustes mesurables au diagnostic et applicables dans la pratique clinique courante.

b- en transplantation rénale

i. identifier les facteurs pronostiques de la survie des reins provenant de donneurs dits à critères élargis dans le but de caractériser des groupes de patients transplantés dont le pronostic des greffons est distinct et ainsi optimiser l'attribution de ces reins.

III Méthodologie

Dans le contexte d'une recherche à visée pronostique, le processus méthodologique d'analyse développé et proposé dans cette thèse peut être résumé en 4 points, ou étapes clefs, qui peuvent être complémentaires et qui impliquent pour chacun des techniques statistiques spécifiques.
Ces 4 étapes à considérer dans le développement d'outils à visée pronostique sont les suivantes :

1- La caractérisation et l'identification du phénotype de la pathologie à étudier ;
2- La recherche de facteurs pronostiques de l'évènement d'intérêt dans la population initialement considérée ;
3- La détermination de la valeur additionnelle apportée par un nouveau facteur pronostique identifié aux facteurs conventionnels déjà connus et fonctionnels dans la pratique clinique ;
4- La construction d'outils à visée pronostique, sur la base de modèles intégratifs, permettant une stratification précise du risque d'évènement, à l'échelle individuelle et aussi populationnelle.

La méthodologie à mettre en œuvre dans ces 4 étapes est présentée en détails ci-après avec, pour chacune d'entre elles, la proposition de standard à considérer d'un point de vue analytique et en particulier statistique.

Une synthèse graphique du processus méthodologique proposé est réalisée dans la figure ci-dessous :

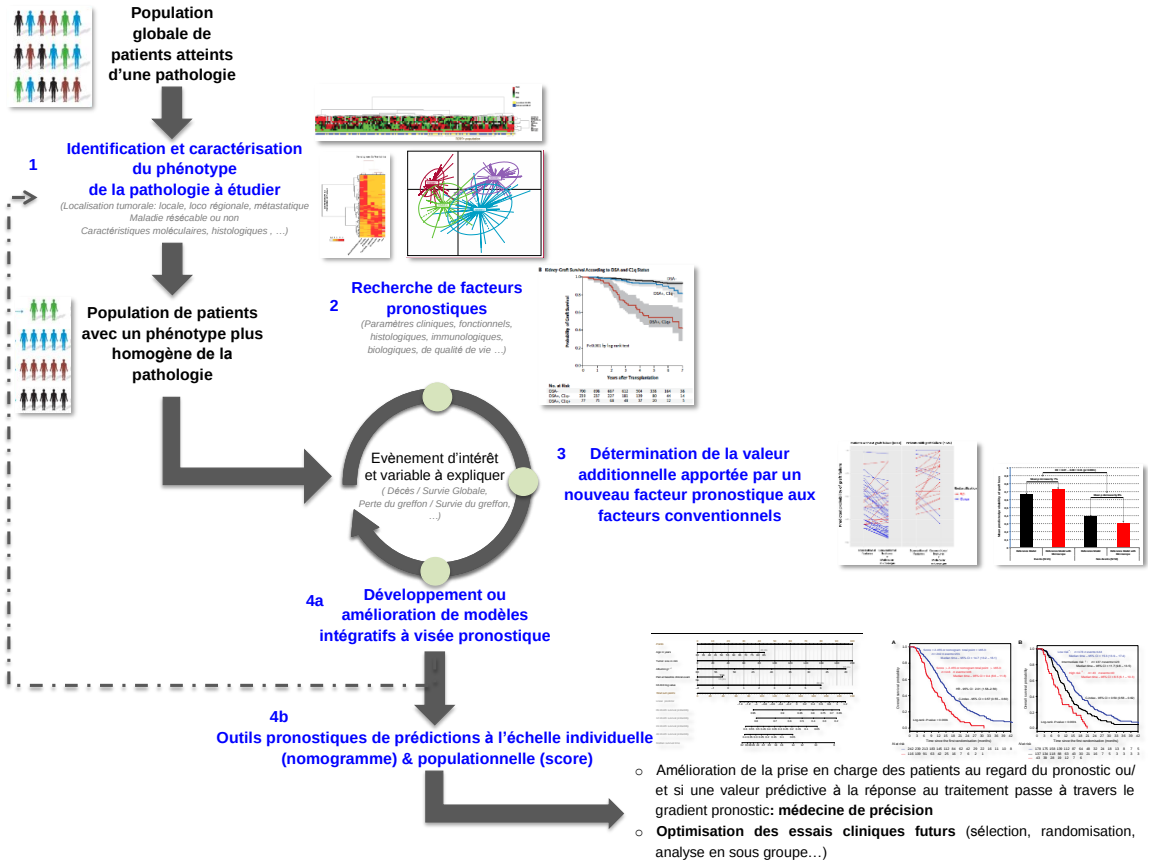


Figure : Identification des étapes clefs d'une recherche à visée pronostique et illustration de la stratégie analytique proposée dans la thèse

III.1 Caractérisation et identification du phénotype de la pathologie à étudier

La première étape de toute recherche et en particulier dans une recherche à visée pronostique consiste en la caractérisation et l'identification précise de la population sur laquelle va porter l'étude comme illustrée dans la figure suivante:.

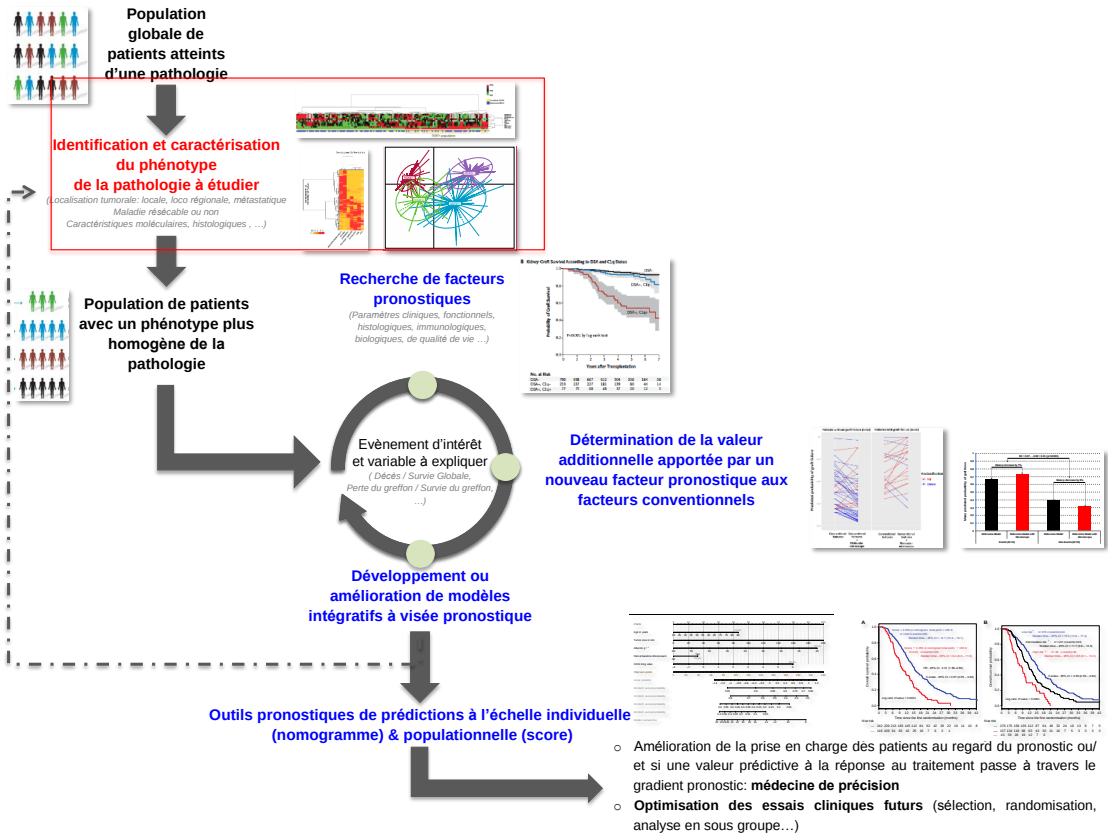


Figure : Première étape d'une recherche à visée pronostique : l'identification et la caractérisation précise de la population d'étude.

Cette étape est primordiale car le résultat obtenu ne pourra porter et être généralisé uniquement sur les patients pouvant être assimilés à la population analysée.

Dans le champ des essais cliniques, la définition des critères d'inclusion et d'exclusion est depuis longtemps considérée avec attention dans les designs

ainsi que dans les publications relatant les résultats. La bonne définition de ces critères pondère la faisabilité de l'étude et l'impact clinique des résultats, à travers la population dans laquelle ils pourront être généralisés. En effet, un phénotype de patients très précis peut rendre la recherche difficilement réalisable par un rythme d'inclusion faible, avec des résultats qui porteront sur une population très sélectionnée et donc pas toujours cliniquement pertinente. A l'opposé des critères de sélection larges faciliteront le déroulement de la recherche puisque le recrutement sera plus aisé et rapide avec une généralisation des résultats facilitée. Cependant, dans ce dernier cas de figure, le résultat pourra véhiculer beaucoup d'hétérogénéité avec l'existence potentielle de sous-groupes de patients pour lesquels l'amplitude et la significativité de l'effet traitement ou intervention testé (interaction statistique) peuvent être sensiblement différents. Aussi, si les critères sont trop larges et caractérisent une population trop hétérogène, avec la présence d'une interaction suggérée mais non significative, en raison par exemple d'un manque de puissance statistique, les résultats sont alors compliqués d'interprétation et mènent souvent à des recherches complémentaires. Ces considérations peuvent tout à fait être transposées à la recherche pronostique. Les critères caractérisant la population dans laquelle une étude pronostique sera réalisée doivent être à ce titre cliniquement significatifs, précis, acceptables éthiquement, univoques et représentatifs d'une population d'intérêt clinique et donc pas trop restrictifs.

III.1.1 Caractérisation : Choix selon des caractéristiques déjà établies

Les pathologies étudiées possèdent pour la plupart une hiérarchisation ou classification et donc des phénotypes de patients déjà bien reconnus, pour lesquels la prise en charge proposée et/ou le pronostic de survie sont clairement différenciés. Dans ce contexte, cette étape est assez simple d'un point de vue méthodologique et se base sur des caractéristiques connues et des classifications consensuelles des pathologies. Par exemple, en oncologie

la stratégie thérapeutique dans le cancer pancréatique est clairement différente selon le stade de la maladie. En effet, dans un contexte de maladie résécable (environ 20% des cas au diagnostic de cancer), la chirurgie représente la pierre angulaire des soins, tandis que pour une maladie non résécable dite avancée (localement avancée ou métastatique) l'option thérapeutique clef est la chimiothérapie. Il est aussi bien connu que le stade de maladie selon la classification TNM est un facteur pronostique majeur. Cette classification est un système international de classement des cancers selon leur extension anatomique. Plusieurs révisions ont été publiées, la dernière étant la septième édition en 2009. Les trois lettres symbolisent la propagation de la maladie cancéreuse au niveau de la tumeur primitive (T), des ganglions lymphatiques (N pour node en anglais) et à distance par métastases (M). Dans ce contexte et dans un objectif de recherche pronostique, il est donc naturel d'identifier ou de considérer séparément dans les études les groupes de patients évoqués ci-dessus (maladie résécable ou avancée). Aujourd'hui les recherches dans un contexte de pathologie avancée ont même tendance à dissocier les patients ayant une maladie localement avancée (non résécable et non métastatique) des patients présentant des métastases. Cette classification des patients distinguant une maladie résécable d'une maladie avancée est aussi la norme dans le cancer colorectal. En oncologie, l'accumulation des connaissances sur la biologie moléculaire des tumeurs par exemple implique une adaptation des options thérapeutiques de plus en plus fréquente et constitue un facteur majeur dans la caractérisation des populations de patients étudiées (recherche de l'instabilité des microsatellites; mutation BRAF ; mutation KRAS ; etc.). D'ores et déjà, et encore plus fréquemment à l'avenir, les recherches vont certainement être réalisées sur des populations plus clairement identifiées par ces caractéristiques de biologies moléculaire.

Dans cette thèse, les travaux qui s'inscrivent dans le domaine de l'oncologie porteront sur les patients atteints d'un cancer pancréatique localement avancé et aussi en situation métastatique ainsi que sur le cancer colorectal métastatique. La sélection de la population d'analyse dans ce

contexte ne représente donc que la première étape de construction du design de l'étude par une caractérisation simple sur la base d'éléments déjà bien établis.

Un enjeu majeur de la transplantation rénale demeure l'identification de déterminants pronostiques de la survie des greffons. A cet effet, les études à visée pronostique considèrent régulièrement l'ensemble de la population de transplantés rénaux comme population d'intérêt. Par ailleurs, l'identification des rejets humoraux médiés par les anticorps comme étant, aujourd'hui, l'une des principales causes de perte de greffons implique l'étude des déterminants de la perte du greffon dans ce contexte spécifique. Des travaux présentés dans cette thèse feront l'objet de l'identification de facteurs pronostiques de la survie des greffons dans la population globale de transplantés et aussi dans la population plus spécifique de patients présentant un rejet humoral post-transplantation. Comme précédemment cette caractérisation de la population d'analyse est simple car elle se base sur des éléments reconnus et n'implique pas de techniques statistiques particulières.

Cependant, la transplantation rénale est une discipline médicale dans laquelle les efforts de recherche sont relativement récents comparés à l'oncologie. Ainsi, la classification internationale de Banff proposée en 1991 qui est une nomenclature et une caractérisation de la pathologie des greffons rénaux est en constante évolution. Aussi, un autre enjeu majeur en transplantation consiste en l'amélioration de cette classification par un phénotypage plus précis des différents types de rejets sur la base de caractéristiques cliniques, histologiques et immunologiques. Dans cette thèse un travail est dédié à cet objectif et repose sur l'identification par des méthodes statistiques non supervisées de différents types de rejets avec des phénotypes distincts. Une fois ces différents types de rejets caractérisés, la question de leur pronostic en termes de survie du greffon est ensuite fondamentale afin d'asseoir leur pathogénicité et donc l'implication clinique de leur identification. Dans ce cas précis de recherche de phénotype basée sur des caractéristiques pathologiques multiples une approche méthodologique

adaptée peut être proposée en impliquant notamment des outils du champ de la statistique descriptive multidimensionnelle.

On désigne par statistique descriptive multidimensionnelle l'ensemble des méthodes de la statistique descriptive ou exploratoire permettant de traiter simultanément un nombre quelconque de variables. En effet, il s'agit d'aller au-delà de l'étude d'une seule ou de deux variables. Ces méthodes sont purement descriptives, c'est-à-dire qu'elles ne supposent, a priori, aucun modèle sous-jacent, de type probabiliste.

III.1.2 Identification par des techniques de statistique descriptive multidimensionnelle

Lorsqu'on veut s'affranchir des classifications connues, l'identification de phénotypes de patients en oncologie ou de rejets en transplantation n'est pas quelque chose d'aisé et doit être réalisée de la façon la moins arbitraire possible. A cet effet, une approche non supervisée par des outils provenant de la statistique descriptive multidimensionnelle est possible notamment avec l'utilisation successive des méthodes de classification hiérarchique et de l'analyse en composantes principales (ACP).

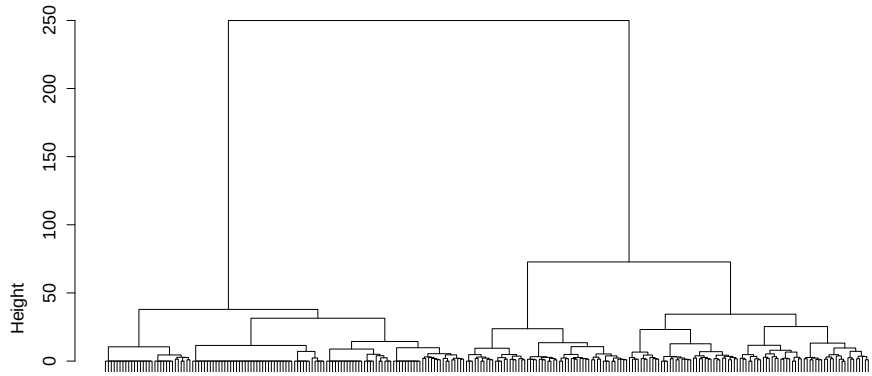
III.1.2.1 Classification hiérarchique par cluster (dendogramme) et illustration par une « Heat map »

L'objectif principal des méthodes de classification est de répartir les éléments d'un ensemble en groupes, c'est-à-dire d'établir une partition de cet ensemble. Pour ce faire différentes contraintes sont imposées. Chaque groupe doit être le plus homogène possible en termes d'éléments mais les groupes doivent être les plus différents possibles les uns des autres. Les méthodes de classification résument donc les données par une variable qualitative. Elles fournissent ainsi des partitions. Cependant, on ne se

contente pas d'une partition puisqu'on cherche une hiérarchie de parties, qui constituent un arbre binaire appelé le dendrogramme. Un arbre est un graphique raciné. Les feuilles sont les parties à un seul élément et sont toujours dans une hiérarchie. La racine est l'ensemble tout entier. Chaque partie n'a qu'un ancêtre, à l'exclusion de la racine qui n'en n'a pas. Si l'arbre est binaire, chaque partie a deux descendants, à l'exclusion des feuilles qui n'en n'ont pas. La représentation graphique d'une classification hiérarchique s'appelle un dendrogramme. La recherche d'une classification hiérarchique s'appuie sur une notion de distance entre individus qui induit une mesure de l'hétérogénéité d'une partie basée sur les distances entre individus qui sont dedans et une mesure de dissimilarité entre deux parties basée sur la distance entre un individu de l'un et un individu de l'autre (Anderberg, 1973; Gordon, 1999; Chessel *et al*, 2014). Il existe une multitude de manière de calculer des distances entre des entités. Une fois la distance définie, il existe à nouveau plusieurs stratégies pour établir une hiérarchie ("ward","single","complete","average", etc.) (Cormack, 1971). A titre d'exemple, celle dite de Ward définit d'abord la valeur d'un regroupement (l'augmentation de l'inertie intra) qui sert de mesure de la différence entre parties ; deux parties sont proches si leur regroupement augmente peu l'inertie intra.

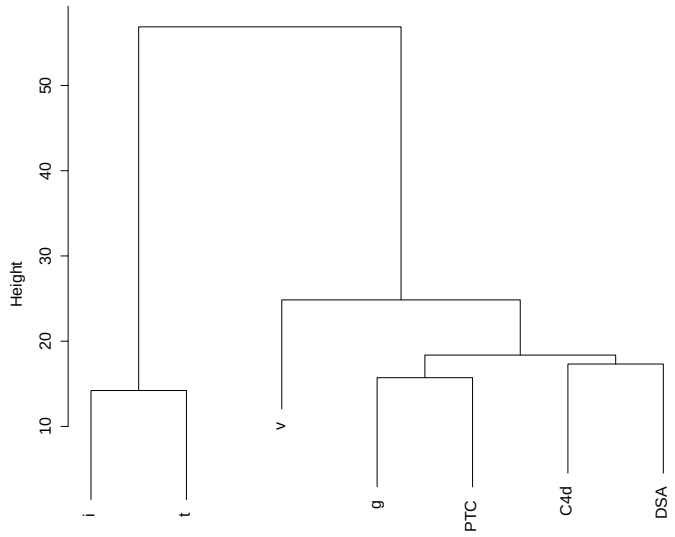
Dans cette thèse, le but n'est pas de rentrer dans les considérations théoriques de l'analyse de classification hiérarchique mais de montrer comment cette méthode peut être utile dans la démarche d'identification de différents phénotypes de maladies ou de patients.

En pratique, si par exemple la question est d'identifier des partitions de rejets qui peuvent se distinguer sur la base des paramètres pathologiques classiquement utilisées dans le diagnostic de rejet aigu (DSA circulants, dépôt de C4d dans le greffon, glomérulite, capillarite péritubulaire, endartérite, tubulite et inflammation interstitielle) il est alors naturel de considérer le rejet comme l'unité d'intérêt de la classification hiérarchique et ainsi de produire un dendogramme orienté sur le rejet comme dans l'illustration ci-dessous :



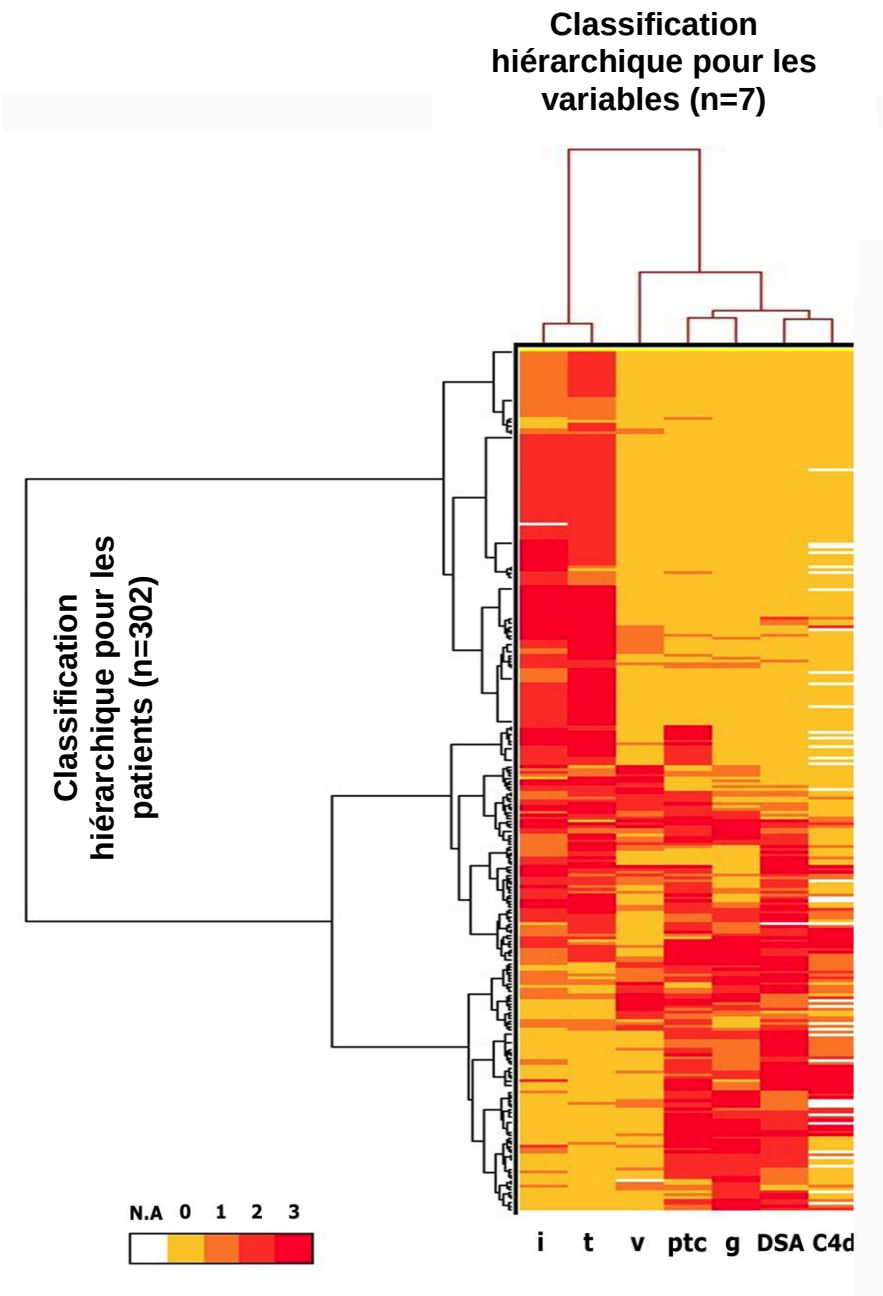
Sur cette figure, sur l'axe des abscisses chaque ligne verticale correspond à un rejet. Cette classification hiérarchique permet donc d'isoler des clusters de rejet qui à l'intérieur du même cluster se ressemblent en termes de caractéristiques pathologiques classiquement utilisées dans le diagnostic de rejet aigu. Plus le trajet qui oppose deux clusters de rejet est important plus ces deux clusters sont différents en termes de caractéristiques et inversement.

Nous pourrions arrêter le raisonnement à ce stade mais cette seule représentation ne permet pas de visualiser les caractéristiques de chacun des groupes de rejets isolés. De plus, on peut changer l'unité de l'analyse et considérer la caractéristique du rejet aigu comme l'unité d'intérêt de la classification hiérarchique et ainsi produire un dendogramme orienté sur les caractéristiques du rejet aigu comme dans l'illustration ci-dessous :



Sur cette figure, sur l'axe des abscisses chaque ligne verticale correspond à une caractéristique du rejet aigu (DSA circulants : DSA, dépôt de C4d dans le greffon : C4d, glomérulite : g, capillarite péritubulaire : PTC, endartérite : v, tubulite et inflammation interstitielle : i et t). Cette classification hiérarchique a donc permis d'isoler des clusters de paramètres utilisés dans le diagnostic de rejet aigu qui à l'intérieur du même cluster se ressemblent en termes de rejet. En effet, les paramètres spécifiques du rejet humoral (g, PTC, C4d et DSA) s'opposent aux paramètres plus caractéristiques du rejet cellulaire (i et t). L'endartérite (v) ségrége seule entre les deux clusters de paramètres précités.

Ainsi, à ce stade nous avons donc identifié des clusters de rejets et aussi des caractéristiques du rejet aigu qui apportent des informations complémentaires mais sur des supports dissociés. La dernière étape consiste donc en la réunion de ces deux analyses à l'aide de ce qu'on appelle une « heat map » ou carte de chaleur. Une heat map est une représentation graphique de données qui fait correspondre un code couleur à l'intensité d'expression d'une variable (Gehlenborg & Wong, 2012). Ainsi les rejets peuvent être réarrangés comme dans l'analyse hiérarchique précédente et constitue l'entrée gauche du heat map. Il en est de même pour les caractéristiques du rejet qui elles constitueront l'entrée supérieure. Nous obtenons ainsi la visualisation suivante :



Ce procédé apporte aux données un aspect visuel plus facile à regarder qu'un tableau de chiffres. En effet, il est important de noter que dans la « heat map » les données sont exactement celles contenues dans la base de données, il n'y pas eu de transformation des données. Seules une classification hiérarchique des rejets et des caractéristiques impliquées dans l'analyse ainsi qu'un code couleur permettant de représenter graphiquement la valeur d'une caractéristique pour un rejet ont été appliqués afin d'obtenir ce résultat. Nous obtenons donc une photographie descriptive du rejet chez les 302 patients inclus dans l'étude en question. Le dendrogramme oriente vers l'existence de différentes signatures de rejet.

-rejet cellulaire sans lésion vasculaire : TCMR/V-
-rejet humoral sans lésion vasculaire : ABMR/V-
-rejet cellulaire avec lésion vasculaire : TCMR/V+
-rejet humoral avec lésion vasculaire : ABMR/V+

Il est alors intéressant de pouvoir confirmer l'existence de ces quatre entités de rejet dans le dendrogramme et de mieux les caractériser. C'est à ce stade qu'une technique nommée analyse en composantes principales peut s'avérer intéressante et complémentaire.

L'exemple proposé précédemment pour illustrer la construction de la « heat map » reposait sur des caractéristiques du rejet toutes considérées comme des variables catégorielles ordinales (0,1,2,3 ou 0,1 pour le DSA). Le code couleur a été choisi pour caractériser le gradient de ses différentes modalités. Il est néanmoins tout à fait possible d'envisager aussi l'analyse exclusive de variables continues ou encore en complément de variables catégorielles ordinales. Dans cette configuration, afin de rendre le code couleur commun à l'ensemble des variables analysées, l'astuce consiste en une normalisation de la valeur de l'expression du paramètre chez les sujets relativement à la médiane d'expression du paramètre dans l'ensemble de la population analysée (ratio valeur du paramètre chez le sujet i / médiane d'expression du paramètre dans la population). Aussi, pour chaque paramètre chacun des sujets a une valeur numérique qui représente une sur-expression si la valeur est supérieure à 1, une sous-expression si la valeur est inférieure à 1 ou encore une expression moyenne si la valeur obtenue est égale à 1. Un code couleur commun à l'ensemble des paramètres analysés peut alors être établi en définissant par exemple un gradient du vert clair au vert foncé pour la sous-expression et du rouge clair au rouge foncé pour la sur-expression.

III.1.2.2 Analyse en Composantes Principales

Cette analyse fait aussi partie de la statistique descriptive multidimensionnelle car elle permet de traiter simultanément un nombre

quelconque de variables et ne requiert pas d'hypothèse sur le type de variables quantitatives analysées telle que la normalité de leur distribution par exemple. Le principe de l'Analyse en Composantes Principales (ACP) est donc d'étudier simultanément un nombre potentiellement important de variables, pour en faire une représentation graphique globale et simplifiée. La difficulté vient du fait que les individus étudiés, s'ils sont décrits par n variables, ne peuvent plus être représentés dans un plan de dimension 2, mais dans un espace de dimension plus importante (par exemple 4). L'objectif de l'ACP est donc de revenir à un espace de dimension réduite (par exemple 2) en déformant le moins possible la réalité, c'est-à-dire en conservant le maximum de l'inertie contenue dans le nuage de point initial. Autrement dit, il s'agit d'obtenir le résumé le plus pertinent possible des données originelles. C'est la matrice des variances-covariances (ou celle des corrélations) qui va permettre de réaliser ce résumé pertinent, car dans l'ACP est analysée essentiellement la dispersion des données considérées. De cette matrice, par un procédé mathématique approprié, les facteurs que l'on recherche (valeurs propres ou axes principaux) sont extraits en petit nombre. Ils vont permettre de réaliser les graphiques dans un espace de deux dimensions (avec les axes principaux retenus), en déformant le moins possible la configuration globale des individus telle qu'elle est définie par l'ensemble des variables initiales. C'est l'interprétation de ces graphiques qui permettra de comprendre la structure des données analysées.

Sur le plan théorique, l'ACP est une méthode relativement complexe, dans la mesure où elle fait appel à des notions mathématiques non élémentaires notamment celles de matrices et des éléments propres (Venables & Ripley, 2003; Dray & Dufour, 2007; Baccini, 2010). La partie théorique de l'ACP n'étant pas l'objet de cette thèse elle ne sera pas présentée. Il n'est pas nécessaire de connaitre ces notions pour comprendre le mécanisme d'une ACP et donc pour l'utiliser correctement.

Néanmoins très simplement, le principe mathématique général de l'ACP est la réduction de la dimension des données initiales (qui est p si l'on considère p variables), en trouvant q facteurs appropriés ($q < p$). L'idée est donc une réduction de dimension de p à q.

Si nous considérons p variables quantitatives, notées X^1, …,X^j, . . . , X^p, observées sur n individus, notés 1, . . . ,i, . . . , n.
Les données initiales se présentent donc sous la forme suivante

	X^1	$\cdots$	X^j	$\cdots$	X^p
1	x_1^1	$\cdots$	x_1^j	$\cdots$	x_1^p
$\vdots$	$\vdots$		$\vdots$		$\vdots$
i	x_i^1	$\cdots$	x_i^j	$\cdots$	x_i^p
$\vdots$	$\vdots$		$\vdots$		$\vdots$
n	x_n^1	$\cdots$	x_n^j	$\cdots$	x_n^p

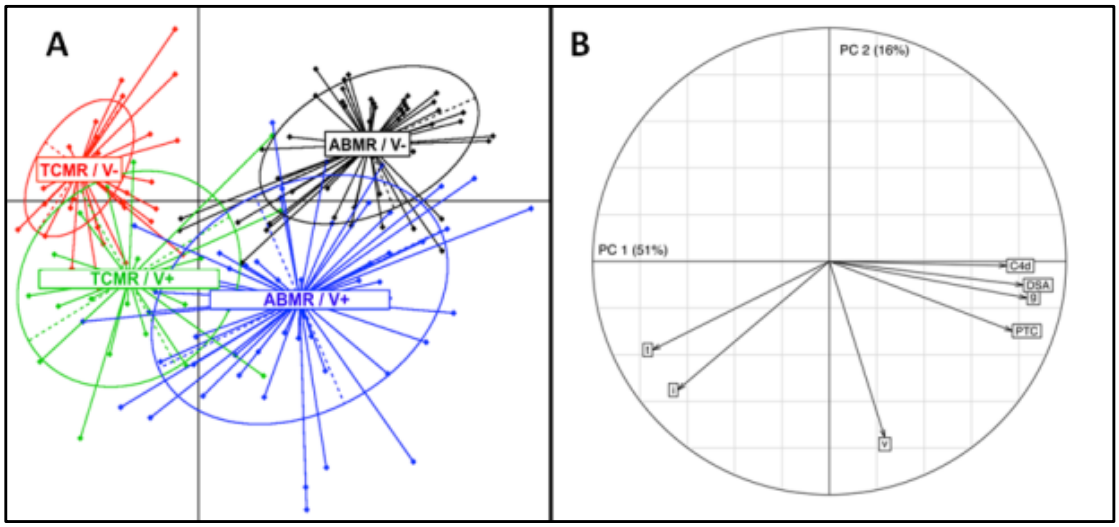
On cherche ensuite à résumer l'information contenue dans ce tableur en construisant des combinaisons linéaires des variables initiales, appelées facteurs, ou encore composantes principales, s´écrivant sous la forme suivante :

$$C^1 = a_1^1 X^1 + a_1^2 X^2 + \cdots + a_1^p X^p$$
$$C^2 = a_2^1 X^1 + a_2^2 X^2 + \cdots + a_2^p X^p$$
$$\cdots$$

C^1 devant contenir un maximum d'information. C'est cette étape qui implique des considérations mathématiques plus complexes que nous n'aborderons pas ici.
On réalise ensuite la même chose pour C^2, en imposant, en plus, que C^1 et C^2 ne soient pas corrélées (pour que l'information apportée par C^2 ne soit pas redondante à celle contenue dans C^1). Et ainsi de suite pour C^3, C^4, …
Enfin, des représentations graphiques en deux dimensions sont faites en générale sur C^1 et C^2 qui expliquent le maximum de l'inertie du tableur.

Sur le plan pratique, en complément de l'analyse de classification hiérarchique précédemment exposée, nous nous sommes appuyés sur l'ACP pour déterminer des entités distinctes de rejet correspondant à des associations différentes de lésions du greffon caractérisant des processus physiopathologiques distincts. Ainsi, nous nous sommes affranchis de l'interprétation diagnostique de la classification de Banff en considérant les

variables histologiques (glomérulite, capillarite péritubulaire, infiltration interstitielle, tubulite, artérite intimale, marquage C4d, et DSA). Cette approche présente l'avantage d'être réalisée sans supervision et permet dans notre cas de s'affranchir de l'interprétation (parfois erronée) de la classification de Banff. Voici dans la figure ci-dessous les résultats de l'ACP ainsi obtenus :



Panel A: Projection des 302 patients avec diagnostic de rejet aigu confirmé histologiquement sur les 2 premiers axes de l'ACP en utilisant 7 variables: DSA, glomérulite, inflammation interstitielle, tubulite, endartérite, capillarite péritubulaire et C4d.
TCMR/V- rejet médié par les lymphocytes T sans endartérite, ABMR/V- rejet médié par les anticorps sans endartérite, TCMR/V+ rejet vasculaire médié par les lymphocytes T et ABMR/V+ rejet vasculaire médié par les anticorps.
Panel B: Le cercle de corrélations permet de caractériser les axes sur lesquels ont été projetés les patients. L'axe horizontal est défini par DSA/glomérulite/capillarite péritubulaire/C4d en opposition à l'inflammation interstitielle/tubulite. L'axe vertical est défini par la présence de lésions d'endartérite.

On retrouve donc les 4 rejets pré-identifiés par la visualisation de la « Heat map ».

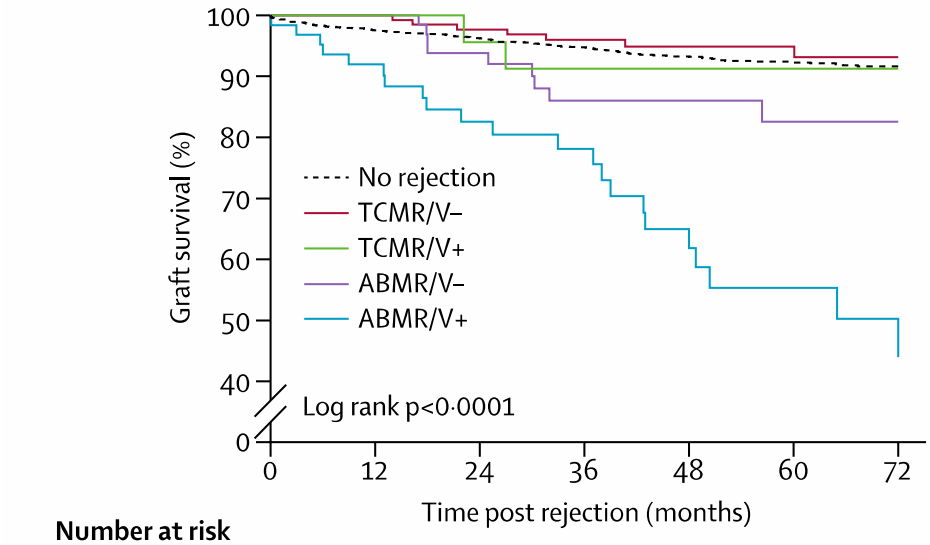
Pour précision, les ellipses du Panel A sont construites de telle sorte que le centre de l'ellipse représente le centre d'inertie du nuage de points appartenant au même groupe et que la probabilité d'être à l'intérieur de l'ellipse pour un point du groupe soit égale à $p = 1 - e^{-0.5 \times k^2}$. Ici k = 1.5 donc

p = 0.67 (~70%) ce qui veut dire que l'intérieur de l'ellipse contient 70% des points du groupe.

Nous venons donc de voir comment nous pouvons par l'utilisation successive de la classification hiérarchique et de l'ACP identifier des phénotypes de patients ou de rejets en s'affranchissant des classifications pré-établies.

Pour finaliser ce raisonnement, mais cela dépasse l'identification et la caractérisation de phénotypes homogènes qui est l'objet de cette partie, nous pouvons faire l'hypothèse que ces différentes entités identifiées puissent se comporter différemment d'un point de vue purement pronostique, c'est-à-dire en termes de survie du greffon.

Comme présenté ci-dessous sur la courbe de Kaplan-Meier, la nouvelle entité de rejet médié par les anticorps avec présence de lésion vasculaire exhibe un profil de survie du greffon dégradée.



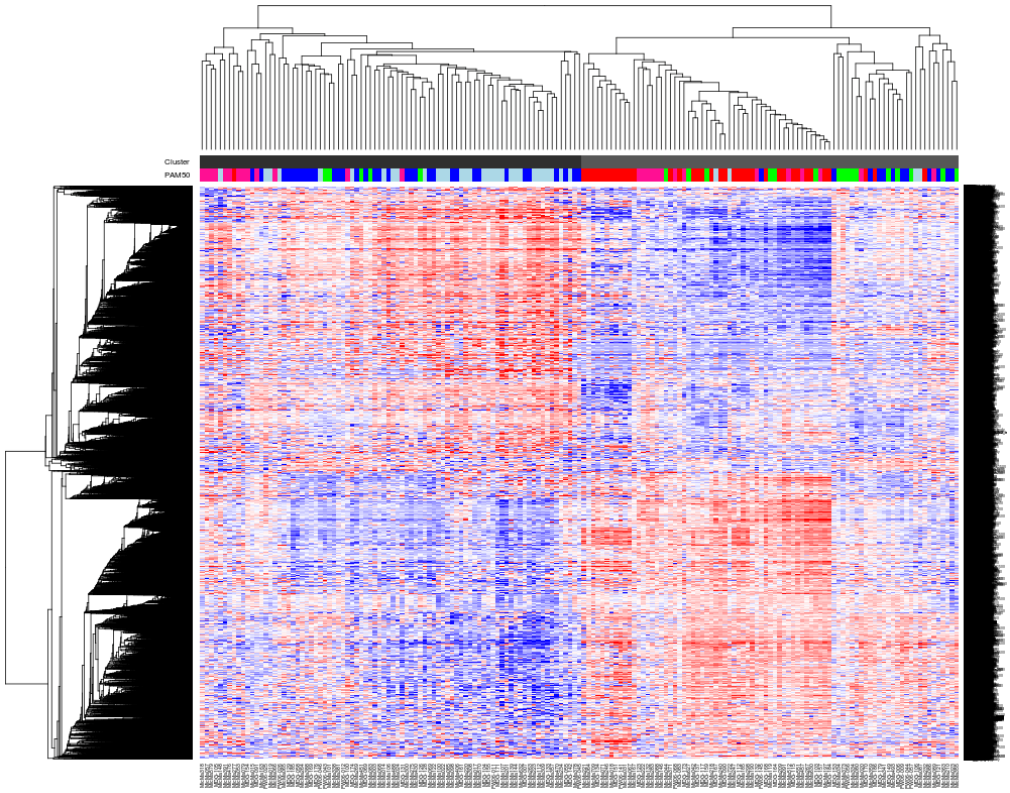
Number at risk

No rejection	1777	1600	1408	1152	933	673	473
TCMR/V–	139	136	121	101	77	54	29
TCMR/V+	26	26	23	20	15	12	4
ABMR/V–	73	68	56	42	28	19	7
ABMR/V+	64	52	41	32	21	12	9

Les résultats seront discutés plus précisément dans la partie résultats de cette thèse où l'étude en question est présentée dans son intégralité.

En conclusion, l'analyse en classification hiérarchique produisant une « Heat map » a permis de visualiser des clusters de rejets qui peuvent se différencier sur la base des caractéristiques usuellement utilisées pour la définition des rejets aigus. Autrement dit la « Heat map » est une photographie des rejets qui permet de mettre en évidence de l'hétérogénéité dans l'expression des rejets. Sur la base d'éléments encourageant observés sur la « Heat map », nous avons ensuite caractérisé les différentes formes de rejets par une ACP. Enfin, nous avons démontré que l'identification de ces phénotypes de rejets distincts est intéressante d'un point de vue clinique par l'observation d'un pronostic de survie du greffon défavorable associé à l'un d'entre eux ; le rejet humoral avec lésion vasculaire – ABMR/V+ ; alors qu'au jour de l'étude il n'était pas considéré comme une entité à considérer spécifiquement.

En sciences médicales, ces approches et méthodes (ACP et Heat map) sont des techniques qui sont classiquement employées dans le domaine de l'épidémiologie génétique lorsque des milliers de gènes sont analysés conjointement. En effet, une « Heat map » permet même avec l'expression de milliers de gènes de voir si plusieurs clusters de patients ou de tissus se distinguent de par leur expression différentielle comme illustré dans la figure ci-dessous :



Cette manière d'identifier des phénotypes de populations ou de pathologies plus homogènes lorsqu'on souhaite s'affranchir des conventions ou classifications usuelles implique des processus statistiques analytiques assez complexes et constitue donc une étape importante dans une recherche à visée pronostique.

L'exemple permettant d'illustrer l'ACP et la Heat map ayant pour objectif l'identification des partitions de rejets qui peuvent se distinguer sur la base des paramètres pathologiques classiquement utilisés dans le diagnostic de rejet aigu (DSA circulants, dépôt de C4d dans le greffon, glomérulite, capillarite péritubulaire, endartérite, tubulite et inflammation interstitielle) est issu d'une publication parue dans le Lancet en 2013.
Cette étude a permis d'identifier une nouvelle forme de rejet aigu en transplantation rénale et est présentée en détail en annexe VIII.1.1 comme publication complémentaire à la thèse servant de support pour la partie méthode.

Lefaucheur C, Loupy A, Vernerey D, Duong-Van-Huyen J-P, Suberbielle C, Anglicheau D, Vérine J, Beuscart T, Nochy D, Bruneval P, Charron D, Delahousse M, Empana J-P, Hill GS, Glotz D, Legendre C, Jouven X
Antibody-mediated vascular rejection of kidney allografts: a population-based study. **Lancet** 381: 313–319, doi:10.1016/S0140-6736(12)61265-3. (2013)
Impact factor: 44.0

III.2 Recherche de facteurs pronostiques de l'évènement d'intérêt

Une fois la population d'intérêt précisément identifiée, s'en suit naturellement la recherche de facteurs pronostiques de l'évènement d'intérêt comme illustré dans la figure ci-dessous.

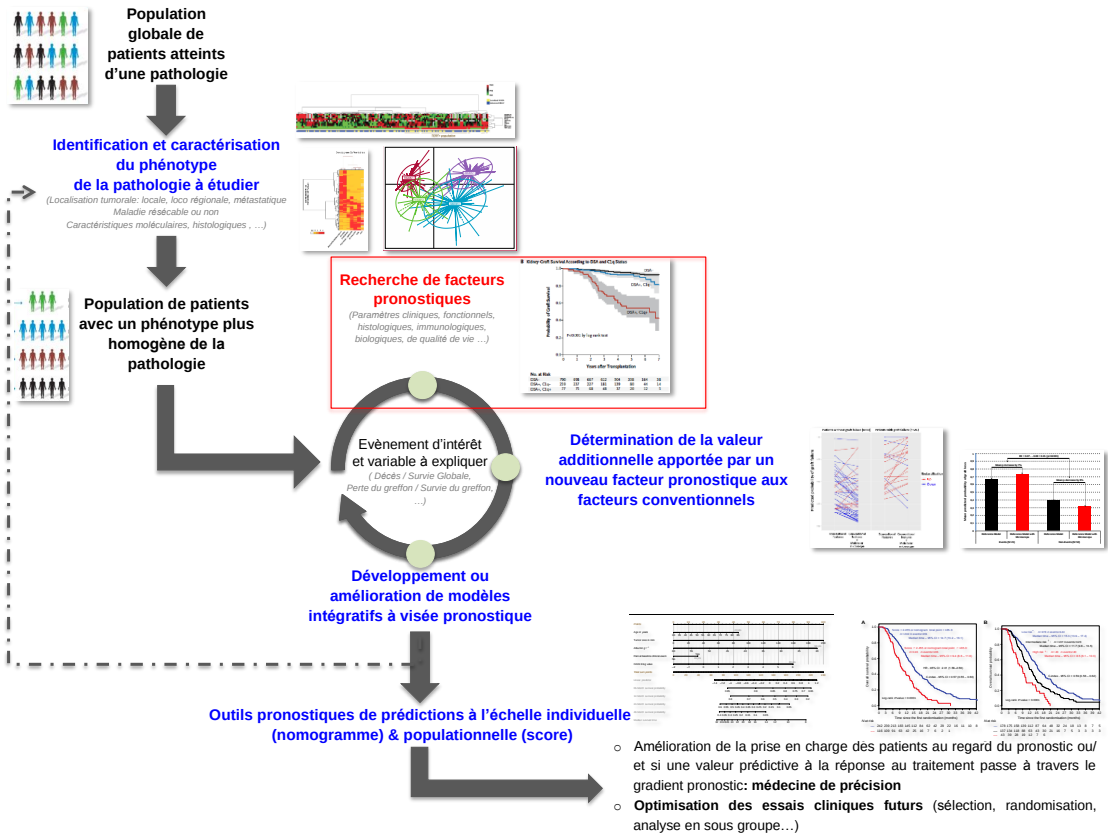


Figure : Seconde étape d'une recherche à visée pronostique : la recherche de facteurs pronostiques de l'évènement d'intérêt.

En statistique le terme pronostic n'a de sens que si l'évènement que l'on cherche à mesurer ou prédire est précisé. Si par exemple, nous considérons la survie globale (SG) comme l'entité à modéliser, nous parlerons alors de la recherche de l'identification de facteurs pronostiques de

la SG. Dans les essais thérapeutiques, cette entité est aussi appelée critère de jugement. C'est en effet sur cette mesure qu'est évaluée l'efficacité de la stratégie thérapeutique expérimentale. En recherche pronostique, nous parlons plus de l'évènement d'intérêt que l'on souhaite prédire ou de la variable à expliquer.

De manière générale en sciences médicales, les évènements que nous cherchons à expliquer et prédire sont des évènements dichotomiques en raison de leur simplicité et pertinence en pratique clinique.

III.2.1 Choix de l'évènement d'intérêt et de la variable à expliquer

Les évènements sont des mesures cliniques et biologiques qui peuvent être classés en deux catégories principales : les « évènements centrés sur le patient » comprenant le décès et la détérioration de la QdV et les « évènements centrés sur la tumeur » comme la progression tumorale.

III.2.1.1 Evènements centrés sur le patient

Les évènements centrés sur le patient sont des variables qui mesurent comment un patient survit ou perçoit son état de santé. Classiquement, de tels critères de jugement comprennent en cancérologie et en transplantation le décès ou une détérioration de QdV. Ils reflètent ainsi une mesure clinique directe pour le patient.

III.2.1.1.1 Le décès

En oncologie, le décès est considéré comme l'évènement de référence par sa précision, son évaluation continue et sa pertinence clinique.
Il est usuel de considérer le décès quelle que soit la cause. Mais parfois la précision de décès lié au cancer, qu'il soit lié à la tumeur primitive ou à un second cancer, peut être apportée. Dans cette situation, les décès liés à un

autre cancer, les décès non liés au cancer et les décès liés au traitement ne sont plus considérés comme un évènement. La difficulté est d'identifier la cause spécifique du décès en l'absence d'autopsie.

III.2.1.1.2 Les évènements reportés par la patient (Patient Reported Outcome) et en particulier la qualité de vie relative à la santé

En 1993, l'organisation mondiale de la santé (OMS) a défini la qualité de vie (QdV) comme « la perception qu'a un individu de sa place dans l'existence, dans le contexte de la culture et du système de valeurs dans lequel il vit, en relation avec ses objectifs, ses attentes, ses normes et ses inquiétudes. Il s'agit donc d'un large champ conceptuel, englobant de manière complexe la santé physique de la personne, son état psychologique, son niveau d'indépendance, ses relations sociales, ses croyances personnelles et sa relation avec les spécificités de son environnement » (WHOQOL, 1993).

Bien qu'il n'existe pas de consensus autour de la définition de la QdV, elle est généralement considérée comme un concept multidimensionnel qui inclut au minimum le bien-être physique, psychologique et social, et aussi les symptômes liés à la maladie et aux traitements.

La QdV entre dans le champ des « Patient-Reported Outcomes » (PROs), c'est-à-dire des mesures de l'état de santé perçue par le patient (Doward & McKenna, 2004). Ces mesures doivent donc être rapportées par le patient lui-même. Généralement, la QdV est mesurée par le biais de questionnaires administrés aux patients. Plusieurs dimensions de la QdV sont adressées par les questionnaires et en général certaines sont mises en avant selon la pathologie et le contexte de l'étude.

L'évènement d'intérêt peut alors être une détérioration d'une dimension de la QdV. Contrairement au décès, où son évaluation peut être considérée comme continue, se pose alors la question du rythme d'évaluation de la QdV. Les évaluations doivent être assez nombreuses pour pouvoir capturer un changement significatif d'un point de vue clinique. La fréquence des évaluations dépend de la pathologie, de la situation thérapeutique ainsi que de la probabilité d'observer un changement de la QdV durant une période

particulière. Un autre élément important est la définition de la différence minimale cliniquement importante (DMCI) qui représente le plus petit changement de score de QdV cliniquement pertinent. Selon la pathologie étudiée le rythme d'évaluation et la DMCI sont des éléments qui sont en général bien définis.

Dans cette thèse le seul évènement centré sur le patient est le décès quelle que soit la cause.

III.2.1.2 Evènements centrés sur la maladie

Les événements centrés sur la maladie (ici la tumeur ou l'organe transplanté) sont des marqueurs biologiques spécifiques de la pathologie étudiée et ainsi ne représentent pas une mesure clinique directe pour le patient.

Dans cette thèse, les évènements centrés sur la maladie sont la progression tumorale en oncologie et la perte du greffon en transplantation.

III.2.1.2.1 La perte du greffon

En transplantation, la perte du greffon est le critère de jugement le plus usuel que nous pouvons considérer centré sur la pathologie. Il est défini par le retour en dialyse du patient.

III.2.1.2.2 La progression tumorale

En oncologie, l'amélioration de la survie des patients, et donc la difficulté d'observer avec un délai acceptable pour le patient et raisonnable financièrement les évènements tels que le décès, a amené à considérer l'évaluation d'évènements intermédiaires tels que la progression tumorale. Ainsi, ces critères intermédiaires qui pourraient être utilisés comme critères de substitution au décès sont de plus en plus étudiés. Néanmoins dans ce contexte de substitution une validation statistique est nécessaire ce qui est rarement le cas et pose souvent des problèmes d'interprétation.

L'évènement de progression tumorale est codifié très précisément par les critères RECIST. Dans les critères RECIST v1.1, la progression est définie comme une augmentation d'au moins 20% de la somme des diamètres des lésions cibles, en prenant comme référence la plus petite somme de l'étude (ce qui inclut la mesure à l'inclusion si elle est la plus petite de l'étude). Cependant, il existe une variabilité dans la définition de la progression (clinique, radiologique (RECIST) ou biologique). Par exemple, la progression clinique non radiologique est souvent subjective, avec un manque de définition standardisée pouvant constituer une source potentielle de biais.
La progression est évaluée par des examens radiologiques à des intervalles réguliers dont le rythme est souvent propre au type de maladie tumorale. Dans la pathologie pancréatique et colorectale ce rythme est usuellement de 3 mois. Il est donc important de noter que la mesure d'une progression tumorale peut être sujette à des erreurs de mesure et à des biais ce qui la rend parfois difficilement comparable d'une étude à l'autre.

Cependant, l'évènement de progression tumorale constitue une mesure objective de l'activité biologique de la maladie ce qui peut s'avérer intéressant dans certains contextes notamment lorsqu'on cherche à démontrer l'activité d'une drogue sur la maladie tumorale. C'est en particulier l'objectif des essais de phase 2.

III.2.1.3 Critère de jugement du type « temps jusqu'à événement » ou censuré à droite

Tous les évènements qu'ils soient centrés sur le patient ou sur la maladie peuvent être considérés à un temps donné et être analysés de la sorte. La recherche pronostique consistera donc en la recherche de déterminants associés à la survenue de l'évènement au cours de la période d'observation qui précède le temps considéré. Cela revient à une étude cas-témoin, les cas ayant expérimenté l'évènement contrairement aux témoins, pour laquelle la méthode d'analyse statistique est classiquement la régression logistique. La régression logistique ou modèle logit est un modèle de régression binomiale

qui est un cas particulier du modèle linéaire généralisé. Il s'agit de modéliser l'effet d'un vecteur de variables aléatoires sur une variable aléatoire binomiale.

Ainsi, des évènements se produisant au tout début de l'étude ou juste avant le temps T seront considérés de la même manière indépendamment de leur occurrence qui peut être temporellement très différente.

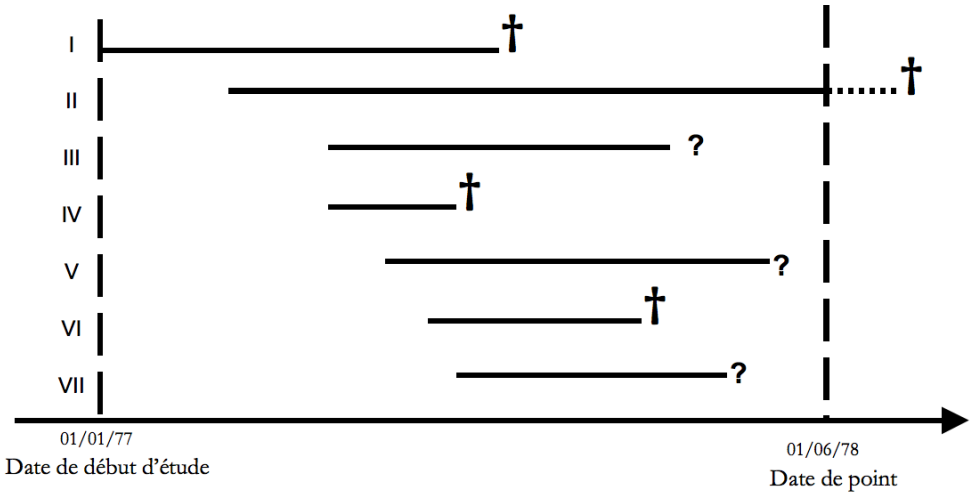
Cette approche peut s'avérer intéressante, mais lorsque les évènements peuvent être évalués longitudinalement sur des périodes importantes, il est plus informatif et surtout plus précis de considérer le temps jusqu'à occurrence de ce dernier. Nous nous retrouvons donc dans un contexte de données censurées à droite. C'est cette démarche qui a été privilégiée dans cette thèse.

Il s'agit alors d'analyse de données de survie. Le terme « durée de survie » désigne génériquement le temps écoulé jusqu'à la survenue d'un évènement. Cet évènement n'est pas nécessairement le décès mais peut être la progression tumorale (temps entre le diagnostic et la progression) ou encore la perte du greffon (temps entre la transplantation et la perte du greffon). L'analyse des données ou durées de survie repose donc sur l'étude du délai de survenue de cet évènement à compter d'une date dite d'origine. On cherche alors à estimer la distribution des temps de survie (fonction de survie) et, dans le cadre de la recherche pronostique à analyser, la manière dont des variables explicatives peuvent modifier cette fonction de survie.

Une des caractéristiques des données de survie est l'existence d'observations incomplètes. En effet, les données sont souvent recueillies partiellement, notamment, à cause du processus de censure. En médecine et en biologie, les situations les plus courantes sont celles où l'on étudie, dans un échantillon de données, des événements à venir afin de faire ultérieurement des prédictions pour d'autres populations. Ceci conduit naturellement à l'observation de données censurées à droite. Une durée de survie, dont on sait seulement qu'elle est supérieure à un délai donné, constitue ce que l'on appelle une observation censurée à droite. Autrement dit en présence de censure à droite, les durées de vie exactes ne sont pas

toutes observées. En effet, pour certaines d'entre elles, on sait seulement qu'elles sont supérieures à une valeur connue.

Schématiquement, dans le cas de l'étude de l'événement « décès » par exemple, les différentes observations possibles peuvent être représentées de la sorte :



Pour les sujets I, IV et VI, l'évènement a été observé avec une durée connue. Le sujet II est considéré comme n'ayant pas expérimenté l'évènement à la date de point, il est donc censuré en termes de suivi à la date de point et son statut est vivant. Les observations III, V et VII représentants quant à elle des sujets dont on ne connaît pas l'état à la date de point, c'est-à-dire des sujets perdus de vue. Ils sont donc censurés à leur date de dernière nouvelle avec un statut vivant. On cherche toujours à réduire le nombre de sujets perdus de vue, d'une part parce que ces observations représentent une perte d'information et d'autre part parce qu'elles sont source de biais. On n'est en effet jamais certain que les sujets perdus de vue ont eu une évolution similaire à celle des autres sujets.

Aussi chaque patient ou observation est défini par une durée et un état au regard de l'évènement considéré. Au sujet i on associe en outre la variable d_i indicatrice de son état au temps t_i. Cette variable vaut 1 si, en t_i, on a déjà observé l'évènement chez le sujet avant la date de point. Elle vaut 0 si le sujet

n'a pas encore eu l'évènement en t_i. C'est le couple (t_i,d_i) qui pour chaque sujet contient l'information minimum nécessaire aux calculs de survie.

Supposons que l'on étudie n sujets, chacun est suivi pendant une période d'observation au maximum égale à L_i (i=1, 2,.., n). Le délai exact de survie du sujet i ne sera connu que si :

$$T_i \leq L_i$$

où T_i est la variable aléatoire temps de survie du sujet i. On observe alors l'évènement chez le sujet i, et la variable indicatrice de l'état aux dernières nouvelles d_i vaut 1. Dans le cas contraire :

$$T_i > L_i$$

l'observation est alors censurée et l'état aux dernières nouvelles d_i vaut 0.

On peut représenter les observations par n couples de valeurs aléatoires : (t_i, d_i) où $t_i = \min(T_i, L_i)$ et $d_i = \begin{cases} 1 & T_i \geq L_i \\ 0 & T_i > L_i \end{cases}$

Plusieurs mécanismes de censure peuvent être envisagés notamment des mécanismes aléatoire et non aléatoire. L'hypothèse d'un mécanisme aléatoire est souvent faite dans la littérature à juste titre mais peut parfois être discutée. En effet, par exemple en transplantation, lorsque nous étudions la survie du greffon, on considère régulièrement que le décès du patient avec un greffon fonctionnel engendre une censure pour la durée de survie du greffon. C'est-à-dire que nous reportons le greffon fonctionnel jusqu'au décès du patient avec une censure du suivi du statut du greffon au moment du décès du patient. Il est plus facile d'admettre que le mécanisme de censure est aléatoire pour les causes accidentelles de décès que pour les autres causes qui pourraient, peut-être, être liées à la transplantation et ses conséquences. C'est en réalité rarement le cas et donc cela ne pose pas de problème de considérer un mécanisme de censure aléatoire mais il est important de se poser la question. De plus dans les cohortes étudiées, les patients qui décèdent avec un greffon fonctionnel ne représente en général qu'une très faible proportion des sujets (<5%). Dans cette situation précise, il est possible de considérer les

évènements décès et perte du greffon comme des évènements compétitifs puisque l'un peut empêcher d'observer l'autre. Une analyse en risque compétitif est alors à envisager mais avec une fréquence faible de l'évènement compétitif (<5%) cela produit souvent des résultats similaires à la censure, c'est en tout cas ce qui a été observé dans les travaux présentés dans cette thèse.

III.2.1.4 Définition des variables du type « temps jusqu'à évènement» utilisées dans cette thèse

III.2.1.4.1 En transplantation

III.2.1.4.1.1 La survie globale du greffon

La survie globale du greffon est définie comme le temps entre la date d'origine ou d'inclusion dans l'étude (transplantation, rejet ou biopsie, etc.) et le retour en dialyse du receveur (débit de filtration glomérulaire entre 8 et 12 mL/min/1.73m2). Les greffons provenant de receveurs perdus de vue ou décédés ont été censurés comme étant fonctionnels au moment de la date des dernières nouvelles ou du décès.

III.2.1.4.1.2 La survie globale du patient

La survie globale du patient est définie comme le temps entre la date d'origine ou d'inclusion dans l'étude (transplantation, rejet ou biopsie, etc.) jusqu'au décès quelle que soit la cause. Les patients perdus de vue ou vivants à la date de point sont censurés comme étant vivants à la date des dernières nouvelles.

Dans la base DIVAT (Données Informatisées et VAlidées en Transplantation)

d'où provient la majeure partie des données de suivi analysées dans cette thèse, les données de survie des patients retournés en dialyse ne sont pas collectées. Ainsi, avec la base de données DIVAT, il n'est pas possible d'étudier la mortalité globale car l'information du décès est le plus souvent manquante pour les patients retournés en dialyse. Aussi lorsque nous étudions la survie globale, il s'agit toujours du décès avec un greffon fonctionnel.

III.2.1.4.2 En oncologie

III.2.1.4.2.1 La survie sans progression

En général, la survie sans progression est le temps entre la date d'origine ou d'inclusion dans l'étude (diagnostic, début du traitement, etc.) et la progression tumorale (clinique, radiologique (RECIST) ou biologique) ou le décès quelle que soit la cause.

Dans la situation où l'entité d'intérêt est composite et inclut différents évènements, il est alors recommandé de respecter des consensus de définitions, s'ils existent, afin de proposer des recherches homogènes et comparables. L'initiative DATECAN (Definition for the Assessment of Time-to-event Endpoints in CANcer trials) 1 et 2 qui vise à obtenir au niveau international un consensus basé sur l'opinion d'experts pour la définition des critères de jugement en cancérologie, propose de telles recommandations pour le pancréas (Bonnetain *et al*, 2014), le rein (Kramar *et al*, 2015), les sarcomes (Bellera *et al*, 2015) et le sein (Gourgou-Bourgade *et al*, 2015).

III.2.1.4.2.2 La survie globale

La survie globale est définie comme le temps entre la date d'origine ou d'inclusion dans l'étude (diagnostic, début du traitement, etc.) jusqu'au décès quelle que soit la cause. Les patients perdus de vue ou vivants à la date de point sont censurés comme étant vivants à la date des dernières nouvelles.

III.2.2 L'analyse des données de survie

III.2.2.1 La distribution de survie

Soit T ($T \geq 0$), la variable aléatoire qui représente la durée de survie définie comme le délai entre la date d'origine et la date de l'évènement d'intérêt.

La probabilité de survivre sans évènement au-delà de la valeur t se nomme alors fonction de survie, courbe de survie ou survie en t.

On a alors: $S(t) = P(T \geq t) = 1 - F(t)$

F(t) est donc la fonction de répartition de la variable aléatoire T qui est la probabilité de décéder entre 0 et t :

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

La fonction de survie S(t) est une fonction monotone décroissante et continue, telle que :

$$S(0) = 1 \qquad \text{et} \qquad \lim_{t \to \infty} S(t) = 0$$

On appelle risque instantané de faire l'évènement, ou plus correctement taux instantané d'évènement, la fonction h(t) définie par :

$$h(t) = \lim_{dt \to 0} \frac{P\left(t \leq T < t + dt \,\middle|\, T > t\right)}{dt}$$

h(t)dt représente donc pour un individu la probabilité d'expérimenter l'évènement entre t et t+dt, conditionnellement au fait que ce sujet n'a pas encore eu l'évènement au temps t. Cette fonction s'appelle aussi fonction de risque (on trouve à tort le terme de hasard, mauvaise traduction du mot anglais "hazard" qui signifie risque).

La fonction h(t) vérifie la relation suivante :

$$h(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = -\frac{d}{dt} Ln\left[S(t)\right]$$

On obtient donc:

$$S(t) = exp\left(-\int_0^t h(u)\, du \right)$$

Ainsi, le risque instantané de faire l'évènement ou la fonction de risque h(t) fournit la description la plus concrète d'une distribution de survie (Hill, 2009)

III.2.2.1.1Estimation de la distribution de survie

III.2.2.1.1.1 Estimation paramétrique

L'estimation paramétrique de la distribution de survie repose sur un modèle dans lequel la fonction de risque h(t) est une fonction mathématique dépendante d'un ou de plusieurs paramètres. Plusieurs fonctions peuvent être considérées (Lawless, 1982).

Le modèle le plus simple est le modèle exponentiel pour lequel la fonction h(t) est constante au cours du temps ($h(t) = h_0$). La fonction de survie est donc une simple fonction exponentielle. Ce modèle ne dépend que du paramètre h_0, on a donc :

$$S(t) = exp\,(-h_0 t)$$

$$f(t) = h_0\, exp\,(-h_0 t)$$

et : $E(T) = 1/h_0$ et $Var(T) = 1/(h_0)^2$

III.2.2.1.1.2 Estimation non paramétrique de Kaplan-Meier

La méthode de Kaplan-Meier repose sur l'idée intuitive suivante : n'avoir pas fait l'évènement après l'instant t, c'est être dépourvu d'évènement avant t et ne pas faire l'évènement à l'instant t (Kaplan & Meier, 1958). Dans la

pratique, le temps n'est pas mesuré avec une précision infinie et les données de survie sont exprimées dans une unité, en fonction de la précision avec laquelle sont relevées les informations. En recherche clinique, la précision habituelle est le jour dans les calculs de la survie.

N'avoir pas fait l'évènement à t jours implique d'avoir atteint le jour 1, le jour 2, ..., puis le jour t sans évènement. Cette affirmation se traduit en termes de probabilité par :

$$S(t) = P(T \geq t)$$
$$= P(T \geq t \mid T \geq t-1)\, P(T \geq t-1)$$
$$= P(T \geq t \mid T \geq t-1) \ldots P(T \geq 1 \mid T \geq 0)\, P(T \geq 0)$$

En notant Q_j la probabilité d'avoir atteint le jour j sans évènement conditionnellement au fait de ne pas avoir eu l'évènement juste avant ce jour. On a :

$$Q_j = Pr\,(T \geq j \mid T \geq j-1)$$
$$S(t) = Q_t\, Q_{t-1} \ldots\ldots Q_1 \times 1$$

On estime S(t) par le produit des estimations q_j de Q_j, où q_j est égal à la proportion observée de sujets ayant atteints le jour j sans évènement parmi ceux qui n'ont pas eu l'évènement juste avant ce jour.

Si m_j évènements ont été observés en j, on a :

$$q_j = (n_j - m_j) / n_j$$

où n_j est le nombre de sujets qui n'ont pas eu l'évènement juste avant j.

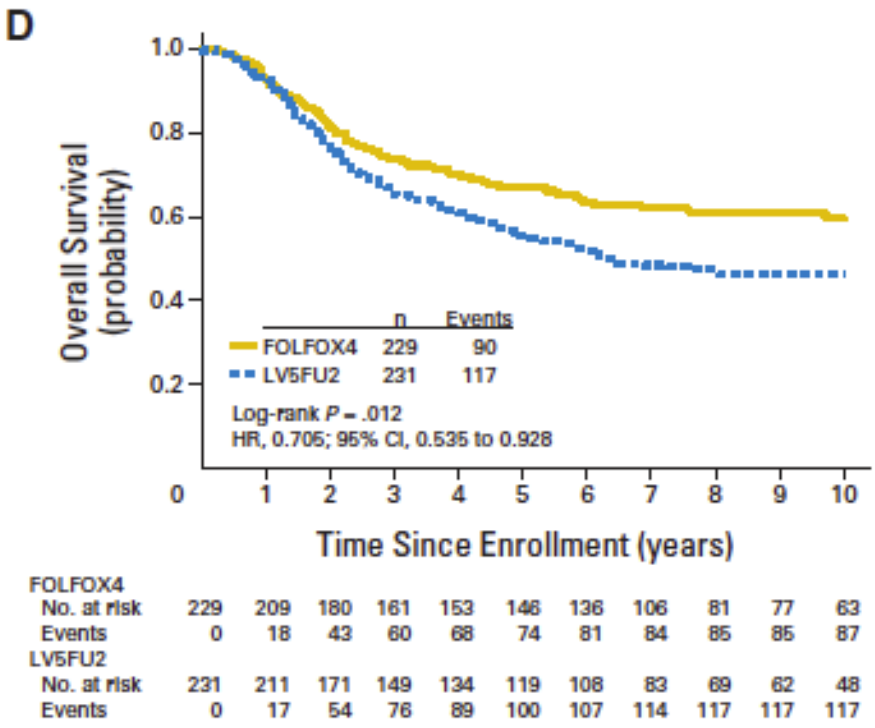
Si, au jour j, aucun évènement n'a été observé qj vaut 1, en conséquence dans les calculs seuls les jours, où des évènements ont été observés, interviennent. L'estimateur de la fonction S(t) est constant entre deux temps d'évènements consécutifs.

La représentation de l'estimation non paramétrique de Kaplan-Meier d'une distribution de survie se fait ainsi par un graphique en marche d'escalier, avec une marche à chaque temps d'évènement observé.

Afin d'être précis, il faut donner le nombre de sujets encore exposés au risque de décès en quelques points de la distribution de survie. On peut représenter sur le graphique les différents temps de censure par des traits verticaux.

On a souvent en fin de courbe un ou plusieurs plateaux qui correspondent en général à la présence d'un petit nombre de sujets suivis pendant une longue période. L'interprétation de cette apparente stabilité des survies à long terme doit être faite avec prudence car elle peut être plus le reflet d'un manque d'information que la démonstration d'une réelle disparition du risque à long terme.

Voici ci-dessous une courbe de Kaplan-Meier estimant la distribution de survie globale des patients atteints d'un cancer colorectal résécable sous deux types de chimiothérapie (LV5FU2 et FOLFOX4) (André *et al*, 2015)



La différence entre les distributions de survie de groupes de patients estimés et représentés sur une courbe de Kaplan-Meier peut être naturellement comparée visuellement et aussi objectivée par des tests statistiques.

III.2.2.1.1.3 Comparaison de la distribution de survie de deux ou plusieurs groupes estimés par la méthode non paramétrique de Kaplan-Meier : le test du log-rank

D'un point de vue statistique, le problème posé est celui de la comparaison de deux ou plusieurs échantillons de durées de survie éventuellement censurées. L'hypothèse nulle à tester est celle de l'identité des distributions de survie à comparer. L'avantage de l'approche non-paramétrique est qu'elle permet de ne faire que peu ou pas d'hypothèses sur la forme de la distribution de survie estimée.

Les différents tests non-paramétriques correspondants sont des extensions des tests de rangs, adaptées à la présence de données censurées.
Le test non paramétrique le plus connu et utilisé est le test qui s'appelle test de Mantel-Haenszel (Mantel, 1966), de Cox (Cox, 1972) ou du log-rank (Peto & Peto, 1972).
Dans le contexte de la comparaison de 2 groupes, sa statistique est la suivante:
Les temps d'occurrence de l'évènement ordonnés des deux échantillons réunis seront notés : t_1, t_2...t_i...t_k.
En t_i, les nombres d'évènements observés dans chacun des groupes A et B seront notés :

m_{Ai} et m_{Bi} avec $m_{Ai} + m_{Bi} = m_i$ et $m_i > 0$

Les nombres de sujets exposés au risque d'évènement en cet instant sont :

n_{Ai} et n_{Bi} avec $n_{Ai} + n_{Bi} = n_i$

soit $e_{Bi} = \frac{m_i n_{Bi}}{n_i}$

et la variance $v_i = m_i \frac{(n_i - m_i)}{(n_i - 1)} \frac{n_{Ai} n_{Bi}}{n_i^2}$

Alors la statistique du log-rank peut alors s'écrire :

$$X^2 = \frac{(O_B - E_B)^2}{\sum_{i=1}^{k} v_i}$$

avec $O_B = \sum_{i=1}^{k} m_{Bi}$ et $E_B = \sum_{i=1}^{k} e_{Bi}$

D'autres tests existent : Gehan (Gehan, 1965), Peto et Prentice (Prentice, 1978) et peuvent présenter des propriétés intéressantes dans des cas particuliers. Le test du log-rank demeure cependant à juste titre le plus utilisé.

La méthode de Kaplan-Meier permet donc d'estimer de façon non paramétrique la distribution de survie d'un échantillon et le test du log-rank rend possible la comparaison statistique toujours non paramétrique des estimations de la distribution de survie lorsque plusieurs groupes de patients sont considérés.

Ces résultats sont souvent les résultats d'intérêt dans des études comme les essais thérapeutiques. En effet, la randomisation répartit, en moyenne de façon identique dans les deux groupes, l'ensemble des paramètres individuels connus et inconnus, pouvant influencer la survenue de l'évènement. Ainsi, si une différence entre les estimations de distributions de survie de ces groupes est mise en évidence, elle sera généralement imputable au seul facteur qui les différencie : le traitement administré.

Si un facteur de confusion venait à demeurer malgré la randomisation la solution consiste alors à comparer les distributions de survie des groupes de traitement après ajustement, ou stratification, sur le facteur incriminé.

Cependant dans le contexte de cette thèse et généralement dans les études pronostiques, la représentation de distributions de survie par la méthode de Kaplan-Meier et leur comparaison statistique par le log-rank ne peut pas constituer un résultat définitif en soit. Elle est plutôt utilisée comme illustration de l'effet d'un paramètre d'intérêt identifié préalablement. En effet, dans les études prospectives où une cohorte est suivie dans le temps, on compare souvent l'occurrence de l'évènement dans le temps entre deux ou plusieurs groupes définis selon certaines caractéristiques des sujets. Mais l'absence de tirage au sort conduit à donner à l'interprétation des résultats un caractère plus descriptif que causal, si on se contente de l'approche de Kaplan-Meier.

On privilégiera donc les approches multivariées pouvant prendre en compte plusieurs paramètres permettant par exemple des ajustements.

III.2.2.1.2 Modélisation de la distribution de survie pour la prise en compte de plusieurs facteurs de risque

A nouveau nous pouvons distinguer deux grands types de modèles, les modèles paramétriques et les modèles non paramétriques.

III.2.2.1.2.1 Modélisation paramétrique

Les modèles paramétriques permettent d'écrire la distribution de survie comme une fonction paramétrique du temps. Celle-ci peut s'exprimer en fonction d'un ou plusieurs facteurs de risque susceptibles d'influer sur la survie. La modélisation paramétrique permet d'évaluer l'importance de l'effet propre de chaque facteur sur la survie.

En pratique, les modèles paramétriques sont rarement utilisés en recherche clinique. En effet, les données disponibles ne permettent généralement pas de justifier un choix quelconque entre plusieurs fonctions de risque paramétriques. On aura tendance à préférer les modèles non

paramétriques et en particulier le modèle semi-paramétrique proposé par Cox au début des années 70 (Cox, 1972) qui repose sur moins d'hypothèses. Ce choix est raisonnable puisqu'il est démontré que l'efficacité relative du modèle de Cox par rapport aux modèles paramétriques est généralement bonne.

III.2.2.1.2.2 Modélisation semi-paramétrique : le modèle de Cox

C'est le modèle le plus souvent rencontré dans le domaine médical car il permet de prendre en compte l'effet de plusieurs variables explicatives pour expliquer la survenue d'un événement (partie paramétrique) sans faire d'hypothèse paramétrique sur les distributions de survie et la fonction de risque instantané de base (partie non-paramétrique).

L'utilité de ce modèle dans de nombreux types d'études : ajustement sur des variables pronostiques dans un essai thérapeutique, identification des associations de variables pertinentes à des fins pronostiques, etc., a fait qu'il est devenu incontournable aujourd'hui.

Les modèles de régression pour des données de survie s'écrivent sous la forme d'une relation entre le risque instantané de faire l'évènement et les covariables considérées. Le modèle le plus souvent utilisé s'écrit :

$$h(t,Z) = h_0(t)\, c(b,Z) \quad \text{où} \quad \begin{cases} b' = [b_1,\ldots,b_p] \\ Z' = [z_1,\ldots,z_p] \end{cases}$$

Ceci revient à dire que le risque de faire l'évènement, pour un sujet de caractéristique Z à l'instant t, est le produit d'une fonction $h_0(t)$ qui ne dépend que du temps et d'une fonction c(b,Z) qui n'en dépend pas.

La fonction c(b,Z) dépend des caractéristiques Z du sujet, la dépendance étant mesurée par le vecteur des coefficients de régression b. On choisit en général:

$$c(b,Z) = exp\ (b' Z)$$

En effet la fonction exponentielle permet d'obtenir une fonction de risque positive sans contrainte sur les valeurs acceptables des coefficients.

Le modèle s'écrit finalement :

$$h(t,Z) = h_0(t)\ exp\ (b' Z)$$

C'est de cette manière que peut s'écrire un modèle de Cox. En effet, Cox a proposé de considérer $h_0(t)$ comme une fonction inconnue qu'on ne cherche pas à modéliser.

Dans le modèle de Cox, la relation entre le risque instantané et les covariables est log-linéaire, c'est-à-dire que *Ln*[h(t,Z)] est une fonction linéaire de Z. On a :

$$Ln\ [h(t,Z)] = Ln\ [h_0(t)] + b' Z$$

ou encore :

$$Ln\ [-Ln\ \{S(t,Z)\}] = Ln\ H_0(t) + b' Z$$

Par ailleurs, le rapport des risques instantanés pour deux sujets i et j de caractéristiques Z_i et Z_j ne dépend que de Z_i et Z_j et ne dépend pas du temps. On a :

$$\frac{h(t,Z_i)}{h(t,Z_j)} = \frac{\exp(b'Z_i)}{\exp(b'Z_j)} = \exp\left(b'\left(Z_i - Z_j\right)\right) = K$$

C'est-à-dire que les fonctions de risque instantané des deux individus i et j sont proportionnelles, et que leur rapport de proportionnalité ne dépend pas du temps t. Le rapport des taux instantanés d'évènements représente le risque relatif à l'instant t des sujets de caractéristiques Z_i par rapport aux sujets de caractéristiques Z_j. C'est le hazard ratio (note HR) des Anglo-Saxons. Le modèle de Cox est donc un modèle semi-paramétrique à risques proportionnels et log-linéaire. L'adéquation du modèle de Cox doit être

discutée au regard de ces caractéristiques.

L'intérêt du modèle multivarié de Cox est de nous donner les moyens de répondre à la question : existe-t-il un effet de la variable z_1 lorsque l'on contrôle les effets des autres variables $z_2,...z_p$?
Le modèle de Cox suppose donc que le modèle des risques instantanés proportionnels est approprié pour (z1, z2,...,zp). Ces hypothèses doivent être vérifiées.

La fonction de survie estimée par un modèle de Cox peut donc s'écrire de la sorte :

$$S(t,Z) = \exp^{\int_0^t h(u)du} = \exp^{\int_0^t h0(u)\exp^{bZ}du}$$

Comme évoqué ci-dessus on peut aussi écrire :

$$\ln\left(-\ln\left(S(t,Z)\right)\right) = \ln\left(H0(t)\right) + bZ$$

$$\exp^{\ln\left(-\ln\left(S(t,Z)\right)\right)} = \exp^{\ln\left(H0(t)\right)+bZ} = \exp^{\ln\left(H0(t)\right)} \times \exp^{bZ}$$

$$-\ln\left(S(t,Z)\right) = H0(t) \times \exp^{bZ}$$

$$\ln\left(S(t,Z)\right) = -H0(t) \times \exp^{bZ}$$

$$\exp^{\ln\left(S(t,Z)\right)} = \exp^{-H0(t)\times\exp^{bZ}}$$

$$S(t,Z) = S_0(t)^{\exp^{bZ}}$$

Plus simplement on écrira :

$$S\left(t,Z_j\right) = S_0(t)^{\exp\left(b'Z_j\right)} = S_0(t)^{\exp\left(\sum_{i=1}^{p} b_i\left(Z_{ij}-\overline{Z_i}\right)\right)}$$

III.2.2.1.2.2.1 Estimation et test des paramètres du modèle de Cox

Les modèles à risques proportionnels ont une interprétation en termes de

rapport de risque (RR). Leur validité repose sur l'hypothèse d'un risque constant au cours du temps. Si *z*1 est une variable à 2 modalités, *z*1 = {1/0}, parmi les *p* variables explicatives, le RR associé à cette variable s'écrit :

$$\mathrm{RR}_{z_1=1/z_1=0} = \frac{h_0(t_i)\ exp(\beta_1)\ exp\left(\sum_{j=2}^{p} \beta_j z_{i,j}\right)}{h_0(t_i)\ exp\left(\sum_{j=2}^{p} \beta_j z_{i,j}\right)} = exp(\beta_1) \quad \forall\ t_i$$

Les patients ayant la variable explicative {*z*1 =1} ont *exp*(*β*1) fois plus de risque d'échec que ceux ayant la variable explicative {*z*1 =0}. Le RR lié à la variable *z*1 est dit ajusté sur toutes les autres variables, autrement dit toutes les autres variables étant constantes par ailleurs. Il est constant au cours du temps puisqu'il ne dépend que du coefficient de régression *β*1.
L'hypothèse *H*0 intéressante à tester est l'égalité des fonctions de risque de chaque modalité de la variable explicative *z*1 :{*β*1 =0 ⇐⇒ RR*z*1=1/*z*1=0 =1}. L'hypothèse alternative *H*1 est le plus souvent bilatérale, c'est-à-dire la non-égalité des fonctions de risque de chaque modalité de la variable explicative *z*1 : {*β*1≠0 ⇐⇒ RR*z*1=1/*z*1=0 ≠1} Si *H*0 est rejetée, alors il semble qu'il y ait une différence de survie entre les 2 populations définies par la variable *z*1.

L'estimateur $\hat{b}$ du maximum de vraisemblance (L) du vecteur b des paramètres du modèle vérifie :

$$\frac{\partial L^*(\hat{b})}{\partial b} = 0$$

C'est ce vecteur b qui quantifie l'association entre les facteurs ou covariables incluses dans le modèle et la distribution de survie.

Tester l'association des covariables avec la distribution de survie revient à tester l'hypothèse nulle H_0 : b=0

Le Test de Wald (ou du maximum de vraisemblance) qui suit asymptotiquement, sous l'hypothèse nulle, une distribution de chi2 à p degrés de liberté si le vecteur b est de dimension p, permet de répondre à cette problématique.

C'est le résultat de ce test qui fournit les HR et p-value usuellement reportée dans la littérature pour quantifier l'association des covariables avec la distribution de survie ainsi que son degré de signification statistique.

III.2.2.1.2.2.2 Adéquation de l'hypothèse des risques proportionnels pour un modèle de Cox

Tous les modèles de survie, y compris le modèle de Cox, reposent sur des hypothèses précises. Ce dernier suppose que le rapport des risques instantanés correspondants à deux populations de caractéristiques différentes ne dépend pas du temps. Le risque instantané peut alors s'écrire comme le produit d'une fonction du temps $h_0(t)$ et d'une fonction g(Z) du vecteur Z de covariables :

$$h(t,Z) = h_0(t)\, g(Z)$$

En notant Z_j et $Z_{j'}$ les valeurs des covariables pour les sujets j et j', la relation :

$$h(t,Z_j) / h(t,Z_{j'}) = g(Z_j) / g(Z_{j'})$$

montre que le rapport des risques instantanés ne dépend pas du temps et donc la différence de leur logarithme est constante par rapport à t. On dit alors que les risques instantanés sont proportionnels.

Pour vérifier l'hypothèse de proportionnalité des risques, on peut utiliser deux méthodes complémentaires : graphique et analytique.

Pour étudier l'adéquation du modèle de Cox, des méthodes graphiques simples ont été proposées (Kay, 1984). Le principe est de considérer un modèle stratifié et d'estimer les logarithmes des fonctions de risque cumulé

pour chaque niveau i de la stratification. On peut alors étudier le graphe des estimations de $Ln[H_i(t)]$ en fonction de t ou de $\ln(-\ln(S_i(t)))$ en fonction de t, ce qui revient au même. Si les risques sont proportionnels, les courbes doivent avoir un écart constant au cours du temps. Si les courbes n'ont pas un écart constant, alors l'effet de la variable *z*1 n'est pas proportionnel et le modèle à risques proportionnels n'est pas adapté. Cette approche graphique est subjective. De plus, elle ne permet de tester qu'une seule variable, sans ajustement.

La seconde méthode consiste en l'analyse des résidus de Schoenfeld (Grambsch & Therneau, 1994). Cette approche analytique est plus objective. Cependant, elle fait des hypothèses sur la forme de la corrélation entre les résidus et le temps. De plus, ce test est peu puissant : on pourra voir des courbes divergentes alors que l'analyse des résidus de Schoenfeld ne permet pas de rejeter la proportionnalité. Dans la suite de ce rapport, en cas d'incohérence entre les résultats graphiques et l'analyse par les résidus de Schoenfeld, on privilégiera l'analyse graphique qui permet d'évaluer la pertinence clinique de prendre en compte ou non une non-proportionnalité.

Au-delà de l'hypothèse des risques proportionnels tous les modèles de survie supposent l'indépendance des temps de censure et des temps de survie. C'est pour cette raison que l'on essaie dans les études de limiter le nombre de patients perdus de vue. En effet, s'il n'y a aucun perdu de vue à la date de point, les censures ne concernant que des sujets de recul insuffisant, le mécanisme de censure peut être considéré comme complètement maîtrisé et indépendant de la survie des malades.

III.2.2.1.2.2.3 Codage des variables à inclure dans le modèle de Cox

Lorsqu'on a écrit que le risque instantané de faire l'évènement dépend des covariables de façon log-linéaire, le modèle est encore très loin d'être précisé. En effet, le codage avec lequel les covariables sont introduites dans le modèle repose aussi sur des hypothèses.

III.2.2.1.2.2.3.1 Variables catégorielles

Le cas des variables catégorielles est relativement simple. Il faut identifier une classe de référence pour que les estimations des rapports de risque (HR) des autres catégories soient homogènes à travers les différentes variables incluses dans le modèle. En général on choisira la classe la moins à risque comme classe de référence. Par exemple, dans l'étude de la survie globale de patients atteints d'un cancer pancréatique, le paramètre N (envahissement ganglionnaire Y : N ou 0/1) peut être considérer plutôt comme protecteur s'il est égal à 0 c'est donc cette modalité qui sera la classe de référence pour cette variable. Lorsqu'il n'y pas d'hypothèse a priori sur la relation de risque entre les différentes modalités d'une variable catégorielle, on choisira la classe avec l'effectif de sujets le plus important. C'est ce qui est par exemple usuellement fait en génétique pour l'étude de polymorphisme bi-allélique (ex : AA AT TT).
Il est aussi important de vérifier que les sujets de la population étudiée sont suffisamment représentés dans chacune des modalités de la variable catégorielle considérée. Si ce n'est pas le cas des regroupements sont nécessaires. Ces derniers peuvent se faire sur la base des connaissances de la maladie et aussi sur la base de l'observation des tendances lors de l'analyse avec la variable non regroupée. Par exemple, si nous considérons la variable T (taille tumorale) telle que proposée par la classification TNM, cette dernière est définie selon 4 modalités (T1, T2, T3, T4). Si les patients ayant une maladie T4 sont trop peu nombreux, alors il sera logique de les regrouper avec les T3.

Evidemment, toute variable catégorielle incluse dans un modèle de Cox doit vérifier l'hypothèse des risques proportionnels à travers les différentes modalités considérées pour son codage.

III.2.2.1.2.2.3.2 Variables quantitatives

Si on étudie la valeur pronostique d'une variable continue, il y a plusieurs manières d'envisager son codage.
Considérons par exemple l'âge calculé en années, le modèle de Cox ($h(t,z) = h_0(t)\ exp\ (bz)$) fait l'hypothèse très forte d'une relation log-linéaire de l'âge sur le risque instantané de faire l'événement. Ceci veut dire qu'une augmentation de 5 ans multiplie le risque par *exp*(5b) que l'on passe de 30 à 35 ans ou de 70 à 75 ans. Il n'est en général pas conseillé d'adopter un modèle aussi fort d'emblée.

En pratique, il est recommandé de commencer par diviser l'âge en k classes et regarder comment varie le risque relatif d'une classe d'âge à l'autre. Notons alors a_{min} et a_{max}, les valeurs extrêmes de l'âge sur la série observée, on adopte la partition suivante :

$$[a_{min}, a_1[\ ;\ [a_1,a_2\ [\ ;\ [a_2,a_3\ [\ ;\ [a_3,a_{max}]$$

$[a_{min}, a_1[$ peut être considéré comme la classe de référence puisque par exemple en oncologie, l'âge est a priori un facteur de mauvais pronostique pour la survie.

Si on introduit la variable âge catégorisée de cette façon en 4 classes avec des estimations de rapport de risque croissants linéairement à travers les catégories, on peut alors être un peu plus confiant quant à un effet log-linéaire de l'âge sur h(t,z).
En réalité les résultats de cette approche sont souvent en défaveur de la considération de la variable en continu sans transformation et difficiles d'interprétation puisqu'ils peuvent être dépendants des bornes des classes dont le choix est souvent arbitraire. Cependant, la catégorisation de l'âge et l'utilisation de variables indicatrices revient à représenter l'association entre l'âge et le risque d'évènement par un escalier, ce qui a peu de chance de correspondre à la réalité et donc ne constitue pas une solution acceptable non plus. Néanmoins, la pratique qui reste la plus courante en épidémiologie

pour inclure une variable quantitative dans un modèle de régression est de la transformer en variable qualitative en faisant des classes.

Les avantages mis en avant en faveur de la catégorisation d'une variable quantitative sont de deux types. Le premier est un argument concernant la simplicité de la présentation et de l'interprétation des résultats, qui en donnant des hazard ratios par catégorie sont plus facilement compréhensibles, notamment pour des non statisticiens. Le second argument est que les classes constituées sont souvent les catégories utilisées de façon habituelle dans les autres publications et facilitent ainsi la comparaison et la discussion des résultats obtenus. De plus d'un point de vue purement méthodologique, c'est une approche non paramétrique qui ne nécessite pas d'hypothèse de log-linéarité sur la forme de la relation entre les variables et la distribution du risque instantané.

En réalité, ces arguments sont démodés et la tendance actuelle est de souligner les inconvénients de la transformation d'une variable continue en catégories (Greenland, 1995; Altman & Royston, 2006; Bouyer, 2012).

Les arguments en défaveur de la catégorisation de variables quantitatives sont multiples. Tout d'abord le choix des seuils pour définir les classes peut conduire à une mauvaise modélisation de la relation entre la variable et le risque de faire l'évènement.

Ainsi, si la variable intervient comme variable d'ajustement dans l'analyse cela peut mener à une mauvaise prise en compte des phénomènes de confusion (Taylor & Yu, 2002).

Ensuite, la catégorisation d'une variable continue peut être assimilée à une erreur de mesure puisqu'on attribue la même valeur à toutes les observations d'une même catégorie (Altman *et al*, 1994). Comme pour toutes les erreurs de mesure, cela induit une perte de puissance pour le test de l'existence d'une association entre X et Y (Moser & Coombs, 2004). La perte d'information ou de puissance liée à la catégorisation devient peu importante au-delà de 5 classes (Connor, 1972) ce qui représente un nombre de catégories importantes au regard du nombre de sujets parfois modeste de certaines études.

Une perte d'information dans une étude dont l'objet est de fournir des estimations précises du risque de faire l'évènement n'est pas acceptable. Les travaux menés dans cette thèse s'inscrivent pour la plupart dans cette tendance, et notamment ceux dont le but est de fournir des outils d'estimation du risque.
Ainsi, malgré des habitudes solidement ancrées de catégorisation des variables quantitatives, il paraît clair que les inconvénients l'emportent. Des solutions méthodologiques ont donc été envisagées, développées et proposées (Greenland, 1995; Royston & Sauerbrei, 2008)

Pour modéliser une variable quantitative sans catégorisation, la relation linéaire garde une place privilégiée en raison de sa simplicité et de sa facilité d'interprétation et de présentation.

En dehors du modèle linéaire, il y a deux grandes catégories de méthodes de modélisation de la relation entre une variable quantitative et le risque de faire l'évènement selon que l'approche est non paramétrique ou paramétrique.

III.2.2.1.2.2.3.2.1 Approche non paramétrique : méthode des splines cubiques restreintes

Il s'agit de découper la courbe en intervalles et de modéliser la relation entre la variable quantitative et le risque de faire l'évènement séparément dans chaque intervalle.
Les fonctions appelées splines s'inscrivent dans ce type d'approche. Le principe général des fonctions splines est résumé par les étapes suivantes :

- L'ensemble des valeurs de la variable quantitative est découpé en (k+1) intervalles. On choisit pour cela k valeurs numériques (appelées nœuds) qui définissent les bornes des intervalles.
- Au sein de chaque intervalle, la relation entre la variable quantitative et le risque de faire l'évènement est modélisé par un polynôme de degré d.
- Les coefficients des polynômes sont choisis de sorte que la courbe totale soit la plus régulière possible (c'est-à-dire sans rupture et lisse), ce qui

se traduit sur le plan mathématique par le fait qu'elle est continue, et dérivable (d-1) fois

L'avantage des fonctions splines est sa flexibilité et donc sa bonne capacité d'adaptation aux données observées. L'inconvénient est qu'il n'y a pas d'approche consensuelle pour faire le choix du nombre et de la limite des intervalles ainsi que pour la forme de la modélisation au sein de chaque intervalle.

En pratique, la valeur de d est le plus souvent fixée à 3 (on parle alors de splines cubiques), ce qui parait le meilleur compromis entre la flexibilité et la complexité de la courbe à représenter pour modéliser la relation entre la variable quantitative et le risque de faire l'évènement.
Il est reconnu que les fonctions splines cubiques ont tendance à être sensibles aux valeurs observées excentriques dans le premier et le dernier intervalle, avec des intervalles de confiance larges. Pour cette raison, on utilise plutôt, les splines cubiques restreintes ("rectricted cubic splines" en anglais). Ce sont les mêmes fonctions que précédemment (d=3), auxquelles on ajoute la contrainte d'être linéaires pour les deux intervalles extrêmes.

Dans le cas de splines cubiques restreintes, Franck Harrell (Harrell, 2001) recommande d'écarter de façon plus importante les nœuds extrêmes de façon à mieux modéliser la courbe pour les petites et grandes valeurs de la variable quantitative. Les emplacements recommandés par F. Harrell sont indiqués dans le tableau ci-dessous.

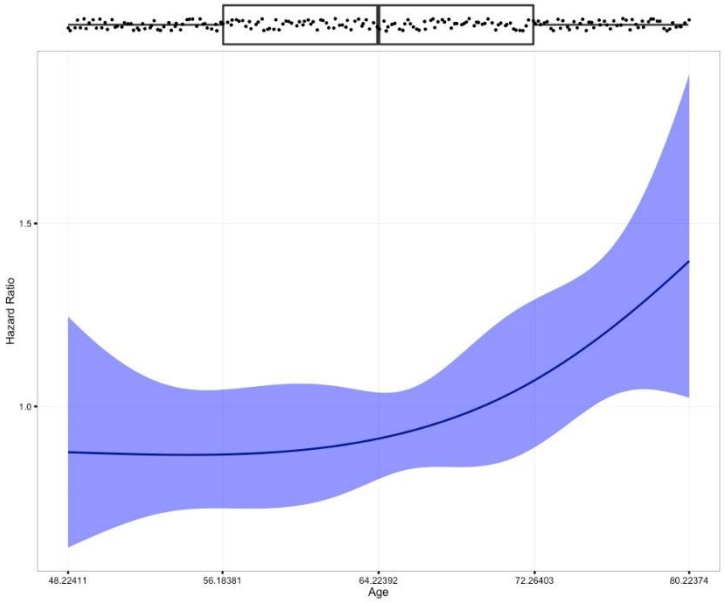
Nombre de nœuds	Percentiles où les nœuds sont placés						
3			10	50	90		
4			5	35	65	95	
5		5	27,5	50	72,5	95	
6	5	23	41	59	77	95	
7	2,5	18,33	34,17	50	65,83	81,67	97,5

III.2.2.1.2.2.3.2.2 Approche paramétrique : méthode des polynômes fractionnaires

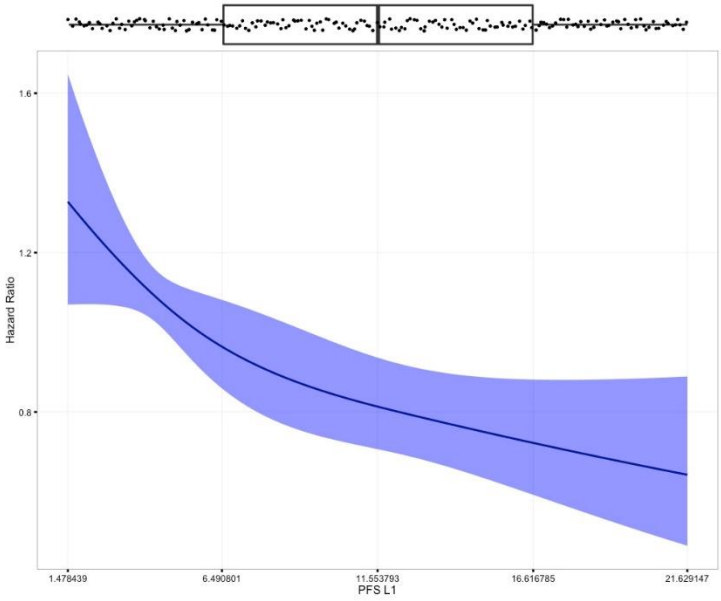
L'approche paramétrique a pour objet de représenter globalement la courbe, modélisant le lien entre la variable étudiée et le risque de faire l'évènement, par une seule fonction de la variable quantitative. La méthode dite des polynômes fractionnaires s'inscrit dans cette idée. Les polynômes fractionnaires ont été introduits dans les années 1990 et largement développés et popularisés par Patrick Royston, Willi Sauerbrei et Douglas G. Altman (Royston & Altman, 1994). Ils sont une extension des polynômes qualifiés d'ordinaires où les exposants peuvent être négatifs et/ou non entiers et sont inclus dans l'ensemble prédéfini de valeurs S = (-2; -1; -0,5; 0; 0,5; 1; 2; 3) avec, par convention, la puissance 0 correspondant à la fonction logarithme. L'ensemble S peut paraître restreint, mais l'expérience montre qu'il est suffisant pour la quasi-totalité des situations rencontrées (Royston & Sauerbrei, 2008). En pratique, on constate qu'il n'est pas nécessaire de considérer des degrés supérieurs à 2 pour représenter de façon satisfaisante la relation entre la variable quantitative et le risque de faire l'évènement.

Dans cette thèse, toutes les variables quantitatives considérées dans les études, dont l'objet était de fournir des outils de prédiction, ont été étudiées selon la méthode des splines cubiques restreintes ainsi que par la méthode des polynômes fractionnaires, pour valider la transformation utilisée et ainsi évaluer avec précision la relation entre les prédicteurs et l'évènement étudié. En effet, nous avons retenu les transformations suggérées par la méthode des polynômes fractionnaires après validation graphique par la méthode des splines cubiques restreintes.
Voici par exemple ci-dessous les figures correspondant à l'application de la méthode des splines cubiques restreintes pour les 4 variables quantitatives impliquées dans la construction d'un modèle de Cox pour la prédiction de la survie globale, dans le cadre d'un traitement de seconde ligne de chimiothérapie pour des patients atteints d'un cancer pancréatique.

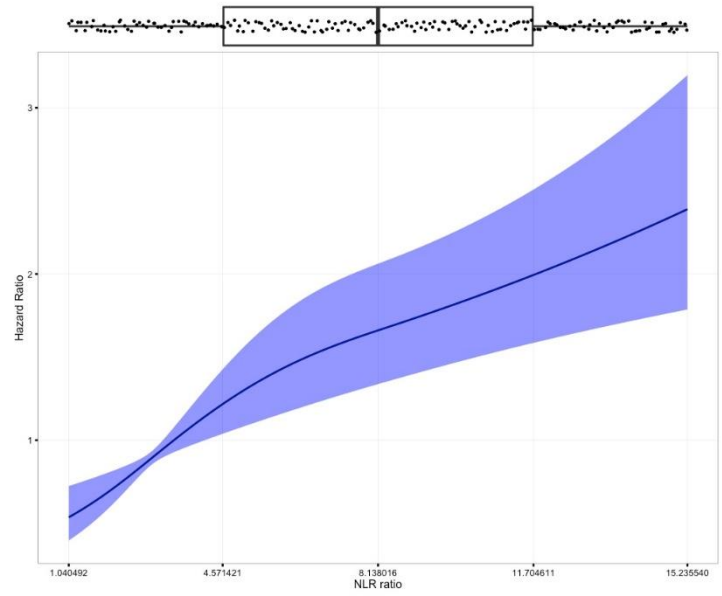
Age



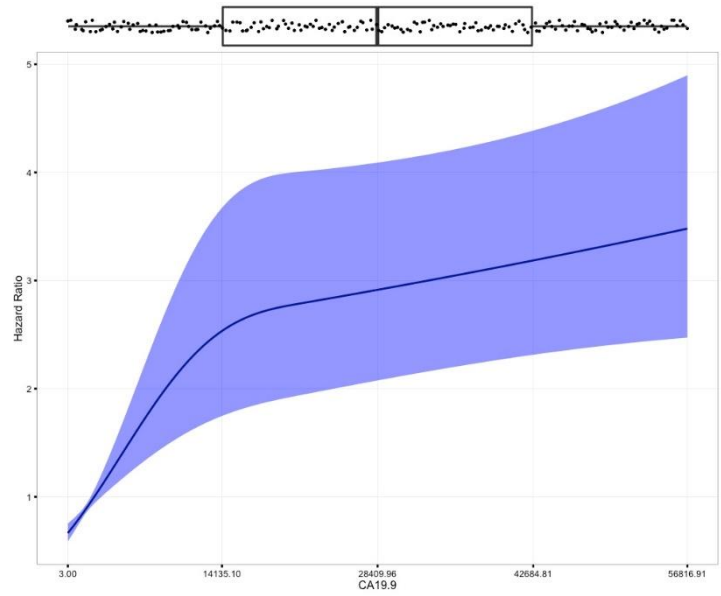
Durée de la première ligne



Ratio Neutrophile - Lymphocyte (NLR)



Antigène Carbohydrate 19-9 (CA 19-9)



Dans cette étude, compte tenu de ces représentations graphiques et des résultats donnés par la méthode des polynômes fractionnaires, une transformation racine carrée a été appliquée pour le ratio neutrophile - lymphocyte (NLR) et le marqueur tumoral (CA 19-9), tandis qu'une

transformation logarithmique a été utilisée pour la durée de la première ligne de chimiothérapie (L1). Toutes les autres variables continues ont été considérées sans aucune transformation.

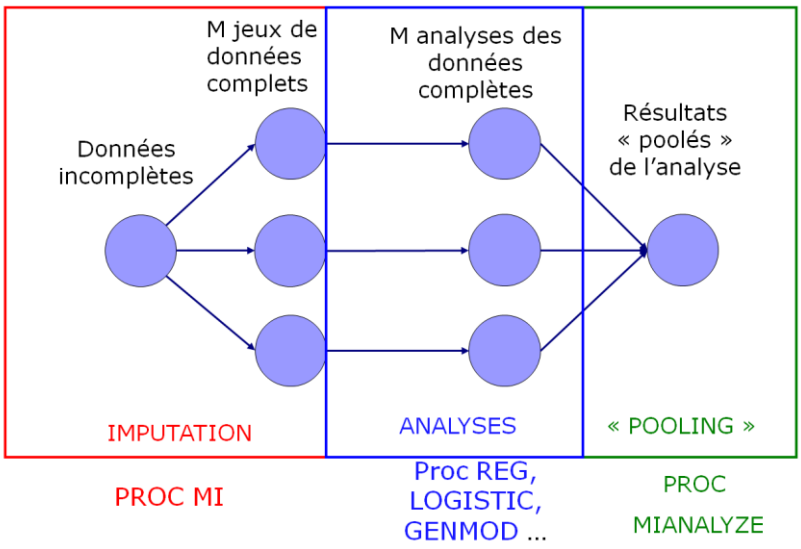
III.2.2.1.2.2.3.3 Considération des valeurs manquantes

La considération d'un sujet dans un modèle qui inclut plusieurs facteurs suppose que tous les facteurs soient mesurés chez ce sujet. En général, certains des renseignements manquent pour une partie de la population. La modélisation semi-paramétrique de Cox ne gère pas les données manquantes, ce qui signifie que si un seul renseignement est manquant pour un sujet, ce sujet est exclu de la modélisation même s'il possède l'ensemble des autres informations. Cette opération n'introduit pas de biais, si et seulement si, les données manquantes sont liées au hasard, c'est-à-dire qu'elles ne sont pas représentatives d'une certaine catégorie de sujets.

La validité de l'information véhiculée par un modèle de Cox repose donc directement sur l'exhaustivité du recueil. Aussi moins il y aura de données manquantes, plus le modèle sera représentatif de l'ensemble de la population.

Il arrive parfois notamment lorsque la nature des données est rétrospective que les informations manquantes ne soient pas récupérables. Il est alors nécessaire d'effectuer des analyses de sensibilité des résultats avec notamment des techniques d'imputations multiples. Si les résultats avec et sans imputation sont convergents alors la validité du modèle n'est pas remise en cause.

Le concept d'Imputations Multiples décrit par Rubin (Rubin, 1987) avec les procédures correspondant dans le software SAS peut être rapidement résumé schématiquement comme proposé dans la figure ci-dessous :



Pour exemple, dans une étude présentée dans cette thèse, un modèle de Cox multivarié pour la prédiction de la survie globale, chez des patients atteints d'un cancer pancréatique localement avancé, a été réalisé sur 358 patients qui possédaient les 5 facteurs impliqués, alors que l'échantillon de départ était composé de 442 patients.

Voici dans le tableau ci-après les résultats obtenus par une analyse d'imputations multiples.

	Analyse des sujets- complets (*N* = 358)			**Analyse avec imputations multiples (MCMC, *N* = 1000 échantillons de données imputés)**		
Paramètres	**ß**	**SE**	***P***	**ß**	**SE**	***P***
Age au diagnostic, années	0.0125	0.0063	0.0478	0.01217	0.00560	0.0298
Douleur	0.2562	0.1193	0.0317	0.23995	0.11055	0.0300
Albumine, g/L	-0.0382	0.0111	0.0006	-0.03911	0.01085	0.0003
Taille tumorale, mm	0.0080	0.0035	0.0214	0.00810	0.00332	0.0148
CA 19-9, UI/mL (transformation log)	0.1587	0.0573	0.0056	0.11556	0.05268	0.0283

Nous ne constatons pas de divergence majeure dans l'estimation des ß ni des p-value entre les résultats portant sur l'analyse de l'ensemble des sujets ayant des données complètes et ceux obtenus avec la procédure d'imputations multiples.

III.2.2.1.2.2.3.4 Considération de la colinéarité des variables explicatives

Il est aussi très important, avant de considérer plusieurs variables dans un même modèle, d'étudier leur potentielle colinéarité. La multi-colinéarité est un phénomène qui peut induire de l'imprécision dans les estimations du modèle et conduire à des conclusions erronées. L'examen de la colinéarité se fait en général par une matrice de corrélation entre les paramètres considérés. Pour une mesure de la corrélation entre deux paramètres, il y a deux entités importantes : l'une est l'amplitude de la corrélation et l'autre sa significativité. Les deux doivent être considérées même si l'amplitude de la corrélation reste le paramètre le plus informatif. Ainsi, sur la base de cette matrice, il est conseillé de prévoir la construction du modèle en éliminant certaines variables colinéaires suspectées de véhiculer une information en partie redondante avec d'autres variables plus intéressantes. Il est bien entendu important de confronter les résultats présentés dans la matrice de corrélation aux résultats de la littérature ainsi qu'au bon sens clinique afin de choisir ou d'éliminer certains paramètres.

Ci-dessous un exemple d'un tableau illustrant la matrice de corrélation entre les variables identifiées pouvant prétendre à être intégrées dans la construction du modèle multivarié provenant d'une étude présentée dans cette thèse et qui montre l'importance de la considération de ces informations.

Table S2: Correlation matrix

	Age	Smoking status	Stages	Number of metastatic sites	Liver metastases	Peritoneal metastases	Isolated lung metastases	Performance status (WHO)	Pain	Jaundice	Ascites	Hemoglobin	Neutrophils	Lymphocytes	Neutrophil-to-lymphocyte ratio	Total bilirubin	Albumin	CA19-9	CEA	Primary tumor resection	Adjuvant chemotherapy	First-line chemotherapy: Number of cures	First-line chemotherapy: RECIST Best response	First-line chemotherapy: Duration of L1	First-line chemotherapy: Reason for discontinuation	First-line chemotherapy: Metastatic progression	Type of L2 regimen
Age	1.00000 261	-0.14746 0.0173 260	-0.03588 0.5639 261	0.06591 0.2888 261	-0.07821 0.2079 261	-0.05167 0.4058 261	0.05491 0.3769 261	0.14152 0.0249 251	-0.06185 0.3301 250	-0.04516 0.4772 250	-0.09459 0.1351 251	-0.02676 0.6845 233	0.01379 0.8392 219	0.01932 0.7797 212	-0.00723 0.9171 210	-0.05639 0.4074 218	-0.05000 0.5288 161	-0.01179 0.8671 204	0.04026 0.6143 159	0.07004 0.2596 261	0.02835 0.6485 261	0.08083 0.1930 261	0.06080 0.3364 252	-0.02969 0.6330 261	-0.04457 0.4734 261	-0.03999 0.5201 261	0.16440 0.0078 261
Smoking status	-0.14746 0.0173 260	1.00000 260	0.02516 0.6863 260	-0.03471 0.5775 260	0.02290 0.7133 260	-0.01591 0.7985 260	0.00000 1.0000 260	0.14604 0.0206 251	0.08743 0.1682 250	-0.02154 0.7347 250	-0.00770 0.9034 251	0.10802 0.1000 233	0.06128 0.3668 219	-0.04696 0.4964 212	0.02644 0.7032 210	-0.01864 0.7844 218	0.02184 0.7833 161	0.11684 0.0961 204	0.04883 0.5411 159	-0.02259 0.7169 260	-0.09312 0.1342 260	-0.11749 0.0585 260	-0.02397 0.7049 252	-0.00369 0.9528 260	0.00719 0.9082 260	0.08874 0.1536 260	-0.15106 0.0148 260
Stages	-0.03588 0.5639 261	0.02516 0.6863 260	1.00000 261	0.24762 <.0001 261	0.29515 <.0001 261	0.07754 0.2118 261	-0.19782 0.0013 261	0.15439 0.0143 251	0.17218 0.0063 250	0.10841 0.0872 250	0.04959 0.4341 251	-0.05577 0.3968 233	0.03659 0.5902 219	0.04132 0.5497 212	0.06777 0.3284 210	0.05225 0.4427 218	-0.08239 0.2988 161	0.24240 0.0005 204	0.17113 0.0310 159	-0.80356 <.0001 261	-0.73948 <.0001 261	-0.07191 0.2470 261	0.12001 0.0571 252	-0.16070 0.0093 261	-0.33994 <.0001 261	-0.01621 0.7944 261	-0.10153 0.1017 261
Number of metastatic sites	0.06591 0.2888 261	-0.03471 0.5775 260	0.24762 <.0001 261	1.00000 261	0.37313 <.0001 261	0.46296 <.0001 261	-0.17898 0.0037 261	0.27300 <.0001 251	0.17469 0.0056 250	0.18116 0.0041 250	0.36348 <.0001 251	-0.11440 0.0814 233	0.14387 0.0333 219	0.07156 0.2997 212	0.16089 0.0197 210	0.17374 0.0102 218	-0.33650 <.0001 161	0.25458 0.0002 204	0.21530 0.0064 159	-0.12688 0.0405 261	-0.13402 0.0304 261	-0.04900 0.4305 261	0.13538 0.0317 252	-0.10673 0.0853 261	-0.34738 <.0001 261	0.36902 <.0001 261	-0.02636 0.6716 261
Liver metastases	-0.07821 0.2079 261	0.02290 0.7133 260	0.29515 <.0001 261	0.37313 <.0001 261	1.00000 261	-0.10922 0.0782 261	-0.27686 <.0001 261	0.08952 0.1573 251	0.05559 0.3815 250	0.18649 0.0031 250	-0.00580 0.9272 251	0.00159 0.9807 233	0.11447 0.0911 219	-0.00603 0.9304 212	0.09283 0.1802 210	0.17078 0.0116 218	-0.22881 0.0035 161	0.30165 <.0001 204	0.24553 0.0018 159	-0.10539 0.0893 261	-0.12126 0.0504 261	-0.06408 0.3024 261	0.27261 0.0003 252	-0.20317 0.0010 261	-0.31909 <.0001 261	0.33558 <.0001 261	-0.02170 0.7271 261
Peritoneal metastases	-0.05167 0.4058 261	-0.01591 0.7985 260	0.07754 0.2118 261	0.46296 <.0001 261	-0.10922 0.0782 261	1.00000 261	-0.17898 0.0037 261	0.25926 <.0001 251	0.14188 0.0249 250	0.14304 0.0237 250	0.47117 <.0001 251	-0.22900 0.0004 233	0.06508 0.3377 219	0.08699 0.2071 212	0.09489 0.1707 210	0.09335 0.1696 218	-0.20211 0.0101 161	0.02311 0.7428 204	-0.00808 0.9195 159	-0.04210 0.4983 261	-0.02556 0.6811 261	0.03491 0.5745 261	-0.00981 0.8769 252	-0.00337 0.9568 261	-0.16829 0.0064 261	0.27864 <.0001 261	-0.02636 0.6716 261
Isolated lung metastases	0.10688 0.0848 261	0.01830 0.7690 260	-0.20557 0.0008 261	-0.31771 <.0001 261	-0.49146 <.0001 261	-0.31771 <.0001 261	1.00000 261	-0.15218 0.0158 251	0.06536 0.3033 250	-0.09916 0.1179 250	-0.11511 0.0687 251	0.06888 0.2951 233	-0.00402 0.9528 219	-0.07871 0.2539 212	-0.04621 0.5054 210	-0.06789 0.3184 218	0.21271 0.0067 161	-0.11784 0.0932 204	-0.06138 0.4421 159	0.18402 0.0028 261	0.18800 0.0023 261	0.01123 0.8567 261	-0.00881 0.8893 252	0.04804 0.4397 261	0.02556 0.3857 261	0.05391 0.3857 261	0.07840 0.2068 261
Performance status (WHO)	0.14152 0.0249 251	0.14604 0.0206 251	0.15439 0.0143 251	0.27300 <.0001 251	0.08952 0.1573 251	0.25926 <.0001 251	-0.05374 0.3965 251	1.00000 251	-0.35032 <.0001 249	0.30444 <.0001 249	0.31981 <.0001 250	-0.18532 0.0045 233	0.28287 <.0001 219	0.11285 0.1013 212	0.26882 <.0001 210	0.23101 <.0001 218	-0.35717 <.0001 161	0.30079 <.0001 204	0.23139 0.0033 159	-0.16280 0.0098 251	-0.19413 0.0020 251	-0.05361 0.3977 251	0.04283 0.5038 246	-0.11750 0.0631 251	-0.20878 0.0009 251	0.13559 0.0318 251	-0.02423 0.7025 251
Pain	-0.06185 0.3301 250	0.08743 0.1682 250	0.17218 0.0063 250	0.17469 0.0056 250	0.05559 0.3815 250	0.14188 0.0249 250	0.00430 0.9461 250	0.35032 <.0001 249	1.00000 250	0.16009 0.0114 249	0.26882 <.0001 250	-0.09799 0.1367 232	0.18040 0.0074 219	0.05594 0.4177 212	0.15138 0.0283 210	0.13250 0.0507 218	-0.20478 0.0094 160	0.09515 0.1769 203	0.20281 0.0104 159	-0.20466 0.0011 250	-0.18369 0.0036 250	-0.04924 0.4382 250	-0.09219 0.1502 245	-0.03607 0.5702 250	-0.17825 0.0047 250	-0.03931 0.5361 250	-0.04366 0.4920 250
Jaundice	-0.04516 0.4772 250	-0.02154 0.7347 250	0.10841 0.0872 250	0.18116 0.0041 250	0.18649 0.0031 250	0.14304 0.0237 250	-0.10297 0.1043 250	0.30444 <.0001 249	0.16009 0.0114 249	1.00000 250	0.21792 0.0005 250	-0.03525 0.5932 232	0.23757 0.0004 218	-0.04586 0.5076 211	0.15755 0.0227 209	0.73813 <.0001 217	-0.25911 0.0009 160	0.24689 0.0004 202	0.33329 <.0001 159	-0.07547 0.2344 250	-0.07315 0.2492 250	0.02998 0.6371 250	-0.05060 0.4304 245	0.00727 0.9089 250	-0.08779 0.1664 250	0.15659 0.0132 250	0.02879 0.6506 250
Ascites	-0.09459 0.1351 251	-0.00770 0.9034 251	0.04959 0.4341 251	0.36348 <.0001 251	-0.00580 0.9272 251	0.47117 <.0001 251	-0.04809 0.4481 251	0.31981 <.0001 250	0.26882 <.0001 250	0.21792 0.0005 250	1.00000 251	-0.18209 0.0053 233	0.19286 0.0042 219	0.09467 0.1696 212	0.20056 0.0035 210	0.14932 0.0275 218	-0.22694 0.0038 161	0.08586 0.2232
203	0.10130 0.2039 159	-0.03489 0.5822 251	-0.08465 0.1813 251	0.02136 0.7363 233	-0.00826 0.9002 219	0.03799 0.5782 212	-0.04580 0.5092 210	0.02329 0.7323 218	0.04158 0.6005 161	-0.15866 0.0234
204	-0.10403 0.1919
159	0.12834 0.0383
261	0.04846 0.4357
261	0.00502 0.9357
261	0.08421 0.1750
261	0.15058 0.0149
261
Hemoglobin	-0.02676 0.6845 233	0.10802 0.1000 233	-0.05577 0.3968 233	-0.11440 0.0814 233	0.00159 0.9807 233	-0.22900 0.0004 233	0.10721 0.1026 233	-0.18532 0.0045 233	-0.09799 0.1367 232	-0.03525 0.5932 232	-0.18209 0.0053 233	1.00000 233	-0.03823 0.5737 219	-0.28315 <.0001 211	-0.19177 0.0053 210	-0.00360 0.9583 214	0.28838 0.0002 159	-0.08633 0.2289
196	-0.14362 0.0746
155	0.09979 0.1288
233	0.12089 0.0655
233	0.04439 0.5002
233	-0.10828 0.1022
229	0.12910 0.0490
233	0.08779 0.1821
233	-0.07288 0.1469
233	0.14689 0.0249
233
Neutrophils	0.01379 0.8392 219	0.06128 0.3668 219	0.03659 0.5902 219	0.14387 0.0333 219	0.11447 0.0911 219	0.06508 0.3377 219	-0.03489 0.6076 219	0.28287 <.0001 219	0.18040 0.0074 218	0.23757 0.0004 219	0.19286 0.0042 219	-0.03823 0.5737 219	1.00000 219	-0.07298 0.2925 210	0.74698 <.0001 210	0.21977 0.0014 209	-0.16769 0.0396 151	0.26700 0.0002 185	-0.13739 0.0596 153	-0.04071 0.5490 219	-0.02762 0.6844 219	-0.07173 0.2906 219	0.11008 0.1075 215	-0.15784 0.0194 219	-0.06500 0.3383 219	0.07184 0.2899 219	0.10766 0.1121 219
Lymphocytes	0.01932 0.7797 212	-0.04696 0.4964 212	0.04132 0.5497 212	0.07156 0.2997 212	-0.00603 0.9304 212	0.08699 0.2071 212	0.05528 0.4233 212	0.11285 0.1013 212	0.05594 0.4177 212	-0.04586 0.5076 211	0.09467 0.1696 212	-0.28315 <.0001 211	-0.07298 0.2925 210	1.00000 212	0.55739 <.0001 210	-0.06244 0.3750 204	-0.11948 0.1439 151	-0.02331 0.7561
180	-0.09540 0.2472
149	-0.11815 0.0861
212	-0.11348 0.1094
212	-0.01582 0.8188
212	-0.08669 0.2131
208	0.04325 0.5312
212	-0.03626 0.5996
212	-0.03737 0.5885
212	-0.05400 0.4341
212
Neutrophil-to-lymphocyte ratio	-0.00723 0.9171 210	0.02644 0.7032 210	0.06777 0.3284 210	0.16089 0.0197 210	0.09283 0.1802 210	0.09489 0.1707 210	0.00824 0.9056 210	0.26882 <.0001 210	0.15138 0.0283 210	0.15755 0.0227 209	0.20056 0.0035 210	-0.19177 0.0053 210	0.74698 <.0001 210	0.55739 <.0001 210	1.00000 210	0.11936 0.0891 204	-0.16140 0.0485 150	0.14222 0.0576
179	0.06026 0.4669
148	-0.13294 0.0544 210	-0.13314 0.0540 210	-0.04580 0.5092 210	0.02680 0.7022 206	-0.07283 0.2935 210	-0.08348 0.2283 210	0.05452 0.4319 210	0.05677 0.4132 210
Total bilirubin	-0.05639 0.4074 218	-0.01864 0.7844 218	0.05225 0.4427 218	0.17374 0.0102 218	0.17078 0.0116 218	0.09335 0.1696 218	-0.06638 0.3293 218	0.23101 0.0006 218	0.13250 0.0507 218	0.73813 <.0001 217	0.14932 0.0275 218	-0.00360 0.9583 214	0.21977 0.0014 209	-0.06244 0.3750 204	0.11936 0.0891 204	1.00000 218	-0.33824 <.0001 152	0.23961 0.0010
185	0.27920 0.0005
150	-0.03298 0.6282 218	-0.01054 0.8770 218	0.02329 0.7323 218	-0.05626 0.4129
214	-0.00158 0.9815
218	-0.02788 0.6823 218	0.11891 0.0798 218	0.10890 0.1089 218
Albumin	-0.05000 0.5288 161	0.02184 0.7833 161	-0.08239 0.2988 161	-0.33650 <.0001 161	-0.22881 0.0035 161	-0.20211 0.0101 161	0.14293 0.0705 161	-0.35717 <.0001 161	-0.20478 0.0094 160	-0.25911 0.0009 160	-0.22694 0.0038 161	0.28838 0.0002 159	-0.16769 0.0396 151	-0.11948 0.1439 151	-0.16140 0.0485 150	-0.33824 <.0001 152	1.00000 161	-0.19866 0.0195
138	-0.29232 0.0018
112	0.07057 0.3737
161	0.05223 0.5106
161	0.04158 0.6005
161	-0.11311 0.1570
158	0.16458 0.0370
161	0.20764 0.0082
161	-0.15539 0.0490
161	-0.12586 0.1116
161
CA19-9	-0.01179 0.8671 204	0.11684 0.0961 204	0.24240 0.0005
204	0.25458 0.0002
204	0.30165 <.0001
204	0.02311 0.7428
204	-0.17131 0.0143
204	0.30079 <.0001
204	0.09515 0.1769
203	0.24689 0.0004
202	0.08586 0.2232
203	-0.08633 0.2289
196	0.26700 0.0002
185	-0.02331 0.7561
180	0.14222 0.0576
179	0.23961 0.0010
185	-0.19866 0.0195
138	1.00000 204	0.48837 <.0001
155	-0.19886 0.0044
204	-0.21588 0.0019
204	-0.15866 0.0234
204	0.23427 0.0008
202	-0.28277 <.0001
204	-0.21302 0.0022
204	0.14171 0.0432
204	-0.02633 0.7085
204
CEA	0.04026 0.6143 159	0.04883 0.5411 159	0.17113 0.0310
159	0.21530 0.0064
159	0.24553 0.0018
159	-0.00808 0.9195
159	-0.09340 0.2416
159	0.23139 0.0033
159	0.20281 0.0104
159	0.33329 <.0001
159	0.10130 0.2039
159	-0.14362 0.0746
155	0.18739 0.0204
153	-0.09540 0.2472
149	0.06026 0.4669
148	0.27920 0.0005
150	-0.29232 0.0018
112	0.48837 <.0001
155	1.00000 159	-0.16833 0.0339
159	-0.17579 0.0267
159	-0.10403 0.1919
159	0.08657 0.2810
157	-0.12113 0.1283
159	-0.17255 0.0296
159	0.13125 0.0991
159	-0.09269 0.2452
159
Primary tumor resection	0.07004 0.2596 261	-0.02259 0.7169 260	-0.80356 <.0001 261	-0.12688 0.0405 261	-0.10539 0.0893 261	-0.04210 0.4983 261	-0.06525 0.2936 261	-0.16280 0.0098 251	-0.20466 0.0011 250	-0.07547 0.2344 250	-0.03489 0.5822 251	0.09979 0.1288 233	-0.04071 0.5490 219	-0.11815 0.0861 212	-0.13294 0.0544 210	-0.03298 0.6282 218	0.07057 0.3737
161	-0.19886 0.0044
204	-0.16833 0.0339
159	1.00000 261	0.86061 <.0001
261	0.12834 0.0383
261	-0.15156 0.0160
252	0.21652 0.0004
261	0.38914 <.0001
261	0.11878 0.0553
261	0.10208 0.0998
261
Adjuvant chemotherapy	0.02835 0.6485 261	-0.09312 0.1342
260	-0.73948 <.0001
261	-0.13402 0.0304
261	-0.12126 0.0504
261	-0.02556 0.6811
261	-0.06282 0.3120
261	-0.19413 0.0020
251	-0.18369 0.0036
250	-0.07315 0.2492
250	-0.08465 0.1813
251	0.12089 0.0655
233	-0.02762 0.6844
219	-0.11348 0.0994
212	-0.13314 0.0540
210	-0.01054 0.8770
218	0.05223 0.5106
161	-0.21588 0.0019
204	-0.17579 0.0267
159	0.86061 <.0001
261	1.00000 261	0.04846 0.4357
261	-0.10549 0.0947
252	0.14050 0.0232
261	0.38008 <.0001
261	0.10440 0.0923
261	0.11957 0.0537
261
First-line chemo therapy: Number of cures	0.08083 0.1930 261	-0.11749 0.0585 260	-0.07191 0.2470
261	-0.04900 0.4305
261	-0.06408 0.3024
261	0.03491 0.5745
261	-0.08473 0.1723
261	-0.05361 0.3977
251	-0.04924 0.4382
250	0.02998 0.6371
250	0.02136 0.7363
251	0.04439 0.5002
233	-0.07173 0.2906
219	-0.01582 0.8188
212	-0.04580 0.5092
210	0.02329 0.7323
218	0.04158 0.6005
161	-0.15866 0.0234
204	-0.10403 0.1919
159	0.12834 0.0383
261	0.04846 0.4357
261	1.00000 261	-0.41322 <.0001
252	0.69973 <.0001
261	-0.00502 0.9357
261	0.08421 0.1750
261	0.15058 0.0149
261
First-line chemo therapy: RECIST Best response	0.06080 0.3364 252	-0.02397 0.7049 252	0.12001 0.0571
252	0.13538 0.0317
252	0.22761 0.0003
252	-0.00981 0.8769
252	-0.17511 0.0053
252	0.04283 0.5038
246	-0.09219 0.1502
245	-0.05060 0.4304
245	-0.06377 0.3192
246	-0.10828 0.1022
229	0.11008 0.1075
215	-0.08669 0.2131
208	0.02680 0.7022
206	-0.05626 0.4129
214	-0.11311 0.1570
158	0.23427 0.0008
202	0.08657 0.2810
157	-0.15156 0.0160
252	-0.10549 0.0947
252	-0.41322 <.0001
252	1.00000 252	-0.69690 <.0001
252	-0.42675 <.0001
252	0.25902 <.0001
252	-0.05724 0.3656
252
First-line chemo therapy: Duration of L1	-0.02969 0.6330 261	-0.00369 0.9528 260	-0.16070 0.0093
261	-0.10673 0.0853
261	-0.20317 0.0010
261	-0.00337 0.9568
261	0.11191 0.0711
261	-0.11750 0.0631
251	-0.03607 0.5702
250	0.00727 0.9089
250	0.00781 0.9020
251	0.12910 0.0490
233	-0.15784 0.0194
219	0.04325 0.5312
212	-0.07283 0.2935
210	-0.00158 0.9815
218	0.16458 0.0370
161	-0.28277 <.0001
204	-0.12113 0.1283
159	0.21652 0.0004
261	0.14050 0.0232
261	0.69973 <.0001
261	-0.69690 <.0001
252	1.00000 261	0.35712 <.0001
261	-0.06440 0.3000
261	-0.06523 0.2938
261
First-line chemo therapy: Reason for discontinuation	-0.04457 0.4734 261	0.00719 0.9082 260	-0.33994 <.0001
261	-0.34738 <.0001
261	-0.31009 <.0001
261	-0.16829 0.0064
261	0.32421 <.0001
261	-0.20878 0.0009
251	-0.17825 0.0047
250	-0.08779 0.1664
250	-0.16244 0.0099
251	0.18888 0.0038
233	-0.06500 0.3383
219	-0.03626 0.5996
212	-0.08348 0.2283
210	-0.02788 0.6823
218	0.20764 0.0082
161	-0.21302 0.0022
204	-0.17255 0.0296
159	0.38914 <.0001
261	0.38008 <.0001
261	-0.00502 0.9357
261	-0.42675 <.0001
252	0.35712 <.0001
261	1.00000 261	-0.43976 <.0001
261	-0.05415 0.3836
261
First-line chemo therapy: Metastatic progression	-0.03999 0.5201 261	0.08874 0.1536 260	-0.01621 0.7944
261	0.36902 <.0001
261	0.33558 <.0001
261	0.27864 <.0001
261	-0.61614 <.0001
261	0.13559 0.0318
251	-0.03931 0.5361
250	0.15659 0.0132
250	0.10709 0.0904
251	-0.07288 0.2679
233	0.07184 0.2899
219	-0.03737 0.5885
212	0.05452 0.4319
210	0.11891 0.0798
218	-0.15539 0.0490
161	0.14171 0.0432
204	0.13125 0.0991
159	0.11878 0.0553
261	0.10440 0.0923
261	0.08421 0.1750
261	0.25902 <.0001
252	-0.06440 0.3000
261	-0.43976 <.0001
261	1.00000 261	0.11396 0.0660
261
Type of L2 regimen	0.16440 0.0078 261	-0.15106 0.0148 260	-0.10153 0.1017
261	-0.02636 0.6716
261	-0.02170 0.7271
261	-0.02636 0.6716
261	-0.10421 0.0930
251	-0.02423 0.7025
251	-0.04366 0.4920
250	0.02879 0.6506
250	0.03316 0.6010
251	0.14689 0.0249
233	0.10766 0.1121
219	-0.05400 0.4341
212	0.05677 0.4132
210	0.10890 0.1089
218	-0.12586 0.1116
161	-0.02633 0.7085
204	-0.09269 0.2452
159	0.10208 0.0998
261	0.11957 0.0537
261	0.15058 0.0149
261	-0.05724 0.3656
252	-0.06523 0.2938
261	-0.05415 0.3836
261	0.11396 0.0660
261	1.00000 261

Pearson correlation coefficients: < 0.2 – 0.2-0.3 – 0.3-0.4 – ≥ 0.4

P-value: > 0.001 – ≤ 0.001

Abbreviations: WHO=World Health Organization, CA 19-9=Carbohydrate Antigen 19-9, CEA=Carcinoembryonic Antigen, L1=first-line chemotherapy, L2=second-line chemotherapy

III.2.2.1.2.2.3.5 Considération du terme d'interaction

Dans une modélisation, il arrive qu'on ne veuille pas seulement évaluer les effets marginaux des variables mais que l'on cherche aussi à mesurer dans quelle mesure une variable va atténuer ou au contraire augmenter l'effet d'une autre variable. En langage statistique, nous parlons de la recherche ou de la considération d'un terme d'interaction, qu'il est possible d'introduire dans une modélisation de Cox.

Par exemple, dans le cancer colorectal de stade 3, les mutations BRAF V600E et KRAS ont été récemment décrites comme étant significativement associées avec une DFS (Disease Free Survival) et une SG plus courtes chez les patients avec un statut microsatelitaire stable, ce qui n'est pas le cas pour les patients dont la tumeur présente une instabilité des microsatellites (Taieb et al, 2016). Typiquement dans cette pathologie des termes d'interactions entre la mutation KRAS et la stabilité des microsatellites (MMR) ainsi qu'entre la mutation BRAF V600E et le statut MMR doivent être considérés dans le futur.

Pour les travaux proposés, dans cette thèse nous n'avons pas exploré de terme d'interaction dans les modélisations en raison de l'absence d'hypothèses cliniques fortes et/ou de documentation dans la littérature.

III.3 Détermination de la valeur additionnelle apportée par un nouveau facteur pronostique identifié aux facteurs conventionnels déjà connus et fonctionnels dans la pratique clinique

Ensuite, il apparaît fondamental de déterminer la valeur additionnelle apportée par un nouveau facteur pronostique identifié aux facteurs conventionnels déjà connus et fonctionnels dans la pratique clinique comme illustré dans la figure ci-dessous

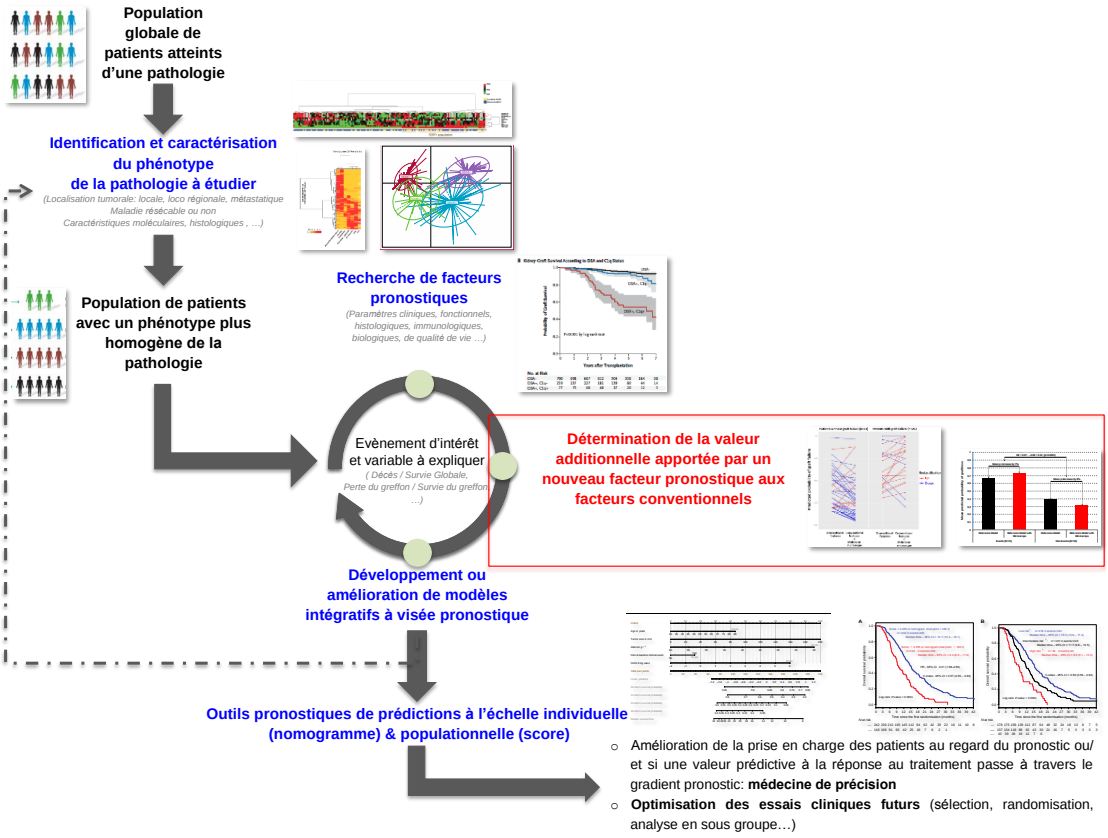


Figure : Troisième étape d'une recherche à visée pronostique : la détermination de la valeur additionnelle apportée par un nouveau facteur pronostique identifié aux facteurs conventionnels déjà connus et fonctionnels dans la pratique clinique.

Un intérêt majeur de la recherche médicale repose sur la quantification de l'information additionnelle apportée par un nouveau marqueur (moléculaire, génétique, radiologique, etc.) aux spectres de paramètres conventionnels.

L'analyse de données de survie par une modélisation de Cox permet donc, comme cela a été présenté précédemment, de tester l'association d'un facteur avec la distribution de survie et aussi de tester des associations indépendantes dans le cas de l'évaluation conjointe dans un même modèle de plusieurs covariables.

Quand on étudie la valeur pronostique additionnelle d'un nouveau facteur à d'autres paramètres, on cherche à savoir si la considération de ce facteur apporte une information pronostique complémentaire à celle dont on dispose déjà avec les facteurs conventionnels.
Théoriquement avec un modèle de Cox cela peut s'écrire :

$$h(t,z) = h_0(t)\ exp\ (b_1 z_1 + b' Z)$$

t représente le temps ; Z est le vecteur des variables pronostiques connues, et z_1 la nouvelle covariable.

L'idée est donc de tester l'hypothèse nulle H_0 : $b_1=0$.

Si l'hypothèse nulle est rejetée alors la variable z1 peut être considérée comme associée à la distribution de survie de manière indépendante aux autres variables et on considère, en général, qu'elle apporte une information pronostique additionnelle dans la prédiction de l'occurrence de l'évènement d'intérêt.
En pratique, si nous incluons cette variable parmi les facteurs considérés dans un même modèle de Cox et que le paramètre estimé pour cette variable est significativement non nul (c'est-à-dire que la p-value correspondant au test de Wald associé à cette variable est significative), on en conclut qu'elle

apporte une information pronostique intéressante au-delà de celle apportée par les autres facteurs. On parle alors d'une association ou d'une corrélation indépendante avec l'occurrence de l'évènement. Cependant, compte tenu d'une utilisation erronée de la terminologie statistique et plus précisément de la notion de prédiction, ce résultat a conduit à de nombreuses surinterprétations et donc à des conclusions trop optimistes quant à l'intérêt d'un nouveau marqueur au-delà des informations conventionnelles (Ioannidis, 2008) (Foucher *et al*, 2012). En effet, au cours de la dernière décennie beaucoup de biomarqueurs prometteurs ont été identifiés comme étant associés de manière indépendante à l'évènement mais très peu se sont avérés réellement intéressants en apportant une information additionnelle aux paramètres déjà connus en termes de stratification du risque de faire l'évènement étudié.

Ainsi, pour qu'un nouveau facteur pronostique soit qualifié de bon prédicteur, une association indépendante avec l'évènement est nécessaire mais doit être complétée par d'autres analyses orientées, notamment sur la stratification du risque de faire l'évènement. Autrement dit, considérer qu'un facteur apporte une information additionnelle pour la prédiction de survenue de l'évènement requiert une association indépendante significative avec la distribution de survie, mais il doit aussi apporter une amélioration de la stratification du risque et donc être discriminant. En pratique clinique un facteur pronostique est pertinent lorsqu'il est utile pour guider les décisions (Windeler, 2000). Il doit à ce titre non seulement être associé à la survenue de l'évènement, fournir une bonne prédiction de cet évènement et ainsi stratifier correctement le risque des patients face à sa survenue.

Dans le cas des données censurées, les indices de discriminations ou indices de concordances constituent une réponse à cette problématique.

III.3.1 L'indice de discrimination ou de concordance

La valeur prédictive ou la capacité discriminante d'un modèle est définie

par son aptitude à discriminer les sujets qui ont fait l'évènement de ceux qui ne l'ont pas fait, et ainsi séparer les sujets de pronostic différents. On peut donc dire qu'un modèle est discriminant s'il produit une probabilité de faire l'évènement plus importante pour un sujet qui fait l'évènement par rapport à celui qui ne le fait pas. Au cours des dernières années, de nouveaux indices statistiques permettant d'investiguer cette notion de discrimination ont été développés.

Dans le contexte d'une modélisation avec un évènement binaire sans prise en compte du temps de survenue de l'évènement (régression logistique), l'indice de discrimination de référence est l'aire sous la courbe ROC (ou area under the ROC curve = AUC). Classiquement, une courbe ROC est un graphique représentant la sensibilité d'un marqueur pour un état de santé (axe des ordonnées) en fonction de 1-spécificité (axe des abscisses). Autrement dit, il s'agit du taux de vrais positifs en fonction du taux de faux positifs. On obtient alors la représentation graphique suivante. La courbe ROC est couramment utilisée pour évaluer la qualité d'un test diagnostic à discriminer les malades des non-malades. L'indice qualitatif utilisé à cet effet est l'AUC qui représente la probabilité qu'un sujet malade ait un score au test plus élevé qu'un sujet non malade. Un test parfait ayant une sensibilité et une spécificité de 1 donnera une AUC de 1 (courbe A sur la Figure). Un test non discriminant ne fera pas mieux que le hasard et donnera une AUC de 0.5 (droite C). En pratique courante, l'AUC se situe entre 0.5 et 1 (courbe B). La courbe ROC est aussi utilisée en épidémiologie analytique afin d'évaluer les capacités discriminantes d'un modèle logistique. L'AUC représente alors la probabilité que le modèle donne une probabilité prédite plus importante pour un cas que pour un non-cas.

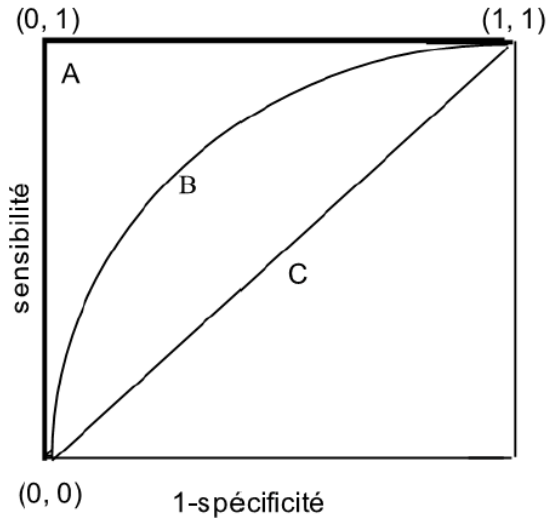
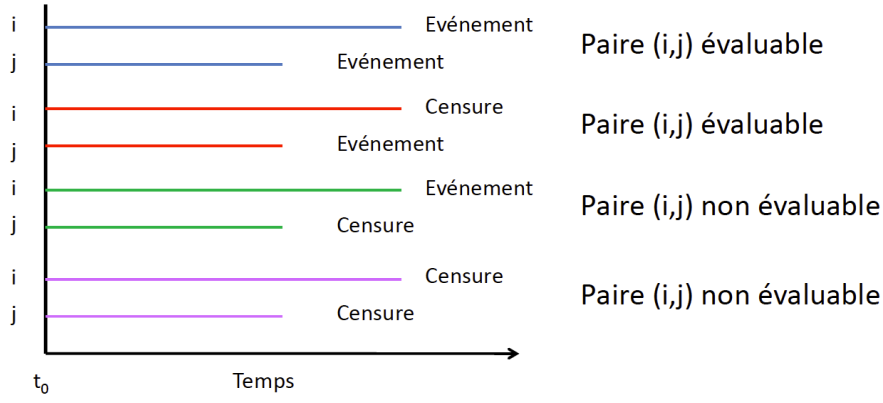


Figure : courbe ROC

L'indice de concordance de Harrell (indice C de Harrell) est une extension de l'AUC adaptée aux données de survie censurées (Harrell *et al*, 1996). Il est défini comme la proportion de paires de sujets comparables pour lesquelles les temps observés d'occurrence de l'évènement (T) et le vecteur de risque (X) sont concordants. En pratique, une paire de sujets i et j est comparable si elle contient au moins un évènement (soit évènement + évènement, soit évènement + non-évènement). La paire n'est pas comparable si un sujet a présenté un évènement et que l'autre sujet est censuré avant la survenue de cet évènement.

Ci-dessous une illustration des situations qu'il est possible de rencontrer dans l'échantillon de patients analysés :



On suppose pour simplifier que la mortalité augmente avec X.

Une paire est donc concordante si pour $X_i < X_j$, on observe $T_i < T_j$.

La probabilité de concordance entre deux variables aléatoires continues X et T est définie par la formule suivante:

$$C=P(X_i>X_k \mid T_i<T_k)$$

où i et k indiquent deux patients choisis au hasard.

Dans le contexte d'un modèle de survie de Cox, la probabilité de concordance C est donc calculée entre les 2 variables aléatoires : la composante prédictive du modèle X = ßz, où X est le vecteur des covariables, et la variable T qui représente la durée de survie.

Autrement dit, plus la composante prédictive de faire l'évènement est élevée, plus le temps d'observation de celui-ci est faible.

Dans le cas d'un modèle avec censure à droite la procédure de Harrell est en fait une estimation de :

$P(X_i<X_j\ T_i<T_k \mid T_i>T_j\ ;\ T_j< \min(C_i,C_j))$

Avec $T'_i=\min(T_i,C_i)$

L'indice C de Harrell s'interprète de la même manière que l'AUC. C'est une probabilité donc une quantité numérique qui varie entre 0 et 1. Un C égal à 0.5 ne confère aucune capacité discriminante au modèle (pas mieux que le hasard) tandis qu'un C égal à 1 caractérisera un modèle qui discrimine parfaitement le risque d'évènement parmi la population analysée.

III.3.1.1 Evaluation de la valeur additionnelle du nouveau facteur pronostique en termes de discrimination.

Pour déterminer l'intérêt d'un nouveau facteur dans la prédiction individuelle

du risque parmi un spectre de paramètres déjà connus, l'idée est d'estimer l'indice C de Harrell d'un modèle dit de référence (M_{ref} modèle dépourvu du nouveau facteur) puis l'indice C du modèle de référence auquel on implémente le nouveau facteur ($M_{ref+facteur}$). A priori, si le nouveau paramètre apporte une information additionnelle, le C du modèle $M_{ref+facteur}$ devrait être plus élevé que celui du modèle M_{ref}. Ainsi, si la différence des deux indices C obtenus dans les deux modèles emboîtés, $C_{ref+facteur} - C_{ref}$ est positive, on peut conclure à l'observation d'une augmentation du pouvoir discriminant du modèle lorsqu'on ajoute le nouveau marqueur.
La différence positive, entre les deux estimations de C issues des modèles sans et avec le nouveau facteur, constitue un élément en faveur de la valeur discriminante additionnelle du modèle $M_{ref+facteur}$, mais ne permet pas de conclure en la significativité statistique de cette observation.

III.3.1.2 Evaluation de la significativité statistique de la valeur additionnelle par procédure de ré-échantillonnage

En effet, comme toujours en statistique, une observation n'a de sens que lorsqu'elle est accompagnée de son intervalle de confiance. Ainsi, si l'intervalle de confiance de la différence des $C = C_{ref+facteur} - C_{ref}$ est supérieur à 0, alors on peut conclure en la significativité statistique de la valeur discriminante additionnelle apportée par le nouveau marqueur aux autres paramètres.

Dans cette thèse, nous avons utilisé des méthodes de ré-échantillonnage aussi appelée bootstrap pour tester la significativité de cette différence et lui conférer un intervalle de confiance (usuellement établi à 95%). Le bootstrap a été décrit la première fois par Bradley Efron (Efron, 1979) et demeure la technique de simulation la plus usuelle. Boostraper signifie générer un grand nombre n de jeux de données de la même taille que l'échantillon d'origine, avec une procédure de ré-échantillonnage qui repose sur un tirage au sort

avec remise (un patient préalablement sélectionné peut être sélectionné à nouveau). En appliquant les deux modèles dans les n jeux de données, on peut alors obtenir n différences $C^i = C^i_{ref+facteur} – C^i_{ref.}$ L'idée est alors d'identifier les percentiles de la distribution pour obtenir un intervalle de confiance de la différence de la quantité $C = C_{ref+facteur} - C_{ref}$ observée dans l'échantillon original.

Le bootstrap permet en effet de créer de la variabilité, ce qui donne la possibilité de construire un intervalle de confiance autour d'une estimation, notamment lorsque cette dernière repose sur une statistique non paramétrique, tel que l'indice de concordance. Un intervalle de confiance de la statistique d'intérêt obtenue suite à une procédure de bootstrap, au niveau $\boldsymbol{\alpha}$, se détermine en identifiant les percentiles de la distribution de la statistique observée dans les n échantillons, en laissant de part et d'autre de la distribution $\boldsymbol{\alpha}$ /2 X 100% des observations. Ainsi, au seuil 5%, l'intervalle de confiance est simplement délimité par les percentiles 2,5% et 97,5% de la distribution des n observations.

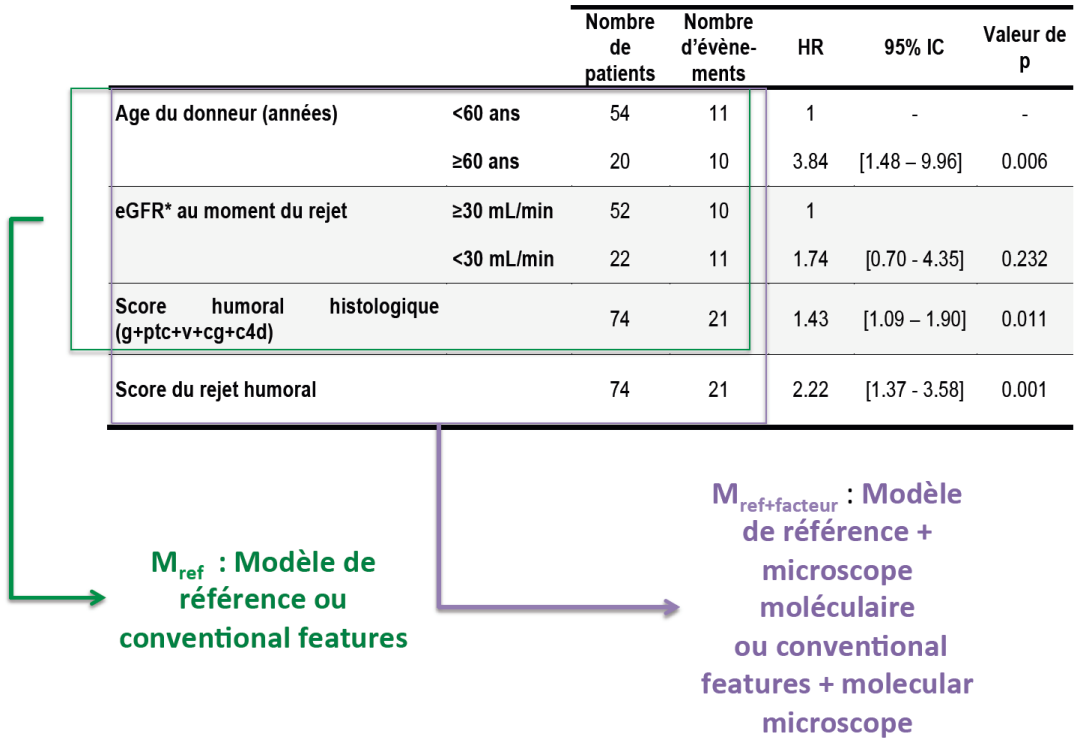
Pour obtenir un intervalle de confiance par bootstrap, un nombre de simulations n nécessite d'être suffisamment grand ; en particulier $n \geq 250$. Il est fréquemment admis qu'au-dessus de 250 échantillons il y une stabilité des résultats. Cependant la performance des moyens informatiques actuels incite à être plus gourmand avec un n souvent fixé à 1 000, en tout cas dans cette thèse.

Une illustration peut être apportée par des résultats concernant un travail de cette thèse qui investigue la valeur additionnelle du microscope moléculaire (score de rejet humoral) pour prédire la perte du greffon rénale.

Le modèle multivarié obtenu dans cette analyse était le suivant :

		Nombre de patients	Nombre d'évène-ments	HR	95% IC	Valeur de p
Age du donneur (années)	**<60 ans**	54	11	1	-	-
	≥60 ans	20	10	3.84	[1.48 – 9.96]	0.006
eGFR* au moment du rejet	**≥30 mL/min**	52	10	1		
	<30 mL/min	22	11	1.74	[0.70 - 4.35]	0.232
Score humoral histologique (g+ptc+v+cg+c4d)		74	21	1.43	[1.09 – 1.90]	0.011
Score du rejet humoral		74	21	2.22	[1.37 - 3.58]	0.001

Il est alors possible pour évaluer l'apport du microscope moléculaire, défini ici par le score de rejet humoral, de définir M_{ref} et $M_{ref+facteur}$ comme ci-dessous :



L'inclusion du score moléculaire de rejet humoral aigu, dans le modèle de référence, permet une augmentation de la capacité discriminative du modèle (c'est-à-dire sa capacité de distinction entre les patients qui ont perdu leur greffon et les autres). Cette augmentation de la C statistique est significative puisqu'elle passe de 0,770 à 0,8102, avec une différence moyenne obtenue

par bootstrap avec 1 000 échantillons de 0,0496505, dont l'intervalle de confiance à 95% est compris entre 0,0473 et 0,0519.

Outre le fait que comme l'AUC, le C de Harrell est basé sur les rangs des sujets en termes de délais prédits et donc ne prend pas en compte la valeur de la probabilité prédite par le modèle, c'est aussi un estimateur qui peut être biaisé et qui ne représente pas la réelle probabilité de concordance puisque toutes les paires de sujets ne sont pas forcément comparables (censure avant évènement ou censure des 2 sujets à la date de point). Enfin, l'indice de concordance de Harrell est sensible à la durée de suivi de l'étude. D'autres méthodes de calcul ont été proposées pour s'affranchir notamment des censures et apportent donc des solutions (Uno *et al*, 2011).

L'indice de concordance est une statistique incontournable pour estimer la capacité discriminante d'un modèle ou d'un paramètre. Cependant, dans le cadre de l'évaluation de la valeur additionnelle d'un facteur pronostique, le C étant une statistique de rang, elle ne donne pas un résultat très intuitif pour comprendre réellement l'information supplémentaire véhiculée par le facteur.

Il est important de préciser qu'il est tout à fait envisageable d'évaluer la valeur discriminante additionnelle apportée par un bloc de paramètres. La seule contrainte dans la procédure est que les modèles comparés possèdent un socle commun de paramètres et donc qu'ils soient emboités. Par ailleurs, chacun de ces deux modèles doit être évalué sur une population qui ne fluctue pas. En effet, le modèle de Cox ne gérant pas les données manquantes, l'introduction d'un paramètre ou d'un bloc de paramètres avec la présence de données manquantes peut modifier la population d'analyse. Afin d'attribuer la différence en termes de C statistique entre les deux modèles à la seule introduction du paramètre ou du bloc, il est important de s'assurer que les deux populations soient similaires et ainsi qu'elles ne puissent pas impacter la différence observée entre les C des deux modèles.

Il est aussi important de noter que d'autres indices statistiques peuvent être employés pour évaluer la performance d'un modèle. Celle de Schemper (Schemper, 2003) permet de quantifier la variabilité de l'évènement modélisé expliquée par les covariables du modèle et peut être employée aussi pour quantifier l'information additionnelle véhiculée par l'ajout d'une nouvelle information.

III.3.2Les indices de reclassification

Face aux limites rencontrées, notamment par le C de Harrell, certains auteurs ont proposé de nouvelles méthodes complémentaires basées sur la quantité et la direction de la reclassification des sujets en termes de probabilité d'évènements, après l'ajout d'un nouveau marqueur de risque à un modèle de prédiction (Pencina *et al*, 2008). Autrement dit, il a été proposé que le niveau de reclassification des probabilités prédites de faire l'évènement dans deux modèles emboités puisse être une mesure de l'utilité additionnelle apportée par le marqueur.

Ces méthodes ont initialement été développées dans le cadre d'études cas-témoins c'est-à-dire en présence d'un échantillon de sujet ayant expérimenté un évènement Y (Y=1) et l'autre non (Y=0). Dans ce contexte précis, une modélisation de la probabilité d'être un cas (P(Y=1)) en fonction de paramètres d'intérêt peut être déduite d'une régression logistique.

Il est tout à fait possible d'adapter ces outils dans un contexte d'analyse de données de survie avec des données censurées à droite. La différence est que le modèle de Cox intègre une notion de temps d'occurrence de l'évènement. Il est par conséquent impératif de définir un ou des temps d'intérêt auxquels l'examen des probabilités prédites par le modèle est intéressant. Ces temps sont, généralement, des moments clefs auxquels des options ou décisions thérapeutiques ou des bilans sont envisagés. C'est donc à ces temps précis qu'il faut évaluer la valeur additionnelle d'une information en termes de reclassification du risque de faire l'évènement dans un contexte d'analyse de données dites de survie.
Les probabilités prédites par le modèle de Cox au moment t peuvent être calculées simplement.
Le modèle de Cox se présentant sous la forme suivante :
$S(t, X_j) = S_0(t)^{\exp(LP_j)}$ avec LPj le prédicteur linéaire pour le patient j et S0(t) est la survie de base au temps t.

Si nous calculons dans l'échantillon de données le risque de base S0 alors il est très simple d'obtenir pour le patient i sa probabilité de survie au moment t $S(t, X_j)$.

A ce stade il est donc possible pour un patient i de connaître la probabilité p_i de faire l'évènement au temps t, issue d'un modèle de Cox, qui est égale à 1-$S(t, X_j)$ et nous obtenons ainsi un vecteur de n observation

$$\text{probabilité predite de faire l'evenement au temps t} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{n-1} \\ p_n \end{pmatrix}$$

III.3.2.1 Net Reclassification Improvement (NRI)

Le NRI représente la proportion de sujets cas (ayant expérimenté l'évènement au temps t ou avant) et non-cas correctement reclassés en termes de probabilité de risque de faire l'évènement au temps t, après l'ajout du nouveau facteur dans un modèle contenant les paramètres conventionnels. Initialement proposé avec des seuils de classification bien établis, le NRI s'est rapidement étendu dans des contextes sans seuils d'usage courant (Pencina *et al*, 2016).

III.3.2.1.1 Le NRI catégoriel

Comme son nom l'indique le NRI catégoriel repose sur la connaissance de catégories de risque connues. Par exemple, si la probabilité de faire l'évènement est inférieure à 10% on parlera de risque faible, si la probabilité se situe entre 10% et 20% on sera dans une situation de risque modéré et si elle dépasse 20% le sujet est considéré à haut risque de faire l'évènement: (<10%: risque faible, 10-20%: risque modéré, >20%: risque élevé). Ceci est un exemple dans le cas de trois catégories de risque différent.

Les sujets sont classés dans une catégorie de risque (faible, modéré ou élevé) selon leur probabilité d'évènement prédite au temps t par un 1er modèle intégrant les facteurs de risque établis (modèle de référence: M_{ref}).

La 2ème étape consiste à reclasser les sujets à partir d'un second modèle de prédiction emboité incluant le nouveau biomarqueur $M_{ref+facteur}$.

Cette reclassification est effectuée séparément chez les cas et les non-cas qui ont expérimentés l'évènement au temps t : un cas est considéré comme correctement reclassé si sa catégorie de risque augmente (modéré → élevé par exemple); à l'opposé un non-cas est correctement reclassé si sa catégorie de risque diminue (élevé → modéré par exemple) après la prise en compte du nouveau facteur.

Pour les événements, nous attribuons donc 1 pour reclassification à la hausse, -1 vers le bas et 0 pour les sujets qui ne changent pas de catégorie de risque. La démarche opposée est faite dans la population n'ayant pas fait l'événement au temps t.

Ensuite, les points obtenus sont sommés et divisés par le nombre de sujets dans chaque groupe. On obtient ainsi en sommant les deux résultats la proportion globale de sujets correctement reclassés:

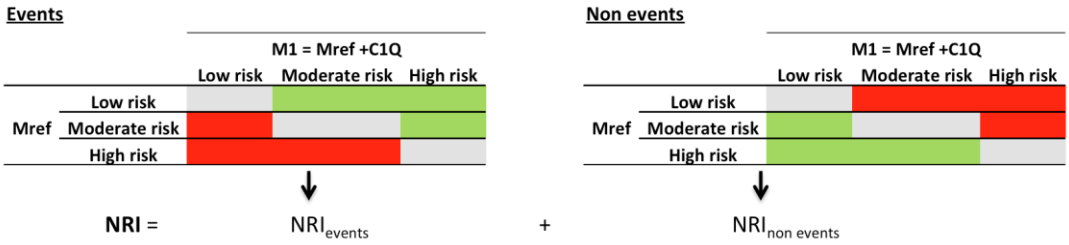
$$NRI_{cas} = \frac{(Nsup, cas - Ninf, cas)}{N\ total\ de\ cas}$$

$$NRI_{non\ cas} = \frac{(Ninf, non\ cas - Ninf, cas)}{N\ total\ de\ non\ cas}$$

$$NRI_{global} = NRI_{cas} + NRI_{non\ cas}$$

Voici ci-après schématiquement le principe du NRI catégoriel:

Les sujets correctement reclassés sont représentés en vert et les sujets mal reclassés en rouge. Les sujets non reclassés sont en gris.



Si le NRI est supérieur à 0, alors l'ajout du facteur au modèle de référence a amélioré la classification des sujets dans des catégories de risque adéquates au regard des évènements observés à un temps t.
La significativité du NRI peut être testée simplement (hypothèse nulle: NRI = 0) à l'aide d'un paramètre z suivant une loi normale

$$z = \frac{\widehat{\mathrm{NRI}}}{\sqrt{\frac{\hat{p}_{\mathrm{up,events}} + \hat{p}_{\mathrm{down,events}}}{\#\ \mathrm{events}} + \frac{\hat{p}_{\mathrm{up,nonevents}} + \hat{p}_{\mathrm{down,nonevents}}}{\#\ \mathrm{nonevents}}}}$$

Une réserve majeure faite au NRI catégoriel est qu'il dépend fortement des catégories de risque choisies pour son calcul. En pratique, cette approche en catégorie suppose l'existence de catégories de risque préétablies dans la littérature. Typiquement, dans la recherche cardiovasculaire avec l'acceptation consensuelle des catégories de risque de Framingham pour la survenue d'évènements cardiovasculaires, le NRI catégoriel est envisageable.

Dans les domaines d'applications abordés dans cette thèse (oncologie digestive et transplantation rénale), cela n'est pas le cas et afin de pallier l'absence de catégories de risque préétablies, des extensions récentes ont été proposées (Pencina *et al*, 2011). Il s'agit du NRI continu. Dans les cas où aucune catégorie de risque n'existe, il est en effet plus approprié d'utiliser le NRI continu. Nous pourrions envisager, par exemple dans le cas de 3 catégories de risque différentes, d'utiliser les tertiles des probabilités de faire l'évènement au temps t du modèle de référence pour définir les seuils de ces catégories, mais ces valeurs sont par définition dépendantes de la distribution obtenue dans l'échantillon analysé (et donc pas toujours reproductibles). Par ailleurs, elles ne permettent pas toujours d'identifier des catégories de patients différentes en termes de prise en charge clinique ce qui va à l'encontre de l'essence même du NRI catégoriel.

Nous reprenons notre exemple précédent concernant la valeur additionnelle du microscope moléculaire (score de rejet humoral) pour prédire la perte du greffon rénale comme illustré ci-dessous.

		Nombre de patients	Nombre d'évène-ments	HR	95% IC	Valeur de p
Age du donneur (années)	<60 ans	54	11	1	-	-
	≥60 ans	20	10	3.84	[1.48 – 9.96]	0.006
eGFR* au moment du rejet	≥30 mL/min	52	10	1		
	<30 mL/min	22	11	1.74	[0.70 - 4.35]	0.232
Score humoral histologique (g+ptc+v+cg+c4d)		74	21	1.43	[1.09 – 1.90]	0.011
Score du rejet humoral		74	21	2.22	[1.37 - 3.58]	0.001

M_{ref} : Modèle de référence ou conventional features

$M_{ref+facteur}$: Modèle de référence + microscope moléculaire ou conventional features + molecular microscope

Nous avons initialement démontré que l'apport de l'information du microscope permettait d'améliorer significativement l'indice de concordance et donc la capacité discriminante du modèle de référence : C_{ref} = 0,770 à $C_{ref+facteur}$ = 0,8102, avec une différence moyenne obtenue par bootstrap sur 1000 échantillons de 0,0496505 (95%IC= 0,0473 – 0,0519) et donc significativement supérieure à 0.

Dans le contexte du NRI catégoriel les patients doivent être classés dans une catégorie de risque de perte du greffon à 8 ans (ici le temps t d'intérêt est donc 8 ans post-transplantation). En transplantation, il n'y a aujourd'hui pas de convention pour définir différentes catégories de patients à risque pour la perte du greffon. Aussi l'utilisation du NRI catégoriel n'est pas vraiment recommandée mais dans un objectif illustratif nous allons considérer les tertiles des probabilités de perdre le greffon à 8 ans pour définir les 3

catégories de risque. Les patients inclus dans l'analyse et la modélisation multivariée sont alors catégorisés en nombre égal dans chacune de ces catégories :

Risque faible: $p < 0.25$ (25%)
Risque intermédiaire: 0.25 (25%) $\leq p \leq 0.53$ (53%)
Risque élevé: $p > 0.53$ (53%)

On peut alors très facilement regarder, lorsque nous incluons l'information apportée par le microscope moléculaire, si les probabilités prédites à 8 ans permettent aux patients de changer de catégorie de risque.

Ainsi, nous obtenons des tables de reclassification comme suit:

<u>Pas de perte du greffon avant 8 ans (n=53)</u>

		Mref+microscope		
		Risque faible	**Risque intermédiaire**	**Risque élevé**
Mref	**Risque faible**	22 42%	3 6%	0 0%
	Risque intermédiaire	4 7%	10 19%	2 4%
	Risque élevé	0 0%	7 13%	5 9%

NRI (population sans évènement) = 0.113

<u>Perte de greffon observée dans les 8 ans post-greffe (n=21)</u>

		Mref+microscope		
		Risque faible	**Risque faible**	**Risque faible**
Mref	**Risque faible**	0 0%	0 0%	0 0%
	Risque intermédiaire	1 5%	5 24%	4 19%
	Risque élevé	0 0%	0 0%	11 52%

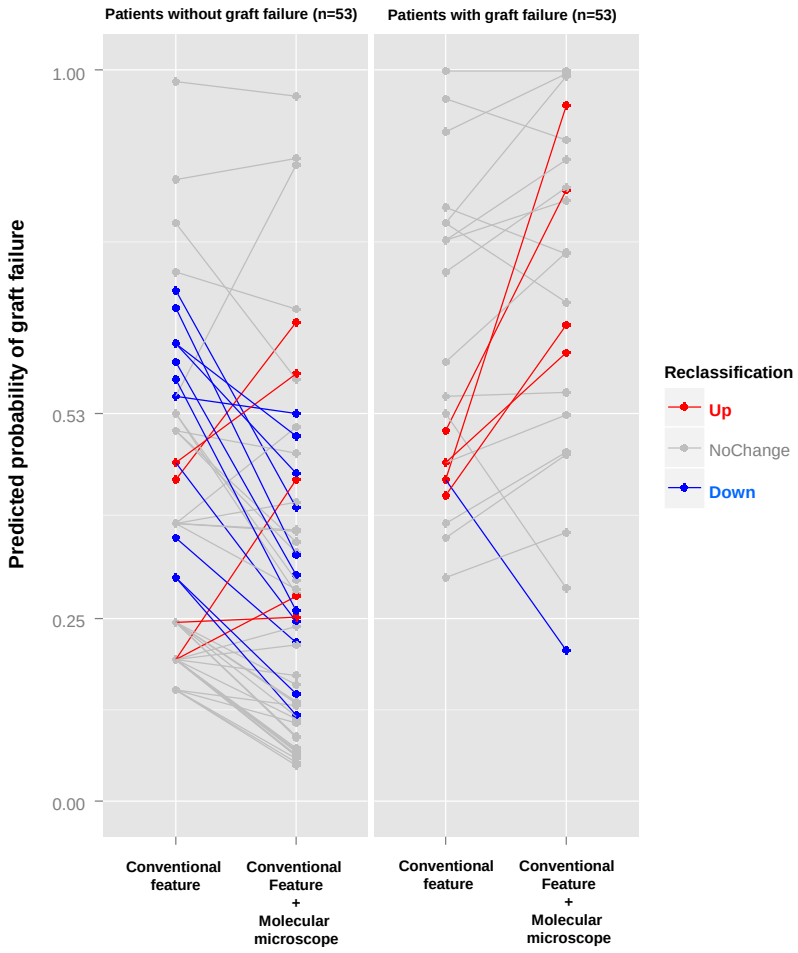
NRI (population avec évènement) = 0.143

NRI = 0.113+0.143=0.256 ; 95%CI (0.0095- 0.5026) P=0.04177

On peut donc conclure que l'apport du microscope moléculaire a permis significativement de reclasser les patients ayant perdu le greffon à 8 ans majoritairement dans des catégories de risque plus élevé, et inversement pour les patients qui n'ont pas encore perdu le greffon à 8 ans.

Ceci est classiquement présenté dans les publications qui utilisent le NRI catégoriel.

Un développement proposé dans cette thèse consiste en une représentation graphique plus didactique et plus informative sur laquelle les mouvements précis de reclassifications sont présentés. Voici le résultat correspondant à l'exemple du microscope moléculaire pour la prédiction de la perte du greffon à 8 ans dans le cas du NRI catégoriel:



Dans ce contexte de catégories de risque a priori bien définies et pertinentes cliniquement (ce qui n'est pas forcément le cas dans cette exemple puisqu'on se base sur les tertiles pour définir les catégories) une reclassification adéquate correspond au changement de catégories qui vont dans le bon sens (c'est-à-dire diminution du gradient de risque pour les patients n'ayant pas perdu le greffon et augmentation du gradient pour les patients qui l'ont perdu). Les lignes grises correspondent aux patients qui ne changent pas de catégories, donc a priori pour lesquels l'ajout de l'information n'a pas de répercutions cliniques directes dans le cas de seuils pertinents.

III.3.2.1.2 Le NRI continu

L'idée du NRI continu repose toujours sur une reclassification effectuée séparément chez les cas et les non-cas, mais considère un cas comme correctement reclassé si sa probabilité de faire l'évènement au temps t augmente après l'ajout du nouveau marqueur. Donc un non-cas est correctement reclassé si sa probabilité diminue après la prise en compte du nouveau facteur et ceci indépendamment de l'amplitude de la reclassification. Nous obtenons ainsi la proportion globale de sujets correctement reclassés.

Le NRI continu permet donc d'apprécier les mouvements des reclassifications des probabilités prédites de faire l'évènement avec l'ajout du nouveau marqueur en comparaison à celle prédite par le modèle de référence.

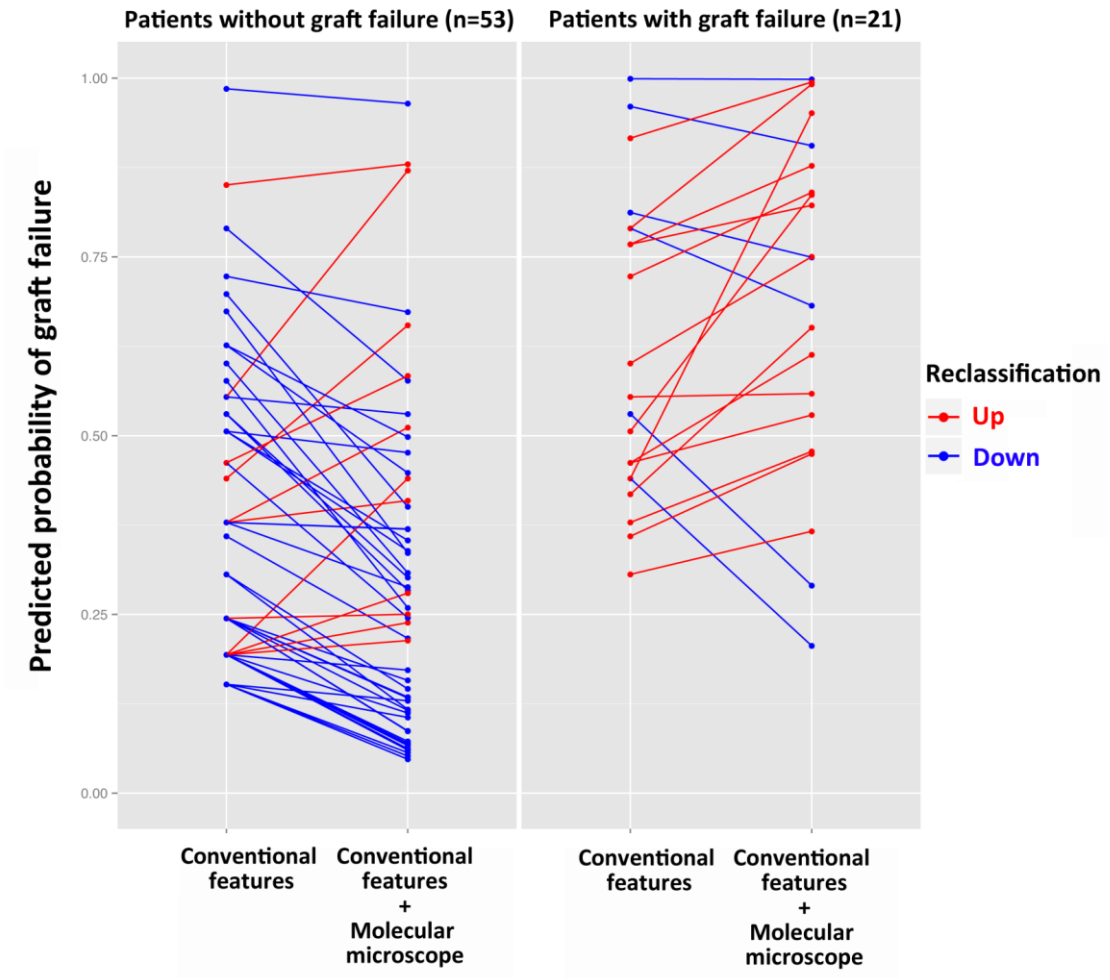
Voici ci-après pour illustration le résultat obtenu pour l'exemple du microscope moléculaire:

	Diminution de la probabilité prédite par le modèle avec le microscope moléculaire comparé au modèle de référence	Augmentation de la probabilité prédite par le modèle avec le microscope moléculaire comparé au modèle de référence
Pas de perte du greffon à 8 ans n=53	42	11
Perte du greffon à 8 ans n=21	6	15

La représentation graphique proposée précédemment peut s'avérer à nouveau plus informative. Voici le résultat correspondant à l'exemple du microscope moléculaire pour la prédiction de la perte du greffon à 8 ans dans un contexte de NRI continu:

Ci-dessous l'illustration de la valeur additive du microscope moléculaire (score du rejet humoral aigu) pour la reclassification du risque de perte du greffon (NRI continu).

Les lignes bleues chez les patients sans perte du greffon indiquent un mouvement de la prédiction du risque par le microscope moléculaire dans la bonne direction (vers le bas) (42/53 = 79.2%). Inversement, les lignes rouges chez les patients avec perte du greffon indiquent une modification correcte, vers le haut, de l'évaluation du risque lorsqu'on utilise le microscope moléculaire (15/21 = 71.4%).



L'inclusion du score moléculaire du rejet humoral dans le modèle de référence permettait donc de reclasser les patients à faible risque (absence de perte du greffon à 8 ans) et les patients à risque élevé (perte du greffon à 8 ans), avec un NRI = 1,0135 (95% IC= 0 ,5696 – 1,4573, p < 0,0001).

La figure proposée ci-dessus permet de montrer l'amplitude des mouvements qui n'est cependant pas prise en compte dans le calcul du NRI continu. Une bonne reclassification pour un sujet qui a subi l'évènement au temps t consiste en une augmentation de sa probabilité de faire l'évènement quelle que soit cette augmentation. Ainsi, l'intégralité de l'information de reclassification n'est pas totalement captée par le NRI continu. C'est pour répondre à cette limitation qu'une autre statistique appelée l'IDI (Integrated Discrimination Index) a été proposée pour apporter une information complémentaire au NRI continu.

III.3.2.2 Integrated discrimination improvement (IDI)

L'IDI peut être considérée comme complémentaire du NRI continu. Il peut se définir comme la différence des moyennes des probabilités estimées par le modèle de référence et le modèle de référence avec l'ajout du facteur chez les cas et chez les non-cas.

Un facteur qui apporte une information additionnelle au modèle de référence en termes de reclassification du risque aura tendance à augmenter la moyenne des probabilités chez les cas (qui ont expérimenté l'évènement au temps t ou avant), et à l'opposé chez les non-cas elle aura tendance à la diminuer.

$$\widehat{\mathrm{IDI}} = (\overline{\hat{p}}_{\mathrm{new,events}} - \overline{\hat{p}}_{\mathrm{new,nonevents}}) - (\overline{\hat{p}}_{\mathrm{old,events}} - \overline{\hat{p}}_{\mathrm{old,nonevents}})$$

De la même manière que le NRI, l'IDI peut être testé simplement à 0 à partir de la statistique de test suivante:

$$z = \frac{\widehat{\mathrm{IDI}}}{\sqrt{(\widehat{\mathrm{se}}_{\mathrm{events}})^2 + (\widehat{\mathrm{se}}_{\mathrm{nonevents}})^2}}$$

En quelque sorte le NRI continu capte la validité des mouvements observés tandis que l'IDI a pour objet de quantifier l'amplitude en population de ces mouvements.

En supplément de la représentation graphique proposée pour le NRI continu (sur laquelle l'information véhiculée par l'IDI peut être évalué visuellement), il est possible d'illustrer aussi de manière didactique l'information captée par la statistique de l'IDI.

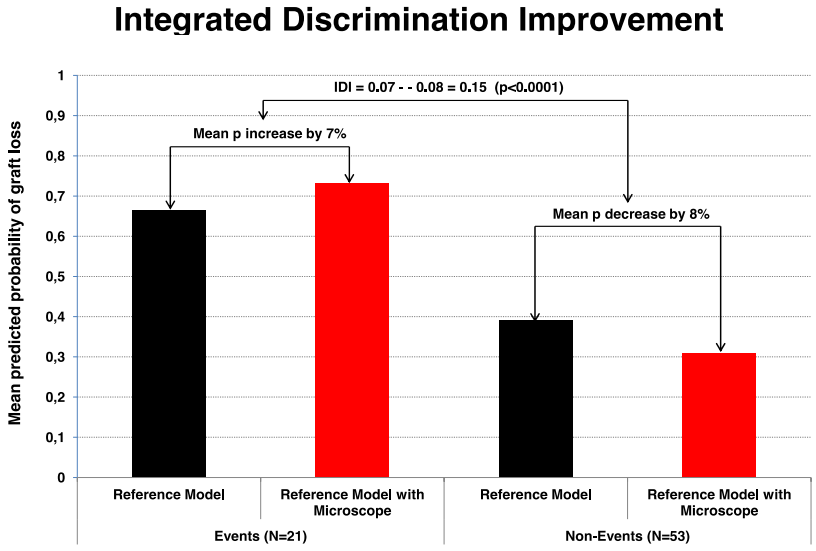
Toujours en guise d'illustration, voici le résultat obtenu pour l'exemple du microscope moléculaire:

L'IDI était de 0,1579, p = 0,00015 allant dans le sens de la considération du microscope moléculaire en complément du modèle de référence pour la prédiction de risque de perte du greffon à 8 ans.

Un autre développement proposé dans cette thèse consiste en une représentation graphique plus didactique et plus informative pour présenter les résultats de l'indice IDI.

Ainsi, voici ci-dessous l'illustration proposée de l'amélioration de la prédiction du risque de perte du greffon à 8 ans en considérant le microscope moléculaire en complément des paramètres conventionnels.

Cette amélioration de la prédiction du risque est évaluée par l'IDI qui repose sur la quantification par la moyenne des probabilités de perte du greffon dans les deux populations caractérisées par l'occurrence d'un évènement ou non. Le changement de moyenne de probabilité prédit avec l'ajout du microscope moléculaire est satisfaisant s'il se fait à la hausse chez les patients qui ont réellement perdu le greffon et à l'opposé à la baisse chez ceux qui n'ont pas perdu le greffon.



En conclusion, il est essentiel de noter que les indices de reclassifications (NRI et IDI) apportent une information additionnelle à l'index de concordance.

Ces informations C-index, NRI et IDI constituent un spectre d'information permettant de valider et d'illustrer la valeur additionnelle apportée par l'ajout d'un facteur à un modèle dit de référence constitué de facteurs pronostiques de l'évènement dit conventionnels.

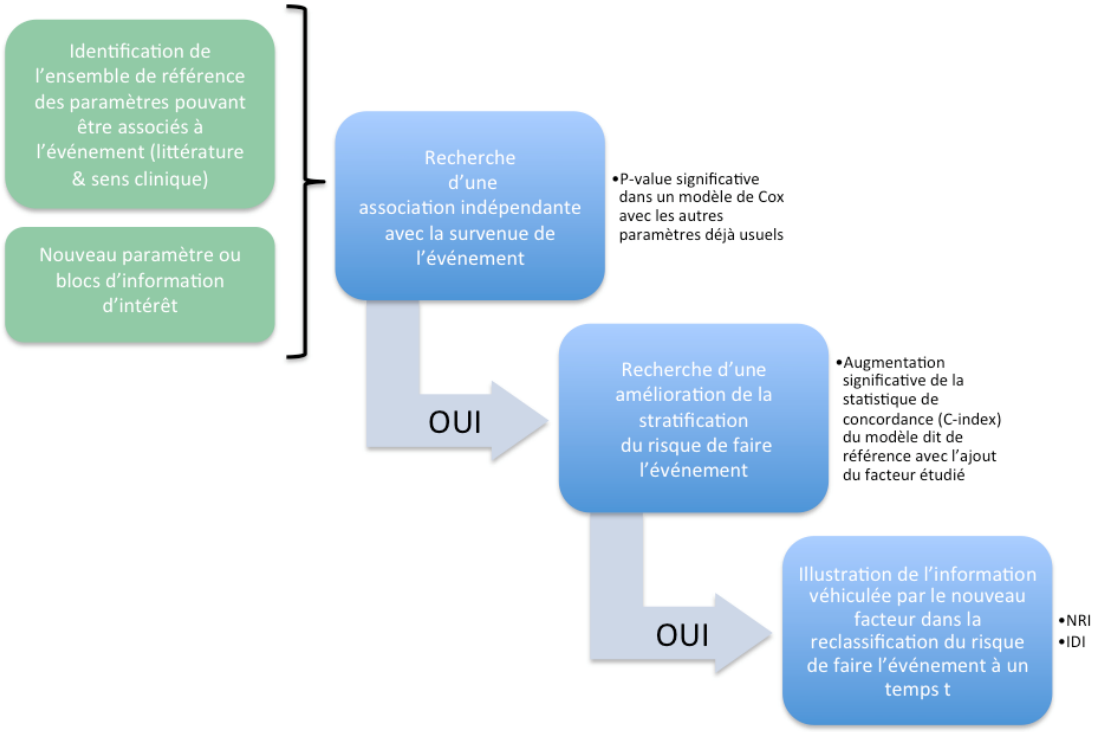
Le C-index permet d'apprécier la capacité discriminante du facteur tandis que le NRI et l'IDI permettent d'apprécier la nature de l'information apportée par ce facteur (marqueur de mauvais pronostic s'il reclasse le risque de faire l'évènement à la hausse chez les cas ou marqueur de bon pronostic s'il

reclasse à la baisse cette probabilité chez les non-cas ou les deux).

Il est aussi important de noter que ces statistiques de reclassification doivent être employées dans un contexte de modèle bien calibré sinon elles peuvent être erronées (Leening *et al*, 2014).

Un autre point intéressant à souligner est que l'ensemble de cette démarche peut s'appliquer à un bloc de paramètres et pas seulement à un seul paramètre. Tant que les modèles sont emboités, donc réalisés sur la même population de patients, et qu'ils possèdent un socle commun de variables, le calcul de la différence des C du NRI et de l'IDI est tout à fait possible. A noter toutefois que si le résultat est positif, et donc en faveur d'une valeur additionnelle du bloc de paramètres implémenté au modèle de référence, le résultat est attribuable au bloc entier et il n'est donc pas possible de quantifier immédiatement l'apport de chaque facteur constituant le bloc.

En guise de synthèse, le schéma ci-dessous illustre une proposition d'un protocole méthodologique pour explorer la potentielle valeur pronostique additionnelle d'un paramètre ou d'un bloc de paramètres pour la prédiction de survenue d'un évènement d'intérêt:



L'exemple permettant d'illustrer l'indice de concordance, le NRI et l'IDI ayant pour objectif la démonstration de l'intérêt pronostique du microscope moléculaire dans la stratification du risque de perte de greffon est issu d'une publication parue dans le JASN en 2013.
Cette étude a permis de confirmer l'intérêt du « microscope moléculaire » dans le rejet humoral pour affiner la prédiction du risque de perte du greffon et de progression de la maladie et est présentée en détail en annexe VIII.1.3 comme publication complémentaire à la thèse servant de support pour la partie méthode.

Loupy A, Lefaucheur C, Vernerey D, Chang J, Hidalgo LG, Beuscart T, Verine J, Aubert O, Dubleumortier S, Duong van Huyen J-P, Jouven X, Glotz D, Legendre C, Halloran PF
Molecular Microscope Strategy to Improve Risk Stratification in Early Antibody-Mediated Kidney Allograft Rejection. **J Am Soc Nephrol JASN** doi:10.1681/ASN.2013111149. (2014)
Impact factor: 9.3

Un autre exemple permettant d'illustrer l'indice de concordance, le NRI et l'IDI ayant pour objectif de démontrer que l'addition de la propriété des DSA à fixer le complément augmente significativement les performances d'un modèle dit de référence dans la stratification du risque de perte du greffon est issu d'une publication parue dans le NEJM en 2013 et est présentée en détail en annexe VIII.1.2 comme publication complémentaire à la thèse servant de support pour la partie méthode.

Loupy A, Lefaucheur C, Vernerey D, Prugger C, Duong van Huyen J-P, Mooney N, Suberbielle C, Frémeaux-Bacchi V, Méjean A, Desgrandchamps F, Anglicheau D, Nochy D, Charron D, Empana J-P, Delahousse M, Legendre C, Glotz D, Hill GS, Zeevi A, Jouven X
Complement-binding anti-HLA antibodies and kidney-allograft survival. **N Engl J Med** 369: 1215–1226, doi:10.1056/NEJMoa1302506. (2013)
Impact factor: 54.4

III.4 Construction d'outils à visée pronostique sur la base de modèles intégratifs

Enfin, en considérant conjointement le spectre de paramètres pronostiques de l'occurrence de l'évènement étudié, il est important de développer des outils de prédiction adéquats permettant une stratification précise du risque d'évènement, à l'échelle individuelle et populationnelle

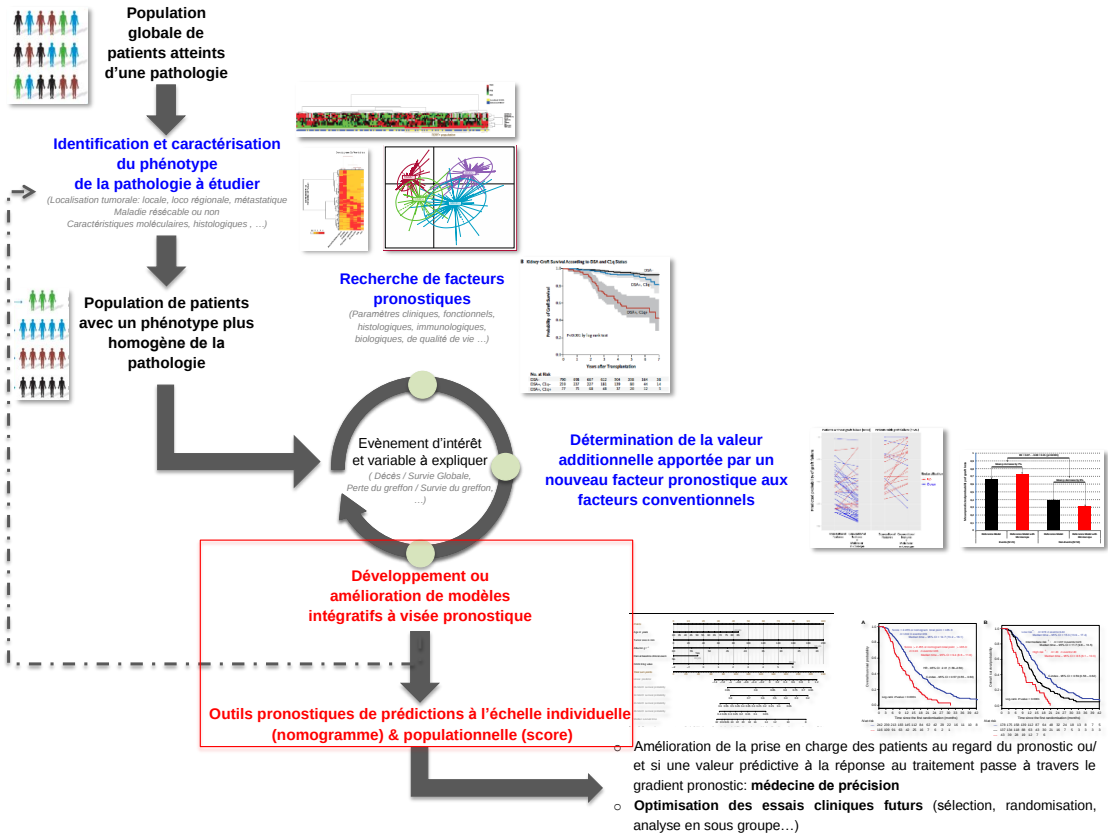


Figure : Quatrième et dernière étape d'une recherche à visée pronostique : le développement des outils de prédiction adéquats permettant une stratification précise du risque d'évènement, à l'échelle individuelle et populationnelle

Après avoir déterminé un sous-ensemble de paramètres pronostiques de la survenue d'un évènement, l'idée qui vient naturellement est de pouvoir estimer pour chaque sujet son risque individuel de faire l'évènement. L'idée sous-jacente est de proposer un résumé de l'information pronostique la plus précise possible.

III.4.1Le modèle de Cox multivarié

Nous avons vu précédemment que c'est le modèle de Cox qui est utilisé pour modéliser l'apparition d'un évènement dans le temps et donc pour analyser des données dites de survie censurées à droite. Dans ce modèle multivarié, la relation de chaque facteur inclus avec le risque de faire l'évènement, est modélisée et pondérée relativement à la force de sa relation avec la distribution de survie. Ces paramètres sont souvent appelés « bêtas ». Ces bêtas estimés par le modèle pour chaque facteur définissent donc leur contribution relative à la distribution de survie.

Il est donc essentiel de prendre en compte cette pondération et donc de ne pas considérer une contribution similaire pour chaque facteur pronostique identifié.

La robustesse des estimations est donc un point clef, ce qui implique un examen attentif des performances et de la validité du modèle de Cox multivarié sur lequel va reposer ces estimations et *in fine* la construction des outils.

III.4.1.1 Performance du modèle

Une recherche visant à développer un nouveau modèle de prédiction doit absolument proposer une évaluation approfondie de ce dernier pour quantifier ses performances, dans l'objectif de fournir des prédictions dans un cadre méthodologique clairement établi. La validité, justesse ou encore précision d'un modèle de prédiction, doit être vérifiée selon deux paramètres : la discrimination et la calibration.

III.4.1.1.1La discrimination

Comme évoqué précédemment la valeur prédictive ou la capacité discriminante d'un modèle est définie par son aptitude à discriminer les sujets qui ont fait l'évènement de ceux qui ne l'on pas fait et ainsi séparer les sujets de pronostic différent (Pencina MJ *et al*, 2015).

Autrement dit, un modèle est discriminant s'il prédit une probabilité d'évènement plus importante à un sujet qui fait effectivement un évènement par rapport à celui qui n'en fait pas ou par rapport à celui qui fait un évènement mais plus tardivement.

Dans le cas de l'analyse de données de survie censurées à droite, c'est l'indice de concordance qui permet d'évaluer la capacité discriminante d'un modèle de Cox. Outre l'estimation de l'indice de concordance, il est impératif de produire un intervalle de confiance autour de cette estimation. A cet effet, nous avons utilisé des méthodes de ré-échantillonnage (bootstrap) pour créer de la variabilité de mesure et ainsi obtenir un intervalle de confiance à 95% autour de l'estimation. Concrètement, cela revient à construire n échantillons de données basés sur la population d'analyse (n > 250) avec un tirage au sort aléatoire avec remise. Sur chacun de ces échantillons de données, il est alors possible de calculer une statistique de concordance. Nous obtenons alors une distribution de n C-statistiques à partir de laquelle, en calculant les percentiles (α/2% et 100 – α /2%), nous obtenons un intervalle de confiance à 95% autour de l'estimation de la C statistique observée sur l'échantillon de départ.

III.4.1.1.2 La calibration

La calibration est un paramètre central à prendre en compte. En effet, un modèle discriminant mais mal calibré aura une utilité clinique très limitée. La calibration ou la qualité d'ajustement du modèle se réfère à sa capacité à produire des prédictions précises non biaisées donc fiables, et ceci dans des groupes de patients dit similaires. La calibration peut être alors perçue comme la concordance entre les probabilités prédites et celles observées dans l'échantillon de données.

La calibration d'un modèle est évaluée visuellement par un graphique appelé « courbe de calibration » et aussi statistiquement par le test de Hosmer-

Lemeshow pour un modèle logistique ou son extension dans le cas d'un modèle de survie ou de Cox.

En pratique, un modèle est considéré comme bien calibré si la différence entre ses prédictions et les observations faites dans l'échantillon de données qui a servi à sa construction est proche de 0 (calibration parfaite).

Pour le test de Hosmer-Lemeshow, n groupes de patients équilibrés sont définis par les quantiles (usuellement les déciles ; n = 10) de la probabilité prédite par le modèle. Le biais mesuré se définit comme la somme des différences entre la moyenne des probabilités prédites et la moyenne des probabilités observées dans chacun des groupes de patients dit similaires. La calibration du modèle peut alors s'apprécier par la p-value du test basé sur cette somme. Lorsque cette somme est inférieure à 0.1, cela indique une calibration non satisfaisante.

Dans le cas de l'évaluation graphique, une courbe de calibration est produite dans un repère orthonormé (0,1)×(0,1) dont l'axe des abscisses correspond à la probabilité prédite par le modèle et l'axe des ordonnées à celle observée dans l'échantillon de données. La courbe de calibration peut être construite de deux manières. L'une, représente les prédictions et observations dans chaque groupe de patients dit similaires et possède ainsi un lien direct avec le test statistique de Hosmer-Lemeshow. Cette représentation véhicule la même limite que le test de Hosmer-Lemeshow qui est dépendant du nombre n de groupes de patients choisis. L'autre, propose une représentation lissée de la courbe de calibration obtenue par une régression entre les valeurs prédites et observées et donc ne souffre pas des limites du test de Hosmer-Lemeshow, en s'affranchissant du choix du nombre de groupes de patients similaires à considérer. La seconde méthode doit donc être la méthode à privilégier. La courbe de calibration permet dans les deux cas d'apprécier la concordance entre la prédiction du modèle et la réalité des données. En pratique, cela revient à s'assurer que la courbe ou les points représentés sont proches de la bissectrice de la zone de représentation.

Ce qui est particulièrement intéressant avec l'évaluation graphique de la calibration, c'est qu'il est possible en un coup d'œil d'apprécier la calibration du modèle pour l'ensemble des patients. Idéalement, un modèle doit

présenter une bonne calibration pour l'ensemble des patients qu'ils soient à risque élevé ou faible de présenter l'évènement et même à risque intermédiaire.
Dans la construction de la courbe de calibration, il est aussi possible d'ajouter un degré supplémentaire d'incertitude pour prendre en compte la fluctuation possible des prédictions du modèle, en créant un intervalle de confiance autour de ces dernières par une procédure de ré-échantillonnage par bootstrap. Avec la présence de cet élément, nous considérons que l'intervalle de confiance doit couper la bissectrice du repère orthonormée pour que nous puissions conclure à une bonne calibration.
Enfin, il est important de garder à l'esprit que la calibration (et aussi la discrimination) d'un modèle ne constitue pas une validation externe puisque ce sont des propriétés intrinsèques à la population de développement qui sont appréciées.

Dans un contexte d'analyse de données de survie du patient ou du greffon avec des données censurées à droite et avec une modélisation de Cox, il est impératif de définir des temps d'intérêt auxquels la bonne calibration du modèle doit être évaluée. Lorsque le modèle est utilisé pour dériver des prédictions individuelles, il est évident que les temps auxquels ces prédictions sont proposées doivent être considérés. Ces temps sont, généralement, des moments clefs auxquels des options ou décisions thérapeutiques ou des bilans sont envisagés. C'est donc à ces temps précis qu'il faut évaluer la bonne calibration du modèle.
En transplantation rénale, les temps d'intérêt pour une modélisation de la survie du greffon en post-transplantation sont en général fixés à 3, 5 puis 7 ou 8 ans post-transplantation. En oncologie digestive, dans un contexte de cancer pancréatique localement avancé, nous les fixerons par exemple à 6, 12, 24 et 48 mois tandis qu'en situation de maladie avancée et en seconde ligne de traitement les temps à privilégier seront 6, 12, 18 et 24 mois.
L'évaluation de la bonne calibration d'un modèle de survie à un temps t peut alors s'effectuer selon les 3 approches évoquées précédemment.

III.4.1.1.2.1 Le test de Hosmer-Lemeshow

Pour mesurer la bonne calibration d'un modèle logistique (modèle reposant sur une variable à expliquer binaire), Hosmer et Lemeshow ont proposé une statistique pour laquelle ils ont montré par une technique de simulation qu'elle pouvait être assimilée à une variable distribuée selon une loi du Chi-2 à n-2 degré de liberté si n groupes de patients sont considérés (Hosmer & Lemesbow, 1980; Hosmer *et al*, 1997; Jr & Lemeshow, 2004).

Le test est construit pour évaluer si les taux d'évènements observés correspondent aux taux d'évènements prédits dans les sous-groupes de patients de la population de développement du modèle. Dans un premier temps, les patients sont ordonnés selon leur probabilité prédite par le modèle de faire l'évènement Y (P(Y=1)). Un nombre n de groupes de patients équilibrés est construit sur la base des quantiles en question.
La statistique du test de Hosmer-Lemeshow est alors donnée par la formulation suivante :

$$H = \sum_{g=1}^{G} \frac{(O_{1g} - E_{1g})^2}{E_{1g}} + \frac{(O_{0g} - E_{0g})^2}{E_{0g}} = \sum_{g=1}^{G} \frac{(O_{1g} - E_{1g})^2}{N_g \pi_g} + \frac{(N_g - O_{1g} - (N_g - E_{1g}))^2}{N_g(1-\pi_g)} = \sum_{g=1}^{G} \frac{(O_{1g} - E_{1g})^2}{N_g \pi_g (1-\pi_g)}.$$

où

O_{1g}, E_{1g}, O_{0g}, E_{0g}, N_g et Π_g correspondent respectivement à l'observation de l'évènement (Y=1), la prédiction de l'évènement (Y=1), l'observation du non-évènement (Y=0), la prédiction du non-évènement (Y=0) et le nombre de sujets dans le g^{ème} groupe de quantile et G le nombre de groupes de patients.

Comme évoqué ci-dessus cette statistique H suit asymptotiquement une loi du Chi2 du Chi-2 à G-2 degré de liberté.

Autrement dit, le test de Hosmer-Lemeshow regroupe les probabilités prédites par le modèle en n groupes et pour chaque groupe on observe l'écart entre les valeurs prédites et celles observées. L'importance de la distance entre ces valeurs est évaluée grâce à une statistique a n-2 degré de liberté qui teste H_0 : Distance faible vs H_1 : Distance élevée.

Dans le cadre d'une modélisation de Cox avec des données censurées à droite, plusieurs façons de construire une extension de cette statistique ont été proposées et discutées dans la littérature.

Pour commencer il faut considérer un temps d'intérêt t. Les probabilités prédites par le modèle au moment t auquel la calibration du modèle est évaluée peuvent être calculées simplement.

Le modèle de Cox se présentant sous la forme suivante :

$S(t, X_j) = S_0(t)^{\exp(LP_j)}$ avec LPj le prédicteur linéaire pour le patient j et S0(t) est la survie de base au temps t.

Si nous calculons dans l'échantillon de données le risque de base S0 alors il est très simple d'obtenir pour le patient i sa probabilité de survie au moment t $S(t, X_j)$.

A ce stade il est donc possible pour un patient i de connaître sa probabilité de faire l'évènement au temps t qui est égale à 1- $S(t, X_j)$ et nous obtenons ainsi un vecteur de n observations.

$$\text{probabilité predite de faire l'evenement au temps t} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{n-1} \\ p_n \end{pmatrix}$$

Plusieurs auteurs (D'Agostino & Nam, 2003; Cook & Ridker, 2009) ont proposé que le nombre d'événements observés à un moment fixé t dans l'échantillon soit basé sur la probabilité de survie estimée par la méthode de Kaplan-Meier afin de tenir compte du phénomène de censures.

La statistique du test de Hosmer-Lemeshow peut donc s'écrire de la manière suivante :

$$\widehat{HL}_{DAN}(t) = \sum_{i=1}^{K} \frac{n_i\left(\left(1-KM_i(t)\right) - \bar{p}_i\right)^2}{\bar{p}_i(1-\bar{p}_i)} \sim \chi^2_{K-1}$$

avec $KM_i(t)$ l'estimation de survie issue de la méthode de Kaplan-Meier au temps t pour le groupe i

n_i le nombre d'observations dans le groupe i

$\bar{p}_i = \left(\frac{1}{n_i}\right)\sum_{j=1, j\epsilon J_i}^{n_j}\left(1 - \left[\widehat{S_0}(t)\right]^{\exp(X_j\widehat{\beta})}\right)$ et K le nombre de groupes

Il faut noter que les auteurs qui ont proposé cette approche mettent en garde sur l'utilisation d'une p-value pour juger de la bonne calibration d'un modèle et préfèrent recommander de regarder la statistique du test en sachant que plus elle est faible plus les prédictions du modèle se rapprochent de celles observées dans l'échantillon et donc plus le test indique une bonne calibration (D'Agostino *et al*, 2001; Marrugat *et al*, 2007).

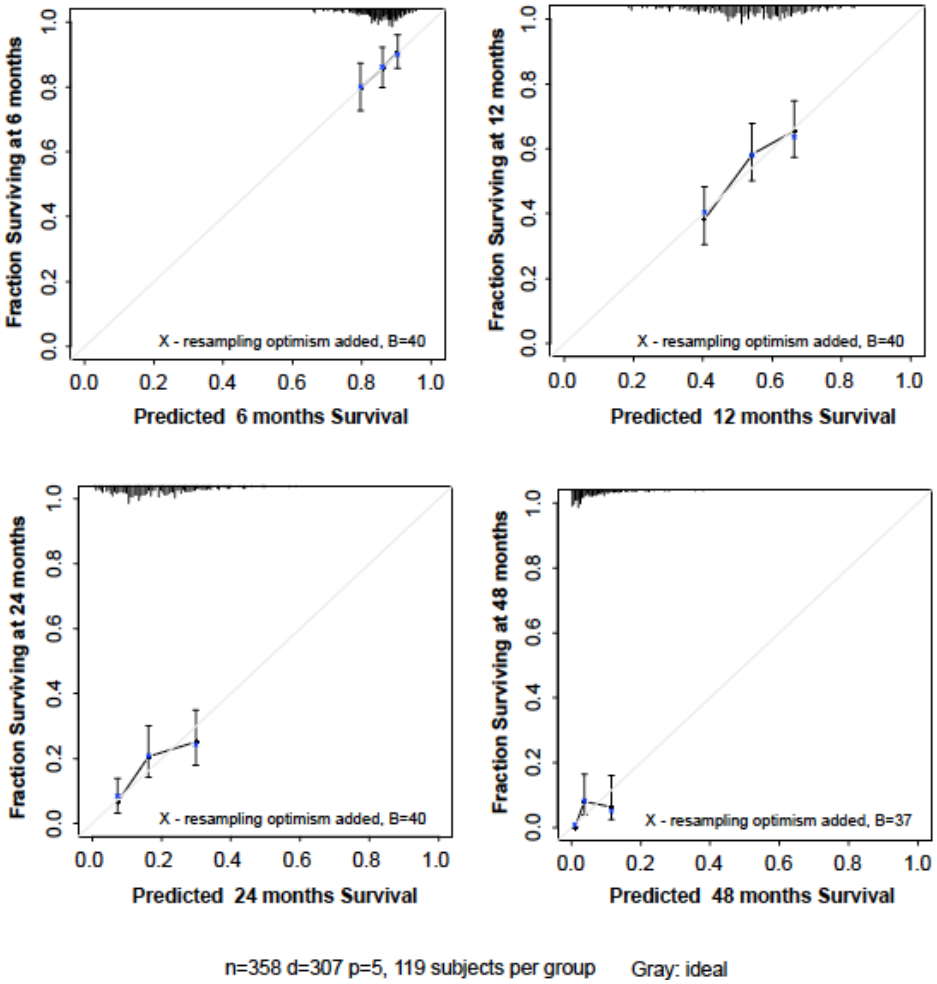
Au final cette recommandation peut être aussi suivie en évaluant la bonne calibration du modèle par une courbe de calibration qui synthétise sur un graphique l'information prise en compte pour le calcul du test de Hosmer-Lemeshow ci-dessus.

III.4.1.1.2.2 Le graphique de calibration issu du test de Hosmer-Lemeshow

Dans chaque groupe de patients semblables déterminé dans l'extension du test de Hosmer-Lemeshow aux données de survie et donc catégorisés selon leur probabilité prédite par le modèle au temps t, il est donc possible de comparer la moyenne des probabilités prédites par le modèle au résultat moyen observé par la méthode de Kaplan-Meier. L'illustration graphique du

test de Hosmer-Lemeshow consiste en la représentation dans un repère orthonormé de chacun de ses points.
Il est important de noter toutefois que le nombre de groupes à constituer et leur définition qui est basée sur les quantiles, demeurent des choix arbitraires qui peuvent rendre la méthode imprécise.

Ci-dessous un exemple de courbe de calibration pour un modèle multivarié de Cox construit dans une étude présentée dans cette thèse et qui avait pour objet de modéliser la survie globale au diagnostic chez des patients atteints d'un cancer pancréatique localement avancé.



Dans la figure ci-dessus une courbe de calibration à 6, 12, 24 et 48 mois est proposée pour le modèle multivarié final.
L'axe vertical reporte la proportion observée des patients survivants au moment t. La ligne grise représente un modèle parfaitement calibré. La ligne

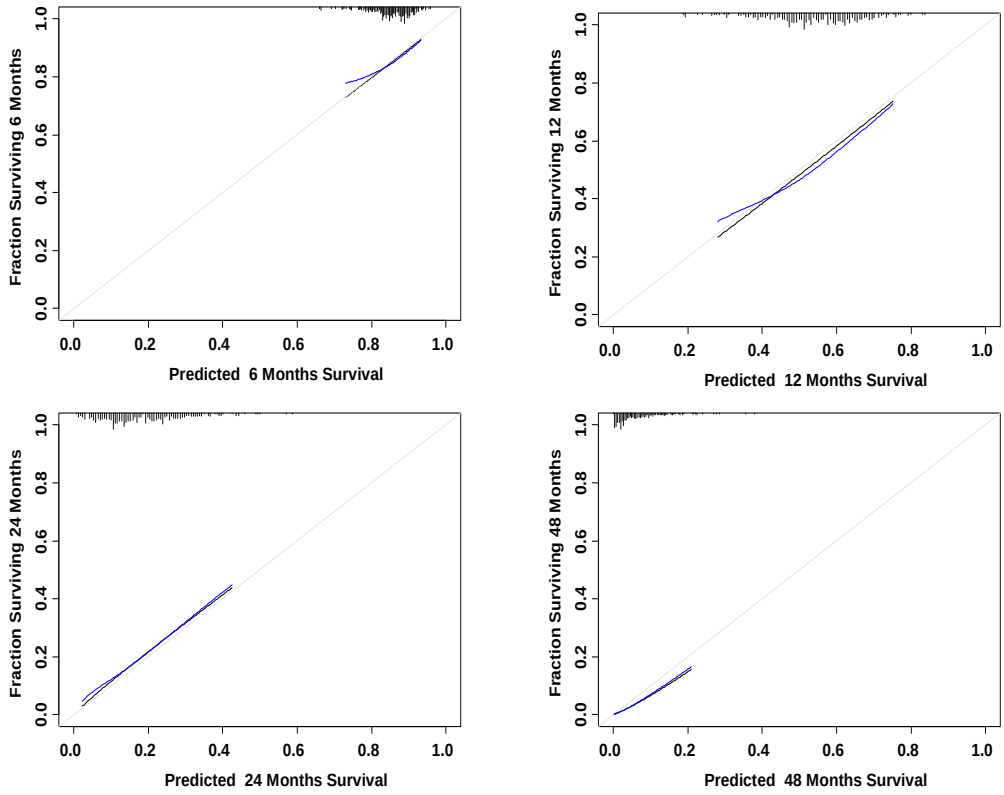
continue illustre la performance actuelle du modèle de prédiction avec des intervalles de confiance à 95% en utilisant une procédure de ré-échantillonnage par bootstrap.

III.4.1.1.2.3 La courbe de calibration lissée

Il est relativement aisé de concevoir que le choix arbitraire du nombre de groupes et ainsi la catégorisation des patients à travers ces groupes peut entrainer une confusion. En effet, deux patients voisins en termes de prédiction de probabilité de faire l'évènement peuvent tout à fait se retrouver dans des groupes différents alors qu'ils possèdent une prédiction quasi-similaire. Une solution acceptable est de choisir un nombre de groupe important (les déciles sont préconisés), mais les résultats peuvent être difficilement interprétables si l'hétérogénéité pronostique de la population est faible ou si la taille d'échantillon de la population analysée est modérée.
Des techniques de lissage (smoothing techniques) sont dans ce cas recommandées et permettent de s'affranchir de la constitution de groupes. En effet, ces techniques peuvent être utilisées pour estimer les probabilités observées en relation avec les probabilités prédites, en utilisant par exemple l'algorithme de lœss (Harrell, 2001).
Ces techniques rendent les résultats beaucoup plus lisibles et s'affranchissent de la catégorisation des sujets qui peut être problématique.

En guise d'illustration voici ci-dessous la courbe de calibration lissée à 6, 12, 24 et 48 mois en reprenant le modèle multivarié évoqué précédemment pour les patients atteints d'un cancer pancréatique.
Il est aussi très important de noter qu'il est essentiel d'adresser la question de la robustesse de la calibration observée. Ceci se fait usuellement en introduisant de la variabilité par des méthodes de ré-échantillonnages (bootstrap) ou par des procédures de pénalisation (régression de lasso) (Steyerberg *et al*, 2001). Ceci permet de considérer que nous sommes peut-

être dans une situation optimale de par la configuration de l'échantillon de données et ainsi conclure de façon trop optimiste que la calibration du modèle est bonne alors qu'elle est en réalité liée à un arrangement particulier des données.



Black : observed Gray: ideal
Blue : optimism corrected B=1000 based on observed-predicted

III.4.1.2 Validation du modèle :

Un modèle pronostique déterminé sur un échantillon de sujets est, par essence, aussi bon que possible pour cet échantillon. Ainsi, il ne faut pas s'attendre à ce qu'il soit automatiquement aussi performant pour de nouveaux sujets. Il est donc important d'étudier la valeur pronostique de ce modèle sur un autre échantillon que celui qui a servi à le définir. En effet, Il n'est pas acceptable de publier le développement d'un modèle pronostique sans validation, en tout cas au moins sans validation interne. Prétendre qu'un

modèle est cliniquement pertinent, est acceptable seulement avec une validation externe indépendante.

III.4.1.2.1Validation interne du modèle

À défaut de posséder une cohorte indépendante pour permettre une validation externe, des méthodes de validation interne sont disponibles, basées sur des techniques de ré-échantillonnage comme le bootstrap.

Le bootstrap est la technique de simulation la plus courante et a été décrite initialement par Bradley Efron. L'idée est que l'ensemble de données analysées constitue un échantillon aléatoire de patients représentatifs d'une population générale. Bootstrapper est le fait de générer de multiples échantillons de données, chacun de la même taille que celui d'origine, par tirage au sort avec remise, ce qui implique mécaniquement qu'un patient préalablement sélectionné peut être sélectionné à nouveau.

Sur la base des résultats obtenus dans les n échantillons générés, des intervalles de confiance peuvent être déterminés pour les paramètres estimés du modèle (bêta, HR et p value) et reflètent ainsi la robustesse des estimations faites dans le modèle final dans l'échantillon de développement.

III.4.1.2.2 Validation interne dans un échantillon provenant d'un fractionnement des données

Quand on se trouve en possession d'un grand jeu de données, il est souvent tentant de le séparer en deux. Lorsque cette approche est employée, l'ensemble du jeu de données disponible est divisé aléatoirement par tirage au sort, stratifié ou non, en 2 échantillons; un échantillon de développement et un échantillon de validation dont le rapport des effectifs peut être 1/2 1/2 ou encore 2/3 1/3 selon l'échantillon de départ. Cette approche a longtemps été mise en avant dans les études pronostiques, cependant elle véhicule des

limites importantes. En effet, il n'y pas de consensus concernant la répartition des données dans les cohortes et les résultats peuvent variés selon le découpage choisis (1/2 1/2 ou encore 2/3 1/3).

De plus les deux ensembles de données seront par construction très similaires puisqu'ils ne varient que par le hasard, ce qui force en quelque sorte la validation des résultats identifiés dans la cohorte de développement. Une meilleure alternative, si la taille de l'échantillon est suffisamment grande, est de le diviser selon l'emplacement géographique du centre par exemple, on parlera alors de validation externe (Collins *et al*, 2015) (Moons *et al*, 2015).

III.4.1.2.3 Validation externe du modèle

Un modèle doit idéalement être validé sur une population indépendante c'est ce qu'on appelle une validation externe. Elle confère au modèle ainsi testé des qualités de reproductibilité et de transportabilité importantes. Ces qualités dépendent évidemment du type de validation externe réalisée. En effet, plusieurs types de validation externe peuvent être distingués. On pourra notamment parler de validation externe temporelle si seule la période d'inclusion est modifiée avec par exemple une cohorte de développement incluant les patients traités entre 2005 et 2010 et la cohorte de validation ceux traités après 2010. Une validation externe spatiale est quant à elle caractérisée par la modification seule du lieu de prise en charge du patient (mêmes investigateurs). Cependant la force de la validation est généralement considérée comme plus forte avec une validation qualifiée de complètement externe qui est caractérisée par un jeu de donnée impliquant d'autres centres, d'autres investigateurs etc.… Cette dernière validation est le gold standard de la validation. Par ailleurs, lorsque le modèle est développé à partir de patients issus d'un essai clinique et donc potentiellement sélectionnés il est important de le valider dans une cohorte de patients représentatifs de la pratique clinique courante afin de pouvoir le généraliser à l'ensemble des patients.

De plus des travaux récents suggèrent qu'une validation externe puisse être au moins composée de 100 évènements d'intérêt (Collins *et al*, 2016). Il faut donc essayer de respecter cette recommandation en sachant évidemment

que selon la pathologie étudiée il n'est pas toujours possible d'atteindre ce seuil et une validation plus petite peut néanmoins être très intéressante.

Dans une validation externe voire interne par fractionnement des données l'idée qui vient naturellement est de répliquer le modèle. Si le modèle est répliqué et que tous les facteurs inclus dans ce dernier demeurent significativement associés à l'occurrence de l'évènement d'intérêt, c'est la situation idéale. Cependant, si certaines estimations fluctuent un peu, ce n'est pas rédhibitoire dans le sens où ce que l'on cherche à valider est plutôt la capacité discriminante du modèle, c'est-à-dire la capacité à identifier correctement les patients qui ont fait l'évènement de ceux qui ne l'ont pas fait, et ainsi hiérarchiser les patients qui ont fait l'évènement selon son temps d'apparition. C'est cette capacité discriminante et donc l'efficacité du modèle dans la stratification du risque de produire l'évènement que l'on cherche à valider.

III.4.1.3 Evaluation de l'intérêt clinique du modèle multivarié : l'analyse par courbe de décision

Apres s'être assuré que le modèle multivarié de Cox construit propose des performances en termes de discrimination et de calibration satisfaisantes et après l'avoir validé en interne et idéalement en externe, il est très important d'évaluer son intérêt clinique.

Un modèle pronostique implémente une combinaison de facteurs potentiellement multiples pour prédire le risque individuel d'évènements futurs chez les sujets étudiés. Un tel modèle est qualifié d'utile s'il fournit des prédictions précises qui informent les sujets et les cliniciens (Kattan, 2003). Un autre rôle important des modèles de prédiction est d'apporter des éléments fiables pour une décision clinique éclairée afin d'améliorer les résultats concernant l'évènement d'intérêt. Une des méthodes qui permet de répondre à cette problématique est l'analyse par courbe de décision (DCA) (Fitzgerald *et al*, 2015) .

La DCA est une méthode qui a été initialement développée pour évaluer les

bénéfices apportés aux patients de l'utilisation de l'information qui découle d'un test diagnostique pour accepter le risque véhiculé par une décision thérapeutique (de sous-traitement ou sur-traitement) en fonction d'une palette de probabilités que le diagnostic soit positif (Vickers & Elkin, 2006). En effet, la décision de traiter dépend, d'une part de la conviction que le clinicien a de la présence de la maladie, d'autre part de l'efficacité et des complications du traitement si la maladie est présente et aussi de la volonté du patient à accepter les risques et les désagréments d'un traitement qui pourrait ne pas être nécessaire. La DCA est une méthode qui permet d'évaluer la valeur des informations fournies par un test de diagnostic, par exemple en considérant une fourchette de probabilités d'être malade. Un concept important est que la DCA identifie un intervalle de probabilités dont la limite basse peut fournir un niveau de certitude diagnostique au-dessus duquel le patient choisira d'accepter la décision thérapeutique. L'étendue de cet intervalle de probabilités dépend bien évidement de la situation clinique. En effet, si la décision thérapeutique est assimilée à l'administration d'un traitement qui a une efficacité élevée avec un coût minime et des effets secondaires maitrisés, le seuil de la borne minimale pour la probabilité d'être malade sera bas. Traiter est une bonne option même les faux positifs. A l'inverse, si le traitement possède une efficacité faible ou est associé à une morbidité importante le seuil de la borne minimale pour la probabilité d'être malade sera élevé. Traiter peut s'avérer problématique, il faut donc limiter les faux positifs et ne traiter que les sujets à fort risque de présenter la maladie.

Pour tenir compte de ces considérations, le bénéfice net évalué dans la DCA est déterminé en calculant la différence entre le bénéfice attendu et le préjudice prévu associé à chaque stratégie de dépistage et de décision thérapeutique envisagée. Le bénéfice attendu est représenté par le nombre de patients qui ont la maladie dans le cas d'un test diagnostic et qui reçoivent un traitement (vrais positifs) en utilisant la stratégie proposée. Le préjudice attendu est représenté par le nombre de patients sans maladie qui seraient traités par erreur (faux positifs) multiplié par un facteur de pondération basé sur le seuil de la probabilité d'être malade, c'est-à-dire que le test diagnostique soit positif. Le facteur de pondération dépend du seuil de

probabilité considéré pour prendre la décision thérapeutique (traiter si la décision est de traiter). Plus précisément, le taux de faux positifs est multiplié par le rapport de la probabilité de seuil, divisé par 1 - la probabilité de seuil. Autrement dit, plus le seuil de probabilité est important, plus le poids de la présence de faux positifs est important dans le préjudice attendu. Intuitivement c'est assez naturel : si un clinicien veut traiter son patient qu'à partir d'un seuil de probabilité de 80% d'être malade, c'est que le traitement n'est pas anodin et qu'il n'a pas envie d'exposer à tort une personne non-malade. Dans ce contexte, on comprend bien que l'idée est de limiter les faux positifs. Les faux positifs pèsent alors plus lourd dans le bénéfice de la décision que les vrais positifs.
Le résultat de la DCA est exprimé graphiquement sous la forme d'une courbe avec un score de bénéfice net sur l'axe des ordonnées et des seuils de probabilité de présenter la maladie ou l'évènement sur l'axe des abscisses. Une courbe est alors établie pour chaque approche qui pourrait être considérée pour établir le diagnostic de maladie ou la présence de l'évènement. Pour un seuil de probabilité fixé, la courbe avec le score de bénéfice net le plus élevé est donc la meilleure stratégie thérapeutique. En guise de limite, dans le contexte de maladies qui ne sont pas bien étudiées, il peut y avoir une insuffisance des connaissances pour déterminer la gamme pertinente de seuil de probabilités à considérer.

Dans le cas d'un modèle pronostique appliqué à des données de survie dont le but est de prédire un évènement dans le temps, il est tout à fait envisageable d'appliquer le même concept. L'idée consiste en l'évaluation des prédictions du modèle en classant les patients comme positifs ou négatifs selon un seuil de probabilité de faire l'évènement et d'attribuer un coût aux faux positifs, au regard des vrais positifs en fonction de ce seuil. La petite astuce réside dans le fait qu'il faut se positionner à un temps clef et utiliser les prédictions du modèle de survie à ce temps. Les prédictions apportées par le modèle au temps retenu pourront donc être mises en perspective avec les évènements observés jusqu'à ce temps. C'est exactement la même approche qu'avec les méthodes de reclassification ou dans l'évaluation de la bonne

calibration. Il faut se placer à un temps d'intérêt et contraster ce que le modèle prédit à ce temps avec l'information d'avoir présenté ou pas l'évènement à ce temps ou avant.

Une illustration concrète peut être donnée par l'analyse d'un modèle multivarié produit dans une des études présentées dans cette thèse. L'objet de l'étude était de déterminer un modèle pronostique de la survie globale sous seconde ligne de chimiothérapie pour des patients atteints d'un cancer pancréatique avancé.

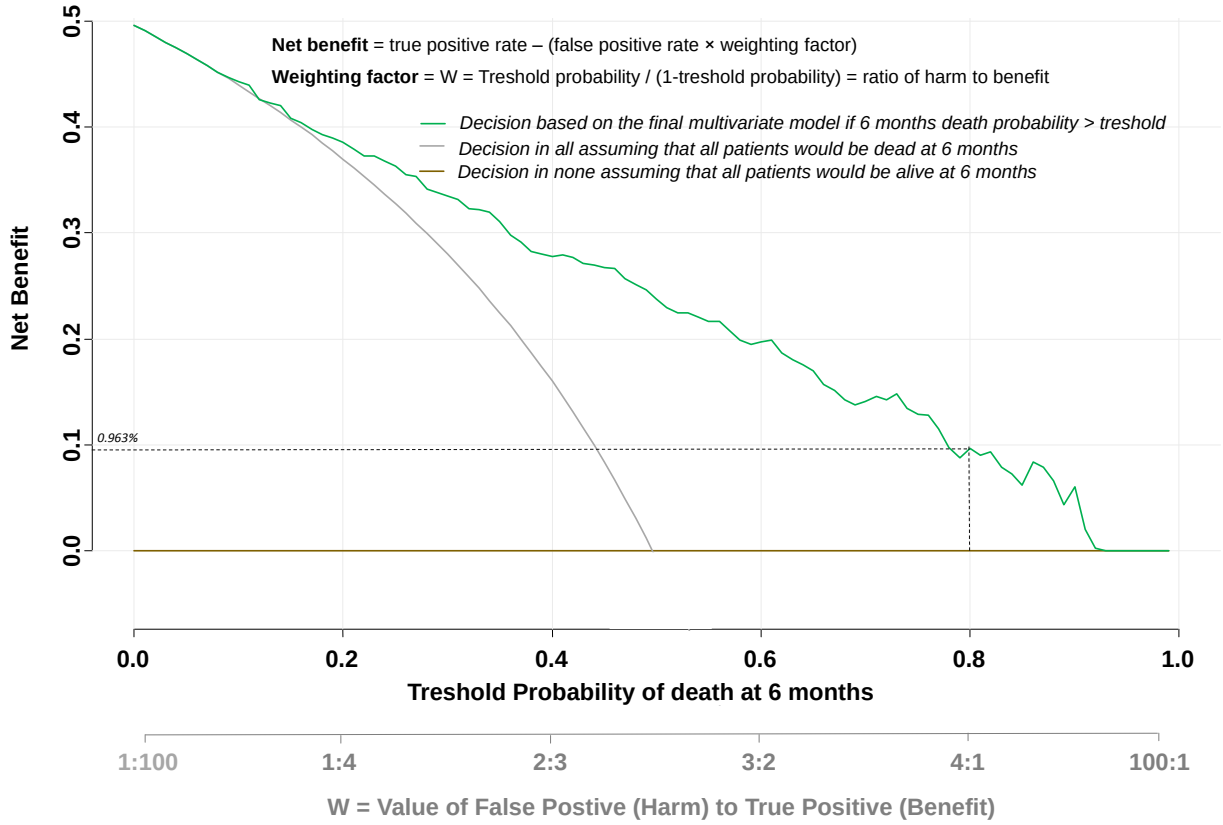
Voici ci-dessous le modèle multivarié de Cox obtenu :

	Nombre de patients	Nombre de décès	Rapport de risque	IC à 95% du rapport de risque	p	Validation interne IC à 95% du rapport de risque par bootstrap	β	Score maximal dans le nomogramme
PS (OMS)								
0	47	40	1					
1	117	114	1.614	1.067 - 2.374		1.00897 - 2.39352	0.47879	25
≥ 2	84	81	3.055	1.935 - 4.822	**<.0001**	1.74622 - 4.72267	1.11671	58
Durée de L1 (transformation log), mois	248	235	0.478	0.325 - 0.702	**.0002**	0.33244 - 0.81427	-0.73868	100
Age, années	248	235	1.019	1.003 - 1.036	**.0168**	1.00237 - 1.03901	0.01921	55
Statut tabagique								
Non-fumeur ou ancien fumeur	190	178	1					
Fumeur actif	58	57	1.491	1.086 - 2.046	**.0134**	1.01718 - 2.13902	0.39946	21
Métastases hépatiques								
Non	99	93	1					
Oui	149	142	2.149	1.620 - 2.850	**<.0001**	1.48348 - 2.93914	0.76502	40
Douleur								
Non	118	108	1					
Oui	130	127	1.411	1.059 - 1.880	**.0186**	0.98883 - 1.85176	0.34442	18
Ictère								
Non	210	198	1					
Oui	38	37	1.958	1.327 - 2.890	**.0007**	1.18897 - 2.98240	0.67192	35
Ascite								
Non	200	188	1					
Oui	48	47	1.792	1.264 - 2.540	**.0011**	1.20746 - 2.76271	0.58335	30
Protocole en L2								
5-FU ± Oxaliplatine et/ou Irinotecan	195	184	1					
À base de Gemcitabine	53	51	1.765	1.268 - 2.457	**.0008**	1.18119 - 2.49164	0.56827	30

Dans cette situation clinique, compte tenu d'un pronostic de survie parfois très sombre on peut se poser la question de l'utilité d'administrer une deuxième ligne de chimiothérapie à un patient qui a une très forte probabilité de décéder à 6 mois. La décision thérapeutique ici sera donc l'absence

d’administration d’une seconde ligne de chimiothérapie. Les seuils de probabilité quant à eux correspondent à la probabilité de faire l’évènement à 6 mois.

Il est alors possible de représenter la courbe issue de la DCA comme suit :



Cette courbe de DCA identifie un intervalle de seuil de probabilité de décès à 6 mois dans lequel les décisions entraîneront plus d'avantages pour les vrais positifs et la décision impliquant les faux positifs sera réduite.
C’est en effet ce qu’on attend du modèle. Etant donné que l’on souhaite éventuellement envisager le non-traitement des patients de très mauvais pronostic, cette décision implique implicitement que nous souhaitons minimiser les faux positifs, c’est-à-dire les patients considérés de mauvais pronostic à tort. Il s’agit des patients encore en vie à 6 mois qui devraient être traités, puisqu’à priori le traitement leur apporte un bénéfice de survie.

Sur la courbe, le bénéfice net pour les décisions basées sur les prédictions issues de notre modélisation multivariée finale est plus intéressant pour le seuil des probabilités supérieur à 20%, comparé aux stratégies qui reposent sur la considération de l'ensemble de la population au même niveau (ne pas traiter personne ou ne pas traiter tout le monde).

Dans l'ensemble, la courbe de décision montre que le bénéfice net de baser sa décision sur notre modèle multivarié final est d'intérêt, en particulier pour les seuils supérieurs à 0,5, ce qui peut être défini comme la zone d'intérêt dans le cancer pancréatique avancé.

III.4.2 Outils pronostiques exploitant l'information issue d'un modèle de Cox

Une fois que le modèle de Cox multivarié, intégrant l'ensemble des facteurs pronostiques d'intérêt pour prédire la survenue de l'évènement considéré, a été étudié en termes de performance (discrimination et calibration), validé en interne et dans l'idéal validé dans une cohorte de validation externe, et aussi que son utilité ou efficacité clinique a été évaluée, il est intéressant de proposer des outils pronostiques pouvant être exploités en pratique clinique. L'utilisation de l'information contenue dans la modélisation de Cox offre la possibilité de proposer aux patients, aux cliniciens et aux chercheurs, une information pertinente et objective, surtout personnalisée en termes de prédiction de l'évènement étudié et aussi prédictive de la réponse à un traitement, si des propriétés prédictives sont démontrées et conférées au modèle.

III.4.2.1 Approche individuelle : le nomogramme

Classiquement afin d'exploiter la valeur relative de chaque facteurs inclus dans le modèle de Cox dans la prédiction de l'évènement, une pondération

basée sur les estimations des paramètres bêtas estimés est réalisée. Une proposition est de diviser chacun de ces coefficients par le plus petit d'entre eux et ainsi obtenir un score dans lequel chaque facteur contribue à sa juste valeur. Cette approche est intéressante mais elle a le défaut de ne pas être visuelle et la tendance actuelle (à juste titre) d'utiliser les variables quantitatives sans les catégoriser rend l'interprétation de leur contribution relative moins intuitive.

Dans différents contextes de cancer, les nomogrammes ont été de plus en plus utilisés pour la prédiction de l'occurrence d'un évènement comme une alternative au système de classification TNM traditionnel (Balachandran *et al*, 2015).

Les nomogrammes sont une représentation graphique d'une formule mathématique complexe (Grimes, 2008). Ces représentations graphiques des modèles tels que celui de Cox ont pour objet de fournir pour un nouveau sujet une probabilité de survenue de l'événement sur la base de la modélisation proposée par le modèle. Un nomogramme intègre donc les variables inclues dans le modèle de Cox pour prédire la survenue de l'évènement et génère donc une probabilité de faire l'événement pour un individu à un ou des temps d'intérêt préalablement choisis.

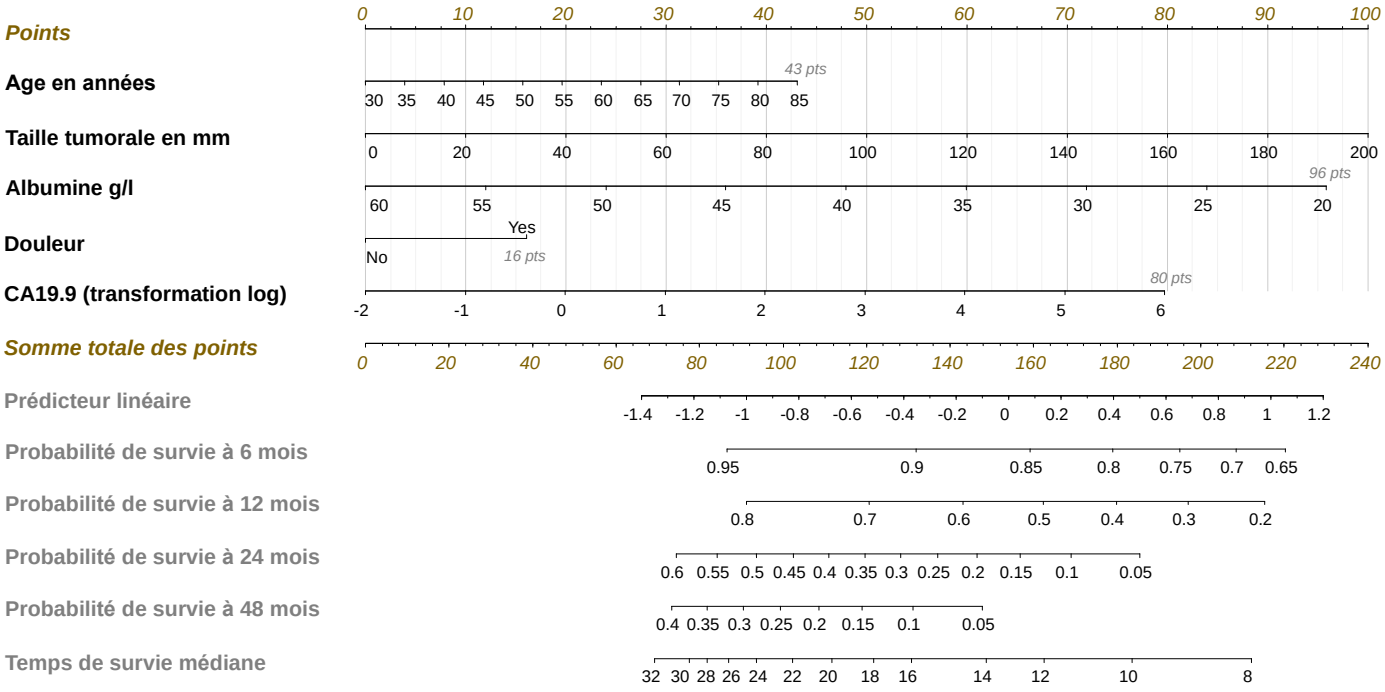
Pour illustration, dans une des études présentées dans cette thèse nous avons modélisé, avec un modèle de Cox multivarié, la survie globale au diagnostic de patients atteints d'un cancer pancréatique localement avancé.

Dans le tableau ci dessous voici le modèle final obtenu pour l'analyse de Cox multivarié avec la SG (*N* = 358)

	Nombre de patients	Nombre de décès	HR	95% IC	*P*	Validation interne BCA HR 95%	Score maximal dans le nomogramme
Age au diagnostic, années	358	307	1.013	1.000 to 1.025	**.0478**	1.000 to 1.025	43
Douleur							
Non	155	130	1	-	**-**		
Oui	203	177	1.292	1.023 to 1.632	**.0317**	1.046 to 1.657	16
Albumine, g/L	358	307	0.963	0.942 to 0.984	**.0006**	0.942 to 0.988	96
Taille tumorale, mm	358	307	1.008	1.001 to 1.015	**.0214**	1.000 to 1.016	100
CA 19-9, UI (valeur logarithmique)	358	307	1.172	1.047 to 1.311	**.0056**	1.039 to 1.306	80

Abréviations: IC: Intervalle de confiance, HR: Hazard Ratio, BCA: accelerated bootstrap confidence interval

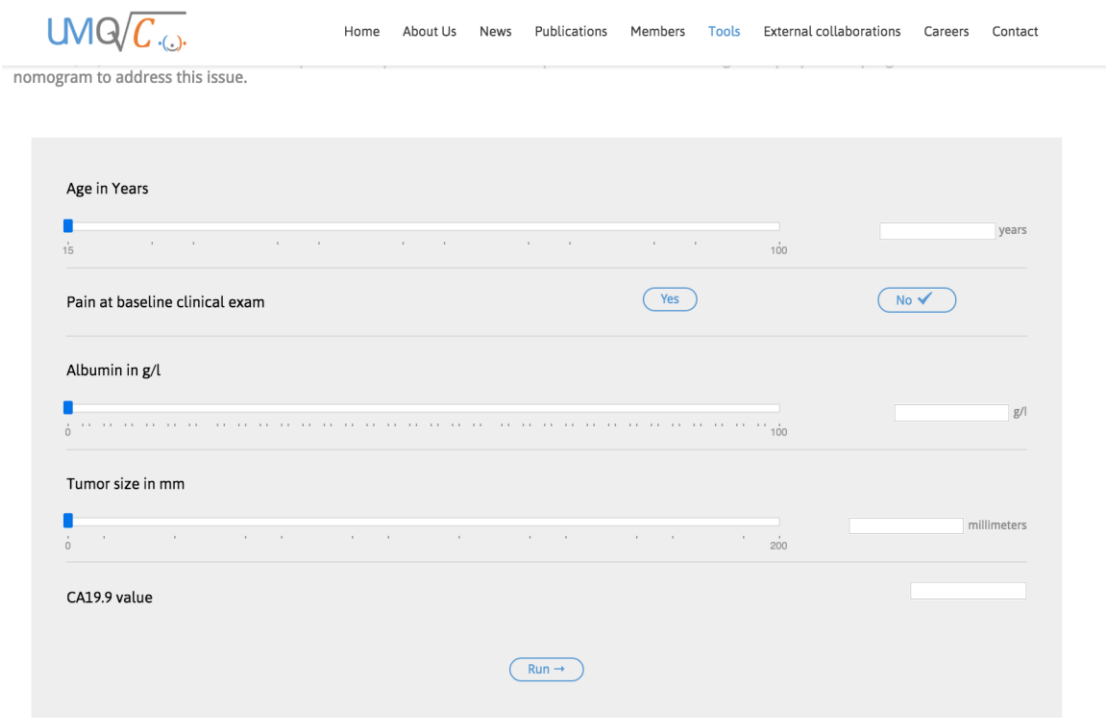
En pratique, le nomogramme pour prédire la survie globale individuelle au diagnostic des patients atteints d'un cancer pancréatique localement avancé correspondant à ce modèle se présente sous la forme suivante :



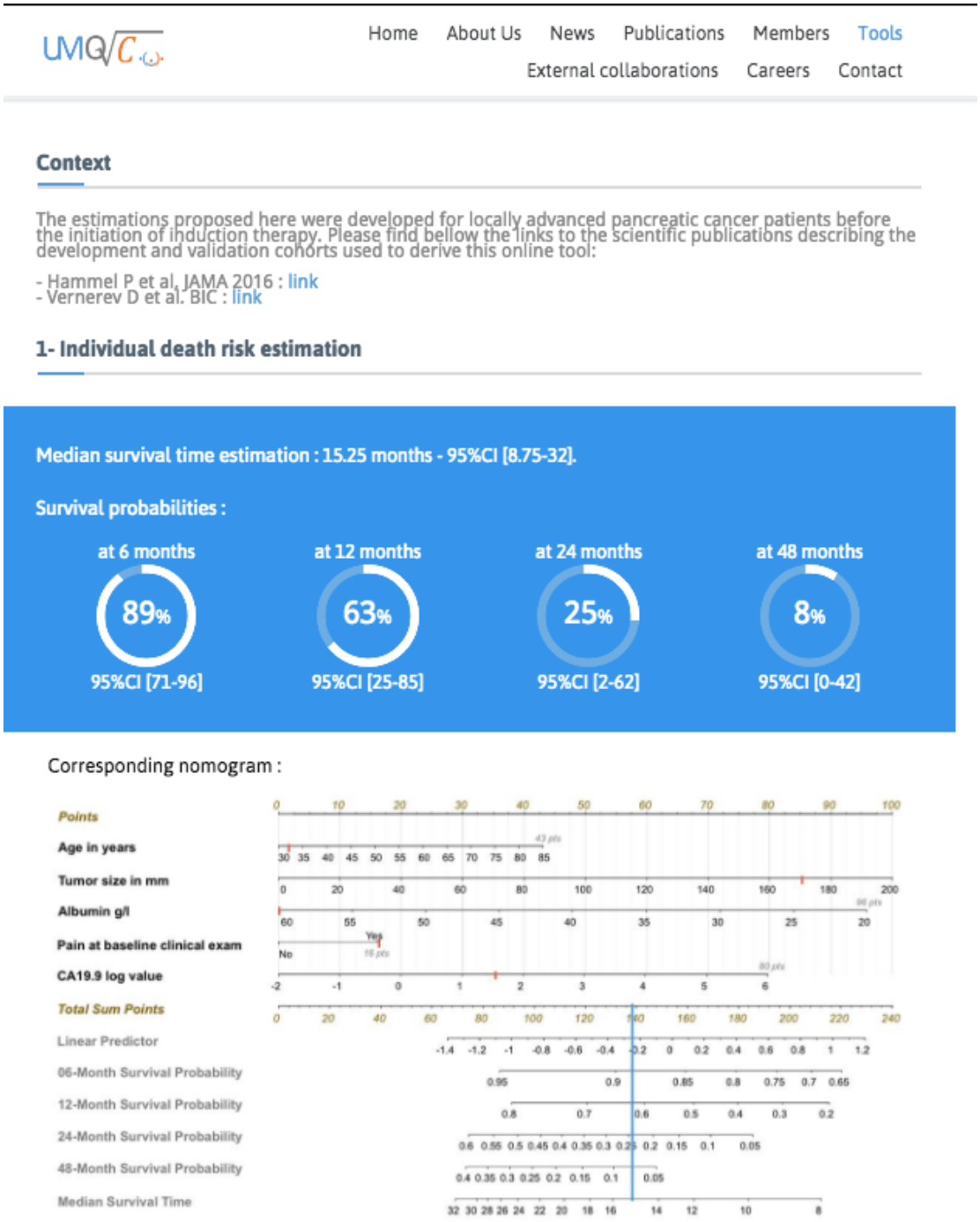
Des points sont donc assignés pour chaque facteur de risque en traçant une ligne verticale vers la ligne correspondant aux points. Le total des points pour les 5 facteurs est ensuite tracé sur la ligne correpondant à la somme totale des points. Une ligne verticale vers le bas est ensuite tracée pour déduire les prédictions correspondantes aux temps d'intérêt ainsi que la médiane de survie du patient.

Il y a deux manières d'utiliser un nomogramme. La première consiste en l'utilisation en version papier de la représentation, mais cela peut s'avérer long au regard du temps clinique si le nombre de variables est important, et surtout cette approche peut véhiculer des sources d'erreur ou encore d'imprécisions.

Dans un contexte contemporain où l'outil informatique est prédominant et avec un accès illimité à internet via les smartphones, il est inévitable de fournir conjointement à la version papier une interface web où les calculs sont automatisés. Le clinicien ou le chercheur n'a alors plus qu'à saisir les informations relatives à chaque facteur et automatiquement les prédictions individuelles de survie sont générées. Il paraît aussi opportun, lorsque l'outil est mis en application dans la pratique clinique, de pouvoir fournir un export des résultats afin d'intégrer ces informations dans le dossier médical.
Voici par exemple l'interface que j'ai pu proposer pour le nomogramme présenté ci-dessus :



http://www.umqvc.org/en/tool/prolap.html
http://www.umqvc.org/en/tool/proscap.html



Cette automatisation requiert bien évidement l'implémentation des calculs en ligne. Ces derniers sont explicités théoriquement dans les parties qui suivent.

III.4.2.1.1 Estimation de la probabilité de survie sans évènement à un temps donné

L'estimation de la survie pour un patient j au temps t, basée sur la modélisation issue d'un modèle de Cox, s'écrit de la manière suivante :

Formule générique:

$$S(t, X_j) = S_0(t)^{exp(\sum_{i=1}^{p} \beta_i X_{ij} - \sum_{i=1}^{p} \beta_i \overline{X_i})} S(t, X_j)$$

$$= S_0(t)^{exp\left(\sum_{i=1}^{p} \beta_i (X_{ij} - \overline{X_i})\right)} = S_0(t)^{exp(LP_j)}$$

où:

- p est le nombre de facteurs de risque impliqués dans le modèle de Cox,
- $ß_i$ est le paramètre de régression estimé pour le paramètre i avec i=1, 2, …, p,
- $S_0(t)$ est la survie de base au temps t,
- X_{ij} correspond à la valeur du paramètre i pour le patient j,
- X_j le vecteur des valeurs pour les facteurs de risque pour le patient j $X_j = (X_{1j}, X_{2j}, …, X_{pj})$,
- $\overline{X_i}$ représente la moyenne de la distribution du facteur i dans la population de patients impliqués dans le modèle multivarié,
- LP_j est appelé le prédicteur linéaire pour le patient j.

Les estimations issues d'un modèle de Cox multivarié peuvent être utilisées pour générer un nomogramme permettant des estimations de survie individuelles à des temps d'intérêt (t_1, t_2, …) et aussi l'estimation d'une médiane de survie.

Le prédicteur linéaire du patient j est obtenu en sommant les produits des estimations $ß_i$ issus du modèle de Cox avec les différences $X_{ij} - \overline{X_i}$ de la valeur du paramètre i pour le patient j (X_{ij}) à laquelle est retranchée la valeur moyenne de ce même paramètre i dans la population d'étude ($\overline{X_i}$) et ceci pour chaque facteur de risque i avec i variant de 1 à p. Autrement dit, le prédicteur

linéaire du patient j LPj peut être interprété comme le vecteur de risque du patient j comparé à celui d'un patient dit de référence ou de risque moyen.

$$\mathrm{LP_j} = \sum_{i=1}^{p} \beta_i \times \left(X_{ij} - \overline{X_i}\right)$$

Si le risque de base au temps t d'intérêt $S_0(t)$ est connu alors, la probabilité de survie d'un patient j au temps t peut s'écrire comme ci-dessous :

$$\mathrm{S(t, X_j)} = \mathrm{S_0(t)}^{\exp\left(\sum_{i=1}^{p} \beta_i (X_{ij} - \overline{X_i})\right)} = \mathrm{S_0(t)}^{\exp(LP_j)}$$

III.4.2.1.2Estimation de la médiane de survie pour le patient j

La survie médiane pour un patient j ($Survie\ mediane_j$) est par définition égale au temps t pour lequel la survie du patient j est estimée à 0.5 ou 50% ($S(t,X_j) = 0.5$). En pratique, ce temps médian est défini comme étant le temps le plus précoce pour lequel $S(t,X_j) \leq 0.5$. L'estimation de la médiane est plus appropriée que celle de la moyenne dans un contexte de données censurées à droite, comme c'est le cas des données de survie.

Avec une formulation mathématique, nous pouvons donc écrire :

$$Survie\ mediane_j = min\left(t \,/\, S(t, X_j) \leq \frac{1}{2}\right)$$

$$S(t, X_j) = S_0(t)^{exp(\mathrm{LP_j})} = \frac{1}{2}$$

$$\Leftrightarrow ln\left(S_0(t)^{exp(\mathrm{LP_j})}\right) = ln\left(\frac{1}{2}\right)$$

$$\Leftrightarrow exp(LP_j) ln(S_0(t)) = ln\left(\frac{1}{2}\right)$$

$$\Leftrightarrow ln(S_0(t)) = ln\left(\frac{1}{2}\right) exp(-LP_j)$$

$$\Leftrightarrow t \,/\, S_0(t) = \exp\left(ln\left(\frac{1}{2}\right) exp(-LP_j)\right)$$

$$Survie\ mediane\ S_j = \min(t) \,/\, S_0(t) = \exp\left(ln\left(\frac{1}{2}\right) exp(-LP_j)\right)$$

III.4.2.1.3 Intervalle de confiance pour les estimations de survie

Une estimation statistique n'est jamais réellement informative sans être accompagnée par un intervalle de confiance dans lequel l'estimation peut fluctuer. Usuellement, une probabilité α (risque de première espèce) fixée à 5% représente le risque de se tromper lorsque nous concluons que l'estimation se trouve dans l'intervalle de confiance défini. L'intervalle de confiance ($IC_{1-\alpha}$) donne donc une fourchette dans laquelle nous pouvons conclure que l'estimation à 1 - α % (usuellement 95%) de chance de se trouver.
Dans ce contexte, il est impératif de produire un intervalle de confiance autour de chaque estimation de survie produite ($S(t,X_j)$ et Survie $médiane_j$).

La survie d'un patient à un temps t ($S(t,X_j)$) et sa survie médiane (Survie $médiane_j$) étant deux entités mathématiques dans lesquelles le seul paramètre qui varie est le prédicteur linéaire du patient (LP_j), la détermination de leur intervalle de confiance repose sur le calcul de celui du prédicteur linéaire LP_j ($(1-\alpha)\%\text{IC } LP_j$) :

Soit a le $\left(100-\frac{\alpha}{2}\right)^{\text{ème}}$ quantile de la distribution normale standard.
Alors nous pouvons écrire :

$$(1-\alpha)\%\text{IC } LP_j$$
$$= \left[LP_j \text{ borne min};\ LP_j \text{ borne max}\right]$$
$$= \left[LP_j - a \times \text{erreur standard}(LP_j);\ LP_j + a \times \text{erreur standard}(LP_j)\right]$$

Le calcul de l'intervalle de confiance du prédicteur linéaire pour le patient j $(1-\alpha)\%\text{IC } LP_j$ implique donc par conséquent le calcul de son erreur standard qui est la racine carrée de la variance de LP_j.

Si nous considérons :

$X_j = \begin{pmatrix} x_{1,j} \\ x_{2,j} \\ \vdots \\ x_{p-1,j} \\ x_{p,j} \end{pmatrix}$ le vecteur des valeurs pour les facteurs de risque du patient j

et

$\overline{X} = \begin{pmatrix} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \\ \overline{X}_{p-1} \\ \overline{X}_p \end{pmatrix}$ le vecteur des moyennes pour les distributions des facteurs de risque dans la population de patients impliqués dans le modèle de Cox multivarié

et

$VCov(X) = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_p) \\ Cov(X_2, X_1) & \ddots & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_p, X_1) & \cdots & \dots & Var(X_p) \end{pmatrix}$ la matrice de variance covariance issue du modèle de Cox

Alors il est possible d'écrire $\text{Variance}\left(LP_j\right) = \left(X_j - \overline{X}\right)' \times VCov(X) \times \left(X_j - \overline{X}\right)$

Et donc :

$$\text{Standard Error}\left(LP_j\right) = \sqrt{\text{Variance}\left(LP_j\right)} = \sqrt{\left(X_j - \overline{X}\right)' \times VCov(X) \times \left(X_j - \overline{X}\right)}$$

A ce stade nous pouvons donc déterminer aisément :

$$(1 - \alpha)\%\text{IC } LP_j$$
$$= \left[LP_j - a \times \text{erreur standard}\left(LP_j\right);\ LP_j + a \times \text{erreur standard}(LP_j)\right]$$
$$= \left[LP_j \text{ borne min};\ LP_j \text{ borne max}\right]$$

Il devient donc très simple de proposer un intervalle de confiance à 1 - α % pour l'estimation de survie du patient j au temps t ($S(t,X_j)$) :

$$(1-\alpha)\%\text{IC}\ S(t,X_j) = (1-\alpha)\%\text{IC}\ S_0(t)^{\exp(LP_j)}$$
$$= \left[S_0(t)^{\exp(LP_j borne\ max)}\ ;\ S_0(t)^{\exp(LP_j\ borne\ min)}\right]$$

III.4.2.1.4 Points du Nomogramme

Pour chaque facteur de risque, le nombre de points maximal attribuable peut être obtenu comme ci-après.
Pour le facteur de risque i:

$$\text{Points_max}_i = 100 \times \frac{\beta_i|\max_i - \min_i|}{\beta_j|\max_j - \min_j|} = 100 \times \frac{\text{Absolute range } \beta_i \text{ value}}{\text{Absolute range } \beta_j \text{ value}}$$

où le facteur j est le facteur qui possède une valeur de « absolute range β » maximal parmi l'ensemble des facteurs
et max et min représentent les valeurs extrêmes possibles pour les paramètres avec un ordonnancement des valeurs de la plus protectrice à la plus à risque pour l'évènement étudié.

La détermination du nombre de points maximal du nomogramme pour les facteurs de risque impliqués dans le modèle multivarié permet, dans un premier temps, d'expliciter l'importance de la contribution des variables dans le modèle puis, dans un second temps, de calculer pour un patient j avec un facteur i égal à b le nombre de points qui lui est attribué pour ce paramètre.

$$Points_{ij} = |b - min_i| \times \frac{Score_max_i}{|max_j - min_j|}$$

Le nombre total de points pour un patient j est tout simplement la somme de ces points pour chacun des facteurs de risque et peut s'écrire comme suit :

$$\text{Total Points}_j = \sum_i \text{Points}_{ij}$$

L'approche individuelle permet donc via la construction d'un nomogramme de fournir des estimations individuelles de probabilité de ne pas avoir expérimenté l'évènement d'intérêt à des temps d'intérêt. Ces informations sont intéressantes mais en pratique clinique, les cliniciens fondent souvent leur raisonnement sur des populations de patients dans lesquelles le risque de faire l'évènement est distinct.

III.4.2.2 Approche populationnelle : Construction d'un score pronostique

Chaque patient j impliqué dans la construction du modèle de Cox multivarié peut via l'utilisation du nomogramme être caractérisé par un nombre total de point : Total Points_j.
Ce nombre total de points permet d'obtenir une estimation individuelle de la survie du patient à un temps t ainsi que sa médiane de survie.
Cependant, il est aussi intéressant de raisonner à l'échelle populationnelle et donc d'identifier des populations de pronostics clairement différents pour *in fine* positionner le patient dans une population ayant un pronostic bien défini.
Ainsi, si le modèle de Cox multivarié est établi sur la base d'un panel de n patients alors nous possédons une distribution de n estimations ordonnant les patients du moins au plus à risque de faire l'événement. Ce vecteur de n observations constitue le score pronostique :

$$\text{Total Points}_j = \begin{pmatrix} \text{Total Point}_1 \\ \text{Total Point}_2 \\ \vdots \\ \text{Total Point}_{n-1} \\ \text{Total Point}_n \end{pmatrix}$$

Pour un patient j, le nombre total de point issu du nomogramme, Total Points_j, possède une correspondance directe avec le prédicteur linéaire LP_j (cf illustration). Il est alors tout à fait possible et similaire d'utiliser les n prédicteurs linéaires comme score pronostique.

III.4.2.2.1 Identification de populations de pronostics différents

Dans le contexte d'une approche populationnelle et sur la base d'un score pronostique qui peut être défini par les points totaux issus du nomogramme, il est essentiel de déterminer dans un premier temps le nombre de populations à mettre en évidence et dans un second temps les bornes du score permettant de définir précisément l'appartenance d'un patient à l'un ou l'autre des groupes à risque. Ces choix doivent être autant que possible non arbitraires.

III.4.2.2.2 Choix du nombre de groupes à risque à identifier

Le choix du nombre de populations peut être déterminé sur la base de plusieurs éléments. Tout d'abord, l'importance de l'hétérogénéité de pronostic pour l'évènement étudié dans la pathologie d'intérêt est évidement un des points clefs. Plus l'hétérogénéité entre les patients est importante plus a priori les groupes de pronostics différents seront nombreux. Inversement, si le pronostic des patients n'est pas très hétérogène, l'identification de 2 voire 3 groupes est déjà d'intérêt. Le second critère de choix repose sur la question clinique posée. Si l'idée est d'identifier des populations extrêmes, c'est-à-dire de très mauvais ou de très bon pronostic, alors fixer un nombre important de groupes est préconisé afin d'identifier et donc d'isoler des populations de risque intermédiaire, permettant de rendre plus homogènes les groupes extrêmes. Cela dit il est bien connu que les groupes dits de pronostics intermédiaires peuvent poser des questionnements en termes de management clinique. En revanche si la question clinique implique plutôt une catégorisation simple binaire des patients, le nombre de groupes doit être égal à 2. Un troisième argument pour faire le choix peut être orienté par la volonté dans le futur d'utiliser le score pronostique dans la construction des essais cliniques. D'un point de vue méthodologique, le design d'un essai clinique peut être optimisé en stratifiant la randomisation sur les classes de pronostics identifiées ce qui permettrait de s'assurer de la maitrise de l'hétérogénéité de pronostic connu au moment de la construction de l'essai et ainsi répartir équitablement les différents patients entre les bras, ce qui évidement permet d'optimiser l'estimation de l'effet traitement ou intervention

expérimentale. Dans ce dernier cas l'identification du nombre de groupes maximum est préconisée.
La considération de ces différents éléments pouvant orienter le choix permet de le rendre le moins arbitraire possible. Evidemment, si nous nous apercevons que le nombre de populations recherchées est surestimé, il convient naturellement de le diminuer.

III.4.2.2.3Choix des seuils du score permettant de définir les populations

Le choix des seuils ou bornes permettant de définir les populations de pronostics bien différenciés est un sujet qui ouvre automatiquement des discussions. L'idée sous-jacente est d'être le moins arbitraire possible afin de pouvoir garantir une généralisation du résultat et donc d'avoir une définition des groupes à risque à des conditions expérimentales plus vastes (autres centres, cliniciens, patients, pays, période etc.).
Deux contextes de recherche peuvent se distinguer et ainsi engager des approches différentes.
Le premier s'inscrit dans le contexte de maladies comme les pathologies cardiovasculaires dans lesquelles une recherche pronostique reconnue, documentée et ayant déjà fournie des outils cliniquement bien implémentés et donc des définitions consensuelles pour définir différentes populations à risque pour lesquelles le management est distinct. Par exemple, il est possible d'estimer le risque absolu vasculaire au niveau coronarien et cérébral chez chaque patient (ce qui correspond pour un patient à la probabilité de présenter un événement cardio-vasculaire sur une période de temps), grâce à des équations établies sur des populations. La plus utilisée est celle de Framingham. Les variables incluses dans l'équation de prédiction du risque d'accident coronarien sont le sexe, l'âge, la valeur de la pression artérielle systolique, la valeur du cholestérol total et du HDL cholestérol, la consommation de tabac (oui/non), un diabète (oui/non), une hypertrophie ventriculaire gauche définie sur un électrocardiogramme (oui/non). Ces différents paramètres sont pondérés par des coefficients traduisant leur impact sur le risque. L'équation de Framingham permet aussi l'évaluation du

risque vasculaire cérébral. Quatre catégories de risque sont définies par les recommandations Nord-Américaines :

- le très haut risque cardio-vasculaire, qui comprend les patients candidats à la prévention secondaire (après infarctus du myocarde ou AVC) et aussi ceux candidats à la prévention primo-secondaire ayant une localisation extra-coronarienne de la maladie athéroscléreuse (atteinte carotidienne symptomatique, artériopathie, anévrisme de l'aorte abdominale) ;
- le haut risque cardio-vasculaire, qui comprend les patients à haut risque cardio-vasculaire absolu (>20% à 10 ans) ou les diabétiques candidats à la prévention primaire ;
- le risque cardio-vasculaire modéré, qui comprend les patients candidats à la prévention primaire ayant un niveau de risque moyen à 10 ans (10-20%) ou ayant 2 facteurs de risque cardio-vasculaire ;
- le risque cardio-vasculaire faible, qui comprend les patients candidats à la prévention primaire n'ayant qu'un seul facteur de risque cardio-vasculaire ou n'ayant pas de facteur de risque.

Dans ce contexte précis, la recherche de populations de risques différents ne peut se faire sans prise en compte de l'outil dit équation de Framingham présenté ci-dessus et les seuils qui en découlent pour l'action clinique.

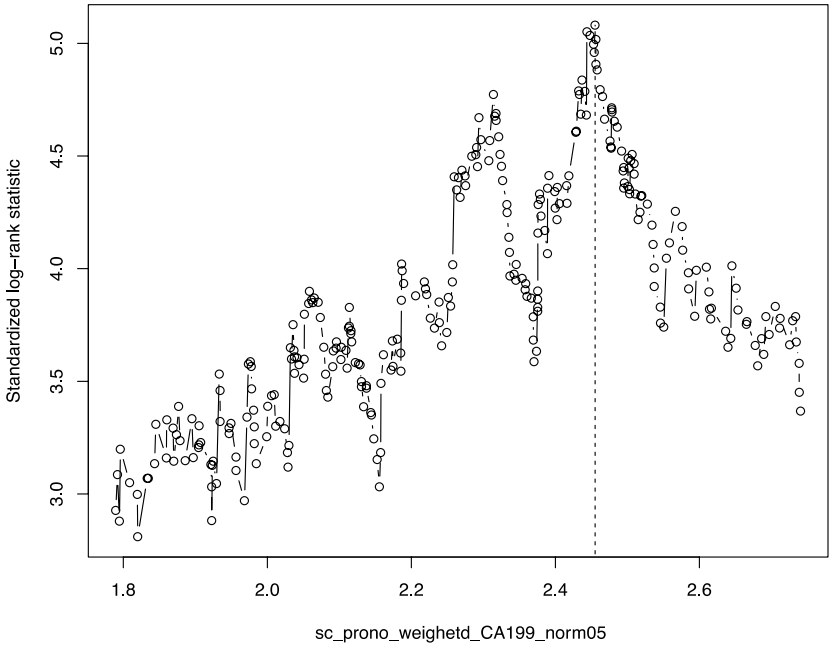
Le second contexte de travail est quant à lui caractérisé par l'absence de recherche pronostique fournissant un outil fonctionnel et bien intégré dans la pratique clinique courante. Dans ce contexte, la recherche de bornes pour définir les populations de risque bien différenciées est relativement libre mais doit être le moins arbitraire possible.

Une approche fréquemment utilisée dans la littérature repose sur le calcul des quantiles qui sont les valeurs qui divisent un jeu de données en intervalles contenant le même nombre de données. Ainsi, par exemple pour la construction de 2, 3 ou 4 groupes les bornes les définissants sont respectivement la médiane, les tertiles ou les quartiles de la distribution du score. Cette approche a le mérite d'être simple et d'identifier des groupes de taille suffisamment importante et pertinente mais elle repose sur la distribution du score observé dans l'échantillon de développement et ainsi n'est pas toujours reproductible dans d'autres échantillons de données. De plus

l'hypothèse que les groupes de pronostics différents soient constitués d'un nombre similaire de patients est questionnable. En effet, intuitivement il n'y a pas de rationnel clair pour que par exemple les patients de mauvais pronostics soient aussi nombreux que ceux de bons pronostics.

Sur la base de ce dernier constat, il paraît opportun de recommander des méthodes plus sophistiquées pour déterminer les bornes délimitant les groupes recherchés. Lorsque le nombre de groupes est fixé à 2, la méthode dite de la borne optimale proposée par Hothorn et Lausen (Hothorn & Lausen, 2003) constitue une solution efficace et flexible. Dans un contexte d'analyse de données de survie, le principe de cette méthode consiste en la recherche de la valeur qui permet de maximiser la statistique du log-rank. Cette recherche de valeur optimale s'accompagne généralement de contraintes concernant la taille minimale et maximale pour l'un ou l'autre des groupes recherchés ce qui garantit l'identification de groupes de patients suffisamment représentatifs.

Ci-dessous une illustration de la méthode de Hothorn et Lausen. La figure représente la valeur de la statistique du log-rank pour chaque valeur du score qui permet de respecter les contraintes fixées.



Comme observé sur la courbe ci-dessus le meilleur seuil se situe vers 2.5 pour maximiser la différence entre les deux groupes en garantissant au moins

10% mais pas plus de 90% des observations dans le groupe le moins à risque.

Une autre méthode qui véhicule l'idée que les patients de pronostics extrêmes (bon ou mauvais) sont généralement moins nombreux que ceux de pronostic intermédiaire et donc que les groupes ne sont pas à priori de taille similaire est la méthode dite de Cox (Cox, 1957). En effet, il est plus efficace d'avoir des groupes plus importants dans le centre de la distribution du score pour rendre les groupes extrêmes encore plus extrême. Un point fort non négligeable de cette méthode est sa capacité à identifier un nombre de groupes supérieur ou égal à 2.

En pratique, les seuils identifiés par la méthode de Cox dépendent de la moyenne et de l'erreur type de la distribution du score, dont on suppose qu'elle est assimilable à une distribution normale. Cox a produit sur la base de simulations une table de valeurs permettant de déterminer, à l'aide de la moyenne (m) et de l'erreur type (err) de la distribution du score, les bornes correspondantes pour un nombre de groupes choisis.

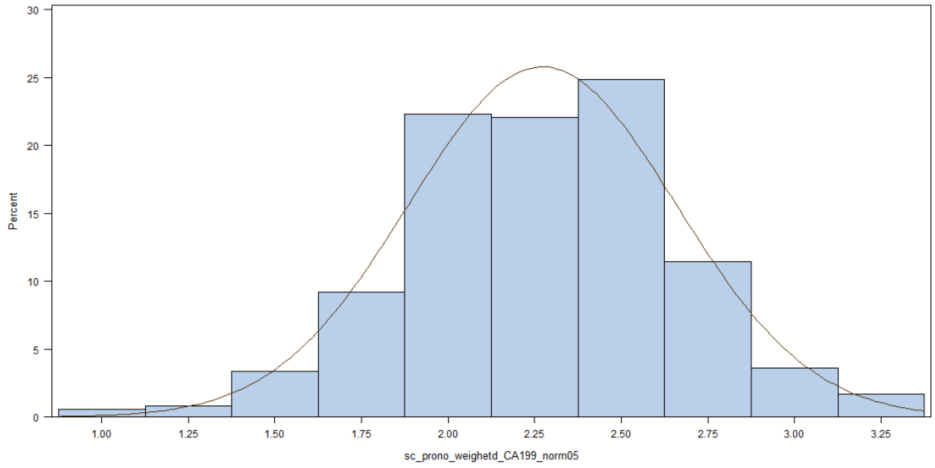
Ci-dessous voici la table de valeurs (a) correspondantes :

OPTIMAL m SIZES AND PROBABILITY OF PURCHASE, STANDARD NORMAL DISTRIBUTION

i	m									
	1	2	3	4	5	6	7	8	9	10
1	0.697	0.046	−0.321	−0.575	−0.767	−0.920	−1.048	−1.156	−1.251	−1.334
2		1.161	0.498	0.117	−0.149	−0.351	−0.513	−0.649	−0.764	−0.865
3			1.440	0.780	0.397	0.128	−0.078	−0.245	−0.384	−0.503
4				1.638	0.985	0.604	0.334	0.127	−0.041	−0.183
5					1.792	1.146	0.767	0.499	0.291	0.122
6						1.916	1.277	0.902	0.635	0.428
7							2.020	1.388	1.016	0.751
8								2.110	1.484	1.116
9									2.189	1.569
10										2.258
F^*/c	0.313	0.471	0.568	0.635	0.683	0.720	0.749	0.772	0.792	0.808

Le calcul des seuils est le suivant : m + a×err.

En guise d'exemple pour le score correspondant à la distribution suivante de moyenne 2.272175 et d'erreur standard 0.386572:

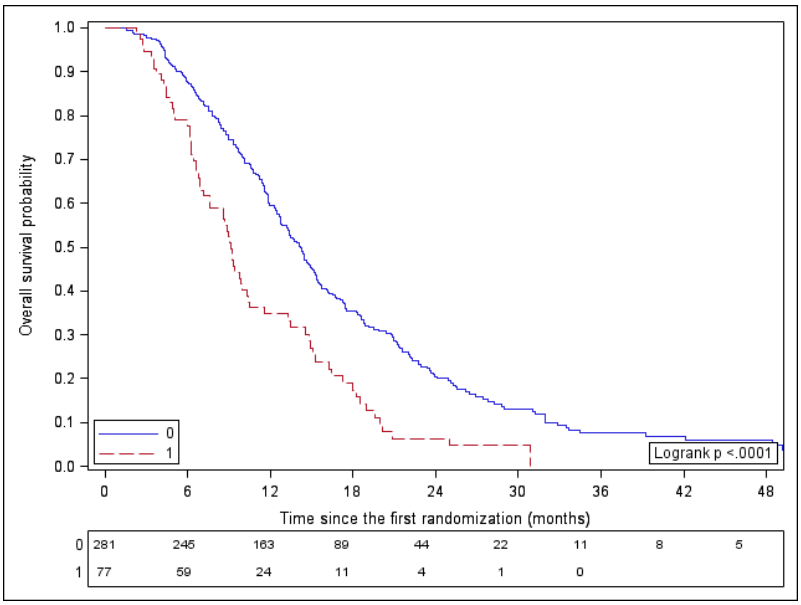


Les seuils des catégories calculés selon la méthode de Cox sont les suivants:

- Pour un choix de 2 groupes:

  x1= 2.2721754+0.38657242×0.697= 2.541616

  Ci-dessous la courbe de Kaplan-Meier pour la survie globale correspondant aux deux groupes ainsi séparés:

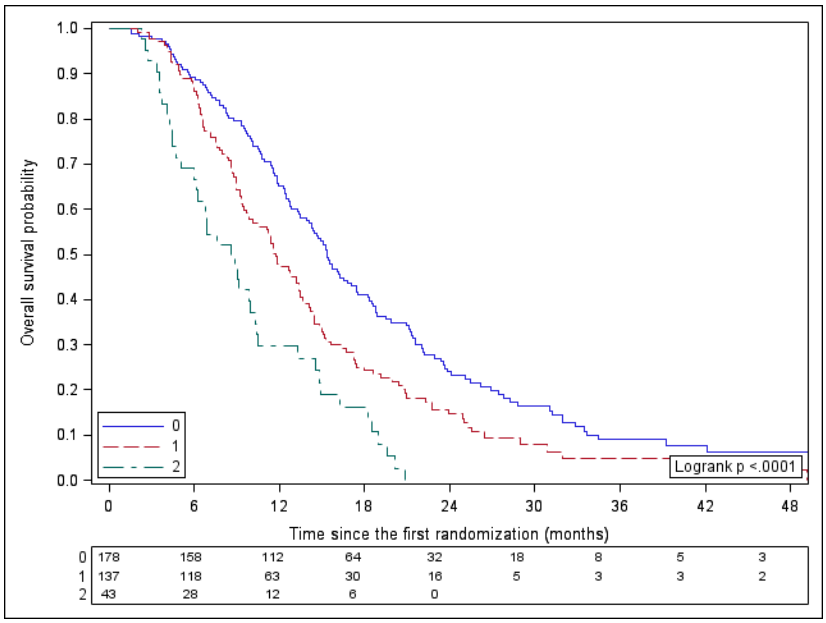


- Pour un choix de 3 groupes:

  x1= 2.2721754+0.38657242×0.046 = 2.289958

  x2 = 2.2721754+0.38657242×1.161 = 2.720986

Ci-dessous la courbe de Kaplan-Meier pour la survie globale correspondant aux trois groupes ainsi définis:

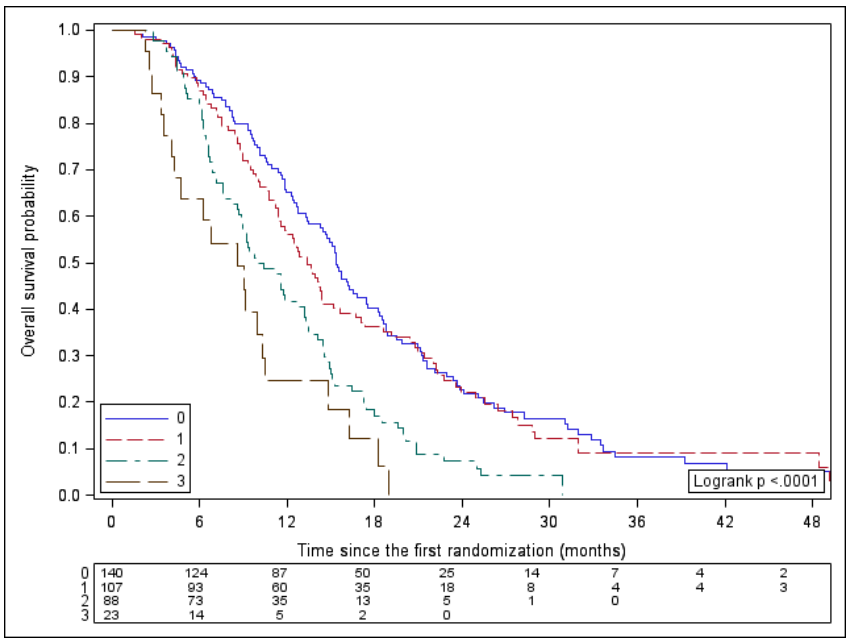


- Pour un choix de 4 groupes:

x1= 2.2721754 + 0.38657242×-0.321 = 2.148086

x2 = 2.2721754 + 0.38657242×0.498 = 2.464688

x3 = 2.2721754 + 0.38657242×1.440 = 2.82884



En considérant que le score pronostique considéré ci-dessus assume une loi normale et que plus le nombre de groupes est important moins nous perdons d'information dans la séparation de la population, il apparait relativement clair que l'approche la plus intéressante réside en la considération de trois groupes distincts dans cet exemple.

Dans le cas où le score pronostique n'est pas assimilable à une loi normale, il faut alors appliquer des méthodes non paramétriques pour déterminer les seuils qui définissent les différentes populations. Cette situation n'a pas été rencontrée dans les travaux proposés dans cette thèse. Une approche nommée « Recursive partitioning » peut constituer une alternative intéressante (Molica *et al*, 2010) (Pan *et al*, 2016). Cette méthode permet de construire un arbre de décision et de maximiser la ressemblance des patients inclus dans le même nœud. Une des limites de cette méthode est qu'elle peut être très échantillon dépendant, ainsi une option consiste en une analyse introduisant de la fluctuation d'échantillon par bootstrap comme dans les analyses dites « random forest » (Strobl *et al*, 2009).

III.4.2.2.4 Possibilité d'évaluation de l'interaction entre les prédictions de survie et le traitement

Dans des domaines comme l'oncologie ou la transplantation rénale, les outils pronostiques sont développés à partir de populations de patients pris en charge dans la pratique clinique ou inclus dans les essais thérapeutiques. Par conséquent, ces patients sont à priori traités par les thérapies dites « gold standard ». On parle alors d'outils pronostiques pour une population de patients bien définie à un moment donné et dans un cadre thérapeutique précis. Certaines situations offrent la possibilité d'évaluer le caractère prédictif de la réponse au traitement à travers les populations de pronostique de survie globale identifiées. En effet, dans les essais thérapeutiques en général deux

bras de traitement sont comparés. Prenons l'exemple de l'essai LAP07 qui compare l'effet de la chimiothérapie à celui de la radio-chimiothérapie chez des patients atteints de cancer pancréatique localement avancé contrôlé après 4 mois de chimiothérapie d'induction à base de gemcitabine. L'effet sur la survie globale a été reconnu comme non significatif et des patients ayant reçu l'un ou l'autre de ces traitements peuvent se trouver dans chacune des populations identifiées de pronostique diffèrent. Il est alors possible de regarder l'effet traitement à travers les groupes de pronostique identifiés (pour lesquelles les prédictions de survie sont différentes) afin de savoir si la réponse au traitement (effet prédictif pour les cliniciens) ou un différentiel de survie selon le traitement peut être identifié.

IV Résultats

IV.1 Axe oncologie

IV.1.1 Développement d'un nomogramme et d'un score pronostique pour prédire la survie globale chez des patients atteints d'un cancer pancréatique localement avancé au diagnostic avant l'initiation du traitement d'induction.

Vernerey D, Huguet F, Vienot A, Goldstein D, Paget-Bailly S, Van Laethem J-L, Glimelius B, Artru P, Moore MJ, André T, Mineur L, Chibaudel B, Benetkiewicz M, Louvet C, Hammel P, Bonnetain F

Prognostic nomogram and score to predict overall survival in locally advanced untreated pancreatic cancer (PROLAP). **Br J Cancer** 115: 281–289, doi:10.1038/bjc.2016.212. (2016)

Impact factor: 5.6

IV.1.1.1 Positionnement du problème

Le cancer pancréatique (PC) est une des premières causes de la mortalité liée au cancer avec un pronostic très sombre illustré par un taux de survie à 5 ans de l'ordre de 8%. Des études épidémiologiques montrent une augmentation de l'incidence de ce cancer et prédisent une progression des estimations de mortalité liée à ce dernier dans le futur. Pour illustration, le nombre de décès par PC en 2014 dans l'union européenne a été estimé à 82 300, ce qui est quasi similaire aux décès attribuables au cancer du sein

chez les femmes (n=89300). Il est donc nécessaire d'encourager la prévention ainsi que la recherche dans cette pathologie.

L'absence de symptômes spécifiques au début de la maladie rend le diagnostic souvent tardif. Ainsi, la majorité des PC sont inopérables (environ 80%) et non résécables en raison d'un envahissement vasculaire, de la présence de métastases locorégionales ou à distance. Au diagnostic, environ 35% des patients présentent un envahissement vasculaire sans métastases caractérisant ainsi un PC localement avancé (LAPC).

Le management des patients atteints d'un LAPC demeure aujourd'hui complexe et controversé. La radiothérapie notamment a été dans un premier temps encouragée pour les patients de bon pronostic sur la base d'études en population rétrospectives. Puis, dans un second temps, elle a été freinée en raison de résultats équivalents à la chimiothérapie seule pour la survie globale (SG) dans l'essai de phase 3 internationale multicentrique LAP07. Cependant, l'absence de données de qualité de vie relative à la santé ainsi qu'une tendance favorable sur la survie sans progression (SSP) amènent certains cliniciens à considérer la radiothérapie comme une option thérapeutique toujours potentiellement intéressante pour les patients LAPC. Il apparaît surtout que le sous-groupe de patients pour lequel la radiothérapie aurait une action bénéfique reste à être identifié plus précisément.

Par ailleurs, l'absence actuelle de consensus, pour la définition de profils de risque spécifique pour la SG chez les patients LAPC, peut induire une confusion dans le développement de stratégies de recherche et aussi dans le management des patients.

Dans ce contexte et avec l'intérêt actuel grandissant pour une médecine de précision, il est donc important de développer des outils pronostiques modernes pour optimiser la sélection des patients LAPC afin de leur proposer des thérapies appropriées et donc in fine un management plus personnalisé. Plus spécifiquement, une meilleure prédiction de la SG au diagnostic de LAPC peut aussi favoriser, dans les futurs essais cliniques, la construction de bras de traitement mieux équilibrés en termes de pronostic et ainsi permettre de mieux mesurer l'effet traitement ou intervention testé dans

l'essai. Une recherche pronostique actualisée chez les patients LAPC permettrait donc d'améliorer le design des essais cliniques futurs.

Les nomogrammes sont des outils de prédiction individuels qui ont connus un développement accru en oncologie et sont proposés comme une alternative intéressante à la traditionnelle classification TNM. Cependant, aucun nomogramme pour prédire la SG n'est disponible et validé pour les patients LAPC.

Cette étude a pour objet, dans un premier temps, de proposer et de valider un nouveau modèle pronostique pour la SG dans le LAPC, en étudiant un spectre important de paramètres disponibles au diagnostic. Dans un second temps, sur la base de ce nouveau modèle pronostique des outils de prédiction individuel (nomogramme) et populationnel (score) seront développés et automatisés pour simplifier leur utilisation par les chercheurs et cliniciens.

IV.1.1.2 Population de l'étude

La cohorte de développement utilisée dans cette étude est constituée des patients inclus dans l'essai de phase III randomisé international et multicentrique LAP07 (ClinicalTrials.gov NCT00634725) qui est à ce jour la cohorte de patients LAPC la plus importante jamais étudiée dans un essai clinique.

La cohorte de validation externe inclus tous les patients LAPC traités au Centre Hospitalier Universitaire de Besançon sur une période de 11 ans. Cette cohorte de validation est particulièrement intéressante puisqu'elle est constituée de patients de la pratique clinique régulière et donc à ce titre elle est non-sélectionnée.

Les patients de la cohorte de développement ont donc été traités selon le protocole proposé dans l'essai LAP07 tandis que les patients de la cohorte de validation ont été pris en charge selon les protocoles cliniques standard en vigueur au moment du diagnostic.

Les données démographiques, liées aux antécédents de cancer, clinico-pathologiques, biologiques et radiologiques ont été collectées au

moment du diagnostic de cancer dans la cohorte de validation et au moment de l'inclusion dans l'essai pour la cohorte de développement.

IV.1.1.3 Approche pronostique méthodologique

Dans la cohorte de développement, la SG a été calculée comme le délai entre la date de randomisation dans l'essai et le décès toutes causes confondues. Les patients en vie à la fin du suivi ont été censurés et renseignés comme étant vivant à cette date. La même définition de la SG a été appliquée dans la cohorte de validation externe à l'exception de la date de randomisation qui a été remplacée par la date de diagnostic du LAPC.

La SG a été estimée par la méthode de Kaplan-Meier et décrite à l'aide de la médiane de survie ou du taux de survie aux temps d'intérêt accompagné de leurs intervalles de confiances à 95%.

Le suivi de l'étude a été estimé selon la méthode du Kaplan-Meier inverse.

Une modélisation de Cox à risque proportionnel a été appliquée pour établir les rapports de risques instantanés (Hazard Ratio ou HR) ainsi que leur intervalles de confiance à 95% pour analyser l'association des paramètres investigués avec la SG.

La construction du modèle de Cox multivarié final pour prédire la SG a d'abord impliqué une étape univariée afin d'identifier les paramètres associés avec la SG puis une étape multivariée dans laquelle tous les facteurs identifiés avec une p-value inférieure à 0.1 en univarié ont été considérés pour être inclus dans un même modèle de Cox. Enfin, les seuls paramètres significativement indépendamment associés avec une p-value < 0.5 ont été retenus pour former le modèle multivarié final. Afin de garder le maximum d'information, toutes les variables quantitatives ont été utilisées dans leur forme continue. Les variables biologiques avec une distribution asymétrique ayant des valeurs très élevées ont été normalisées par une transformation logarithmique. Les transformations réalisées ont été validées par les deux approches statistiques que sont les splines cubiques restreintes et les polynômes fractionnaires. En l'absence d'hypothèse clinique et/ou de

documentation dans la littérature, aucune interaction n'a été considérée dans la construction du modelé multivarié final.

La validité et les performances du modèle multivarié ont été évaluées selon deux paramètres : la discrimination et la calibration.

La valeur prédictive et la capacité à discriminer la SG, et donc à séparer les patients avec un pronostic diffèrent du modèle, ont été évalué avec l'indice de concordance de Harrell (C de Harrell). Mille échantillons aléatoires de la population de développement ont été obtenus par procédure de ré-échantillonnage (bootstrap) et ont été ensuite utilisés pour construire un intervalle de confiance à 95% autour du C de Harrell.

La calibration du modèle, quant à elle, se réfère à sa capacité à fournir des prédictions non biaisés dans des groupes de patients considérés comme similaires. Elle a été analysée graphiquement par une courbe de calibration.

Une validation interne du modèle a été proposée à l'aide d'une procédure de ré-échantillonnage en produisant des intervalles de confiance autour des estimations du modèle.

Le modèle multivarié final a été ensuite utilisé pour produire un nomogramme permettant l'estimation, pour un patient donné, de sa médiane et de ses probabilités individuelles de survie à 6, 12, 24 et 48 mois post-diagnostic. Une approche populationnelle a aussi été proposée en dérivant un score pronostique à partir des points du nomogramme. L'identification de deux et trois groupes de pronostics différents a été définie par des valeurs seuils déterminées à l'aide respectivement, de la méthode du seuil optimal proposé par Hothorn et Lausen en 2003 pour l'approche à deux groupes, et de la méthode dite de Cox pour l'approche à 3 groupes. La capacité discriminante des deux approches (score pronostique à 2 et 3 groupes) a été évaluée par le C de Harrell en considérant la classification des patients dans les différents groupes à risque.

Enfin, en guise de validation externe, le même algorithme pour l'identification des groupes issus de la cohorte de développement a été appliqué dans la cohorte de validation. La capacité discriminante du modèle

multivarié final et du score pronostique a été confirmée dans la cohorte de validation externe par le C de Harrell en considérant respectivement la prédiction de survie issue du modèle et la classification des patients dans un groupe à risque.

IV.1.1.4 Résultats

442 patients traités dans l'essai thérapeutique LAP07 entre février 2008 et décembre 2011 ont été inclus dans la cohorte de développement. 106 patients LAPC pris en charge consécutivement au Centre Hospitalier Universitaire de Besançon ont quant à eux été considérés dans la cohorte de validation externe.
Les caractéristiques de ces deux populations sont présentées dans la table 1.

Table 1 : Caractéristiques des patients au moment du diagnostic dans les cohortes de développement et de validation

		Cohorte de développement – Essai LAP07 (N = 442)	Cohorte de validation externe – Centre Hospitalier Universitaire de Besançon (N = 106)	*P*
Age, années*		63.5 ± 9.6	67.1 ± 10.2	**.0007**
Genre, N (%)	*Homme*	228 (52%)	57 (54%)	
	Femme	214 (48%)	49 (46%)	.7456
Localisation, N (%)	*Tète/Tète et Corps*	302 (69%)	71 (70%)	
	Autres (corps et/ou queue)	138 (31%)	31 (30%)	.9059
	Manquant	2	4	
Grade histologique, N (%)	*Bien différencié*	107 (48%)	9 (26%)	
	Modérément différencié	75 (34%)	17 (50%)	
	Mal différencié	41 (18%)	8 (24%)	**.05556**
	Manquant	219	72	
Invasion des ganglions lymphatiques, N (%)	*Oui*	179 (41%)	-	
	Non	258 (59%)	-	
	Manquant	5	-	
OMS-PS, N (%)	*0*	197 (46%)	31 (30%)	
	1	203 (47%)	59 (56%)	
	2	31 (7%)	14 (13%)	
	3	0 (0%)	1(1%)	**.004892**
	Manquant	11	1	
Taille tumorale (mm)*		44.0 ± 18.0	38.3 ± 14.7	**.00385**
	Manquant	20	9	
Taille tumorale (mm), N (%)	*<30*	66 (16%)	22 (23%)	
	30-50	226 (54%)	56 (58%)	
	≥50	130 (31%)	19 (20%)	**.04676**
	Manquant	20	9	
Envahissement vasculaire, N (%)	*artère mésentérique supérieure*	144 (33%)	-	
	tronc coeliaque	188 (43%)	-	
	artère hépatique	25 (6%)	-	
	Non	79 (18%)	-	
	Manquant	6	-	
Médiane de suivi, mois [95% IC]		34.3 [27.6 to 43.7]	Tous les patients ont été suivis jusqu'au décès (Temps maximum de suivi = 76.6) à l'exception de 5 patients toujours en vie avec une médiane de suivi de 10 mois	

* moyennes ± écart-type

Abréviations: PS: performance status, OMS: Organisation Mondiale de la Santé, IC: intervalle de confiance.

Parmi les 30 paramètres collectés au diagnostic de cancer et analysés, seulement 8 ont été trouvés comme étant significativement associés à la SG dans l'analyse de Cox univarié. Ces paramètres sont l'âge, le grade histologique, le performance status, la pression artérielle systolique, la douleur, l'albumine, la taille de la tumeur et le marqueur tumoral CA 19-9.

Le modèle multivarié final de Cox se compose de 5 facteurs indépendamment associés à la SG : l'âge, la douleur, l'albumine, la taille de la tumeur et le marqueur tumoral CA 19-9 (Table 2).

Table 2: Analyse de Cox multivarié des facteurs associés avec la SG (*N* = 358)

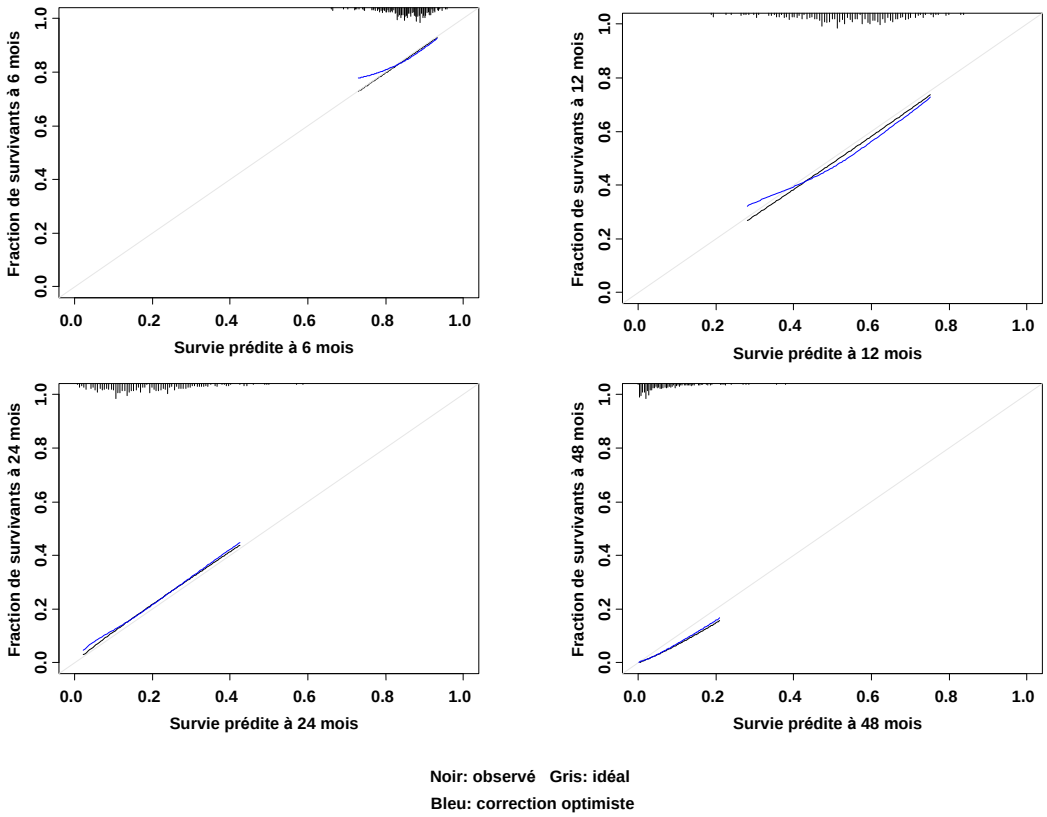
	Nombre de patients	**Nombre de décès**	**HR**	**95% IC**	***P***	**Validation interne BCA HR 95%**	**Score maximal dans le nomogramme**
Age au diagnostic, années	358	307	1.013	1.000 to 1.025	**.0478**	1.000 to 1.025	43
Douleur							
Non	155	130	1	-	**-**		
Oui	203	177	1.292	1.023 to 1.632	**.0317**	1.046 to 1.657	16
Albumine, g/L	358	307	0.963	0.942 to 0.984	**.0006**	0.942 to 0.988	96
Taille tumorale, mm	358	307	1.008	1.001 to 1.015	**.0214**	1.000 to 1.016	100
CA 19-9, UI (valeur logarithmique)	358	307	1.172	1.047 to 1.311	**.0056**	1.039 to 1.306	80

Abréviations: IC: Intervalle de confiance, HR: Hazard Ratio, BCA: accelerated bootstrap confidence interval

Le modèle multivarié final est caractérisé par une capacité discriminante acceptable (C-index = 0.60; 95% IC par bootstrap, 0.57–0.64) et une bonne calibration aux temps d'intérêt (6, 12 ,24 et 48 mois) comme reportée sur la figure 1.

Les intervalles de confiance à 95% sont produits autour des HR en guise de validation interne.

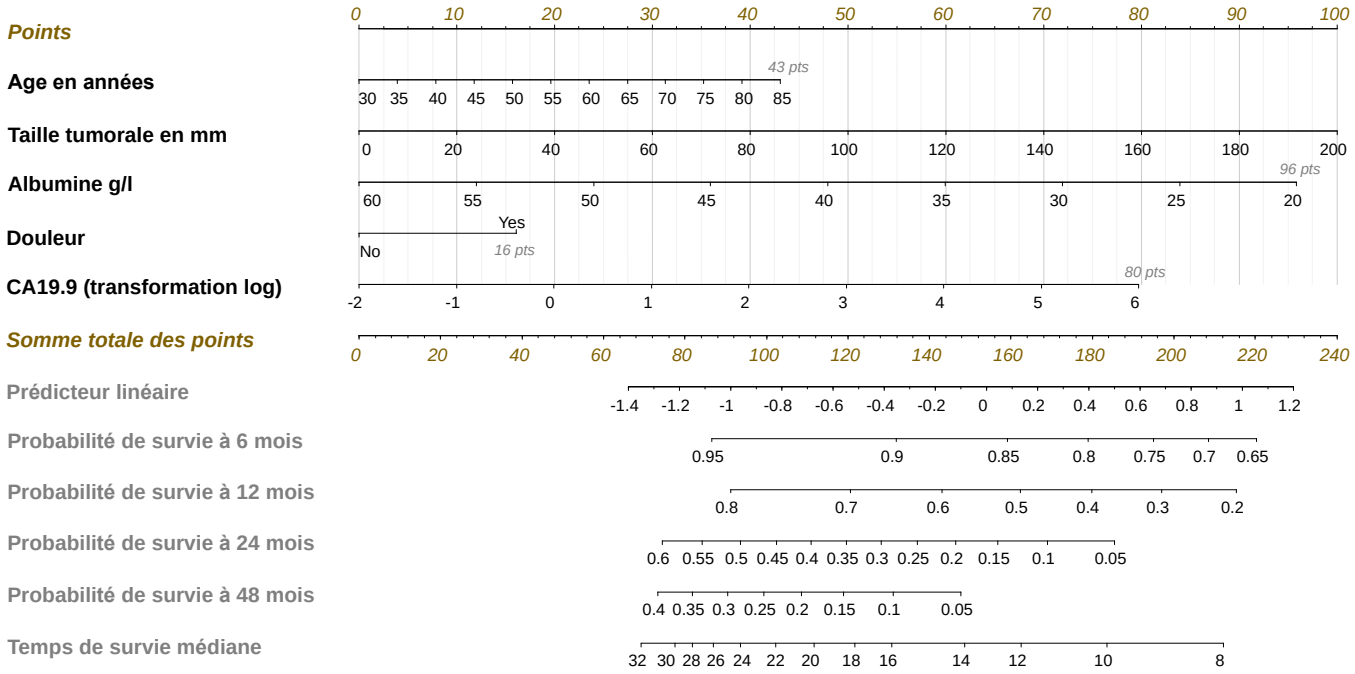
Figure 1: Courbes de calibration à 6, 12, 24, et 48 mois pour le modèle multivarié final



L'axe vertical représente la proportion de patients survivants observés aux temps d'intérêt tandis que l'axe horizontal représente la prédiction issue du modèle. La ligne grise représente un modèle parfaitement calibré.

Sur la base du modèle multivarié final, un nomogramme a été construit permettant pour un patient donné de calculer sa médiane de survie et aussi ses probabilités individuelles de survie à 6, 12, 24 et 48 mois post-diagnostic de cancer (Figure2).

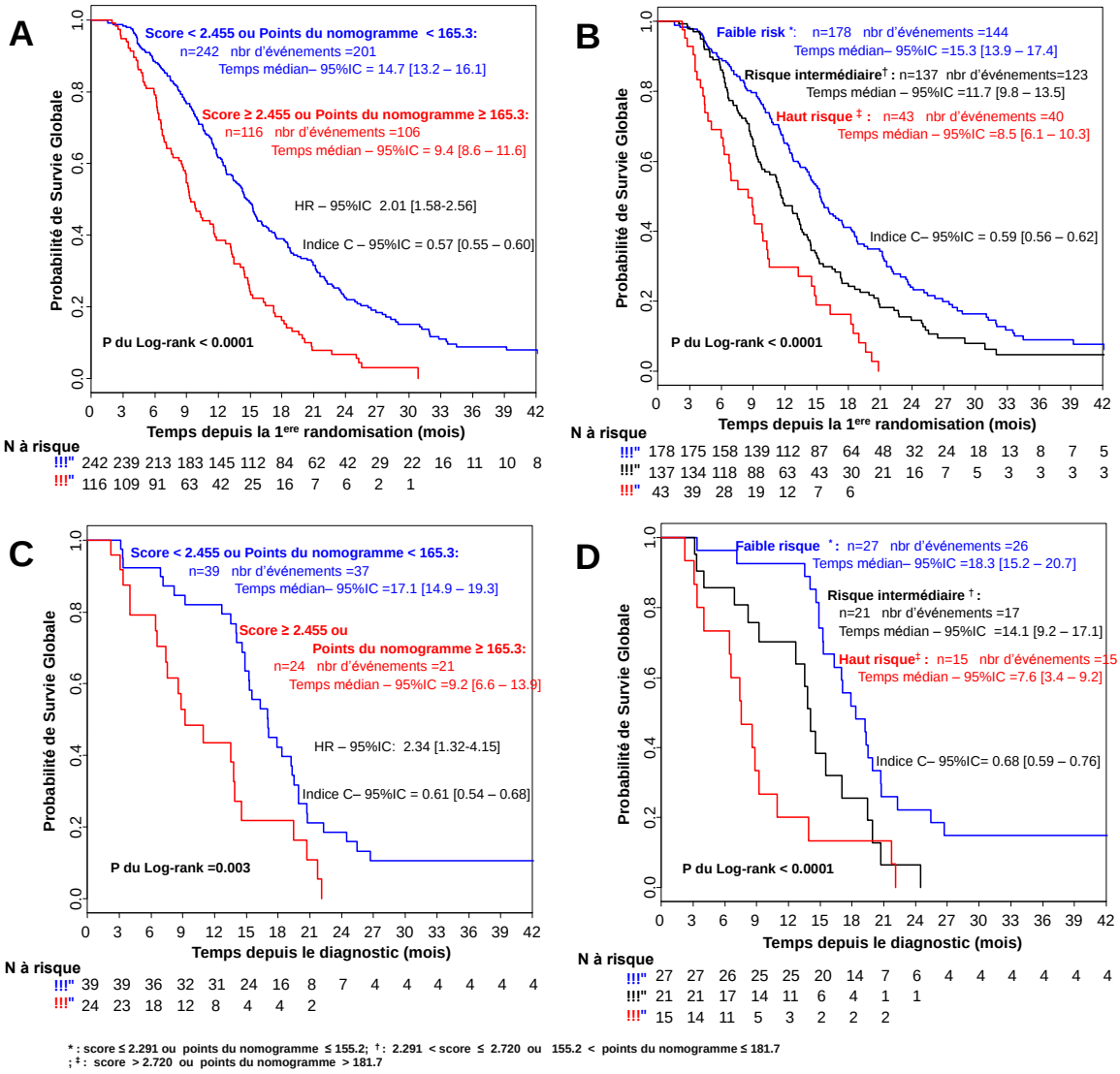
Figure 2: Nomogramme pour prédire la survie globale individuelle pour des patients LAPC



Les points sont assignés pour chaque facteur de risque en tracant une ligne verticale vers la ligne correspondant aux points. Le total des points pour les 5 facteurs est ensuite tracé sur la ligne correpondant à la somme totale des points. Une ligne verticale vers la bas est alors tracée pour deduire les predictions correspondantes aux temps d'intérêt ainsi que la médiane de survie du patient.

Pour les 358 patients impliqués dans le modèle multivarié, il a donc été possible d'obtenir un nombre total de points issu du nomogramme. Ce vecteur de 358 estimations a constitué donc notre score pronostique. En appliquant la méthode du seuil optimal de Hothorn et Lausen, la valeur seuil de 165.3 permet donc de construire 2 groupes de pronostics différents (Figure 4 A) tandis que la méthode dite de Cox nous fournit les seuils (155.2 et 181.7) pour isoler 3 groupes de pronostics différents (Figure 4 B). La pertinence des populations isolées dans les approches à 2 et 3 groupes a été validée en externe dans la cohorte constituée de patients de la pratique clinique courante (Figure 4 B et D).

Figure 4 : Courbes de Kaplan-Meier de la survie globale pour (A, B) la cohorte de développement et (C, D) la cohorte de validation externe selon la catégorisation en 2 groupes avec la méthode de Horton (A, C) et en 3 groupes par la méthode de Cox (B, D).



* : score ≤ 2.291 ou points du nomogramme ≤ 155.2; † : 2.291 < score ≤ 2.720 ou 155.2 < points du nomogramme ≤ 181.7 ; ‡ : score > 2.720 ou points du nomogramme > 181.7

Parmi les 106 patients de la cohorte de validation externe, le score a été applicable pour les 63 patients qui possédaient l'ensemble des paramètres nécessaires à son calcul (Figure 4).

Globalement le modèle multivarié final identifié dans la cohorte de développement propose une meilleure discrimination la SG dans la cohorte de validation (indice C = 0.68; 95% percentile par bootstrap IC, 0.59–0.77).

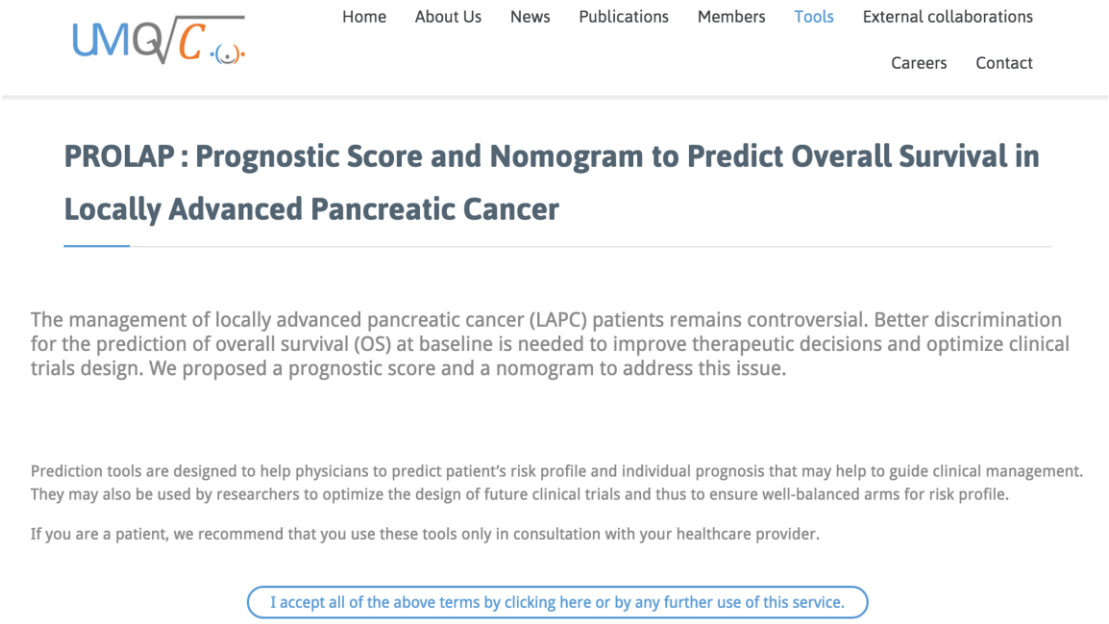
Cette excellente validation se confirme aussi dans l'approche à 2 et 3 groupes avec des estimations de l'indice C améliorée (0.61; 95%IC 0.54 - 0.68 et 0.68; 95%IC 0.59–0.76).
Le score qui a été développé dans une population de patients provenant d'un essai thérapeutique (LAP07) conserve donc d'excellentes propriétés dans une population de patients non sélectionnée représentative de la pratique clinique courante.

Enfin, une application en ligne sur le web (compatible avec les téléphones modernes) a été développée pour fournir des estimations de survie individualisées (nomogramme) et en population (score) pour les cliniciens et chercheur.
L'adresse de cette application est la suivante : http://www.umqvc.org/en/tool/prolap.html.

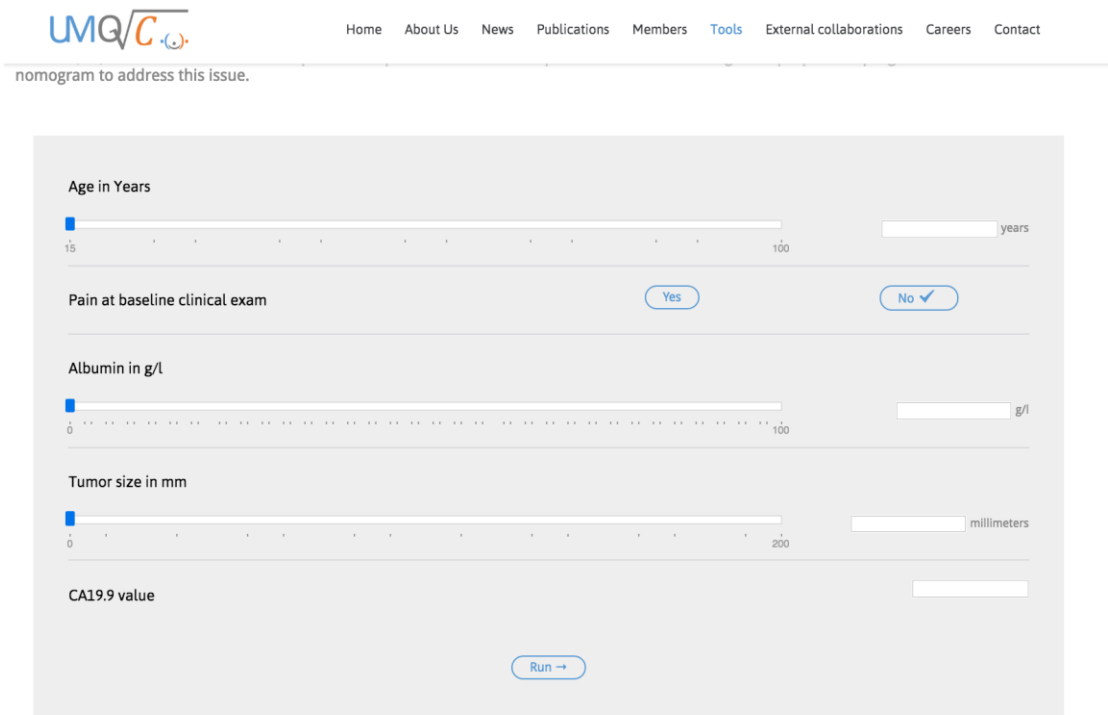
Cette étape s'est avérée primordiale afin de rendre fonctionnels ces outils. En effet, l'utilisation des variables en continu ainsi que leur transformation pour garder le plus d'information possible et modéliser correctement le lien entre ces variables et la distribution de survie rend l'utilisation des outils impossible sans automatisation des calculs.

Concrètement, l'utilisateur rentre les données de son patient, et obtient un rapport qui lui fournit instantanément les estimations du nomogramme et du score.

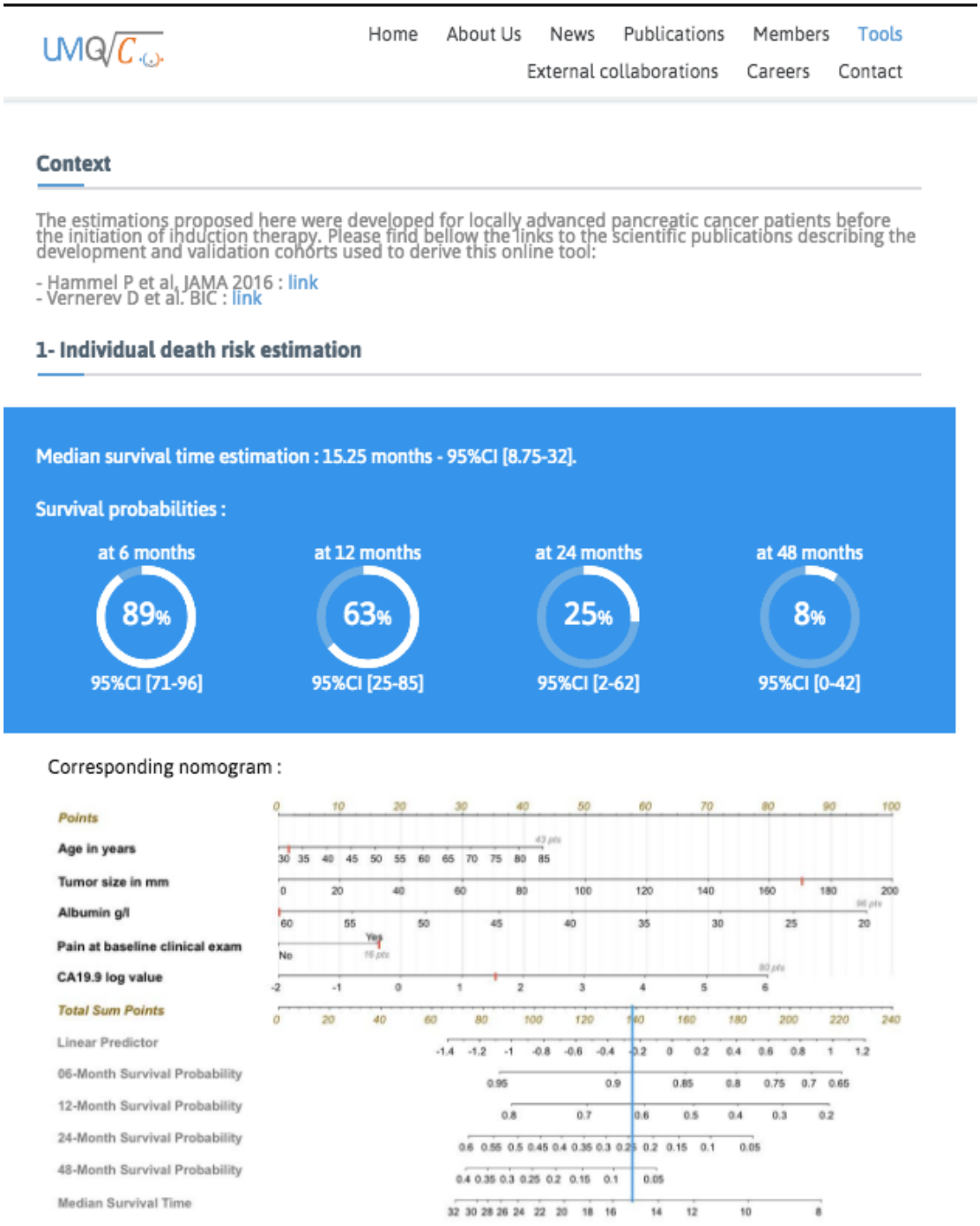
La page d'accueil de l'outil est une page d'information pour situer l'outil dans son contexte.



Ensuite les données du modèle multivarié propres au patient doivent être saisies.



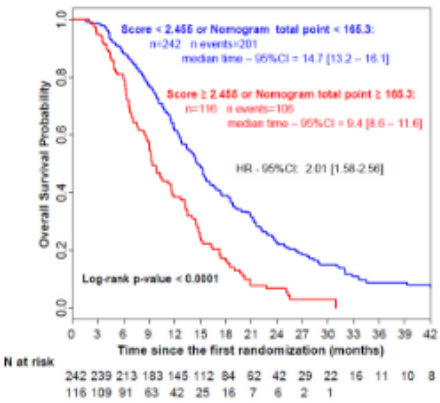
Les résultats issus du nomogramme et du score sont immédiatement visualisables sur la page de résultat comme ci-dessous :



2- Population death risk classification

2-A Two risk-groups approach

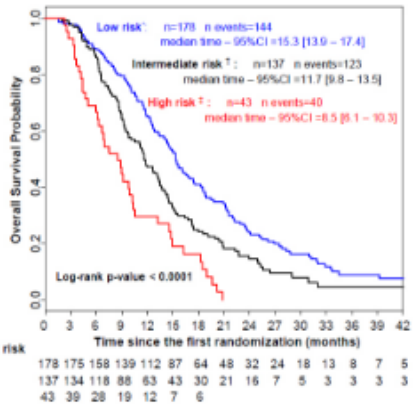
Corresponding risk population
2.01 (Blue curve)



Kaplan–Meier curves for overall according to the prognostic score group.

2-B Three risk-groups approach

Corresponding risk population for the pragmatic approach : 2.01 (Blue curve)



Kaplan–Meier curves for overall according to the prognostic score group.

* : Low risk group score : score ≤ 2.291 or Nomogram total point ≤ 155.2;
† : Intermediate risk group : 2.291 < score ≤ 2.720 or Nomogram point total 155.2 < score ≤ 181.7;
‡ : High risk group : score > 2.720 or Nomogram total point > 181.7

3- For the following data

- CA19.9 use : Yes
- Age in Years : 32 / So 1.56 points
- Tumor size in mm : 170mm / So 85 points
- Albumin in g/l : 71.3 / So 0 points
- Pain at baseline clinical exam : 1 / So 16 points
- CA19.9 value : 34 (log-value : 1.5314789170423) / So 35.31 points

Total of points : 138 points

Results extraction in PDF file :

Enfin une extraction de ces résultats peut être demandée pour insérer les estimations dans un dossier patient par exemple.

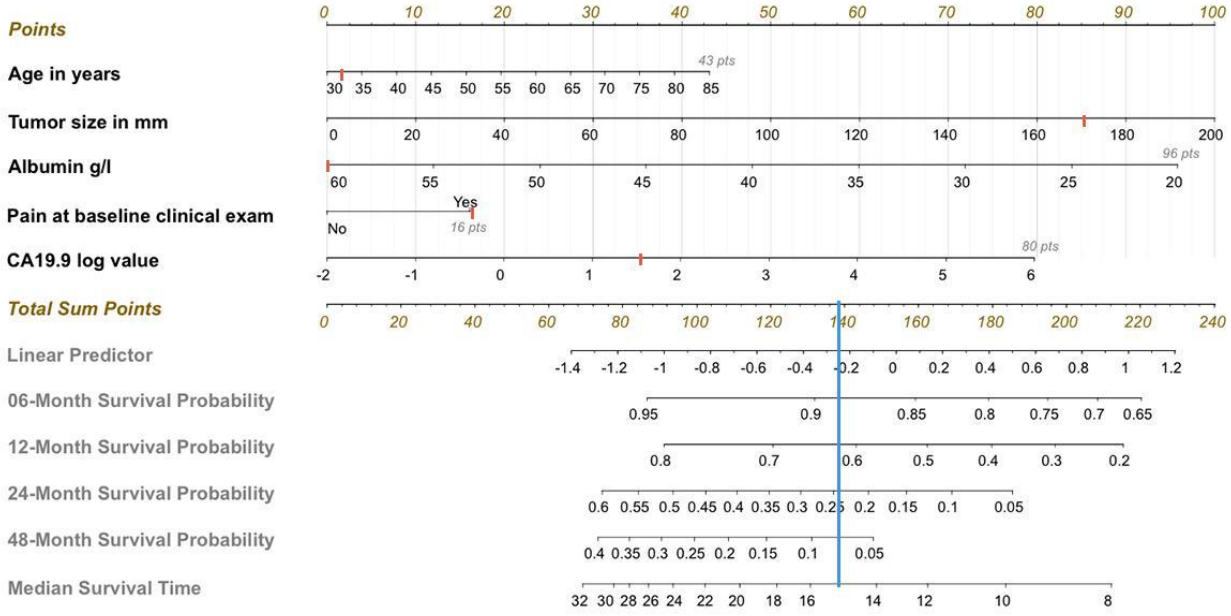
L'extraction en question se présente sous la forme suivante :

1- Individual death risk estimation

Median survival time estimation : 15.25 months // 95%CI [8.75-32].

Survival probabilities at 6 months : 89% // 95%CI [71-96].
Survival probabilities at 12 months : 63% // 95%CI [25-85].
Survival probabilities at 24 months : 25% // 95%CI [2-62].
Survival probabilities at 48 months : 8% // 95%CI [0-42].

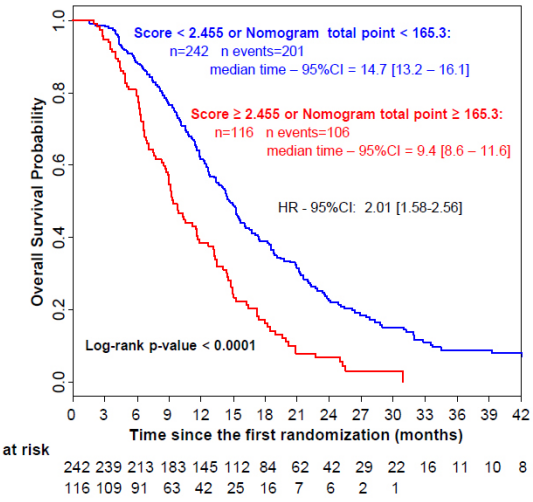
Corresponding nomogram :



2- Population death risk classification

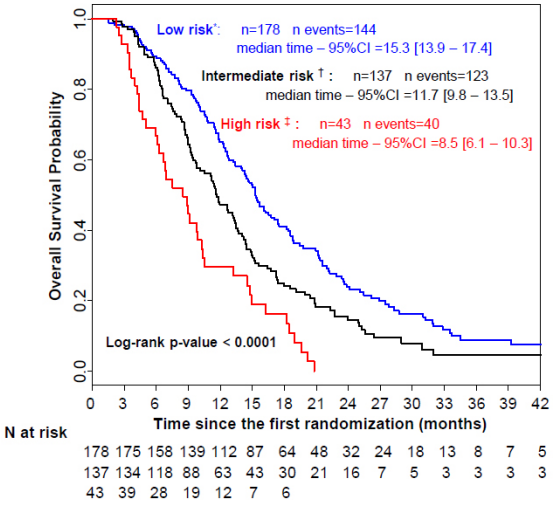
2-A Two risk-groups approach

Corresponding risk population :
2.01 : (Blue curve)



2-B Three risk-groups approach

Corresponding risk population :
2.01 : (Blue curve)



3- For the following data

CA19.9 use : Yes
Age in Years : 32 / So 1.56 points
Tumor size in mm : 170mm / So 85 points
Albumin in g/l : 71.3 / So 0 points
Pain at baseline clinical exam : 1 / So 16 points
CA19.9 value : 34 (log-value : 1.5314789170423) / So 35.31 points

IV.1.1.5 Conclusion de l'étude

Nous proposons donc dans cette étude un nouveau modèle de prédiction de la SG pour des patients atteints d'un LAPC au moment du diagnostic avant le traitement d'induction.

Nous avons, sur la base de ce nouveau modèle, développé des outils pronostiques fonctionnels à l'échelle individuelle (nomogramme) et populationnelle (score) dont l'utilisation a été facilitée avec le développement d'une interface web qui automatise les calculs et synthétise les résultats (http://www.umqvc.org/en/tool/prolap.html).

Au-delà des outils proposés cette étude met en évidence une considérable hétérogénéité de pronostic pour la SG au diagnostic chez les patients LAPC.

Il est important de rappeler que les outils ont été développés à l'aide d'une population de patients provenant d'un essai thérapeutique et validé en externe sur une population de patients rencontrés dans la pratique clinique courante et donc non sélectionnés.

Cette étude fournit des éléments qui permettent d'optimiser le design de futurs essais cliniques et peut offrir l'opportunité de définir des stratégies de management adaptées aux différents profils de risque.

IV.1.1.6 Manuscrit scientifique

FULL PAPER

BJC

British Journal of Cancer (2016), 1–9 | doi: 10.1038/bjc.2016.212

Keywords: prognostic score; nomogram; overall survival; locally advanced pancreatic cancer; chemotherapy

Prognostic nomogram and score to predict overall survival in locally advanced untreated pancreatic cancer (PROLAP)

Dewi Vernerey*[,1], Florence Huguet[2,3,16], Angélique Vienot[4,16], David Goldstein[5,6,16], Sophie Paget-Bailly[1], Jean-Luc Van Laethem[7], Bengt Glimelius[8], Pascal Artru[2,9], Malcolm J Moore[10], Thierry André[2,11], Laurent Mineur[12], Benoist Chibaudel[2,13], Magdalena Benetkiewicz[2], Christophe Louvet[2,14], Pascal Hammel[2,15] and Franck Bonnetain[1,2]

[1]*Methodological and Quality of Life in Oncology Unit, EA 3181, University Hospital of Besançon, 3 Boulevard Alexandre Fleming, Besançon 25030, France;* [2]*Oncology Multidisciplinary Research Group (GERCOR), 151 rue du Faubourg Saint Antoine, Paris 75011, France;* [3]*Department of Radiotherapy, Tenon Hospital (AP-HP), 4 rue de la Chine, Paris 75020, France;* [4]*Department of Gastroenterology, University Hospital of Besançon, 3 Boulevard Alexandre Fleming, Besançon 25030, France;* [5]*Department of Medical Oncology, Prince of Wales hospital and Prince of Wales Clinical school, UNSW, Sydney, New South Wales 2031, Australia;* [6]*AGITG (Australasian Gastrointestinal Trials Group), 119-143 Missenden Rd, Camperdown, New South Wales 2050, Australia;* [7]*Department of Gastroenterology, Erasme University Hospital, Route de Lennik 808, Brussels 1070, Belgium;* [8]*Department of Radiology, Oncology and Radiation Science, University of Uppsala, Uppsala 75105, Sweden;* [9]*Department of Gastroenterology, Jean Mermoz Hospital, 55 avenue Mermoz, Lyon 69008, France;* [10]*Department of Medical Oncology, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada;* [11]*Department of Medical Oncology, Saint-Antoine Hospital (AP-HP), 184 rue du Faubourg Saint Antoine, Paris 75011, France;* [12]*Department of Radiotherapy and Medical Oncology, Sainte-Catherine Institute, 250 Chemin de Baigne Pieds, Avignon 84918, France;* [13]*Department of Medical Oncology, Franco-British Hospital Institute, 3 Rue Barbès, Levallois-Perret 92300, France;* [14]*Department of Medical Oncology, Institute Mutualiste Montsouris, 42 Boulevard Jourdan, Paris 75014, France and* [15]*Department of Digestive Oncology, Beaujon Hospital (AP-HP), 100 boulevard du General Leclerc, Clichy 92110, France*

Background: The management of locally advanced pancreatic cancer (LAPC) patients remains controversial. Better discrimination for overall survival (OS) at diagnosis is needed. We address this issue by developing and validating a prognostic nomogram and a score for OS in LAPC (PROLAP).

Methods: Analyses were derived from 442 LAPC patients enrolled in the LAP07 trial. The prognostic ability of 30 baseline parameters was evaluated using univariate and multivariate Cox regression analyses. Performance assessment and internal validation of the final model were done with Harrell's C-index, calibration plot and bootstrap sample procedures. On the basis of the final model, a prognostic nomogram and a score were developed, and externally validated in 106 consecutive LAPC patients treated in Besançon Hospital, France.

Results: Age, pain, tumour size, albumin and CA 19-9 were independent prognostic factors for OS. The final model had good calibration, acceptable discrimination (C-index = 0.60) and robust internal validity. The PROLAP score has the potential to delineate three different prognosis groups with median OS of 15.4, 11.7 and 8.5 months (log-rank $P < 0.0001$). The score ability to discriminate OS was externally confirmed in 63 (59%) patients with complete clinical data derived from a data set of 106 consecutive LAPC patients; median OS of 18.3, 14.1 and 7.6 months for the three groups (log-rank $P < 0.0001$).

Conclusions: The PROLAP nomogram and score can accurately predict OS before initiation of induction chemotherapy in LAPC-untreated patients. They may help to optimise clinical trials design and might offer the opportunity to define risk-adapted strategies for LAPC management in the future.

*Correspondence: D Vernerey; E-mail: dvernerey@chu-besancon.fr
[16]These authors contributed equally to this work.

Received 18 February 2016; revised 15 June 2016; accepted 20 June 2016

© 2016 Cancer Research UK. All rights reserved 0007 – 0920/16

Pancreatic cancer (PC) is one of the leading causes of cancer-related mortality worldwide, with 5-year relative survival of only 8%. The predicted death rates have progressively increased over the past years (Siegel *et al*, 2016). This unfavourable trend emphasises the importance of giving priority to research in PC prevention and treatment. Number of deaths by PC in 2014 in the European Union (EU) has been estimated to 82 300 cases equally distributed in men and women. By analogy, a similar prediction has been reported for women with breast cancer ($N = 89\,300$) (Malvezzi *et al*, 2014).

Most PCs are unresectable due to the presence of either distant metastases or locoregional metastases, including vascular invasion (locally advanced PC: LAPC). At the time of diagnosis, up to 35% of patients with PC present with locally advanced disease (Philip, 2011). The management of those patients is still controversial, specifically the role of radiotherapy remains uncertain (Huguet *et al*, 2007; Krishnan *et al*, 2007; Chauffert *et al*, 2008; Loehrer *et al*, 2011; Philip, 2011; Hajj and Goodman, 2015).

Currently, there is a lack of staging system and consensus regarding specific risk profile for overall survival (OS) in LAPC, which can lead to confusion in research strategies development and potentially inappropriate management of patients. Given this aspect and the current interest in precision medicine, there is a need for prognostic tools to optimise patients' selection for appropriate treatment therapy in order to achieve more personalised management. Specifically, more accurate prediction of OS at diagnosis may help to ensure well-balanced arms in clinical trials and to a better management patient care in the future. Ultimately, such a tool may lead to improve patient's health-related quality of life (HRQoL), chemotherapy (CT) cost-effectiveness ratio, and design of future clinical trials.

Nomograms have been increasingly used for survival prediction in different cancer settings as an alternative to traditional TNM staging system or a new standard (Balachandran *et al*, 2015). Unfortunately, no single study evaluating a nomogram for survival prediction in LAPC patients is available.

We aimed to develop and validate a new prognosis model, nomogram and score for OS using a broad spectrum of parameters currently available at baseline. This model was developed based on a large cohort of LAPC patients enrolled in the phase III LAP07 trial and validated in patients encountered in daily clinical practice.

PATIENTS AND METHODS

Patients. The development cohort was obtained from an international, multicentre, randomised phase III LAP07 trial (ClinicalTrials.gov NCT00634725). The complete design and trial procedures of the LAP07 study was previously described (Hammel *et al*, 2016). Briefly, the LAP07 study assessed whether chemoradiotherapy
(CRT) administered to LAPC patients whose tumour has been controlled after 4 months of gemcitabine alone or plus erlotinib induction CT increases OS compared with continuation of the same CT. In this trial, CRT did not increase OS compared with the continuation of CT alone. Demographics, cancer history, clinicopathological, biological and radiological parameters at baseline as well as treatment outcomes were collected.

The external validation cohort included 106 consecutive LAPC patients treated at the University Hospital of Besançon, France between January 2003 and December 2013.

Main eligibility and exclusion criteria along with treatment and procedures for the development and validation cohorts are specified in Supplementary Appendix A1.

Follow-up. In the development cohort, patients were observed at 2, 4, 7, 9 and 11 months and every 2 months thereafter. At each follow-up visit, a detailed clinical history and a complete physical examination were carried out. Database was locked on 23 February 2014.

In the validation cohort, patients were treated using standard clinical protocols. A radiological evaluation was performed every 3 months. Database was locked on 13 November 2015.

Statistical analysis. The primary analysis was performed on the LAP07 intent to treat population. Mean (s.d.) values and frequency (percentage) were provided for the description of continuous and categorical variables, respectively. Means and proportions were compared using Student's t-test and χ^2-test (or Fisher's exact test, if appropriate), respectively.

In the development cohort, OS was calculated from the date of study randomisation to the date of death from any cause. Patients last known to be alive were censored at the time of their last follow-up assessment (Bonnetain *et al*, 2014). In the external validation cohort, the starting point for OS was the date of diagnosis of LAPC. OS was estimated using the Kaplan–Meier method and described using median or rate at specific time points with 95% confidence intervals (95% CI). Follow-up was calculated using a reverse Kaplan–Meier estimation when feasible (Schemper and Smith, 1996).

Cox-proportional-hazard models were performed to estimate the hazard ratio (HR) and 95% CI for factors associated with OS. The association of the baseline parameters with OS was first assessed using univariate Cox analyses, and then those with $P < 0.10$ were entered into a final multivariate Cox regression model. Significantly skewed continuous variables, such as neutrophil count, total bilirubin, alkaline phosphatase, aspartate transaminase, alanine aminotransaminase, lactate dehydrogenase, albumin and cancer antigen 19.9 (CA 19-9), were normalised by logarithmic transformation. Hazard proportionality was checked by plotting log-minus-log survival curves and by the cumulative martingale process plots. Log-linearity assumption for continuous variables was also checked graphically by making a categorical variable. For continuous variables involved in the final multivariate model, a fractional polynomial method was applied to validate the transformation used and to assess the relationship between predictors and survival. In the absence of clinical hypothesis and/or documentation in the literature, no interaction term was considered in the model construction.

Accuracy of the final model was verified regarding two parameters: discrimination and calibration. The predictive value and the discrimination ability (that is, the ability to separate patients with different prognosis) of the final model were evaluated with the Harrell's concordance index (C-index) (Harrell *et al*, 1996). One thousand random samples of the population were used to derive 95% CI for the C-statistic. Calibration and goodness of fit refers to the ability to provide unbiased survival predictions in groups of similar patients. These were assessed by visual examination of calibration plot and tested with an extension of the Hosmer–Lemeshow test for survival data. Internal validation of the final multivariate model was performed with a bootstrap sample procedure.

The final model was used to derive a nomogram allowing the estimation of median and individual post-diagnosis OS probabilities at 6, 12, 24 and 48 months. For estimation at population level, a prognostic score was constructed and weighted with β-coefficients estimations in the final model, and normalised between 0 and 5. To give a reasonable spread of risk, we chose to work on two and three prognostic risk groups. In such way, different cut points for the risk score from the development set were explored. These were based on the optimal cut point (Hothorn and Lausen, 2003) and Cox's (Cox, 1957) methods for the two- and three-risk groups approach, respectively. Median and tertiles were also considered. The discrimination abilities of the different prognostic scores produced were assessed with the Harrell's concordance index (C-index) by considering risk groups classification.

Patients' baseline key characteristics were compared between prognostic risk groups using Fisher-exact test and the Kruskal–Walis test for categorical and quantitative parameters, respectively.

To explore the performance of the final model, sensitivity analyses were performed by forcing the treatment assigned at first randomisation (gemcitabine *vs* gemcitabine plus erlotinib) to enter into the model and by including baseline histological grade and systolic blood pressure; two parameters that were not initially selected in the multivariate analysis process due to their high rate of missing data. To assess potential bias arising from missing data for parameters involved in the multivariate final model on their significance (*P*-value) and estimates (β and its standard error), a multiple imputation procedure with a Markov chain Monte Carlo method was performed using SAS MI and MIANALYZE procedure.

The final multivariate and the prognostic score discrimination abilities were confirmed in an external validation cohort and evaluated with the Harrell's concordance index (C-index) with survival predictions from the model and from risk groups classification, respectively. To identify risk groups and to determine their survival profile, the same development cohort-derived risk predictive algorithm was applied.

All analyses were performed using SAS version 9.3 (SAS Institute) and R software version 2.15.2 (R Development Core Team; http://www.r-project.org). Values of $P<0.05$ were considered statistically significant and all tests were two-sided. Details on the interpretation of important statistical concepts are given in Supplementary Appendix A2.

RESULTS

Patient characteristics. A total of 442 LAPC patients treated in the LAP07 trial between February 2008 and December 2011 formed the development cohort. The median follow-up time was 34.3 months (95% CI, 27.6–43.7). Overall, 106 LAPC patients treated in the University Hospital of Besançon between January 2003 and December 2013 were included in the external validation cohort. Patient baseline characteristics according to the cohort set are summarised in Table 1. As expected, given its daily clinical nature the validation cohort differed from the LAP07 population with respect to, age, histological grade, WHO-PS, tumour size and treatment regimens.

Table 1. Baseline characteristics according to the cohort set

	Development set cohort–LAP07 trial (*N*=442)	External validation set cohort–University Hospital of Besançon cohort (*N*=106)	*P*
Age, years[a]	63.5 ± 9.6	67.1 ± 10.2	**0.0007**
Gender, *N* (%)			
Male	228 (52%)	57 (54%)	
Female	214 (48%)	49 (46%)	0.7456
Localisation, *N* (%)			
Head/head and body	302 (69%)	71 (70%)	
Other (body and/or tail)	138 (31%)	31 (30%)	0.9059
Unknown	2	4	
Grading, *N* (%)			
Well differentiated	107 (48%)	9 (26%)	
Moderately differentiated	75 (34%)	17 (50%)	
Poorly differentiated	41 (18%)	8 (24%)	**0.05556**
Missing	219	72	
Lymph nodes, *N* (%)			
Yes	179 (41%)	—	
No	258 (59%)	—	
Missing	5	—	
WHO-PS, *N* (%)			
0	197 (46%)	31 (30%)	
1	203 (47%)	59 (56%)	
2	31 (7%)	14 (13%)	
3	0 (0%)	1(1%)	**0.004892**
Missing	11	1	
RECIST tumour size (mm)[a]	44.0 ± 18.0	38.3 ± 14.7	**0.00385**
Missing	20	9	
RECIST tumour size (mm), N (%)			
<30	66 (16%)	22 (23%)	
30–50	226 (54%)	56 (58%)	
⩾50	130 (31%)	19 (20%)	**0.04676**
Missing	20	9	
Vascular invasion, *N* (%)			
Superior mesenteric artery	144 (33%)	—	
Coeliac trunk	188 (43%)	—	
Hepatic artery	25 (6%)	—	
No	79 (18%)	—	
Missing	6	—	
Median follow-up time, months (95% CI)	34.3 (27.6–43.7)	All patients were followed until death (maximum time observed = 76.6) except five patients alive with a median follow-up equal to 10	

Abbreviations: CI = confidence interval; NA = not available; PS = performance status; WHO = World Health Organization. Bold values refer to significant *P*-values reflecting statistical significant difference for the parameter considered between the 2 cohorts.
[a]Plus–minus values are means ± s.d.

Prognostic factors for OS. Only 8 out of 30 evaluated baseline parameters were associated with OS in the univariate Cox analyses with a $P<0.1$. These were age, histological grade, performance status (the World Health Organization scale), systolic blood pressure, pain, albumin level, tumour size and CA 19-9 (Table 2). The multivariate Cox analysis ($N=358$) showed five independent risk factors for OS: age; pain; albumin; tumour size; and CA 19-9 (Table 3).

Performance assessment and internal validation of the final model. The multivariate model exhibited acceptable discrimination ability (C-index 0.60; 95% bootstrap percentile CI, 0.57–0.64)

Table 2. Univariate analysis of factors associated with overall survival

	Number of patients	Number of events	HR	95% CI	*P*
Demographic parameters					
Age at inclusion, years	442	379	1.013	1.003–1.024	**0.0148**
Sex					
Male	228	198	1	—	—
Female	214	181	0.915	0.748–1.120	0.3897
Pathological parameters					
Primary tumour site					
Head–head and body	302	260	1		
Body and/or tail	138	117	0.951	0.764–1.183	0.6519
Unknown	2				
Histological grade					
Well differentiated	107	89	1	—	—
Moderately differentiated	75	62	1.106	0.799–1.530	—
Poorly differentiated	41	40	1.939	1.331–2.825	**0.0021**
Missing	219				
Regional lymph node invasion					
No	258	222	1	—	—
Yes	179	152	0.954	0.776–1.174	0.6584
Missing	5	—	—	—	—
Arterial involvement					
No	10	7	1	—	—
Yes	430	370	0.635	0.299–1.345	0.2352
Missing	2	—	—	—	—
Clinical parameters					
PS (WHO)					
0	197	162	1	—	—
1	203	179	1.252	1.012–1.550	—
2	31	30	1.731	1.171–2.558	**0.0092**
Missing	11	—	—	—	—
Systolic blood pressure, mmHg	252	222	1.011	1.003 to 1.019	**0.0048**
Diastolic blood pressure, mmHg	260	222	1.006	0.997–1.015	0.2176
Diarrhoea					
No	408	351	1	—	—
Yes	31	27	0.763	0.516–1.129	0.1763
Missing	3	—	—	—	—
Pain					
No	186	155	1	—	—
Yes	253	223	1.199	0.976–1.472	**0.0834**
Missing	3	—	—	—	—
Jaundice					
No	387	334	1	—	—
Yes	51	44	0.971	0.709–1.331	0.8550
Missing	4	—	—	—	—
Body Mass Index, $kg\,m^{-2}$	434	372	0.993	0.968–1.019	0.5831
Body mass index, $kg\,m^{-2}$					
Normal weight (18.5–25)	249	216	1		
Underweight <18.5	37	32	1.041	0.718–1.510	
Overweight (25–30)	114	97	0.979	0.770–1.244	
Obese $\geqslant 30$	34	27	0.944	0.632–1.409	0.9822
Missing	8				
Percentage of weight loss	385	329	1.009	0.993–1.025	0.2653
Biological parameters					
Haematological function					
Neutrophils, mm^3 (log-value)	437	374	1.473	0.818–2.654	0.1973
Haemoglobin, $g\,dl^{-1}$	442	379	0.954	0.887–1.025	0.1976
Platelets, mm^3	438	376	1.000	1.000–1.000	0.1442

Table 2. (Continued)

	Number of patients	Number of events	HR	95% CI	P
Biochemical					
Creatinine, μmol l^{-1}	440	377	0.996	0.990–1.002	0.2153
Creatinine clearance rate, ml min^{-1}	313	263	1.000	0.996–1.005	0.9169
Potassium, mmol l^{-1}	422	363	0.985	0.784–1.237	0.8950
Total bilirubin, μmol l^{-1} (log-value)	440	377	1.028	0.769–1.375	0.8517
Alkaline phosphatase, UI l^{-1} (log-value)	432	371	1.147	0.789–1.667	0.4727
Aspartate transaminase, UI l^{-1} (log-value)	438	375	1.297	0.832–2.021	0.2510
Alanine aminotransferase, UI l^{-1} (log-value)	436	373	0.866	0.623–1.205	0.3930
Lactate dehydrogenase, UI l^{-1} (log-value)	313	268	1.089	0.599–1.983	0.7792
Albumin, g l^{-1}	393	336	0.963	0.944–0.982	**0.0002**
Proteinuria, g l^{-1}	313	272	0.988	0.968–1.007	0.2152
Hypercalcemia, mmol l^{-1}	347	301	1.885	0.755–4.707	0.1748
CA 19-9, UI (log-value)	411	359	1.171	1.057–1.296	**0.0025**
Radiological parameters					
Tumour size, mm	422	362	1.008	1.002–1.015	**0.0081**

Abbreviations: CI = confidence interval; HR = hazard ratio; PS = performance status; WHO = World Health Organization. Bold values refer to *P*-values <0.1 highlighting parameters involved in the final multivariate construction as described in the statistical section.

Table 3. Multivariate analysis of factors associated with overall survival (*N* = 358)

	Number of patients	Number of death	HR	95% CI	P	Internal validation BCA HR 95%	Max scores in nomogram
Age at diagnosis, years	358	307	1.013	1.000–1.025	**0.0478**	1.000–1.025	43
Pain							
No	155	130	1	–	–		
Yes	203	177	1.292	1.023–1.632	**0.0317**	1.046–1.657	16
Albumin, g l^{-1}	358	307	0.963	0.942–0.984	**0.0006**	0.942–0.988	96
Tumour size, mm	358	307	1.008	1.001–1.015	**0.0214**	1.000–1.016	100
CA 19-9, UI (log-value)	358	307	1.172	1.047–1.311	**0.0056**	1.039–1.306	80

Abbreviations: BCA = accelerated bootstrap confidence interval; CI = confidence interval; HR = hazard ratio. Bold values reflect that *P*-values are all significant <0.05.

and a good calibration at 6, 12, 24 and 48 months ($P = 0.99$, $P = 0.55$, $P = 0.30$, and $P = 0.08$, respectively). The calibration plots showed an optimal agreement between model prediction and actual observation for predicting OS probability at 6, 12, 24, and 48 months (Figure 1).

In the internal validation, uncertainties around HR measured with the bootstrapping procedure (Table 3) reflected the robustness of the final model.

Sensitivity analyses. In the sensitivity analysis, the performance of the final model was investigated by forcing the allocated treatment (gemcitabine *vs* gemcitabine plus erlotinib). Associations remained unchanged (Supplementary Appendix Table A1A).

After adjustment of the final model for histological grading and systolic blood pressure, main associations of the final multivariate model remained unchanged except for CA 19-9 (Supplementary Appendix Table A1B).

The final multivariate model was based on 358 (81.0%) patients (complete-subject analysis) for whom the five parameters were available. A multiple imputation analysis based on 1000 imputed data sets for all the variables provided similar results to the complete-subject analysis in term of significance for the association (*P*-value) and for coefficient estimations (β and SE) (Supplementary Appendix Table A1C).

Prognostic nomogram for OS. Nomogram integrating all significant independent factors for OS was built (Figure 2). An illustration and details on the nomogram development are provided in Supplementary Appendix A3.

Prognostic score for death-risk stratification. A PROLAP (**PR**ognostic score for **O**verall survival in **LAP**C) score ranging from 0 to 5 based on the final model was calculated for each patient. All factors included in the score were weighted according to the β-parameter estimation issued from the final Cox model. The contribution of each individual factor was summed to generate a raw score that was finally applied to normalisation between 0 and 5. This score construction and equation are provided as Supplementary Appendix A3.

On the basis of an optimal cut point determined by Hothorn and Lausen method (Supplementary Appendix Figure A1) patients were categorised into two-risk groups with significantly different prognostic profiles (Figure 3A): a low-risk group ($N = 242$, score < 2.455) and a high-risk group ($N = 116$, score $\geqslant 2.455$, HR = 2.01 95% CI, 1.58–2.56; $P < 0.0001$). The median OS was 14.7 and 9.4 months for the low-risk group and the high-risk group, respectively ($P < 0.0001$).

Considering that the prognostic score assumes a normal distribution (mean = 2.27; s.d. = 0.39) for the three-risk groups approach, Cox demonstrated that the loss of information from grouping would be minimised by selecting cut points equal to 2.291 and 2.720. Such grouping achieved a clear separation of the Kaplan–Meier curves (Figure 3B). The median OS was 15.3, 11.7 and 8.5 months for the low, intermediate and the high-risk groups, respectively ($P < 0.0001$).

Results for classification of patient's into two- and three-risk groups related to the median and tertiles approach are provided in Supplementary Appendix Figure A2. These, however, were not selected as the main result given the low efficiency of this approach.

Baseline characteristics of patients according to the risk groups. For the two-risk group approach, in the high-risk group, 75% of patients presented with pain, were older, with moderate albumin level and had tumour of large size. In contrast, in the low-risk group, 48% of patients reported pain, were younger, had normal albumin levels and had tumour of small size (Supplementary Appendix Table A1A). A similar trend across groups was observed in the three-risk groups approach (Supplementary Appendix Table A1B).

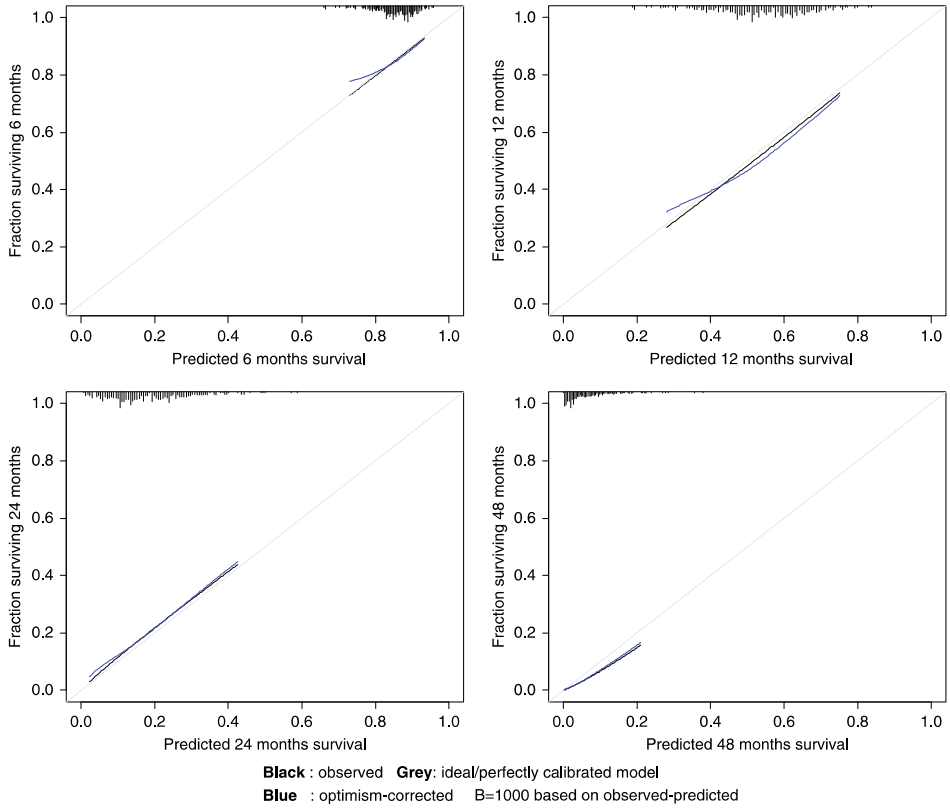


Figure 1. Calibration plots at 6, 12, 24 and 48 months for the final multivariate model. Vertical axis is the observed proportion of patients surviving at time of interest.

Agreement between nomogram prediction and the PROLAP score. Given that the prognostic nomogram and score were built from the same multivariate final model parameters, they can be connected. Characterisation of risk groups with the total number of points derived from the nomogram is provided in the Kaplan–Meir curves (Figure 3A and B).

Predictive value of treatments on OS according to the risk groups. In the two- and three-risk group approaches, there were no differential effects on OS of the gemcitabine-erlotinib combination across groups. In addition, no significant OS difference was observed in favour of CRT over CT across groups.

Interestingly, a total of 121 (68%) low-risk, 85 (62%) intermediate-risk and 19 (44%) high-risk patients reached the second randomisation ($P = 0.0027$). A similar significant trend was also observed in the two-risk group analysis. Detailed results are provided in Supplementary Appendix A4.

Validation of the discriminatory ability of the PROLAP score. A total of 63 (59%) of patients from the external validation cohort had all the information required to be eligible for the score calculation using five baseline parameters (pain, CA 19-9, albumin, tumour size and age). Similar characteristics at diagnosis and outcomes were observed between eligible and non-eligible patients (Supplementary Appendix Table A2).

Overall, the final multivariate model identified in the development set exhibited a better discrimination ability in the validation cohort (C-index 0.68; 95% bootstrap percentile CI, 0.59–0.77).

In the two-risk group approach, the discrimination ability of the risk score algorithm developed in the main analysis was externally confirmed, with median OS estimated to 17.1 months for the low-risk group and 9.2 months for the high-risk group (HR = 2.3; 95% CI, 1.3 to 4.2; log-rank $P = 0.003$; Figure 3C) and with a better C-index estimation (0.61; 95% CI 0.54 to 0.68) than in the development cohort.

Similarly in the three-risk group approach, the discrimination ability of the risk score was externally confirmed, with median OS estimated to 18.3 months for the low-risk group, 14.1 months for the intermediate-risk group and 7.6 months for the high-risk group (log-rank $P < 0.0001$; Figure 3D) and with a better C-index estimation (0.68; 95% CI 0.59–0.76) than in the development cohort.

DISCUSSION

Currently, there is no efficient prognostic staging system that could help to estimate OS at diagnosis in LAPC patients. On the basis of the largest phase III clinical trial of LAPC, we established a novel easy-to-use survival prediction model built on five key parameters (age, albumin, tumour size, CA 19-9 and pain) and provided a prognostic nomogram and score.

Nomograms are an important component of modern medical decision-making (Balachandran *et al*, 2015). The main goal of such staging systems is to provide the most accurate predictions for specified end points. In this way, parameters to consider should be easily available and measurable. The proposed nomogram fulfils these requirements and produces individual survival estimations in LAPC patients. The baseline prognostic score can be considered as a complementary tool for clinical practice by producing a classification

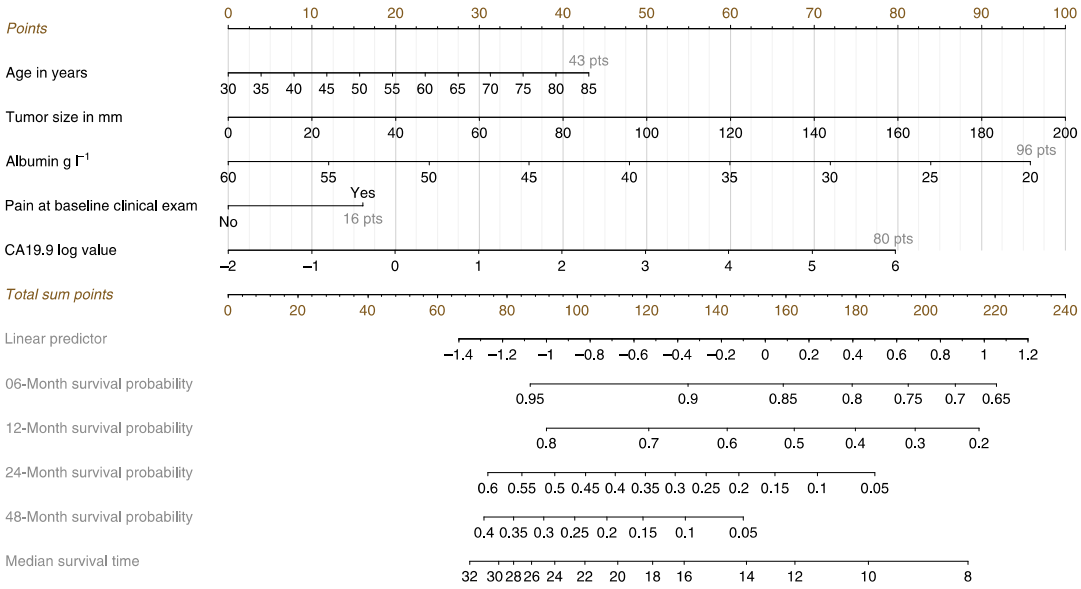


Figure 2. Prognostic nomogram to predict individual overall survival probability in patients with locally advanced pancreatic cancer. Points are assigned for each risk factor by drawing a line upward from the corresponding values to the 'point' line. The total sum of points for four risk factors is plotted on the 'total points' line. A line is drawn down to read the corresponding predictions of 6-, 12-, 24- and 48-month-survival probability and median survival time. For example, a patient with the following characteristics–63 years old (26 points), tumour measurement equal to 105 mm (53 points), albumin at 40.9 g l^{-1} (46 points), CA 19-9 at 89 (log-value = 1.95; 39 points) and the presence of pain at baseline clinical exam (16 point) will have a total points of 180, which corresponds to 6-, 12-, 24- and 48-month-survival probabilities of 80% (95% CI 71–86), 40% (95% CI 25–55), 7% (95% CI 2–17) and 1% (95% CI 0–4) and a predicted median survival time of 10.5 months (95% CI 8.75–13.5). An online web-based application smartphone compatible was developed that provides individualised survival estimates from the nomogram: http://www.umqvc.org/en/tool/prolap.html.

of the patients into well-identified death-risk populations. We have developed an open-access web site smartphone compatible to automatically compute the prognostic nomogram and score estimations. (http://www.umqvc.org/en/tool/prolap.html).

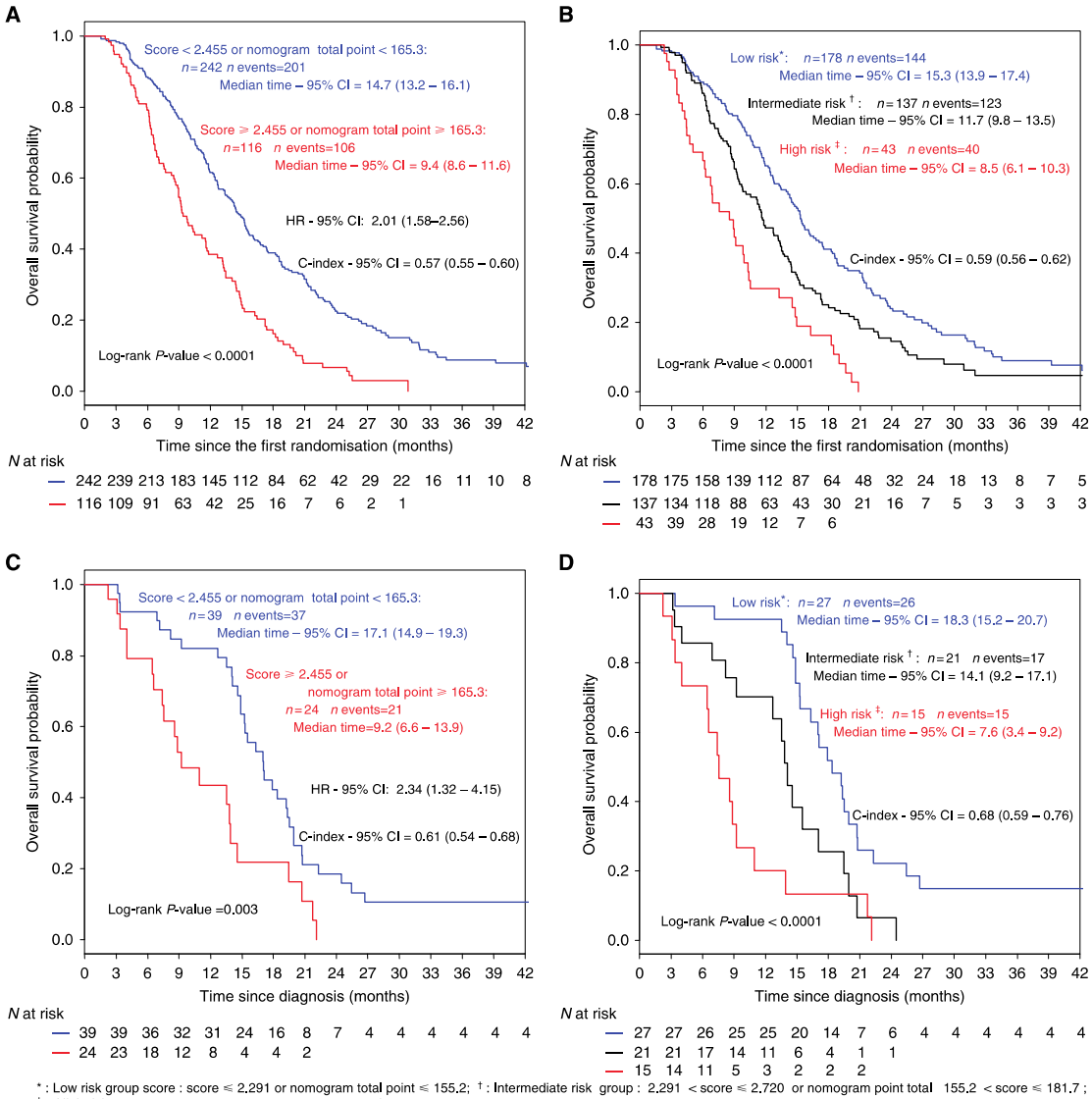
The present work revealed a considerable heterogeneity among LAPC patients regarding their OS-risk profiles. The recognition of risk groups with different prognostic outcomes may offer the opportunity to define risk-adapted strategies for LAPC management in the future. Indeed, while our analysis was based on OS, but not on progression-free survival, we accurately identified patients more likely not to be eligible for the second randomisation. Overall, 32% of the low-risk, 38% of the intermediate-risk and 56% of the high-risk patients failed to reach the second randomisation due to progression.

Our study supports the role of the nomogram and/or the PROLAP score as tools to be used both for patient selection and/or for patients' stratification in future clinical trials. These may help to reduce heterogeneity among treatment arms in terms of risk profile and thus better estimate the treatment effect. Indeed, if patients from the risk groups identified are distributed unequally among arms this difference can lead to a distortion in the outcome analysis. It seems therefore reasonable to recommend pre-planned subgroup analyses according to the level of such a score in future trials.

It is worth noting that two of the five factors included in our model, pain and albumin are objective surrogates for measurement of HRQoL (Burris *et al*, 1997). The importance of these parameters in the model and the recognition of HRQoL as an additional prognostic value for OS in metastatic disease (Gourgou-Bourgade *et al*, 2013) and in other tumour types (Diouf *et al*, 2014; Fiteni *et al*, 2016; Paquette *et al*, 2016) suggest that HRQoL needs to be evaluated in the future prognostic models for LAPC.

Our study has several strengths. The score was constructed in the framework of a large international multicentre trial with a broad spectrum of parameters available at diagnosis. Moreover, discrimination, calibration and internal validation underlined satisfactory performance and validity of the model. We built our model in a rigorous methodological framework respecting a recent recommended checklist for nomograms (Balachandran *et al*, 2015) and providing a transparent reporting of the multivariate model as suggested in the TRIPOD statement (Collins *et al*, 2015). The internal validation ensured a reliable performance for patients similar to those of the development cohort. Our prognostic score was fully externally replicated in the 'real-world' patients encountered in daily clinical practice: different centres; clinicians; treatments; and periods. This external validation population is quite different to the development set and underlines the robustness of our results in the clinical setting. A validation in a larger cohort of patients will be performed with the International Pancreas Database Program ARCAD metabase, which is currently under development. Finally, the five parameters identified in our study reflect the general condition of the patient and disease activity and therefore provide clinically relevant information in LAPC. These factors are admitted prognostic factors in LAPC patients by clinicians and accordingly enhance the relevance of the tools developed.

The present study has also several limitations. The role of pain status could be further refined by its assessment at diagnosis before administration of analgesic or by documenting whether the patient was receiving opioids or not at the time of assessment. In case of baseline cholestasis, the CA 19-9 level should be better estimated after biliary drainage whenever possible in order to minimise overestimation. Despite the broad spectrum of parameters analysed and the acceptable performance of the score, we recognise



* : Low risk group score : score ⩽ 2.291 or nomogram total point ⩽ 155.2; † : Intermediate risk group : 2.291 < score ⩽ 2.720 or nomogram point total 155.2 < score ⩽ 181.7 ;
‡ : High risk group : score > 2.720 or nomogram total point > 181.7

Figure 3. Kaplan–Meier curves of overall survival for (**A**, **B**) the development set cohort and (**C**, **D**) the external validation set cohort according to prognostic score group for the Hothorn & Lausen optimal cut point (**A**, **C**) and Cox's methods (**B**, **D**) for the two- and three-risk groups approach, respectively. An online web-based smartphone compatible application was developed that provides risk-group classification for OS from the PROLAP score: http://www.umqvc.org/en/tool/prolap.html.

that additional potentially relevant variables such as a neutrophil to lymphocyte ratio could not be assessed and merit future consideration (Proctor *et al*, 2012; Stotz *et al*, 2013; Goldstein *et al*, 2015) given that novel biologic insights have recently been validated through the pancreatic genome project (Bailey *et al*, 2016).

In conclusion, we present here a novel OS prediction model based on five key independent prognostic factors at the time of diagnosis. This allowed us to develop and externally validate the first prognostic staging system for OS in LAPC and then to provide functional tools at individual (nomogram) and population (score) levels: http://www.umqvc.org/en/tool/prolap.html. In this study, we revealed considerable heterogeneity among LAPC patients regarding their OS-risk profiles with the recognition of clearly different risk groups. The prognostic nomogram and the score proposed may help to optimise clinical trials design and might offer the opportunity to define risk-adapted strategies for LAPC management in the future.

ACKNOWLEDGEMENTS

We thank Florian Limousin (graphic designer and independent web developer in Pearlweb society) for the web development application support. The LAP07 trial was supported in part by a grant from Roche. Data from the INTERNATIONAL INTERGROUP STUDY LAP07 (France: GERCOR, FFCD & FNCLCC (PRODIGE)-SFRO; Belgium: BGDO; Australia: AGITG; Sweden: Nordic Group).

CONFLICT OF INTEREST

PA and TA have received honoraria and consulting fee from Roche. PH has received consulting fee from Roche. FB has received honoraria and research funding from Roche. The remaining authors declare no conflict of interest.

REFERENCES

Bailey P, Chang DK, Nones K, Johns AL, Patch A-M, Gingras M-C, Miller DK, Christ AN, Bruxner TJC, Quinn MC, Nourse C, Murtaugh LC, Harliwong I, Idrisoglu S, Manning S, Nourbakhsh E, Wani S, Fink L, Holmes O, Chin V, Anderson MJ, Kazakoff S, Leonard C, Newell F, Waddell N, Wood S, Xu Q, Wilson PJ, Cloonan N, Kassahn KS, Taylor D, Quek K, Robertson A, Pantano L, Mincarelli L, Sanchez LN, Evers L, Wu J, Pinese M, Cowley MJ, Jones MD, Colvin EK, Nagrial AM, Humphrey ES, Chantrill LA, Mawson A, Humphris J, Chou A, Pajic M, Scarlett CJ, Pinho AV, Giry-Laterriere M, Rooman I, Samra JS, Kench JG, Lovell JA, Merrett ND, Toon CW, Epari K, Nguyen NQ, Barbour A, Zeps N, Moran-Jones K, Jamieson NB, Graham JS, Duthie F, Oien K, Hair J, Grützmann R, Maitra A, Iacobuzio-Donahue CA, Wolfgang CL, Morgan RA, Lawlor RT, Corbo V, Bassi C, Rusev B, Capelli P, Salvia R, Tortora G, Mukhopadhyay D, Petersen GM. Australian Pancreatic Cancer Genome InitiativeMunzy D, Fisher WE, Karim SA, Eshleman JR, Hruban RH, Pilarsky C, Morton JP, Sansom OJ, Scarpa A, Musgrove EA, Bailey U-MH, Hofmann O, Sutherland RL, Wheeler DA, Gill AJ, Gibbs RA, Pearson JV, Waddell N, Biankin AV, Grimmond SM (2016) Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 531: 47–52.

Balachandran VP, Gonen M, Smith JJ, DeMatteo RP (2015) Nomograms in oncology: more than meets the eye. *Lancet Oncol* **16**: e173–e180.

Bonnetain F, Bonsing B, Conroy T, Dousseau A, Glimelius B, Haustermans K, Lacaine F, Van Laethem JL, Aparicio T, Aust D, Bassi C, Berger V, Chamorey E, Chibaudel B, Dahan L, De Gramont A, Delpero JR, Dervenis C, Ducreux M, Gal J, Gerber E, Ghaneh P, Hammel P, Hendlisz A, Jooste V, Labianca R, Latouche A, Lutz M, Macarulla T, Malka D, Mauer M, Mitry E, Neoptolemos J, Pessaux P, Sauvanet A, Tabernero J, Taieb J, van Tienhoven G, Gourgou-Bourgade S, Bellera C, Mathoulin-Pélissier S, Collette L (2014) Guidelines for time-to-event end-point definitions in trials for pancreatic cancer. Results of the DATECAN initiative (Definition for the Assessment of Time-to-event End-points in CANcer trials). *Eur J Cancer* **50**: 2983–2993.

Burris HA, Moore MJ, Andersen J, Green MR, Rothenberg ML, Modiano MR, Cripps MC, Portenoy RK, Storniolo AM, Tarassoff P, Nelson R, Dorr FA, Stephens CD, Von Hoff DD (1997) Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial. *J Clin Oncol* **15**: 2403–2413.

Chauffert B, Mornex F, Bonnetain F, Rougier P, Mariette C, Bouché O, Bosset JF, Aparicio T, Mineur L, Azzedine A, Hammel P, Butel J, Stremsdoerfer N, Maingon P, Bedenne L (2008) Phase III trial comparing intensive induction chemoradiotherapy (60Gy, infusional 5-FU and intermittent cisplatin) followed by maintenance gemcitabine with gemcitabine alone for locally advanced unresectable pancreatic cancer. Definitive results of the 2000-01 FFCD/SFRO study. *Ann Oncol* **19**: 1592–1599.

Collins GS, Reitsma JB, Altman DG, Moons KGM (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Cancer* **112**: 251–259.

Cox DR (1957) Note on Grouping. *J Am Stat Assoc* **52**: 543–547.

Diouf M, Chibaudel B, Filleron T, Tournigand C, Hug de Larauze M, Garcia-Larnicol M-L, Dumont S, Louvet C, Perez-Staub N, Hadengue A, de Gramont A, Bonnetain F (2014) Could baseline health-related quality of life (QoL) predict overall survival in metastatic colorectal cancer? The results of the GERCOR OPTIMOX 1 study. *Health Qual Life Outcomes* **12**: 69.

Fiteni F, Vernerey D, Bonnetain F, Vaylet F, Sennélart H, Trédaniel J, Moro-Sibilot D, Herman D, Laizé H, Masson P, Derollez M, Clément-Duchêne C, Milleron B, Morin F, Zalcman G, Quoix E, Westeel V (2016) Prognostic value of health-related quality of life for overall survival in elderly non-small-cell lung cancer patients. *Eur J Cancer Oxf Engl 1990* **52**: 120–128.

Goldstein D, El-Maraghi RH, Hammel P, Heinemann V, Kunzmann V, Sastre J, Scheithauer W, Siena S, Tabernero J, Teixeira L, Tortora G, Van Laethem J-L, Young R, Penenberg DN, Lu B, Romano A, Von Hoff DD (2015) nab-Paclitaxel plus gemcitabine for metastatic pancreatic cancer: long-term survival from a phase III trial. *J Natl Cancer Inst*, e-pub ahead of print 31 Jan 2015; doi:10.1093/jnci/dju413.

Gourgou-Bourgade S, Bascoul-Mollevi C, Desseigne F, Ychou M, Bouché O, Guimbaud R, Bécouarn Y, Adenis A, Raoul J-L, Boige V, Bérille J, Conroy T (2013) Impact of FOLFIRINOX compared with gemcitabine on quality of life in patients with metastatic pancreatic cancer: results from the PRODIGE 4/ACCORD 11 randomized trial. *J Clin Oncol* **31**: 23–29.

Hajj C, Goodman KA (2015) Role of radiotherapy and newer techniques in the treatment of GI cancers. *J Clin Oncol* **33**: 1737–1744.

Hammel P, Huguet F, van Laethem J-L, Goldstein D, Glimelius B, Artru P, Borbath I, Bouché O, Shannon J, André T, Mineur L, Chibaudel B, Bonnetain F, Louvet C (2016) Effect of chemoradiotherapy vs chemotherapy on survival in patients with locally advanced pancreatic cancer controlled after 4 months of gemcitabine with or without erlotinib: the LAP07 Randomized Clinical Trial. *JAMA* **315**: 1844–1853.

Harrell Jr FE, Lee KL, Mark DB (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* **15**: 361–387.

Hothorn T, Lausen B (2003) On the exact distribution of maximally selected rank statistics. *Comput Stat Data Anal* **43**: 121–137.

Huguet F, André T, Hammel P, Artru P, Balosso J, Selle F, Deniaud-Alexandre E, Ruszniewski P, Touboul E, Labianca R, de Gramont A, Louvet C (2007) Impact of chemoradiotherapy after disease control with chemotherapy in locally advanced pancreatic adenocarcinoma in GERCOR phase II and III studies. *J Clin Oncol* **25**: 326–331.

Krishnan S, Rana V, Janjan NA, Varadhachary GR, Abbruzzese JL, Das P, Delclos ME, Gould MS, Evans DB, Wolff RA, Crane CH (2007) Induction chemotherapy selects patients with locally advanced, unresectable pancreatic cancer for optimal benefit from consolidative chemoradiation therapy. *Cancer* **110**: 47–55.

Loehrer PJ, Feng Y, Cardenes H, Wagner L, Brell JM, Cella D, Flynn P, Ramanathan RK, Crane CH, Alberts SR, Benson AB (2011) Gemcitabine alone versus gemcitabine plus radiotherapy in patients with locally advanced pancreatic cancer: an Eastern Cooperative Oncology Group trial. *J Clin Oncol* **29**: 4105–4112.

Malvezzi M, Bertuccio P, Levi F, La Vecchia C, Negri E (2014) European cancer mortality predictions for the year 2014. *Ann Oncol* **25**(8): 1650–1656.

Paquette B, Vernerey D, Chauffert B, Dabakuyo S, Feuvret L, Taillandier L, Frappaz D, Taillia H, Schott R, Ducray F, Fabbro M, Tennevet I, Ghiringhelli F, Guillamo J-S, Durando X, Castera D, Frenay M, Campello C, Dalban C, Skrzypski J, Chinot O, Anota A, Bonnetain F (2016) Prognostic value of health-related quality of life for death risk stratification in patients with unresectable glioblastoma. *Cancer Med*. e-pub ahead of print 1 June 2016; doi:10.1002/cam4.734.

Philip PA (2011) Locally advanced pancreatic cancer: where should we go from here? *J Clin Oncol* **29**: 4066–4068.

Proctor MJ, McMillan DC, Morrison DS, Fletcher CD, Horgan PG, Clarke SJ (2012) A derived neutrophil to lymphocyte ratio predicts survival in patients with cancer. *Br J Cancer* **107**: 695–699.

Schemper M, Smith TL (1996) A note on quantifying follow-up in studies of failure time. *Control Clin Trials* **17**: 343–346.

Siegel RL, Miller KD, Jemal A (2016) Cancer statistics, 2016. *CA Cancer J Clin* **66**: 7–30.

Stotz M, Gerger A, Eisner F, Szkandera J, Loibner H, Ress AL, Kornprat P, AlZoughbi W, Zoughbi WA, Seggewies FS, Lackner C, Stojakovic T, Samonigg H, Hoefler G, Pichler M (2013) Increased neutrophil-lymphocyte ratio is a poor prognostic factor in patients with primary operable and inoperable pancreatic cancer. *Br J Cancer* **109**: 416–421.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)

IV.1.1.7 Manuscrit à destination des professionnels spécialisés en Oncologie dans la revue ONKO+

— ANALYSE D'ARTICLE —

Digestif

ESTIMATION DE LA SURVIE GLOBALE CHEZ DES PATIENTS ATTEINTS D'UN CANCER PANCRÉATIQUE LOCALEMENT AVANCÉ

Nomogramme et score pronostiques (PROLAP score)

Dewi Vernerey[1], Dr Florence Huguet[2], Angélique Vienot[3], Pr Pascal Hammel[4], Pr Franck Bonnetain[1]

Référence : *Vernerey D, Huguet F, Vienot A et al. Prognostic nomogram and score to predict overall survival in locally advanced untreated pancreatic cancer (PROLAP). Br J Cancer 2016 ; 115 : 281-9.*
*L'article d'*ONKO+ *est issu du travail de recherche qui a été publié dans le* British Journal of Cancer.

Résumé

La prise en charge des patients atteints d'un cancer pancréatique localement avancé (LAPC) demeure aujourd'hui complexe. Une meilleure estimation de la survie globale (SG) au moment du diagnostic avant le début du traitement est nécessaire pour mieux évaluer le pronostic des patients. L'analyse a porté sur une cohorte de 442 patients atteints d'un LAPC, inclus dans l'essai international de phase III **LAP07.** 106 patients traités pour un LAPC au centre hospitalier universitaire de Besançon ont constitué la cohorte de validation. Parmi les 30 paramètres collectés au diagnostic et analysés, cinq avaient une association significative et indépendante avec la SG dans l'analyse de Cox multivariée : l'âge, la douleur, l'albumine, la taille tumorale et le marqueur tumoral CA 19-9. Sur la base du modèle de Cox multivarié final, un nomogramme a été construit, permettant le calcul de probabilités individuelles de survie à 6, 12, 24 et 48 mois post-diagnostic de cancer. Un score a aussi permis d'isoler trois groupes de pronostics différents. Ces outils pourront optimiser la conception des futurs essais cliniques et offrent l'opportunité de définir une meilleure prise en charge basée sur le pronostic du patient.

Abstract

Prognostic nomogram and score to predict overall survival in locally advanced untreated pancreatic cancer (PROLAP)

The management of locally advanced pancreatic cancer (LAPC) patients remains complex. Better discrimination for overall survival (OS) at diagnosis before initiation of induction chemotherapy in LAPC untreated patients is needed. The development sample and analyses were derived from 442 LAPC patients enrolled in the LAP07 international phase III trial. An external validation cohort of 106 consecutive LAPC patients treated in Besançon Hospital, France, was also considered. Among 30 baseline parameters evaluated at cancer diagnosis, five were found to be independent prognostic factors for OS in the Cox multivariate regression analysis: age, pain, tumour size, albumin and CA 19-9. On the basis of the final model, a prognostic nomogram was developed to predict individual OS probabilities at 6, 12, 24, and 48 months post-diagnosis. The score has the potential to delineate up to 3 different prognosis groups with clearcut OS pronostic profile. The nomogram and score can accurately predict OS before initiation of induction chemotherapy in LAPC untreated patients. They may help to optimise clinical trials design and might offer the opportunity to define risk-adapted strategies for LAPC management in the future.

ESTIMATION DE LA SURVIE GLOBALE CHEZ DES PATIENTS ATTEINTS D'UN CANCER PANCRÉATIQUE LOCALEMENT AVANCÉ

INTRODUCTION

Le cancer du pancréas est la quatrième cause de décès par cancer dans les pays industrialisés et est en mesure de devenir la deuxième cause en 2030 si les tendances actuelles se confirment. Pour illustration, le nombre de décès par cancer du pancréas en 2014 dans l'Union européenne a été estimé à 82 300, ce qui est quasi similaire aux décès attribuables au cancer du sein chez les femmes (89 300) (1). Le pronostic du cancer du pancréas est très sombre avec un taux de survie à 5 ans de l'ordre de 5-7 %. Il est donc nécessaire d'encourager la prévention ainsi que la recherche dans cette pathologie.

L'absence de symptômes spécifiques au début de la maladie rend le diagnostic souvent tardif. Ainsi, la majorité des cancers du pancréas (environ 80 %) sont non résécables en raison d'un envahissement vasculaire, de la présence de métastases locorégionales ou à distance. Au diagnostic, environ 35 % des patients présentent un envahissement vasculaire sans métastases, caractérisant ainsi un cancer du pancréas localement avancé (LAPC).

La prise en charge des patients atteints d'un LAPC demeure aujourd'hui complexe et controversée (2). L'utilisation de la chimioradiothérapie notamment a été dans un premier temps encouragée sur la base d'études prospectives anciennes (3). Puis, dans un second temps, elle a été freinée en raison de résultats équivalents à la chimiothérapie seule pour la survie globale (SG) dans l'essai de phase III internationale multicentrique **LAP07** (4). Cependant, l'amélioration du contrôle local ainsi qu'une tendance favorable sur la survie sans progression amènent certains cliniciens à considérer la chimioradiothérapie comme une option thérapeutique restant potentiellement intéressante pour les patients LAPC. Il apparaît surtout que le sous-groupe de patients pour lequel la radiothérapie aurait une action bénéfique reste à être identifié plus précisément. Par ailleurs, l'absence actuelle de consensus pour la définition de profils de risque spécifiques pour la SG chez les patients atteints de LAPC peut induire une confusion dans le développement de stratégies de recherche et également dans la prise en charge des patients.

Dans ce contexte, et avec l'intérêt actuel grandissant pour une médecine de précision, il est donc important de développer des outils pronostiques modernes pour optimiser la sélection des patients atteints de LAPC afin de leur proposer des thérapies appropriées. Plus spécifiquement, une meilleure prédiction de la SG au diagnostic peut aussi favoriser, dans les futurs essais cliniques, la construction de bras de traitement mieux équilibrés en termes de pronostic et ainsi permettre de mieux mesurer l'effet du traitement testé dans l'essai. Une recherche pronostique actualisée chez les patients avec un LAPC permettrait aussi d'améliorer le design de futurs essais cliniques.

Les nomogrammes sont des outils de prédiction individuelle qui ont connu un développement accru en oncologie et sont proposés comme une alternative intéressante à la traditionnelle classification TNM (5). Cependant, aucun nomogramme pour prédire la SG n'est disponible et validé pour les LAPC.

Cette étude a pour objet, dans un premier temps, de proposer et de valider un nouveau modèle pronostique pour la SG dans le LAPC en étudiant un spectre important de paramètres disponibles au diagnostic. Dans un second temps, sur la base de ce nouveau modèle pronostique des outils de prédiction individuelle (nomogramme) et populationnelle (score) sont développés et automatisés pour simplifier leur utilisation par les chercheurs et cliniciens.

POPULATION DE L'ÉTUDE

La cohorte de développement utilisée dans cette étude est constituée des patients inclus dans l'essai de phase III randomisé international et multicentrique **LAP07** (ClinicalTrials.gov NCT00634725), qui est à ce jour la cohorte de patients ayant un LAPC la plus importante jamais étudiée dans un essai clinique.

La cohorte de validation externe inclut tous les patients LAPC traités au centre hospitalier universitaire de Besançon sur une période de 11 ans. Cette cohorte de validation est particulièrement intéressante puisqu'elle est constituée de patients issus de la pratique clinique régulière et donc à ce titre elle n'est pas sélectionnée.

Les données démographiques, clinicopathologiques, biologiques et radiologiques ont été collectées au moment du diagnostic de cancer dans la cohorte de validation et au moment de l'inclusion dans l'essai pour la cohorte de développement.

[1]Unité de méthodologie et de qualité de vie en cancérologie, Centre hospitalier universitaire de Besançon
[2]Service d'oncologie radiothérapie, hôpital Tenon, hôpitaux universitaire Est Parisien, Université Pierre et Marie Curie, Paris
[3]Service de gastroentérologie, centre hospitalier universitaire de Besançon
[4]Service d'oncologie digestive, hôpital Beaujon, Clichy

APPROCHE MÉTHODOLOGIQUE

Nous avons construit le modèle multivarié d'intérêt dans un cadre méthodologique rigoureux, comme recommandé dans deux importantes et récentes publications concernant la construction de modèles pronostiques (5, 6).

RÉSULTATS

442 patients traités dans l'essai thérapeutique LAP07 entre février 2008 et décembre 2011 ont été inclus dans la cohorte de développement. 106 patients pris en charge consécutivement au centre hospitalier universitaire de Besançon pour un LAPC ont quant à eux été inclus dans la cohorte de validation externe.
Les caractéristiques de ces deux populations sont présentées dans le *tableau 1*.

Parmi les 30 paramètres collectés au moment du diagnostic et analysés, seulement huit étaient significativement associés à la SG dans l'analyse de Cox univariée. Ces paramètres sont l'âge, le grade histologique, le performance status, la pression artérielle systolique, la douleur, l'albumine, la taille tumorale et le marqueur tumoral CA 19-9.
Le modèle multivarié final de Cox se compose de cinq facteurs indépendamment associés à la SG : l'âge, la douleur, l'albumine, la taille tumorale et le marqueur tumoral CA 19-9 *(Tab. 2)*.
Sur la base du modèle multivarié final, un nomogramme a été construit, permettant pour un patient donné de calculer sa médiane de survie et ses probabilités individuelles de survie à 6, 12, 24 et 48 mois post-diagnostic de cancer *(Fig. 1)*.

Les points sont assignés pour chaque facteur de risque en traçant une ligne verticale vers la ligne correspondant aux points. Le total des points pour les cinq facteurs est ensuite tracé sur la ligne qui correspond à la somme totale des points. Une ligne verticale vers le bas est alors tracée pour déduire les prédictions correspondantes aux temps d'intérêt ainsi que la médiane de survie du patient.
Pour les 358 patients impliqués dans le développement du modèle multivarié, il a été possible d'obtenir un nombre total de points issu du nomogramme. Ce vecteur de 358 estimations a constitué notre score pronostique. En appliquant la méthode du seuil optimal de Hothorn et Lausen, la valeur seuil de 165,3 a permis de construire deux groupes de pronostics différents *(Fig. 2A)*, tandis que la méthode de Cox nous a fourni les seuils (155,2 et 181,7) pour isoler trois groupes de pronostics différents *(Fig. 2B)*.
Parmi les 106 patients de la cohorte de validation externe, le score a été applicable pour les 63 patients qui possédaient l'ensemble des paramètres nécessaires à son calcul *(Fig. 2C et D)*.
Globalement, le score qui a été développé dans une population de patients provenant d'un essai thérapeutique (LAP07) conserve d'excellentes propriétés dans une population de patients non sélectionnée représentative de la pratique clinique courante.
Enfin, une application en ligne sur le web (compatible avec les téléphones modernes) a été développée pour fournir des estimations de survie individualisées (nomogramme) et en population (score) pour les cliniciens et chercheurs. L'adresse de cette application est la suivante : www.umqvc.org/en/tool/prolap.html.
Cette étape s'est avérée primordiale afin de rendre fonctionnels ces outils. En effet, l'utilisation des variables en continu, ainsi que leur transformation pour garder le plus d'information possible et modéliser correctement le lien entre ces variables et la distribution de survie, rend l'utilisation de ces outils impossible sans automatisation des calculs.
La page d'accueil est une page d'information pour situer l'outil dans son contexte. Ensuite, les données du modèle multivarié propres au patient doivent être saisies. Les résultats issus du nomogramme et du score sont immédiatement visualisables sur la page de résultat *(Fig. 3)*.
Enfin, une extraction de ces résultats peut être demandée pour insérer les estimations dans un dossier patient par exemple.

DISCUSSION

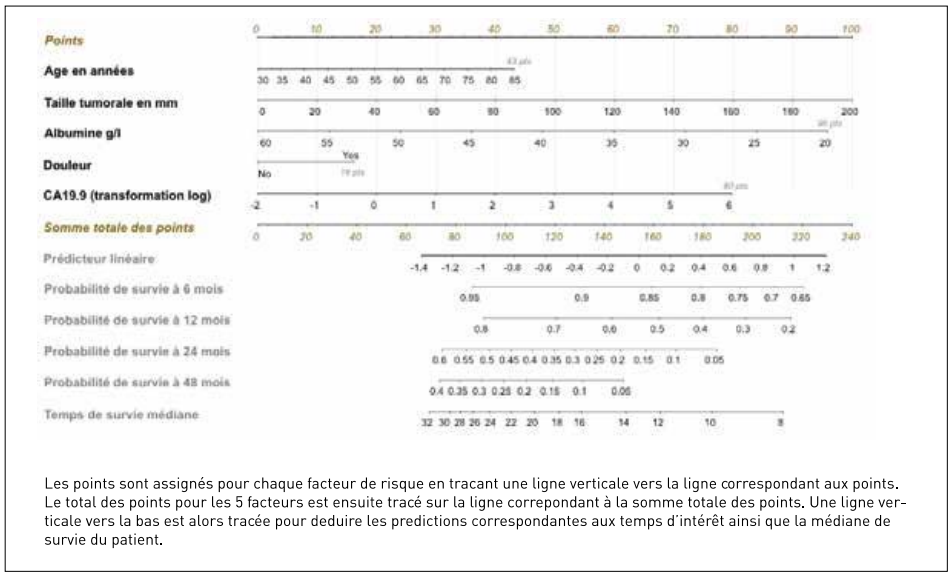
Le modèle multivarié final de Cox sur lequel a reposé la construction de ces outils se compose de cinq facteurs indépendamment associés à la SG : l'âge, la douleur, le taux d'albumine, la taille tumorale et le marqueur tumoral CA 19-9. Ces paramètres sont des facteurs aujourd'hui bien reconnus dans cette pathologie et illustrent le bon sens clinique des outils proposés.
Nous avons pu mettre en évidence l'existence d'une importante hétérogénéité de pronostic de survie au sein de cette population de patients qui doit être prise en compte dans le futur. En effet, les développements issus de cette étude vont permettre d'optimiser la conception des essais cliniques à deux niveaux. Le premier consiste en la

TABLEAU 1 - CARACTÉRISTIQUES DES PATIENTS AU MOMENT DU DIAGNOSTIC DANS LES COHORTES DE DÉVELOPPEMENT ET DE VALIDATION

		Cohorte de développement – Essai LAP07 (N = 442)	Cohorte de validation externe – Centre Hospitalier Universitaire de Besançon (N = 106)	P
Age, années*		63,5 ± 9,6	67,1 ± 10,2	0,0007
Genre, N (%)	*Homme*	228 (52 %)	57 (54 %)	
	Femme	214 (48 %)	49 (46 %)	0,7456
Localisation, N (%)	*Tête/Tête et Corps*	302 (69 %)	71 (70 %)	
	Autres (corps et/ou queue)	138 (31 %)	31 (30 %)	0,9059
	Manquant	2	4	
Grade histologique, N (%)	*Bien différencié*	107 (48 %)	9 (26 %)	
	Modérément différencié	75 (34 %)	17 (50 %)	
	Peu différencié	41 (18 %)	8 (24 %)	0,05556
	Manquant	219	72	
Invasion des ganglions lymphatiques, N (%)	*Oui*	179 (41 %)	-	
	Non	258 (59 %)	-	
	Manquant	5	-	
OMS-PS, N (%)	*0*	197 (46 %)	31 (30 %)	
	1	203 (47 %)	59 (56 %)	
	2	31 (7 %)	14 (13 %)	
	3	0 (0 %)	1 (1 %)	0,004892
	Manquant	11	1	
Taille tumorale (mm)*		44,0 ± 18,0	38,3 ± 14,7	0,00385
	Manquant	20	9	
Taille tumorale (mm), N (%)	*<30*	66 (16 %)	22 (23 %)	
	30-50	226 (54 %)	56 (58 %)	
	≥ 50	130 (31 %)	19 (20%)	0,04676
	Manquant	20	9	
Envahissement vasculaire, N (%)	*artère mésentérique supérieure*	144 (33 %)	-	
	tronc cœliaque	188 (43 %)	-	
	artère hépatique	25 (6 %)	-	
	Non	79 (18 %)	-	
	Manquant	6	-	
Médiane de suivi, mois [95% IC]		34,3 [27,6-43,7]	Tous les patients ont été suivis jusqu'au décès (temps maximum de suivi = 76,6) à l'exception de 5 patients toujours en vie avec une médiane de suivi de 10 mois	

TABLEAU 2 - ANALYSE DE COX MULTIVARIÉE DES FACTEURS ASSOCIÉS AVEC LA SG (N = 358).

	Nombre de patients	Nombre de décès	Rapports de risque	95 % IC	P
Âge au diagnostic (années)	358	307	1,013	1,000-1,025	0,0478
Douleur					
Non	155	130	1	-	-
Oui	203	177	1.292	1,023-1,632	0,0317
Albumine (g/l)	358	307	0.963	0,942-0,984	0,0006
Taille tumorale (mm)	358	307	1.008	1,001-1,015	0,0214
CA 19-9 (UI, valeur logarithmique)	358	307	1.172	1,047-1,311	0,0056



Les points sont assignés pour chaque facteur de risque en tracant une ligne verticale vers la ligne correspondant aux points. Le total des points pour les 5 facteurs est ensuite tracé sur la ligne correpondant à la somme totale des points. Une ligne verticale vers la bas est alors tracée pour deduire les predictions correspondantes aux temps d'intérêt ainsi que la médiane de survie du patient.

FIGURE 1 - Nomogramme pour prédire la survie globale individuelle pour des patients LAPC.

meilleure sélection des patients dans le cas de critères d'éligibilité basés sur le pronostic de survie et le second repose sur la bonne maîtrise de la répartition des patients de pronostics différents à travers les groupes lors du processus de randomisation, ce qui permettra in fine une meilleure estimation de l'effet traitement testé. Nous recommandons aussi des analyses en sous-groupes dans les essais futurs sur la base des différents groupes à risque identifiés. Par ailleurs, le nomogramme et le score vont peut-être offrir l'opportunité de définir, dans le futur, des stratégies de prise en charge adaptées aux différents profils de risque. Dans cette étude, nous n'avons pas réussi à conférer aux outils pronostiques développés une valeur prédictive de la réponse aux traitements, notamment celle de la chimioradiothérapie. En effet, la cohorte de développement analysée se compose de la population de patients inclus dans l'essai thérapeutique de phase III internationale LAP07, dont l'objectif principal était d'évaluer le bénéfice de la chimioradiothérapie comparée à la chimiothérapie seule en termes de SG chez des patients avec un LAPC contrôlé après une chimiothérapie

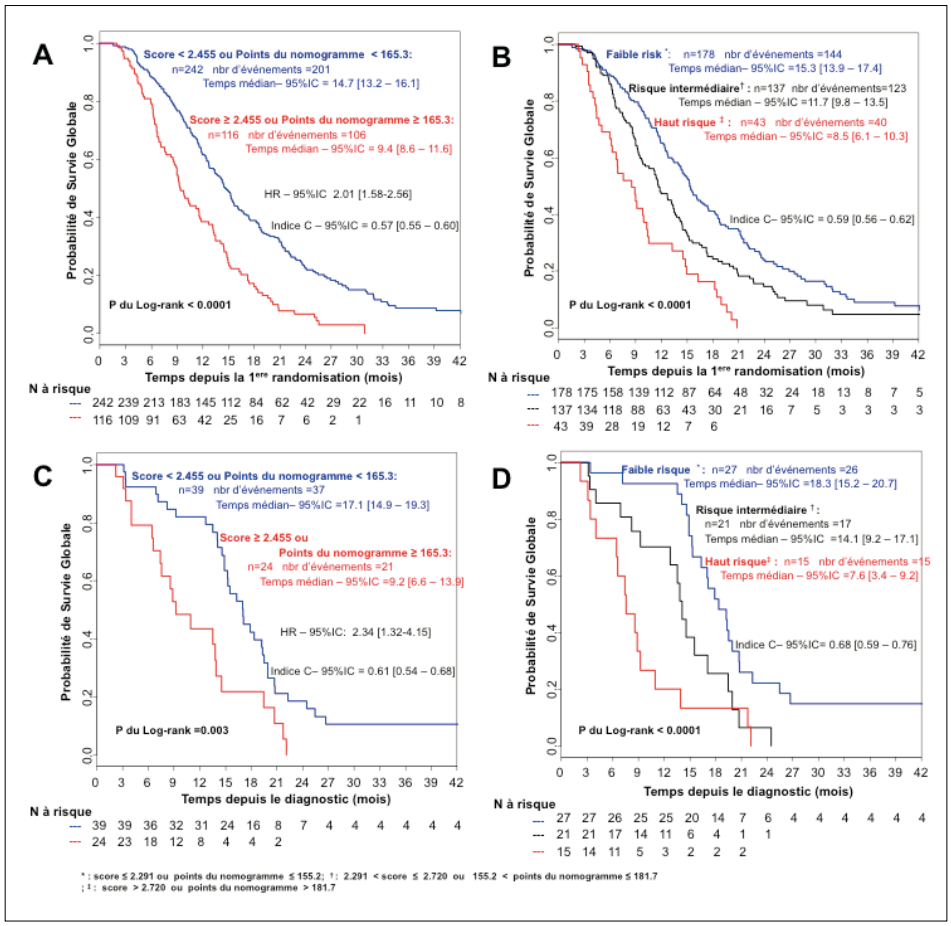


FIGURE 2 - Courbes de Kaplan-Meier de la survie globale pour la cohorte de développement (A, B) et la cohorte de validation externe (C, D) selon la catégorisation en deux groupes avec la méthode de Horton (A, C) et en trois groupes par la méthode de Cox (B, D).

d'induction à base de gemcitabine de 4 mois. Globalement, il n'y avait pas de bénéfice dans cet essai en termes de SG pour les patients traités par chimioradiothérapie, en comparaison avec ceux traités par chimiothérapie seule. Le rationnel de cet essai était basé sur plusieurs séries rétrospectives et études prospectives en faveur de la chimioradiothérapie pour des patients estimés de bon pronostic. Ces observations ont guidé la construction de l'essai LAP07 en proposant la randomisation chimiothérapie versus chimioradiothérapie uniquement pour les patients contrôlés après 4 mois de chimiothérapie d'induction, donc a priori de bon pronostic. Devant le résultat négatif observé dans l'essai, un des enjeux du travail proposé avec la construction du nomogramme et du score pronostique était de pouvoir identifier un effet différentiel de la chimioradiothérapie à travers les groupes de pronostics différents identifiés. Une tendance favorable a été suggérée dans le groupe de bon pronostic, mais n'a pas été trouvée comme statistiquement significative. Le modèle et les outils proposés constituent ainsi une base de travail pour le futur. En effet, le modèle final ne propose qu'un

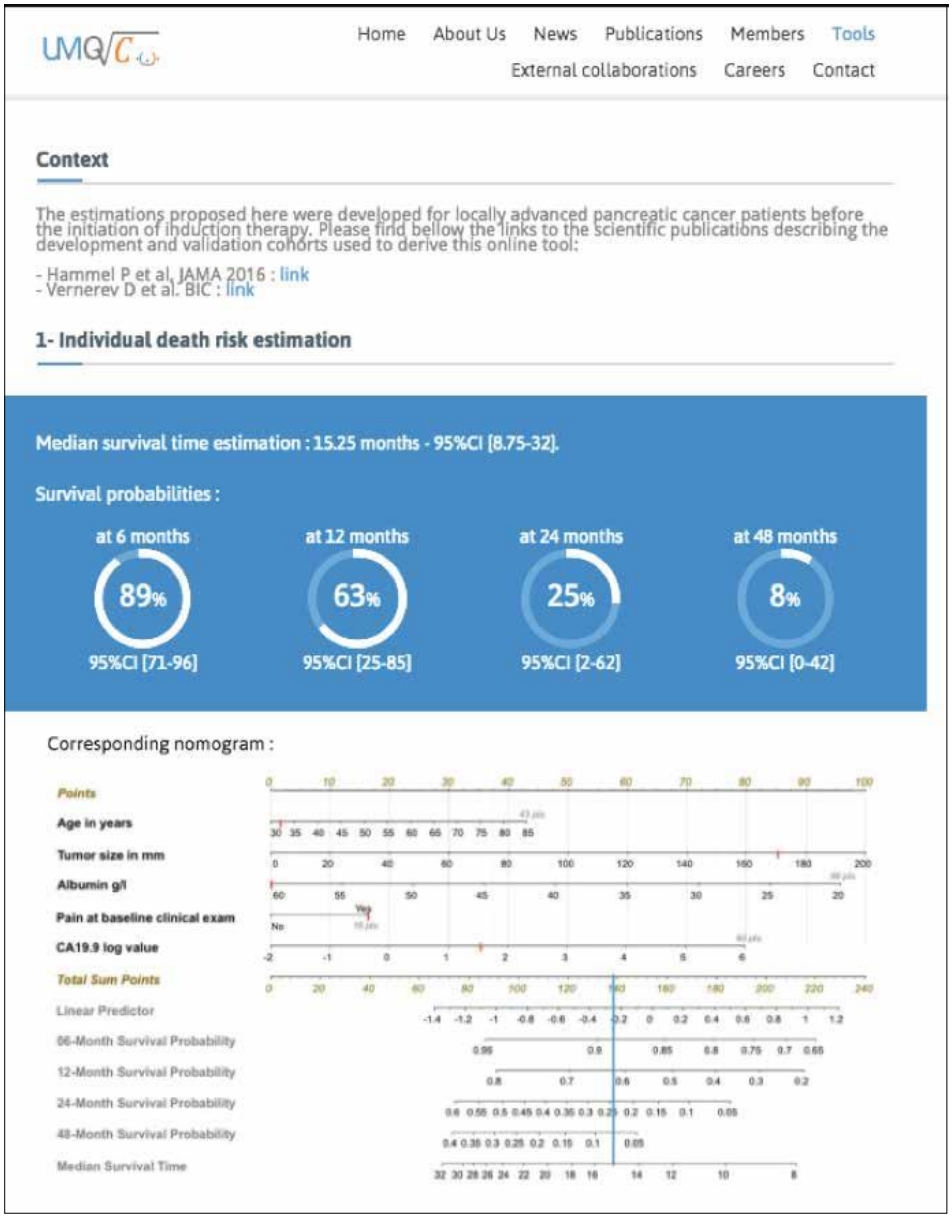


FIGURE 3 - Pages de résultats proposées par l'application en ligne.

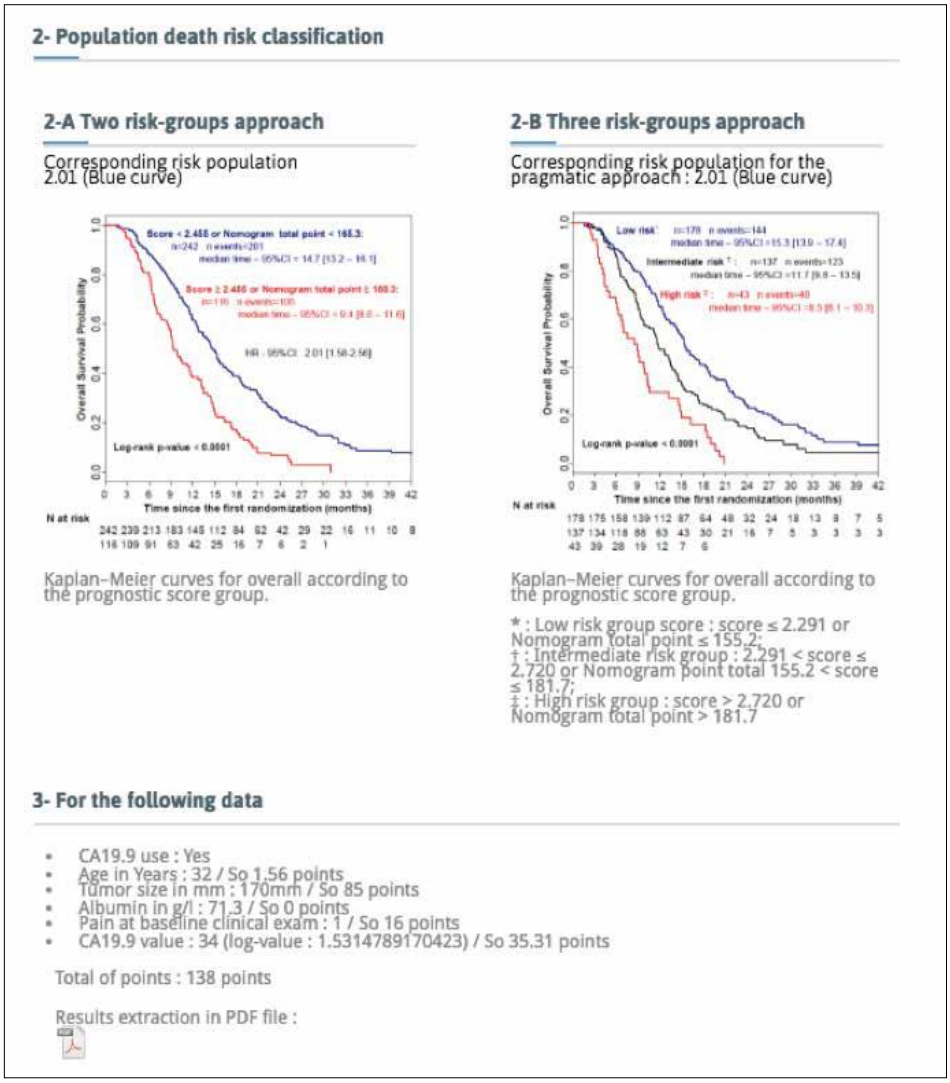


FIGURE 3 (SUITE) - Pages de résultats proposées par l'application en ligne.

indice de discrimination de 0,6, ce qui est acceptable, mais très largement améliorable. Des paramètres importants, par exemple le ratio neutrophile-lymphocyte (7, 8) et aussi la qualité de vie (9), devront dans le futur être intégrés dans le but d'améliorer le pouvoir discriminant du modèle et ainsi peut-être lui conférer un pouvoir prédictif de la réponse aux traitements.

CONCLUSION

Nous proposons donc dans cette étude un nouveau modèle de prédiction de la SG pour des patients atteints d'un LAPC au moment du diagnostic avant le traitement

d'induction. Nous avons, sur la base de ce nouveau modèle, développé des outils pronostiques fonctionnels à l'échelle individuelle (nomogramme) et populationnelle (score).
Au-delà des outils proposés, cette étude met en évidence une considérable hétérogénéité de pronostic chez les patients atteints de LAPC. Cette étude fournit des éléments qui permettront d'optimiser le design de futurs essais cliniques et offriront l'opportunité de définir des stratégies de prise en charge adaptées aux différents profils de risque. ■

✖ *L'auteur déclare ne pas avoir de lien d'intérêts.*

MOTS-CLÉS
Nomogramme et score pronostique, Survie globale, Cancer du pancréas localement avancé, Chimiothérapie, Chimioradiothérapie

KEYWORDS
Prognostic score, Nomogram, Overall survival, Locally advanced pancreatic cancer, Chemotherapy, Chemoradiotherapy

Bibliographie

1. Malvezzi M, Bertuccio P, Levi F et al. European cancer mortality predictions for the year 2014. Ann Oncol 2014 ; 25 : 1650-6.
2. Philip PA. Locally advanced pancreatic cancer: where should we go from here? J Clin Oncol 2011 ; 29 : 4066-8.
3. Huguet F, André T, Hammel P et al. Impact of Chemoradiotherapy After Disease Control With Chemotherapy in Locally Advanced Pancreatic Adenocarcinoma in GERCOR Phase II and III Studies. J Clin Oncol 2007 ;25 : 326-31.
4. Hammel P, Huguet F, van Laethem JL et al. Effect of Chemoradiotherapy vs Chemotherapy on Survival in Patients With Locally Advanced Pancreatic Cancer Controlled After 4 Months of Gemcitabine With or Without Erlotinib: The LAP07 Randomized Clinical Trial. JAMA 2016 ; 315 : 1844-53.
5. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. Lancet Oncol 2015 ; 16 : e173-180.
6. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Br J Cancer 2015 ; 112 : 251-9.
7. Proctor MJ, McMillan DC, Morrison DS et al. A derived neutrophil to lymphocyte ratio predicts survival in patients with cancer. Br J Cancer 2012 ; 107 : 695-9.
8. Goldstein D, El-Maraghi RH, Hammel P et al. Nab-Paclitaxel plus gemcitabine for metastatic pancreatic cancer: long-term survival from a phase III trial. J Natl Cancer Inst 2015 ; 107.
9. Gourgou-Bourgade S, Bascoul-Mollevi C, Desseigne F et al. Impact of FOLFIRINOX compared with gemcitabine on quality of life in patients with metastatic pancreatic cancer: results from the PRODIGE 4/ACCORD 11 randomized trial. J Clin Oncol 2013 ; 31 : 23-9.

IV.1.2 Etude de la valeur pronostique pour la survie globale du Syndecan-1 (CD138) au diagnostic chez des patients atteints d'un cancer colorectal métastatique initialement non résécable.

Jary M, Lecomte T, Bouché O, Kim S, Dobi E, Queiroz L, Ghiringhelli F, Etienne H, Léger J, Godet Y, Balland J, Lakkis Z, Adotevi O, Bonnetain F, Borg C, Vernerey D

Prognostic value of baseline seric Syndecan-1 in initially unresectable metastatic colorectal cancer patients: a simple biological score. **Int J Cancer** doi:10.1002/ijc.30367. (2016)

Impact factor: 5.5

IV.1.2.1 Positionnement du problème

Aujourd'hui, diverses options thérapeutiques sont disponibles en 1ère ligne pour les patients atteints d'un cancer colorectal métastatique (mCRC). Dans ce contexte, l'identification au diagnostic de biomarqueurs permettant une stratification du risque de décès est une problématique cruciale afin de mieux connaître le pronostic de la maladie, d'individualiser plus précisément les différentes stratégies de traitement et par conséquent améliorer le rapport coût-efficacité des interventions.

Les deux scores pronostiques de la survie globale (SG) actuellement les plus utilisés pour les patients mCRC sont le score de Kohne et celui du Groupe Coopérateur Multidisciplinaire en Oncologie GERCOR. Le score de Khone repose sur l'évaluation du « performance status » (PS), du nombre de globules blancs, du nombre de sites métastatiques, et du taux de phosphatases alcalines. Cependant, une étude a montré que dans un contexte de thérapie ciblée seulement 2 de ces paramètres (le PS et le

nombre de globules blancs) demeurent associés à la SG. Le score du GERCOR quant à lui s'articule aussi sur le PS auquel il associe l'évaluation des lactates déshydrogénases (LDH), un paramètre biologique dosé en routine. Le score du GERCOR tend à être aujourd'hui le score de référence puisqu'il est plus discriminant pour le SG que celui de Khone et est de plus bien adapté au contexte des thérapies ciblées.

Ces deux scores sont donc basés sur le PS qui est un paramètre ayant pour objet d'évaluer l'état fonctionnel d'un patient. A cet effet, il est établi sur la base d'un interrogatoire qui peut véhiculer la subjectivité du clinicien et parfois ne pas être reproductible. Le PS peut être influencé par de multiples facteurs liés aux patients (anorexie, état psychologique) et/ou liés à la maladie tumorale (inflammation, taille tumorale, anémie, etc.).

La substitution du PS par un paramètre biologique plus reproductible et aussi plus spécifique et représentatif de la tumeur ou de l'état métabolique du patient pourrait être intéressant pour mieux individualiser les différentes options thérapeutiques au regard de leur mode d'action. Si nous considérons le profil moléculaire de la tumeur, le statut mutationnel BRAF peut constituer un paramètre pronostique clef mais ne caractérise que 8 à 10% des patients mCRC.

Dans cette pathologie, un enjeu majeur réside donc dans l'identification et la validation de biomarqueurs spécifiques de la maladie tumorale dans le but d'améliorer la stratification du risque de décès. L'invasion, l'angiogenèse et le contrôle du microenvironnement tumoral sont des caractéristiques communes de la progression du cancer et peu de biomarqueurs spécifiques de ces caractéristiques sont disponibles au moment du diagnostic. Nous avons récemment montré que l'angiopoïétine-2, qui est impliquée dans l'angiogenèse, est un paramètre pronostique intéressant pour la SG chez des patients mCRC. Le Syndecan-1 (SDC-1, aussi appelé CD138) est un protéoglycane transmembranaire de type héparane sulfate impliqué dans l'angiogenèse, la progression tumorale et aussi dans la transition épithélio-mesenchymateuse. Le CD138 constitue un biomarqueur d'intérêt puisqu'il a déjà été décrit comme jouant un rôle important dans la séquence adénome-

adénocarcinome et son rôle pronostique a déjà été reconnu comme intéressant dans plusieurs cancers solides et hématologiques.
Dans les cancers colorectaux, l'expression de CD138 a été décrite dans le stroma des tumeurs primitives comme étant corrélée à des caractéristiques histopathologiques de mauvais pronostique lorsqu'il existe une perte d'expression membranaire et une augmentation de l'expression stromale en immunohistochimie. Ex vivo, les cellules cancéreuses issues de patients présentant des taux sériques élevés de CD138 présentent une moins bonne sensibilité à la chimiothérapie. Une hypothèse est donc l'augmentation de la chimiorésistance en cas de CD138 sérique élevé, mais celle ci reste à démontrer.

Cependant, à ce jour, le rôle pronostique du CD138 dans sa forme soluble sur la SG n'a jamais été investigué chez les patients mCRC. Cette étude a donc pour objet (i) d'évaluer et de valider le rôle pronostique du CD138 soluble chez les patients mCRC (ii) de déterminer sa valeur discriminante additionnelle aux paramètres déjà connus pour stratifier le risque de décès des patients et (iii) de proposer un score pronostique de la SG basé sur des paramètres biologiques robustes mesurables au diagnostic et applicables dans la pratique clinique courante.

IV.1.2.2 Population de l'étude

Cette étude est basée sur des données de patients mCRC qui proviennent de deux cohortes prospectives dans lesquelles tous les patients ont reçu un traitement en première ligne à base de Bevacizumab.
Dans la cohorte de développement les patients avaient tous des métastases hépatiques non résécables et ont été recrutés entre 2007 et 2010 (NCT00489697).
La cohorte de validation, quant à elle, est un essai de phase II multicentrique monobras dans lequel ont été inclus des patients traités par FOLFIRI-3 plus Bevacizumab entre 2007 et 2009 (NCT00544011). Dans cet échantillon de patients, la chirurgie des métastases était autorisée après 6 cycles de

traitement.
Les données démographiques, cliniques, pathologiques et biologiques ont été collectées à l'inclusion dans chaque cohorte.
Des échantillons de sang ont été prélevés à l'inclusion avant l'initiation du traitement. Le plasma et le sérum ont immédiatement été conditionnés à -80°C comme requis.
Le CD138 a été mesuré rétrospectivement par ELISA (Enzyme-linked immunosorbent assays) pour les patients des deux cohortes.

IV.1.2.3 Approche pronostique méthodologique

La SG, le critère de jugement principal, a été calculée comme le délai entre la date d'inclusion dans l'essai et le décès toutes causes confondues. Les patients en vie à la fin du suivi ont été censurés et renseignés comme étant vivant à cette date. La survie sans progression (SSP), critère de jugement secondaire, a quant à elle été mesurée comme le délai entre la date d'inclusion dans l'essai et la date de première progression clinique ou radiologique ou le décès quelle que soit la cause. Si aucune progression n'est reportée et en l'absence de décès, le délai est censuré à la fin du suivi et le patient est renseigné comme étant vivant à cette date. La SG et la SSP ont été estimées par la méthode de Kaplan-Meier et décrites à l'aide de la médiane de survie ou du taux de survie aux temps d'intérêt accompagnés de leurs intervalles de confiance à 95%.

Le suivi de l'étude a été estimé selon la méthode du Kaplan-Meier inverse.

Une modélisation de Cox à risque proportionnel a été appliquée pour établir les rapports de risques instantanés (Hazard Ratio ou HR) ainsi que leurs intervalles de confiance à 95% pour analyser l'association des paramètres investigués avec la SG. Dans un premier temps, l'association avec la SG des paramètres recueillis à l'inclusion a été investiguée de manière univariée. Dans un second temps et dans le but d'établir un modèle multivarié final, une procédure d'élimination pas à pas descendante a été

proposée en incluant les facteurs identifiés comme étant significativement ($p < 0.05$) associés à la SG dans la phase univariée. Lorsque le CD138 est utilisé en continu dans la modélisation de Cox nous avons considéré une transformation logarithmique compte tenu de sa distribution asymétrique. L'hypothèse des risques proportionnels requise lors des modélisation de Cox a été contrôlée à l'aide d'un représentation graphique du log(-(log(Survie)).

La validité et les performances du modèle multivarié ont été évaluées selon deux paramètres : la discrimination et la calibration. La valeur prédictive et la capacité à discriminer la SG, et donc à séparer les patients avec un pronostic différent du modèle, ont été évaluées avec l'indice de concordance de Harrell (C de Harrell). Mille échantillons aléatoires de la population de développement ont été obtenus par procédure de ré-échantillonnage (bootstrap) et ont été ensuite utilisés pour construire un intervalle de confiance à 95% autour du C de Harrell.

La calibration du modèle, quant à elle, se réfère à sa capacité à fournir des prédictions non biaisées dans des groupes de patients considérés comme similaires. Elle a été analysée graphiquement par une courbe de calibration.

Une validation interne du modèle a été proposée à l'aide d'une procédure de ré-échantillonnage en produisant des intervalles de confiance autour des estimations du modèle.

La valeur prédictive additionnelle du CD138 pour la SG lorsqu'il est inclus dans un modèle de référence (modèle dit du GERCOR puisqu'il repose sur la LDH et le PS) a été évaluée à l'aide du C de Harrell. Afin d'évaluer la significativité de l'amélioration de la discrimination de la SG apportée par le CD138 sur les paramètres conventionnels, cette analyse a été réalisée 1 000 fois dans des échantillons de données issus d'une procédure de ré-échantillonnage afin d'estimer un intervalle de confiance à 95% autour de la différence des C de Harrell du modèle de référence avec et sans l'ajout du CD138.

La possibilité de produire une implémentation simple du dosage de CD138 en pratique clinique a été investiguée avec la détermination d'une valeur seuil

obtenue en comparant la distribution de l'expression de CD138 des volontaires en bonne santé avec celle des patients mCRC impliqués dans l'étude.

Un score biologique simple a donc été proposé sur la base des deux paramètres biologiques identifiés, le LDH et le CD138, en les considérants dans leur forme binaire.

La capacité à discriminer du score biologique a été évaluée à l'aide du C de Harrell.

Une validation externe de l'association indépendante du CD138 et de sa capacité discriminante via le score biologique proposé a été réalisée.

IV.1.2.4 Résultats

126 patients ont été inclus dans la cohorte de développement et 51 dans la cohorte de validation. Au total 177 patients sont éligibles pour l'analyse. Les caractéristiques de ces patients sont présentées dans la table 1

Table 1: Caractéristiques des patients au moment de l'inclusion dans les cohortes de développement et de validation

Caractéristiques	**Population globale (N =177)**		**Cohorte de développement (N =126)**		**Cohorte de validation (N =51)**		P†
	N		N		N		
Age — années ‡	177		126		51		
≤ 65		93 (53%)		66 (52%)		27 (53%)	
>65		84 (47%)		60 (48%)		24 (47%)	p = 0.9461
Patient de sexe masculin — *no. (%)*‡	177	74 (42%)	126	46 (37%)	51	28 (55%)	**p = 0.0246**
Performance status ‡	167		119		48		
0— no. (%)		91 (55%)		64 (54%)		27 (56%)	
>0— no. (%)		76 (45%)		55 (46%)		21 (44%)	p = 0.7719

Site de la tumeur primitive ‡	177		126		51		
Colon — no. (%)		120 (68%)		88 (70%)		32 (63%)	
Rectum — no. (%)		57 (32%)		38 (30%)		19 (37%)	p = 0.3602
Localisation des métastases ‡	177		126		51		
Foie et autres — no. (%)		98 (55%)		81 (64%)		17 (33%)	
Foie seul — no. (%)		68 (39%)		45 (36%)		23 (45%)	
Autre seul — no. (%)		11 (6%)		0 (0%)		11 (22%)	**p <0.0001**
Moment d'apparition des métastases ‡	173		125		48		
Synchrone — no. (%)		119 (69%)		89 (71%)		30 (63%)	
Métachrone — no. (%)		54 (31%)		36 (29%)		18 (37%)	p = 0.2688
Chirurgie de la tumeur primitive — no. (%)‡	175	110 (63%)	126	82 (65%)	49	28 (57%)	p = 0.3293
Leucocytes (x10^6/mL)*	137	8.1 ± 3.2	88	8.2 ± 3.0	49	8.1 ± 3.5	p = 0.8926
Lymphocytes (x10^6/mL)*	132	1.5 ± 0.7	87	1.4 ± 0.6	45	1.7 ± 0.7	p = 0.0550
LDH ‡	115		82		33		
≤350 (ULN)		56 (49%)		36 (44%)		20 (61%)	
>350 (ULN)		59 (51%)		46 (56%)		13 (39%)	p = 0.1050
CD138 (ng/mL)**	177	75.3 (48.8-139.7)	126	91.9 (61.8-169.1)	51	37.5 (25.7-70.2)	**p <.0001**
Décès ‡	177	130 (73%)	126	94 (75%)	51	36 (71%)	p = 0.5838
Médiane du temps de suivi en mois 95%IC		56.0 (50.3-59.0)		51.8 (41.9-59.8)		57.9 (53.1-60.3)	**p = 0.0295**

* Moyennes ± écart-type et le test de student non apparié a été utilisé pour la comparaison des variables entre les cohortes

** Médiane Q2 et intervalle des quartiles Q1 et Q3 sont décrits et le test de Wilcoxon a été utilisé pour la comparaison des variables entre les groupes

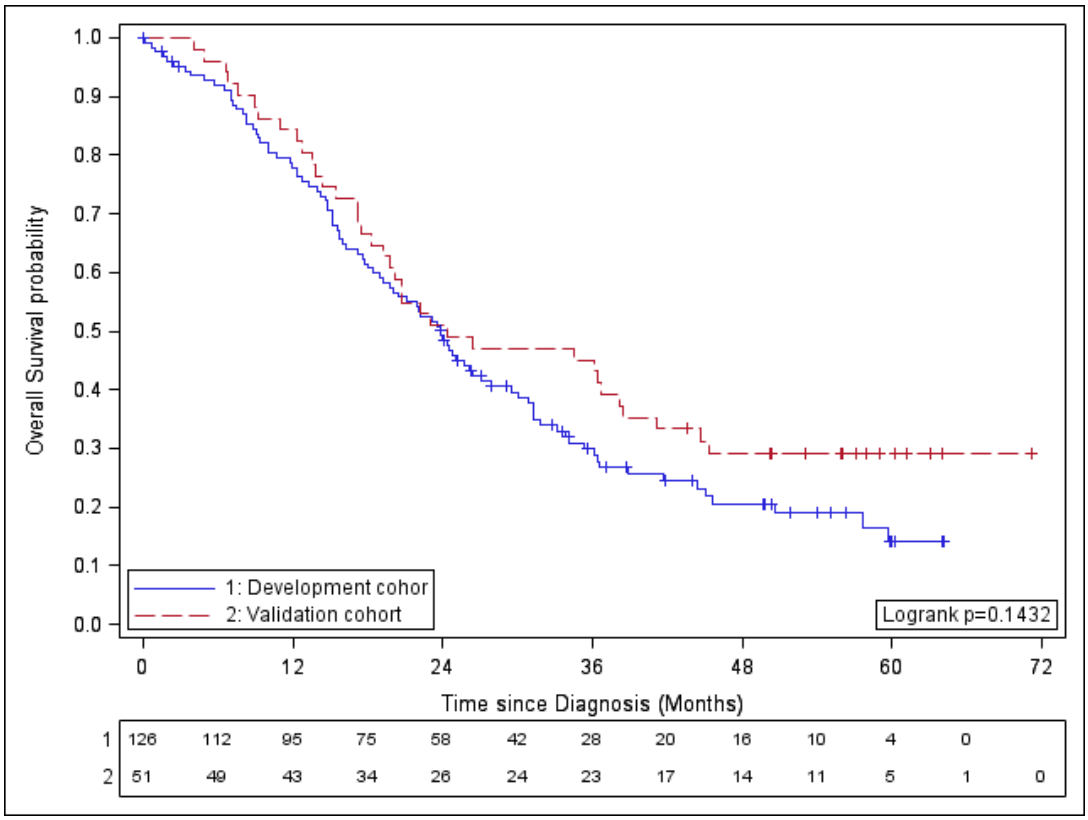
† P values pour la comparaison entre la cohorte de développement et de validation

‡ Les tests du Chi-2 ou de Fisher exact ont été utilisés pour la comparaison des variables catégorielles.

Abréviations: PS: performance status; ULN: Upper Limit of Normal

Les courbes de Kaplan-Meier pour la SG montrent un profil pronostique similaire dans les deux cohortes (HR = 0.750, 95% CI 0.509–1.104, p = 0.14; Figure 1).

Figure 1 : Courbes de Kaplan-Meier de la survie globale selon la cohorte considérée



L'analyse univariée par modèle de Cox a identifié 5 paramètres significativement associés à la SG. L'approche multivariée n'a quant à elle retenue que deux paramètres associés de manière indépendante à la SG : le LDH et le CD138 (Table 2).

Table 2: Analyse de Cox univariée et multivariée pour la survie globale dans la cohorte de développement

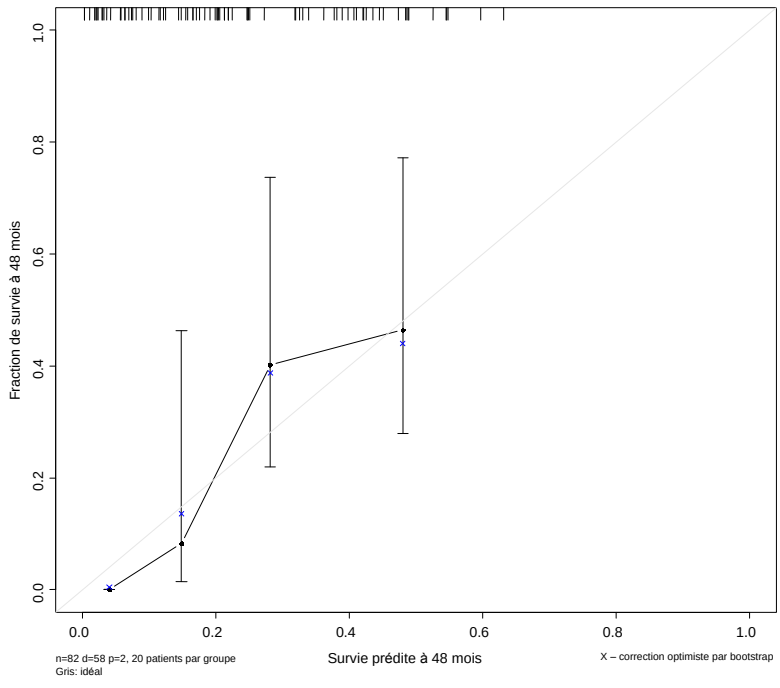
	Patients	Décès	Analyse univariée (n=126)			Analyse multivariée (n=82)		
			HR	95%CI	P	HR	95%CI	P
Age — années								
<= 65	66	47	1	-	-			
>65	60	47	1.309	[0.872; 1.964]	0.1938			
Sexe								
Homme	46	32	1	-	-			
Femme	80	62	1.362	[0.887; 2.093]	0.1580			
Performance Status								
0	64	45	1	-	-			
>0	55	43	1.183	[0.778; 1.799]	0.4315			
Site de la tumeur primitive								
Colon	88	67	1	-	-			
Rectum	38	27	1.114	[0.712; 1.744]	0.6358			
Moment d'apparition des métastases								
Métachrone	36	26	1	-	-			
Synchrone	89	67	1.291	[0.819; 2.035]	0.2708			
Localisation des métastases								
Foie seul	45	28	1	-	-			
Foie et autres	81	66	1.966	[1.260; 3.069]	**0.0029**			
Autres seul	-	-	-	-	-			
Chirurgie de la tumeur primitive								
Oui	44	37	1	-	-			
Non	82	57	1.969	[1.290; 3.004]	**0.0017**			
Leucocytes ($x10^6$/mL)	88	66	1.084	[0.997; 1.178]	**0.0601**			

Lymphocytes (x10^6/mL)	87	65	0.3406	[0.536; 1.240]	0.3406			
LDH								
≤350 (ULN)	36	21	1	-	-	1		
>350 (ULN)	46	37	2.337	[1.356; 4.030]	**0.0022**	2.165	[1.241; 3.778]	**).0066**
Log_CD138 (ng/mL)	126	94	2.321	[1.652; 3.260]	**<.0001**	1.706	[1.092; 2.666]	**).0190**

Abréviations: PS: performance status; ULN: Upper Limit of Normal

Le modèle multivarié final identifié est caractérisé par une capacité discriminante acceptable (C-index 0.644; 95%IC par bootstrap, 0.570- 0.718) et une bonne calibration aux temps d'intérêt comme reporté sur la figure 2.

Figure 2: Courbe da calibration à 48 mois pour le modèle multivarié final



L'axe vertical est la proportion observée des patients survivants à 48 mois. L'axe horizontal représente lui les prédictions du modèle. La ligne grise correspond à un modèle parfaitement calibré. La ligne continue noire illustre la performance actuelle du modèle. Les intervalles de confiance à 95% autour des points observés sont obtenus par une procédure de ré-échantillonnage (bootstrap).

Chez les 79 patients de la cohorte de développement pour lesquels le PS, le LDH et le CD138 étaient disponibles nous montrons que la capacité discriminante du modèle basé sur les paramètres du score GERCOR (0.615) est significativement améliorée lorsqu'on y implémente le CD138 (0.663) (différence moyenne obtenu par bootstrap : 0.048; 95% IC: 0.076–0.087).
En considérant que le PS n'est pas associé significativement à la SG dans le modèle et qu'il semble être péjoratif pour la capacité discriminante du modèle la combinaison de paramètre la plus parcimonieuse pour discriminer la SG est donc le LDH et le CD138 (Table 3).

Table 3: Capacité discriminante des différents modèles

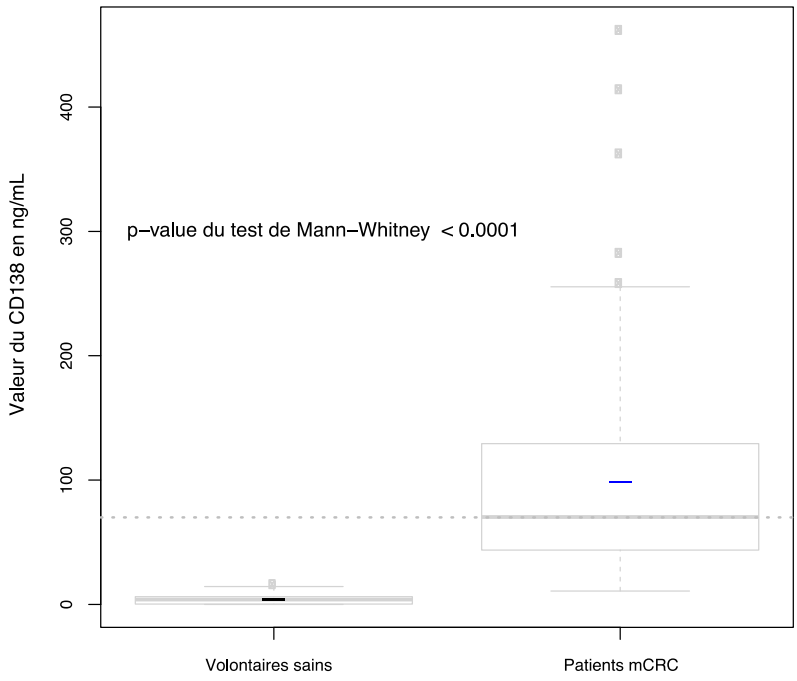
	Patients	Décès	Modèle du GERCOR (n=79)			Modèle du GERCOR + CD138 (n=79)			Modèle multivarié final de l'étude (n=79)		
			HR	**95%IC**	**P**	**HR**	**95%IC**	**P**	**HR**	**95%IC**	**P**
PS											
0	41	27	1	-	-		-	-			
>0	38	29	1.089	[0.641; 1.848]	0.7528	1.051	[0.619; 1.785]	0.8541			
LDH											
≤350 (ULN)	34	19	1	-	-	1			1		
>350 (ULN)	45	37	2.524	[1.433; 4.446]	**0.0013**	2.354	[1.319; 4.201]	**0.0038**	2.366	1.328; 4.213]	**0.0035**
Log_CD138 (ng/mL)	79	56				1.858	[1.181; 2.923]	**0.0074**	1.862	[1.184; 2.927]	**0.0071**
Performance du modèle											
C de Harrell				0.615			0.663			0.664	

Abréviations: ULN: Upper Limit of Normal

Sur la base de cette combinaison de paramètres biologiques identifiés (LDH et CD138), un score biologique a été proposé.
Le choix d'une valeur seuil de 75 ng/mL qui correspond à la médiane d'expression du CD138 pour les 177 patients mCRC inclus dans cette étude est un choix non arbitraire intéressant, puisque par définition il divise la population en deux groupes de patients équilibrés. De plus, ce seuil est très

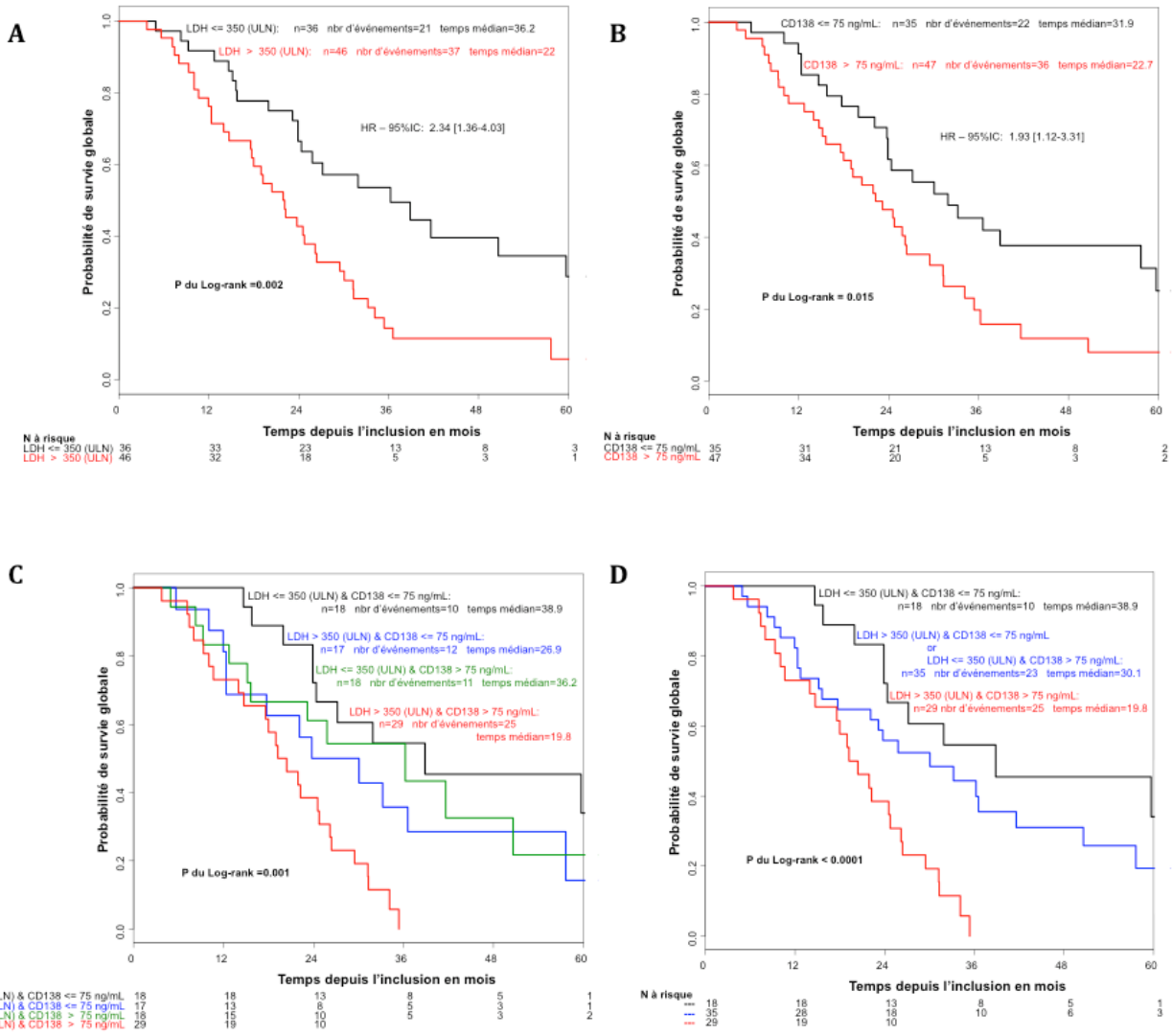
proche des valeurs déjà utilisées dans la littérature et surtout aucun des 40 volontaires « sains » étudiés n'a atteint ce niveau d'expression. (Figure 3)

Figure 3 : Distribution de l'expression du CD138 chez 40 volontaires « sains » (n=40) et chez les 177 patients atteints d'un cancer colorectal métastatique inclus dans l'étude



Sur la base de la variable CD138 et en considérant une valeur seuil de 75 ng/mL ainsi que de la mesure de la LDH considérée comme inferieure ou supérieure à la limite supérieure de la normale, le score biologique LDH-CD138 a donc été proposé comme ci-dessous et donne la possibilité d'identifié 3 groupes de patients de profils pronostiques de la SG très différents (figure 4).

Figure 4 : Courbe de survie globale des patients selon la méthode de Kaplan-Meier selon l'expression de (A) la LDH, (B) le CD138, (C, D) la combinaison de la LDH – CD138

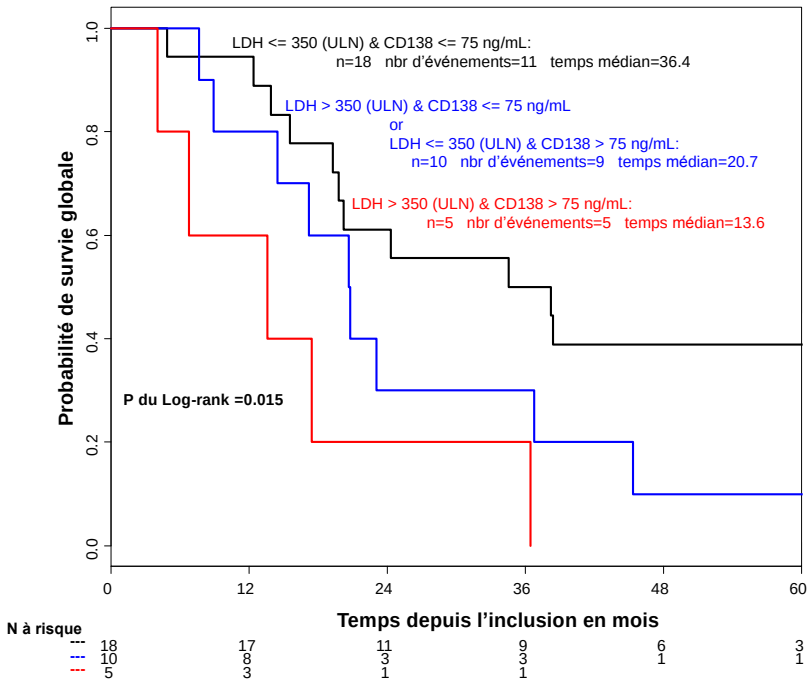


En guise de validation interne du modèle multivarié final construit dans la cohorte de développement, une procédure de ré-échantillonnage produisant 1000 échantillons de données nous a permis d'apprécier la fiabilité des estimations produites dans le modèle multivarié finale en produisant des intervalles de confiances à 95% robustes (LDH: HR = 2.165, 95% CI percentile : 1.149–3.929 et CD138 (transformation log): HR = 1.706, 95% CI percentile : 1.031–2.954).

Dans la cohorte de validation externe, le modèle final multivariée de Cox identifié dans la cohorte de développement permet de valider l'association indépendante du CD138 pour la SG (HR = 1.69, 95% CI: 1.049–2.723, p = 0.0309).

La capacité discriminante du score biologique a aussi été confirmée dans la cohorte de validation externe avec une indice C de 0.633 et une courbe de Kaplan-Meier identifiant trois groupes clairement distincts (p = 0.015, Figure 5).

Figure 5 : Courbes de la survie globale selon la méthode de Kaplan-Meier en fonction de combinaison LDH - CD138 dans la cohorte de validation.

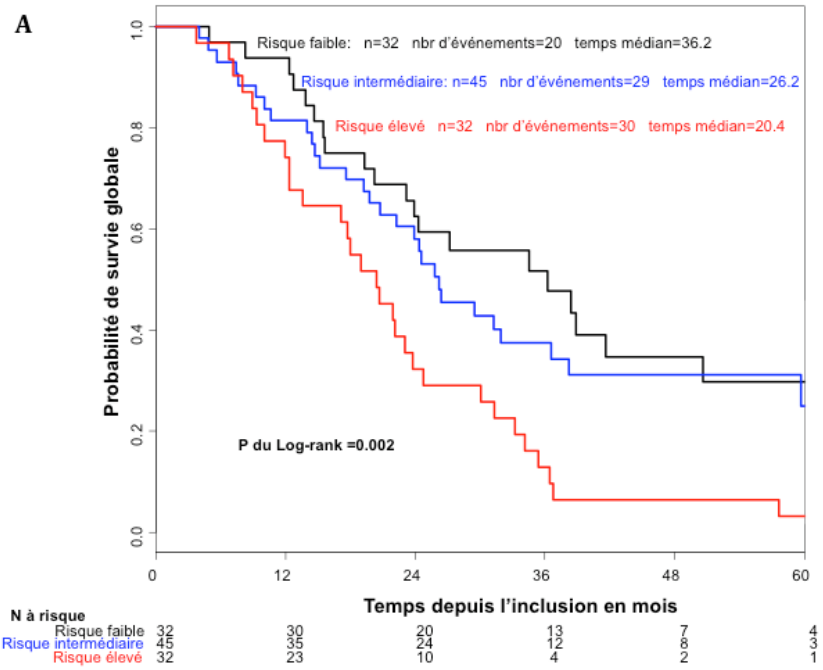


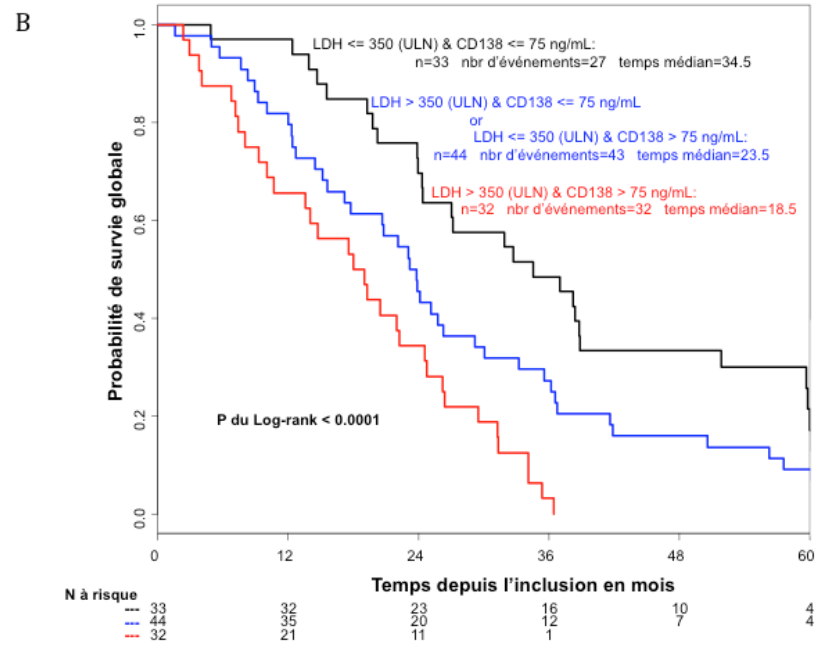
Nous avons à ce stade démontré et validé l'intérêt de la valeur pronostique du CD138 pour les patients atteints de mCRC et par conséquent produit un score biologique simple.

Aujourd'hui, le score du GERCOR est considéré comme un outil bien validé et repose aussi sur deux paramètres clefs simples que sont la LDH et le PS.

Afin d'appliquer et d'apprécier les deux approches avec le plus grand matériel disponible, nous avons décidé de mettre en commun les deux cohortes participant à l'étude (n=177). Parmi les 177 patients, 109 (62%) avaient des valeurs disponibles pour le LDH, le PS et le CD138. La Figure 6 présente les résultats obtenus avec ces deux approches.

Figure 6 : Courbe de la survie globale des patients selon la méthode de Kaplan-Meier dans la population « poolée » et selon (A) l'approche du GERCOR ou (B) la considération de la combinaison LDH - CD138





De manière intéressante, dans notre étude, la combinaison du CD138 avec le LDH offre une meilleure capacité discriminante que l'approche du GERCOR, puisque les groupes sont plus distincts avec le score biologique. Cette observation est renforcée avec l'obtention d'un meilleur indice de discrimination pour le score basé sur CD138 et le LDH de 0,629 comparé à 0,594 pour l'autre approche.

Enfin, nous avons aussi observé une différence très significative entre les médianes de SSP parmi les trois groupes identifiés par le score biologique (log-rank $p < 0.0014$) suggérant que les trois groupes de pronostics de survie globale différents ont aussi un profil de progression tumorale distinct.

IV.1.2.5 Conclusion de l'étude

Le CD138 soluble est donc un facteur pronostique indépendant de la SG chez les patients atteints de mCRC qui apporte une valeur additionnelle aux autres paramètres conventionnels en termes de discrimination du risque de décès.
L'évaluation de la combinaison biologique CD138-LDH au diagnostic offre une amélioration de la capacité de stratification du risque de décès et ainsi peut servir de base pour adapter des interventions thérapeutiques.

En guise de perspectives, les résultats de ce travail sont en train d'être mis à profit dans le cadre d'un projet interventionnel visant à étudier l'impact de l'intensification thérapeutique dans une population de patients de mauvais pronostic, sur la base du score pronostique biologique basé sur le dosage sérique du Syndecan-1 (CD138) et LDH. Ceci est discuté plus en détail dans la section discussion de la thèse.

IV.1.2.6 Manuscrit

IJC
International Journal of Cancer

Prognostic value of baseline seric Syndecan-1 in initially unresectable metastatic colorectal cancer patients: a simple biological score

Marine Jary[1,2,3], Thierry Lecomte[4,5], Olivier Bouché[6], Stefano Kim[1,3], Erion Dobi[1], Lise Queiroz[2], Francois Ghiringhelli[7], Hélène Etienne[1,8], Julie Léger[9], Yann Godet[2], Jérémy Balland[2], Zaher Lakkis[10], Olivier Adotevi[1,2], Franck Bonnetain[2,3,11], Christophe Borg[1,2,3]* and Dewi Vernerey[2,3,11]*

[1] Department of Medical Oncology, University Hospital, Besançon, France
[2] INSERM, Unit 1098, University of Bourgogne- Franche Comté, Besançon, France
[3] Clinical Investigation Center 1431, EFS Bourgogne-Franche Comté, Besançon, France
[4] CNRS, Unit 7292, University François-Rabelais, Tours, France
[5] Department of HepatoGastroenterology and Digestive Oncology, University Hospital, Tours, France
[6] Department of HepatoGastroenterology and Digestive Oncology, University Hospital Robert Debré, Reims, France
[7] Department of Medical Oncology, Leclerc Anticancer Center, Dijon, France
[8] Department of Gastroenterology, University Hospital, Besançon, France
[9] INSERM, Clinical Investigational Center CIC 1415, Tours, France
[10] Department of Digestive Surgery and Liver Transplantation, University Hospital, Besançon, France
[11] Methodological and Quality of Life in Oncology Unit, EA 3181, University Hospital, Besançon, France

In first-line metastatic colorectal cancer (mCRC), baseline prognostic factors allowing death risk and treatment strategy stratification are lacking. Syndecan-1 (CD138) soluble form was never described as a prognostic biomarker in mCRC. We investigated its additional prognostic value for overall survival (OS). mCRC patients with unresectable disease at diagnosis were treated with bevacizumab-based chemotherapy in two independent prospective clinical trials (development set: $n = 126$, validation set: $n = 51$, study NCT00489697 and study NCT00544011, respectively). Serums were collected at baseline for CD138 measurement. OS determinants were assessed and, based on the final multivariate model, a prognostic score was proposed. Two independent OS prognostic factors were identified: Lactate Dehydrogenase (LDH) high level ($p = 0.0066$) and log-CD138 high level ($p = 0.0190$). The determination of CD138 binary information (cutoff: 75 ng/mL) allowed the assessment of a biological prognostic score with CD138 and LDH values, identifying three risk groups for death (median OS= 38.9, 30.1 and 19.8 months for the low, intermediate and high risk groups, respectively; $p < 0.0001$). This score had a good discrimination ability (C-index = 0.63). These results were externally confirmed in the validation set. Our study provides robust evidence in favor of the additional baseline soluble CD138 prognostic value for OS, in mCRC patients. A simple biological scoring system is proposed including LDH and CD138 binary status values.

Key words: CD138, Syndecan-1, colorectal cancer, biomarker, prognostic score

Abbreviations: 95% CI: 95% confidence interval; Ang-2: angiopoietin-2; ELISA: enzyme-linked immunosorbent assays; EMT: epithelial-mesenchymal transition; HR: hazard ratio; HSPGs: heparan sulfate proteoglycans; IHC: immunohistochemical; IQR: interquartile range; LDH: lactate dehydrogenase; mCRC: metastatic colorectal cancer; ORR: objective response rate; OS: overall survival; PFS: progression-free survival; PS: World Health Organization Performance Status, SDC-1: Syndecan-1; ULN: upper limit of normal

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

*CB and DV equally contributed to this study

Conflict of interest: TL, OB, FB, CB: Advisory board for Roche Laboratory.

Authors'contributions: MJ, TL, OB, SK, ED, FG, ZL, CB, data acquisition; JB, YG, CB, biological analyses; TL, OB, FG and CB provided critical biological material, MJ, FB, CB, DV data interpretation; FB, DV statistical analyses; MJ, CB, DV, manuscript drafting; MJ, TL, OB, SK, ED, LQ, FG, HE, JL, YG, JB, ZL, OA, FB, CB, DV, manuscript revision and final approval.

Grant sponsor: INCa (Institut National du Cancer), Cancéropôle Grand-Ouest, Regional Council of Franche-Comte and the French National Society of Gastroenterology; **Grant sponsor:** French Higher Education and Research Ministry; Grant number: "Investissements d'avenir" Grant Agreement: LabEx MAbImprove ANR-10-LABX-53-01

DOI: 10.1002/ijc.30367

History: Received 8 Jan 2016; Accepted 1 July 2016; Online 29 July 2016

Correspondence to: Marine Jary, Department of Oncology. Jean Minjoz University Teaching Hospital, 3 Boulevard Alexander Fleming, Besancon F-25030, France. Tel.: +33-3-8166-8705/Fax: +33-3-8166-8708, E-mail: mjary@chu-besancon.fr

What's new?
Prognostic scores so far validated in unresectable metastatic disease are based on patient performance status, which can be influenced by many factors. For the first time, the authors explored Syndecan-1 (CD138), a molecule involved in angiogenesis, cancer invasion and epithelial-mesenchymal-transition, as a candidate biomarker in first-line metastatic colorectal cancer patients. They validated CD138 independent Overall Survival (OS) prognostic value at baseline and constructed a CD138-Lactate Dehydrogenase (LDH) biological score that discriminates three patient groups with different prognosis profiles. Combining LDH and soluble CD138 measurement offers an easy and reproducible staging system to predict OS, avoiding genomic tumor heterogeneity confounding risks.

Introduction

Nowadays, various therapeutic interventions are available in metastatic colorectal cancer (mCRC) patients.[1] In this context, the identification of prognostic biomarkers at diagnosis allowing risk death stratification is a critical issue to improve mCRC prognosis knowledge, to individualize more accurately treatment strategies and optimize cost-effectiveness ratio.

Two validated prognostic scores have been proposed in unresectable metastatic disease: the Kohne and the *Groupe Coopérateur Multidisciplinaire en Oncologie* GERCOR scores.[2,3] The Kohne score is based on the assessment of World Health Organization Performance Status (PS), white blood cell count (WBC), number of metastatic sites and Alkaline Phosphatase level.[2] However, when applied with targeted biotherapies, the two independent prognostic factors for OS remain PS and WBC.[4] Moreover the GERCOR score, which monitors PS and Lactate DeHydrogenase (LDH) status, exhibits a better discrimination ability and is well adapted for targeted biotherapies.[4] Both Kohne and GERCOR scores are based on performance status. Alteration of PS might be the consequence of multiple parameters linked to the patient (anorexia, psychological status) or to the tumor (inflammation, tumor burden, anemia…). The substitution of PS by biological parameters, more precisely related to the tumor or patient metabolic status, might also help to better individualize different therapeutic options according to their mode of action. Considering tumor molecular profile, BRAF V600E mutation is another key prognostic parameter, but characterizing only 8–10% of mCRC patients.[5–8] Then, the development and validation of additional biomarkers allowing the improvement of risk death stratification at diagnosis remains a major concern.

Invasion, angiogenesis and control of the tumor microenvironment are common features of cancer progression. Few specific biomarkers of these hallmarks are available at diagnosis. We have recently shown that angiopoietin-2 (Ang-2), which is involved in angiogenesis, is an interesting prognostic parameter for OS in mCRC patients.[9] We next sought to investigate the prognostic value of syndecan-1 (SDC-1, also named CD138), a type-1 transmembrane heparan sulfate proteoglycans (HSPGs), involved in angiogenesis, but also in cancer invasion.[10–12] CD138 has been described to have an important role in the colorectal adenoma-adenocarcinoma sequence and in the epithelial–mesenchymal transition (EMT), using immunohistochemical (IHC) staining.[13,14] The prognostic value of CD138 was first investigated in myeloma and is now acknowledged as a promising biomarker in several solid and hematological cancers.[15,16] In colorectal cancers, CD138 expression in the stroma of primary cancers was correlated to worse histopathological features and the shedding of CD138 was proposed as a mechanism conferring resistance to chemotherapy.[17,18] However, the CD138 shed soluble form prognostic value was never investigated in mCRC patients.

Then, the aims of this study were (*i*) to assess and validate the independent prognostic value of baseline serum CD138, (*ii*) to determine the additional value of soluble CD138 among other classical clinical parameters to stratify patients' risk of death and (*iii*) to construct a prognostic score based on robust biological parameters measurable at the baseline and applicable for the clinical practice.

Patients and Methods

Population

Individual patient data were collected from two previous independent prospective cohorts, both containing bevacizumab treatment in first-line mCRC patients. Details of these studies have been extensively described elsewhere.[19,20] In both studies, data were anonymized and all patients were followed for death event.

The development set included 126 patients treated by chemotherapy and bevacizumab between 2007 and 2010. All patients had unresectable liver metastases in this development set.[19] The validation set enrolled 51 patients between October 2007 and October 2009 treated by FOLFIRI-3/bevacizumab, in a pilot, single-arm, multicenter, phase II trial.[20] In this cohort, surgery of metastases was allowed after six cycles of treatment (see Supporting Information Methods).

The flow chart describing the different population studied in this study is available in Supporting Information Figure 1.

Data extraction

The following data were collected at baseline for each patient in the two cohorts: center and patient identification, age, sex, performance status, primary tumor site (colon, rectum), site of metastases (liver-limited, liver and other, other), time of

metastasis diagnostic (synchronous, metachronous), primary tumor resection, LDH level at baseline (normal value was considered if LDH were below 350 UI/L), lymphocyte and leucocyte counts, survival status and date of last news or death.

Serum sample measurements

Blood samples were drawn at baseline, before treatment initiation and immediately processed for plasma and serum freezing at −80°C. Samples frozen >4 hr following venous blood collection were excluded from the analysis. *Enzyme-linked immunosorbent assays* (ELISA) were used to measure CD138 in serum samples (Diaclone®) according to the manufacturer's instructions. Each sample was analyzed in duplicate.

CD138 serum sample was dosed at baseline, before treatment and was considered, in this study, as the primary parameter of interest. Ang-2 measurement was also available in these two cohorts and was then analyzed, with an exploratory approach.

Statistical analysis

We provided the mean (SD) values and frequencies (percentages) for the description of continuous and categorical variables, respectively. The means and the proportions were compared using Student's *t*-test and the χ^2 test (or Fisher's exact test, if appropriate), respectively. Due to the skewed CD138 distribution, we used for its description the median, and the interquartile range for the dispersion measurement, as recommended.[21] Wilcoxon rank-sum test was performed for CD138 distribution comparison among the cohort set.

OS was calculated from the date of study enrolment to the date of death from any cause. Alive patients were censored at the last follow-up. Progression free survival (PFS) was calculated from the date of study to the date of progression, defined by clinical or radiological assessment or the date of death or last news, whichever occurred first. OS and PFS were estimated using Kaplan–Meier method and describe using median or rate at specific time points with 95% confidence intervals (CIs). Follow-up was calculated using reverse Kaplan–Meier estimation.[22] A Cox proportional hazard model was performed to estimate the hazard ratio (HR) and 95% CI for the factors associated with OS.

The association of parameters at enrolment with OS was first assessed using univariate Cox analysis and then included (for those with $p < 0.05$) in a final multivariate Cox regression model with stepwise backward elimination. When used in continuous in the Cox modelization, CD138 variable had to be normalized by logarithmic transformation, considering its skewed distribution. Hazard proportionality was checked by plotting log-minus-log survival curves.

Accuracy of the model was checked regarding two parameters: discrimination and calibration.[23] The predictive value and the discrimination ability (*i.e.*, the ability to separate patients with different prognosis, see Supporting Information Methods) of the model was evaluated with Harrell's Concordance (C)-index. One thousand random samples of the population were used to derive 95% CI for the Harrell's Concordance statistic. Calibration and goodness of fit of the model were assessed by visual examination of calibration plot.

The predictive value that CD138 parameter added to a reference risk model (GERCOR model) was assessed with the use of C-statistic. This analysis was repeated 1000 times using bootstrap samples to derive 95% CI for the difference in the C-statistics between the two models in order to finally assess the improvement in discrimination of CD138 parameter among the other conventional parameters.

The possibility to provide a simple implementation of CD138 monitoring in clinical practice was investigated with the determination of a cut-off value, by contrasting CD138 level distributions in boxplot among healthy volunteers and mCRC patients. By using CD138 binary information, a prognostic score was proposed.

The discrimination ability of CD138-LDH biological score to the GERCOR score was evaluated with the use of C statistic.

We confirmed the internal validity of the model by using a bootstrap sample procedure. Several approaches have been proposed to assess the performance in samples of the same population (internal validation, see Supporting Information Methods). A thousand datasets were generated from the resampling of the original dataset to estimate the biased corrected 95% bootstrap CIs for HRs.

The objective response rate (ORR, defined by the combination of complete and partial responses) distribution was described among the risk groups identified by the prognostic score.

We used a prospective phase II clinical trial as an additional fully external validation cohort set to assess CD138 independent prognostic value identified in the final multivariate Cox model of the development set and to validate the discrimination ability of the CD138-LDH score.[20]

Finally, in an exploratory approach, we first investigated the correlation between CD138 and Ang-2 expression in a continuous and binary approach, with Spearman correlation coefficient and χ^2 test, respectively. Then, the potential independent Ang-2 association with survival among the staging system issued from CD138 analysis was assessed by a Cox modelization. The exploratory analysis of Ang-2 was privileged in this study due to the lack of power to explore adequately Ang-2 and CD138 altogether.

Analyses were conducted using SAS 9.2 (Statistical Analysis System, Cary, NC) and R 3.0.2 (R foundation for Statistical Computing). All statistical tests were 2-sided, and probability values ≤ 0.05 were regarded as significant.

Results

Patients' characteristics

Characteristics of the overall population and according to the two cohort sets are given in Table 1. A total of 177 patients were eligible.

Table 1. Baseline characteristics of patients, according to the cohort set

Characteristics	Overall population (*N* = 177)		Development set (*N* = 126)		Validation set (*N* = 51)		p^3
	N		*N*		*N*		
Age (years)[4]	177		126		51		
≤ 65		93 (53%)		66 (52%)		27 (53%)	
>65		84 (47%)		60 (48%)		24 (47%)	0.9461
Patient male sex — *no. (%)*[4]	177	74 (42%)	126	46 (37%)	51	28 (55%)	**0.0246**
Performance status[4]	167		119		48		
0, no. (%)		91 (55%)		64 (54%)		27 (56%)	
>0, no. (%)		76 (45%)		55 (46%)		21 (44%)	0.7719
Primary tumor site[4]	177		126		51		
Colon, no. (%)		120 (68%)		88 (70%)		32 (63%)	
Rectum, no. (%)		57 (32%)		38 (30%)		19 (37%)	0.3602
Metastases localization[4]	177		126		51		
Liver and other, no. (%)		98 (55%)		81 (64%)		17 (33%)	
Liver alone, no. (%)		68 (39%)		45 (36%)		23 (45%)	
Other alone, no. (%)		11 (6%)		0 (0%)		11 (22%)	**<0.0001**
Timing of metastases[4]	173		125		48		
Synchrone, no. (%)		119 (69%)		89 (71%)		30 (63%)	
Metachrone, no. (%)		54 (31%)		36 (29%)		18 (37%)	0.2688
Surgery of the primary tumor, no. (%)[4]	175	110 (63%)	126	82 (65%)	49	28 (57%)	0.3293
Leucocytes ($x10^6$/mL)[1]	137	8.1 ± 3.2	88	8.2 ± 3.0	49	8.1 ± 3.5	0.8926
Lymphocytes ($x10^6$/mL)[1]	132	1.5 ± 0.7	87	1.4 ± 0.6	45	1.7 ± 0.7	0.0550
LDH[4]	115		82		33		
≤350 (ULN)		56 (49%)		36 (44%)		20 (61%)	
>350 (ULN)		59 (51%)		46 (56%)		13 (39%)	0.1050
CD138 (ng/mL)[2]	177	75.3 (48.8–139.7)	126	91.9 (61.8–169.1)	51	37.5 (25.7–70.2)	**<0.0001**
Death event[4]	177	130 (73%)	126	94 (75%)	51	36 (71%)	0.5838
Median F-up time in months 95% CI All patients were F-up until death		56.0 (50.3–59.0)		51.8 (41.9–59.8)		57.9 (53.1–60.3)	**0.0295**

[1]Plus–minus values are means ±SD and the unpaired *t*-test was used for the comparison of variable among groups.
[2]Median and Interquartile range are described and the Wilcoxon rank-sum test was used for the comparison of variable among groups.
[3]*p* Values are for the comparison between cohort set-1 and cohort set-2 populations.
[4]Chi-square or Fisher's exact tests were used for the comparison of categorical variables.
Abbreviations: PS: performance status; ULN: Upper Limit of Normal.

One hundred and twenty six (71%) patients were enrolled in the development cohort set. Primitive tumor was colon in 88 (70%) patients and surgery of the primary tumor was performed in 82 (65%) patients. Eighty-nine (71%) patients had synchronous liver metastases, while 45 (36%) patients had liver exclusive metastatic sites.

Baseline characteristics of the 51 patients recruited in the external validation cohort set were compared with those of the development cohort in Table 1. Patient's characteristics of these two cohorts were similar, except for sex, liver metastatic involvement (more liver alone in the validation set) and CD138 value, which was lower in the validation cohort ($p = 0.0246$, $p < 0.0001$, $p < 0.0001$, respectively).

Kaplan–Meier curves for OS show similar prognostic profiles between these two cohorts (HR = 0.750, 95% CI = 0.509–1.104, $p = 0.14$; Supporting Information Fig. 2).

Methodologically, in order to robustly demonstrate the prognostic value of soluble CD138, we first analyzed the development cohort set and second replied our main results in the validation set.

OS CD138 prognostic value

First, we assessed the association of the pre cited collected parameters with OS in the development set (Table 2). Four variables were identified as statistically significant ($p < 0.05$) prognostic parameters for OS in the univariate analysis: liver

Table 2. Univariate and multivariate analyses for OS in the development set

			Univariate analysis Development set ($n = 126$)			Multivariate analysis Development set ($n = 82$)		
	Patients	Deaths	HR	95% CI	*P*	HR	95% CI	*p*
Age (years)								
<= 65	66	47	1	-	-			
>65	60	47	1.309	[0.872; 1.964]	0.1938			
Sex								
Male	46	32	1	-	-			
Female	80	62	1.362	[0.887; 2.093]	0.1580			
Performance status								
0	64	45	1	-	-			
>0	55	43	1.183	[0.778; 1.799]	0.4315			
Primary tumor site								
Colon	88	67	1	-	-			
Rectum	38	27	1.114	[0.712; 1.744]	0.6358			
Timing to metastasis								
Meta-chronous	36	26	1	-	-			
Synchronous	89	67	1.291	[0.819; 2.035]	0.2708			
Metastases localization								
Liver alone	45	28	1	-	-			
Liver and other	81	66	1.966	[1.260; 3.069]	**0.0029**			
Other alone	-	-	-	-	-			
Surgery of the primary tumor								
Yes	44	37	1	-	-			
No	82	57	1.969	[1.290; 3.004]	**0.0017**			
Leucocytes (x10^6/mL)	88	66	1.084	[0.997; 1.178]	**0.0601**			
Lymphocytes (x10^6/mL)	87	65	0.3406	[0.536; 1.240]	0.3406			
LDH								
≤350 (ULN)	36	21	1	-	-	1		
>350 (ULN)	46	37	2.337	[1.356; 4.030]	**0.0022**	2.165	[1.241; 3.778]	**0.0066**
Log_CD138 (ng/mL)	126	94	2.321	[1.652; 3.260]	**<.0001**	1.706	[1.092; 2.666]	**0.0190**

Abbreviations: PS: performance status; ULN: Upper Limit of Normal.

and other metastatic sites, absence of resection of the primary site, high LDH level and log CD138. Interestingly, leucocyte count was found to have a borderline association with OS ($p = 0.06$) and then was also implemented with those < 0.05 in the Cox multivariate construction regarding its previously recognized prognostic impact.[4]

Then, the Cox multivariate analysis exhibited two significant independent predictors of OS: high LDH level and log CD138 levels (Table 2).

Accuracy of the model was checked regarding two parameters: discrimination and calibration, which measure the ability to separate patients with different prognosis and to provide unbiased survival predictions in groups of similar patients, respectively. Our final multivariable Cox model exhibited good calibration as shown in the calibration plot (Supporting Information Fig. 3) and acceptable discrimination (C-statistic 0.644, 95% CI: 0.570 0.718).

CD138 and OS scoring system

Some prognostic scores are yet available in routine practice and GERCOR score is a validated one, compatible with the clinical practice and of potent interest with the use of biotherapy. Accordingly, we decided to assess if CD138 might impact the discriminative ability of the GERCOR model in 79 patients of the development set in whom LDH, PS and CD138 were available.

As presented in the Table 3 the discriminative ability of GERCOR model was found to be efficient with a C-statistic equal to 0.615. The inclusion of the CD138 parameter significantly improved its discriminative ability, since the C statistic

Table 3. Discriminative ability of different models

			GERCOR model ($n = 79$)			GERCOR model + CD138 ($n = 79$)			Study final multivariate model ($n = 79$)		
	Patients	Deaths	HR	95% CI	p	HR	95% CI	p	HR	95% CI	p
Model estimation											
OMS											
0	41	27	1	-	-	-	-	-			
>0	38	29	1.089	[0.641; 1.848]	0.7528	1.051	[0.619; 1.785]	0.8541			
LDH											
≤350 (ULN)	34	19	1	-	-	1			1		
>350 (ULN)	45	37	2.524	[1.433; 4.446]	**0.0013**	2.354	[1.319; 4.201]	**0.0038**	2.366	[1.328; 4.213]	**0.0035**
Log_CD138 (ng/mL)	79	56				1.858	[1.181; 2.923]	**0.0074**	1.862	[1.184; 2.927]	**0.0071**
Model performance			0.615			0.663			0.664		
Harrell C-index											

Abbreviations: ULN: Upper Limit of Normal.

increased significantly from 0.615 to 0.663 (bootstrap mean difference = 0.048; 95% CI: 0.076–0.087).

Considering that the PS parameter was not significantly associated with OS in our multivariate analysis (Table 2), we then explored the involvement of PS parameter exclusion of the GERCOR + CD138 multivariate model. PS status was found to be detrimental for the discriminative ability of the model, since the C statistic increased significantly from 0.663 to 0.664 (bootstrap mean difference=-0.004; 95% CI: −0.005 to −0.003) highlighting the study final multivariate model with LDH and log CD138 parameters as the most parsimonious one. The combination of CD138 and LDH values was then named "final multivariate model".

In line with these results, we next explored the potential benefit of CD138 and LDH values for routine practice.

CD138 clinical implementation

Use of a CD138 binary cut-off for routine practice. In order to pinpoint the clinical interest of these results, a cut-off value of soluble CD138 was determined to assess its potential value to stratify colon cancer patients' risk of death.

Most of the previously published soluble CD138 measurements showed low median rate value in healthy volunteers, as reported by Jilani *et al.* study where CD138 median levels were 19.86 ng/ml (range, 14.49–33.14), in 32 healthy volunteers.[24] In tumoral context, the median can vary between 40 and 100 ng/mL, with two studies showing median values higher than 100 ng/mL.[18,25–28]

When using the overall material involved in this study, CD138 median value was 75.3 ng/mL (Interquartile range IQR: 48.8–139.7) in the 177 mCRC patients (Supporting Information Fig. 4).

A complementary set of experiments was then done by dosing the CD138 value in the serum of 40 healthy volunteers, who were blood donors in French Blood Bank (*EFS, Bourgogne Franche-Comté*). These volunteers were major (>50 years old), had signed an informed consent and were randomly chosen. Monitoring soluble CD138 in the plasma by ELISA showed a median value of 4.0 ng/mL (IQR= 0.5–6.2) in this control group (Supporting Information Fig. 4) with an observed higher value of 16.2 ng/mL.

In consequence, since no healthy donors had CD138 serum level above 75 ng/mL and according to previous studies, the soluble CD138 median value observed in our cohort was selected as a non-arbitrary cut-off for clinical use.

Among the 177 patients involved in the cut off determination analysis, 90 (51%) had levels of CD138 above 75 ng/mL. Increased CD138 level was associated with metastatic localization ($p = 0.0019$), more leucocyte count ($p = 0.0249$), more lymphopenia ($p = 0.0067$), higher LDH levels ($p = 0.0186$, Supporting Information Table 1).

According to the CD138 value, the median OS of the 177 mCRC patients involved to determine the C138 threshold was better in patients with low levels of CD138 than in patients with high levels of CD138 [median OS= 33.2 months (95% CI:

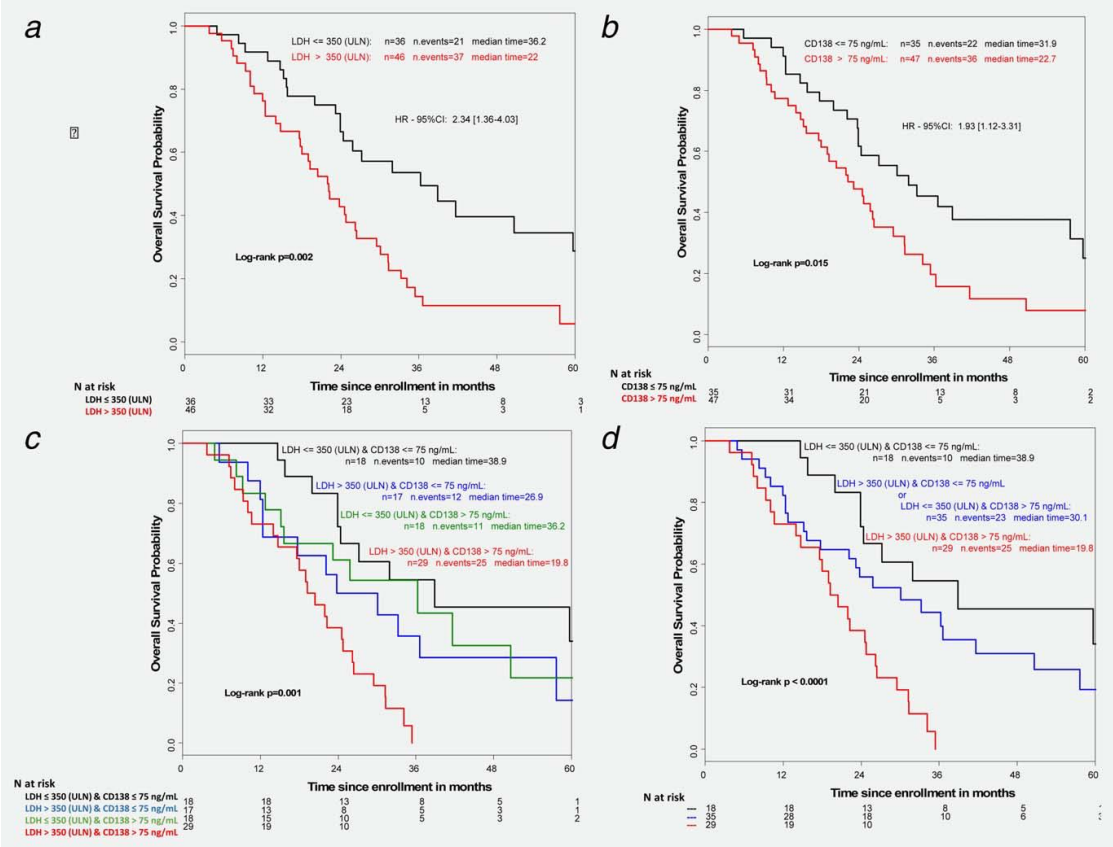


Figure 1. Kaplan–Meier curves for overall survival according to LDH levels (*a*), CD138 levels (*b*), CD138-LDH 4 subgroups (*c*), or CD138-LDH biological score (*d*).

23.9–38.9) *vs.* 18.0 months (95% CI: 15.1–23.2) in the CD138 low and CD138 high levels, respectively, HR = 2.081, 95% CI: 1.464–2.959, $p < 0.0001$, Supporting Information Fig. 5].

New scoring system based on CD138 and LDH values. The independent association of LDH and CD138 values on OS offers the possibility to assess the prognosis of mCRC patients based on easily measurable serum biomarkers and with a potent reproducibility compared to clinical parameters. Then, we investigated the interest to combine CD138 and LDH to predict mCRC patients' prognosis at baseline in the development cohort.

According to the LDH value, the median OS was significantly better in patients with low levels of LDH than in patients with high levels of LDH (median OS= 36.2 months, 95% CI: 23.9–59.7 *vs.* 22.0 months, 95% CI: 17.6–26.2, respectively, HR = 2.34, 95% CI: 1.36–4.03, $p = 0.002$, Fig. 1a).

In the 82 patients with LDH available and according to the CD138 value, the median OS was significantly better in patients with low levels of CD138 than in patients with high levels of CD138 (median OS = 31.9 months (95% CI: 23.8–57.6) *vs.* 22.7 months (95% CI: 15.6–26.4), respectively, HR = 1.93, 95% CI: 1.12–3.31, $p = 0.017$, Fig. 1b).

A new prognostic scoring system was then proposed, by combining LDH and CD138 information.

Four groups were first identified: LDH $\leq$ 350 (ULN) and CD138 $\leq$ 75 ng/mL: Group 0; LDH > 350 (ULN) and CD138 $\leq$ 75 ng/mL: Group 1; LDH $\leq$ 350 (ULN) and CD138 > 75 ng/mL: Group 2, and LDH > 350 (ULN) and CD138 > 75 ng/mL: group 3. Kaplan Meier analysis according to these groups is summarized in Figure 1c.

Based on these data, a prognostic score was then proposed, by pooling the Group 1 and Group 2, characterized by the presence of only one risk factor, and showing similar prognostic profiles. For each patient, three risk groups for death were identified: a low-risk group (Group 0), an intermediate-risk group (Groups 1 and 2) and a high-risk group (Group 3). The median OS was 38.9 months (95% CI: 23.9-NA) in the low-risk group ($n = 18$, 22%), 30.1 months (95% CI: 15.6–41.6) in the intermediate-risk group ($n = 35$,

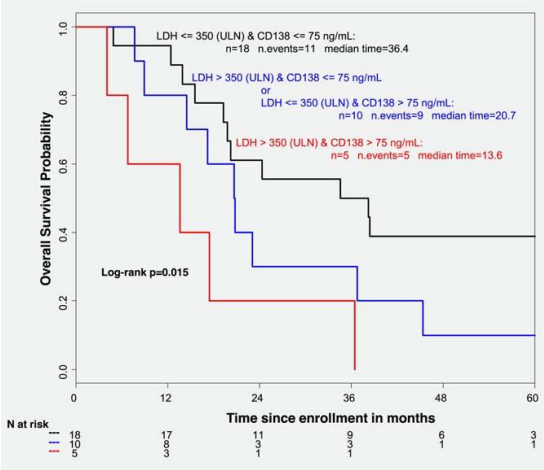


Figure 2. Kaplan–Meier curves for overall survival according to the CD138-LDH biological score in the validation set.

43%) and 19.8 months (95% CI: 14.0–24.8) in the high-risk group ($n = 29$, 35%) (Fig. 1d).

Among the 82 patients of the development set and involved in the scoring system, 78 (95%) were assessed for ORR. Interestingly, among the 17, 34 and 27 patients of the low, intermediate and high risk groups, eight (47%), eleven (32%) and six (22%) presented an ORR, respectively (p= 0.228).

Internal and external validations

Internal validation. A bootstrapping procedure with 1000 samples from the original development dataset confirmed the internal validity and robustness of the final model (biased corrected 95% CI and accelerated bootstrap HRs).

Bootstrapping results for the internal validation reflect the robustness of the final model (LDH: HR = 2.165, HR 95% CI percentile: 1.149–3.929 and CD138 (log value): HR = 1.706, 95% CI percentile for 1.031–2.954).

External validation. First, the final multivariate Cox model proposed in the development set was applied in the external validation set. Log CD138 high level appears to be a significant independent predictor of poorer prognosis for OS (HR = 1.69, 95% CI: 1.049–2.723, $p = 0.0309$).

Second, we applied the CD138-LDH score in the validation set. Similar discriminative ability was observed (C stat = 0.633) and Kaplan–Meier curves showed a statistically difference between the three risk groups with median OS equal to 36.4 months (95% CI: 19.3-NA), 20.7 months (95% CI: 7.6–36.7) and 13.6 months (95% CI: 4.0–36.4), for the low, intermediate and high risk groups, respectively ($p = 0.015$, Fig. 2).

Pooled population investigations

CD138-LDH biological score improves death risk stratification analysis. We demonstrated and externally validated the OS prognostic value interest of CD138 and established a simple prognostic score incorporating this parameter and LDH value (Figs. 1d and 2).

Nowadays, GERCOR score is considered as a well-validated tool, also based on two key parameters (PS and LDH).

In order to apply these two approaches with the largest available material, we decided to pool the two cohorts involved in the study ($n = 177$). Among the 177 patients, 109 (62%) had available LDH, PS and CD138 values. Figure 3 exhibited the results for these two approaches.

Interestingly, in our study the combination of CD138 with LDH allowed a better discrimination than the GERCOR approach, since the median OS of the low, intermediate and high risk groups were 36.2 5, 26.2 and 20.4 months ($p = 0.002$) and 38.4, 23.8 and 19.1 months (($p < 0.0001$) in the GERCOR and in the C138-LDH biological scores, respectively. This observation was reinforced with a better discrimination index for the score based on CD138 and LDH (C= 0.629 *vs.* 0.594).

ORR and PFS assessments among risk groups identified by the biological score. Chemotherapy response could be also assessed in the overall population. Among the 115 patients of the overall population set and involved in the scoring system, 111 (97%) were assessed for ORR. Interestingly, among the 35, 44 and 32 patients of the low, intermediate and high-risk groups, 19 (54%), 16 (36%) and 9 (28%) had an ORR, respectively (p= 0.0778). We also observed a strong significant difference in median PFS among the three prognostic groups identified in the OS analysis: 22 (95% CI: 15.3–34.5), 12 (95% CI: 8.7–15.2) and 8.4 (95% CI: 5.3–11.7) months in the low, intermediate and high risk groups, respectively (log-rank $p < 0.0014$, Supporting Information Fig. 6).

Exploratory analysis of Ang-2 measurement on CD138. Ang-2 was previously described to be associated with mCRC patients'OS.[9] While CD138 expression is related to cancer invasion and extracellular matrix regulation, Ang-2 is a biomarker of neoangiogenesis. Then, we next decided to address the potential interactions between these biomarkers among the 177 patients enrolled in the two cohorts.

Unexpectedly, we first observed a positive, but not complete correlation between CD138 and Ang-2 (Spearman correlation coefficient equal to 0.56; $p < 0.0001$; Supporting Information Fig. 7). When considering Ang-2 and CD138 parameters in a binary approach (Ang-2:±5 ng/mL and CD138:±75 ng/mL), we observed 39 patients (22%) with high CD138 and low Ang-2 values. Inversely, 16 patients (9%) had low CD138 levels when Ang-2 was measured above 5 ng/mL, suggesting that these two markers could provide complementary information.

Regarding the limited number of patients included in the present study, we cannot definitively rule out a possible additive prognostic value of Ang-2. In an exploratory analysis, we observed a favourable tendency for Ang-2 independent association with OS, when adding Ang-2 information to the CD138-LDH biological score (Supporting Information Table 2, $p = 0.07$). The precise additive value of Ang-2 monitoring

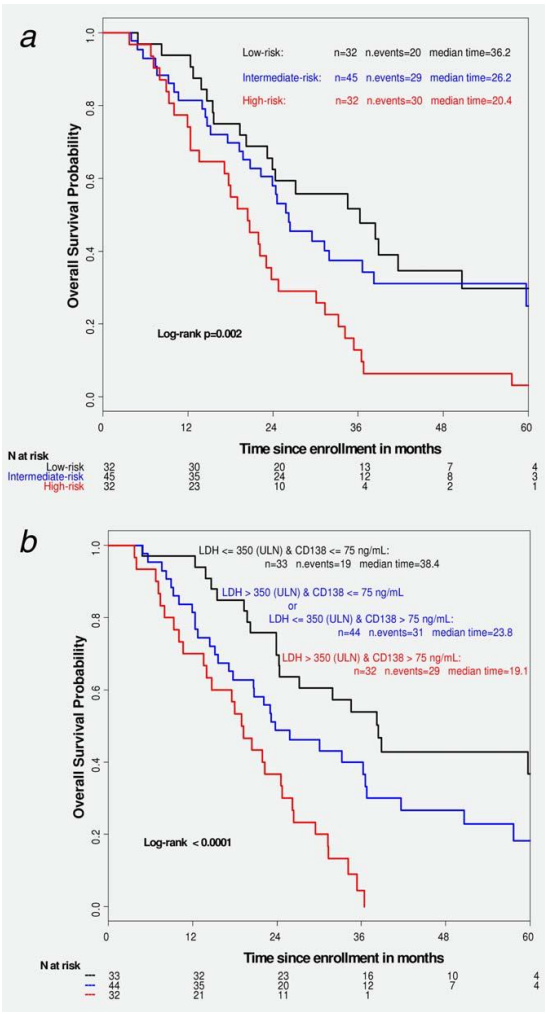


Figure 3. Kaplan–Meier curves for overall survival in the pooled cohorts. Patients were categorized into three risk groups with significantly different prognostic. Overall survival prediction in the pooled cohorts according to the GERCOR score (panel ***a***, Harell C-index = 0.593): low ($N = 32$; reference group), intermediate ($N = 45$; HR = 1.31, 95% CI, 0.76–2.26; $p = 0.328$) and a high-risk group ($N = 32$; HR = 2.55, 95% CI, 1.46–4.45; $p = 0.001$). Overall survival prediction in the pooled cohorts according to the CD138 – LDH biological score (panel ***b***, Harell C-index =0.629): low ($N = 33$; reference group), intermediate ($N = 44$; HR = 1.80, 95% CI, 1.02–3.20; $p = 0.04$) and a high-risk group ($N = 32$; HR = 3.85, 95% CI, 2.10–7.08; $p < 0.0001$).

combined to CD138-LDH score is currently addressed in a prospective study.

Discussion

The present study provides concrete answer for the unmet need of the soluble CD138 prognostic value in mCRC patients. Moreover, the CD138-LDH biological score enables the identification of a high risk group involving about 30% of the entire population without any alive patients at 36 months.

In 52 localized colorectal cancer patients, Wang and colleagues recently described that the disease-free survival of patients with high preoperative serum CD138 was significantly shorter than the disease-free survival of patients with low soluble CD138.[18] Our results are in line with this pioneering study by demonstrating the association of serum baseline CD138 with OS and its discrimination ability, in mCRC patients. Moreover, we took advantage of the reproducibility and robustness of baseline serum CD138 dosage to propose a biological prognostic score, which might improve the ability of oncologists to stratify mCRC patients' risk of death.

However, our study exhibits some limitations. First, we could not perform any comparison with the molecular signatures recently described.[6,29] These new CRC molecular classifications include immune system and angiogenesis components, thus taking into account the tumoral microenvironment. Although these molecular classifications greatly improved our understanding of colon cancer oncogenesis, these studies mainly included localized colon cancers and further investigations should explore how these molecular features of colon cancers are to take into account regarding the efficacy of chemotherapy. In addition, RAS and BRAF status were not fully available in our study. Many studies, and a recent meta-analysis, showed the negative impact of BRAF (V600E) mutation on OS and PFS of mCRC patients.[8] However, BRAF mutation analysis allows OS prediction in <10% of mCRC patients, while CD138-LDH biological score enables the identification of a high-risk group involving about 30% of the entire population. RAS mutations might also be taken into account to discriminate mCRC patients' OS.[30] Some studies have evidenced that CD138 shedding activates RAS/MapK signalling.[18] Then, it would be of interest to characterize the influence of soluble CD138 according to the molecular status of mCRC. The EGFR status, microsatellite assessment or cytokine measurement were not available as well in the present study. A prospective clinical trial is ongoing to address this issue (Epitopes-CRC02, *EudraCT* A01377-36). Our preliminary results among the 58 patients with available RAS mutation status showed a mutation rate of 56% and 48% in patients with high and low soluble CD138 levels, respectively. One BRAF mutation was identified in each group. Then, a correlation between RAS status and soluble CD138 levels seems unlikely.

Second, all patients in our cohorts were treated with bevacizumab and chemotherapy, excluding the possibility to analyze the predictive value of CD138. Recently, LDH soluble value was suggested to be predictive of bevacizumab efficacy, in first line mCRC patients, as bevacizumab administration was correlated to PFS in patients having high LDH levels.[31] Other studies also suggested a correlation between LDH levels with bevacizumab efficacy.[32,33]

Third, comorbidities were not considered in this study although they could also influence survival.

In spite of these limitations, the validation of soluble CD138 independent association with OS in mCRC patients opens the way to complementary explorations.

Compared to tissue biomarkers, soluble prognostic factors, easily measurable in routine practice, can provide relevant information for the clinical decision-making process regardless tumor heterogeneity. Moreover, serum biomarkers can be monitored during the course of the disease, to predict clinical outcomes or treatment effectiveness without requiring iterative invasive biopsies. In line with this decision's making objective, the CD138 measurement was done at baseline, before treatment initiation.

Most studies addressing the prognostic value of soluble CD138 selected thresholds ranging from 40 to 100 ng/mL.[25–27,34] All these cutoff values were established in studies pooling localized and metastatic solid tumors. Here, we demonstrated that normal volunteers [mean age of 55.6 (range 50–60)], never displayed soluble CD138 levels above 20 ng/mL (Supporting Information Fig. 4). We then established the 75 ng/ml value as a relevant cutoff to discriminate mCRC patients for OS (Supporting Information Fig. 4), which makes this threshold of clinical interest.

In the present study, we propose a new scoring system, based on two biological parameters easy to measure at diagnosis, and without any clinical factor. Other studies recently addressed the seric assessment of mCRC patient's prognostic, which have to be compared with this new biological scoring.[35] The CD138-LDH biological score was found to be efficient with good discrimination ability (Harrell's C index = 0.63), at least equivalent to the GERCOR score in term of discriminative ability, since not close to the C = 1 perfectly discriminating model. The importance for the implementation of the score in the future, in order to prevent patients risk profile misclassification, is then underlined. The biological score then separates three groups of low (30%), intermediate (40%) and high (30%) risk patients, having median OS of 38.9, 30.1 and 19.8 months, respectively. In clinical trials with selected patients (PS 0 and 1) it could be of interest to guide survival prediction by biological parameters only, or when available, with the addition of quality of life parameter.

One surprising observation is the reclassification of LDH high patients in a better prognostic group when CD138 soluble level is low, while none of the patients with high LDH and CD138 levels are alive after 3 years of follow-up. A better understanding of the mechanisms underlying CD138 influence on clinical outcomes might allow further stratification of mCRC patients for individualized therapies. LDH is known to be involved in tumor initiation and metabolism and can be a tumor burden reflects. Its LDH-A subunit can also be produced during angiogenesis, since HIF1α binds LDH-A promoter and enhances its transcription.[36] The increased levels of soluble CD138 observed in patients with high LDH (Supporting Information Table 1) might suggest an implication of CD138 shedding with angiogenesis in mCRC. The correlation between CD138 and the angiogenic marker Ang-2, first described, to our knowledge, in this study, also deserve that hypothesis. Previous studies have reported that the shedding of CD138 by heparanase is associated with enhanced levels of angiogenesis in multiple myeloma. One of the mechanism accounting for CD138 role in angiogenesis is the capture of VEGF by CD138 leading to the formation of VEGFR-stimulating complexes anchored on the extracellular matrix.[37]

Another possible hypothesis supporting the improved prognosis of mCRC associated with low baseline CD138 values might be that these tumors exhibit less invasive properties. Indeed, CD138 displays an integrin binding-site which engage αvβ3 integrins both on cancer and endothelial cells leading to IGFR1 docking and activation. This CD138 mediated integrin regulation was shown to promote invasion and angiogenesis.[38] The potential role of CD138 in EMT was reported in other studies.[14,17,39,40]

Lastly, the poor outcomes observed in mCRC patients with high levels of soluble CD138 might be related to a decreased effectiveness of chemotherapy. A previous report explored this hypothesis by treating primary colon cancer samples taken from patients with high or low preoperative serum CD138 levels. Cancer cells derived from patients with high levels of soluble CD138 were less sensitive to chemotherapy *ex vivo*.[18] One hypothesis concerning CD138 shedding worst prognostic could then be chemo-resistance. Accordingly, it would be interesting to validate our preliminary data concerning the influence of soluble CD138 levels on the overall response rate (ORR), in future clinical trials. This hypothesis will also be addressed in the Epitopes-CRC02 trial.

In conclusion, soluble CD138 is an independent prognostic factor in mCRC patients. The assessment of CD138-LDH biological score at baseline offers an improved death risk stratification capacity and could provide a basis for early and adapted therapeutic interventions. The basis underlying CD138 worst prognosis is not clearly understood and has to be further addressed.

References

1. Fakih MG. Metastatic Colorectal Cancer: Current State and Future Directions. *J Clin Oncol* 2015; 33:1809–1824.
2. Kohne CH, Cunningham D, Di Costanzo F, et al. Clinical determinants of survival in patients with 5-fluorouracil-based treatment for metastatic colorectal cancer: results of a multivariate analysis of 3825 patients. *Ann Oncol* 2002; 132:308–317.
3. Chibaudel B, Bonnetain F, Tournigand C, et al. Simplified prognostic model in patients with oxaliplatin-based or irinotecan-based first-line chemotherapy for metastatic colorectal cancer: a GERCOR study. *Oncologist* 2011; 169:1228–1238.
4. Desot E, de Mestier L, Volet J, et al. Prognostic factors in patients with non resectable metastatic colorectal cancer in the era of targeted biotherapies: relevance of Kohne's risk classification. *Dig Liver Dis* 2013; 454:330–335.
5. Schirripa M, Bergamo F, Cremolini C, et al. BRAF and RAS mutations as prognostic factors

in metastatic colorectal cancer patients undergoing liver resection. *Br J Cancer* 2015; 11212:1921–1928.
6. Popovici V, Budinska E, Tejpar S, et al. Identification of a poor-prognosis BRAF-mutant-like population of patients with colon cancer. *J Clin Oncol* 2012; 3012:1288–1295.
7. Sorbye H, Dragomir A, Sundstrom M, et al. High BRAF Mutation Frequency and Marked Survival Differences in Subgroups According to KRAS/BRAF Mutation Status and Tumor Tissue Availability in a Prospective Population-Based Metastatic Colorectal Cancer Cohort. *PLoS One* 2015; 106:e0131046
8. Venderbosch S, Nagtegaal ID, Maughan TS, et al. Mismatch repair status and BRAF mutation status in metastatic colorectal cancer patients: a pooled analysis of the CAIRO, CAIRO2, COIN, and FOCUS studies. *Clin Cancer Res* 2014; 2020: 5322–5330.
9. Jary M, Vernerey D, Lecomte T, et al. Prognostic Value of Angiopoietin-2 for death risk stratification in Patients with Metastatic Colorectal Carcinoma. *Cancer Epidemiol Biomarkers Prev* 2015; 24:603–12.
10. Mitselou A, Skoufi U, Tsimogiannis KE, et al. Association of syndecan-1 with angiogenesis-related markers, extracellular matrix components, and clinicopathological features in colorectal carcinoma. *Anticancer Res* 2012; 329:3977–3985.
11. Yang N, Mosher R, Seo S, et al. Syndecan-1 in breast cancer stroma fibroblasts regulates extracellular matrix fiber organization and carcinoma cell motility. *Am J Pathol* 2011; 1781:325–335.
12. Maeda T, Desouky J, Friedl A. Syndecan-1 expression by stromal fibroblasts promotes breast carcinoma growth in vivo and stimulates tumor angiogenesis. *Oncogene* 2006 259:1408–1412.
13. Day RM, Hao X, Ilyas M, et al. Changes in the expression of syndecan-1 in the colorectal adenoma-carcinoma sequence. *Virchows Arch* 1999; 4342:121–125.
14. Leppa S, Vleminckx K Van Roy F, eta. Syndecan-1 expression in mammary epithelial tumor cells is E-cadherin-dependent. *J Cell Sci* 1996; 109: 1393–1403.
15. Seidel C, Sundan A, Hjorth M, et al. Serum syndecan-1: a new independent prognostic marker in multiple myeloma. *Blood* 2000; 952:388–392.
16. Akl MR, Nagpal P, Ayoub NM, et al. Molecular and clinical profiles of syndecan-1 in solid and hematological cancer for prognosis and precision medicine. *Oncotarget* 2015;
17. Hashimoto Y, Skacel M, Adams JC. Association of loss of epithelial syndecan-1 with stage and local metastasis of colorectal adenocarcinomas: an immunohistochemical study of clinically annotated tumors. *BMC Cancer* 2008; 8:185
18. Wang X, Zuo D, Chen Y, et al. Shed Syndecan-1 is involved in chemotherapy resistance via the EGFR pathway in colorectal cancer. *Br J Cancer* 2014; 11110:1965–1976.
19. Goujon G, Palanchon P, Ladam-Marcu V, et al. Evaluation precoce et quantitative de l'effet pharmacodynamique du bevacizumab au moyen de l'echographie de contraste dans les metastases hepatiques de cancer colorectal. *JFHOD* 2010; 200:
20. Kim S, Dobi E, Jary M, et al. Bifractionated CPT-11 with LV5FU2 infusion (FOLFIRI-3) in combination with bevacizumab: clinical outcomes in first-line metastatic colorectal cancers according to plasma angiopoietin-2 levels. *BMC Cancer* 2013; 13:611
21. Schuster DPPW, Castro M, et al. Translational and Experimental Clinical Research. *Lippincott Williams & Wilkins* 2005;
22. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials* 1996; 174:343–346.
23. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; 154: 361–387.
24. Jilani I, Wei C, Bekele BN, et al. Soluble syndecan-1 (sCD138) as a prognostic factor independent of mutation status in patients with chronic lymphocytic leukemia. *Int J Lab Hematol* 2009; 311:97–105.
25. Joensuu H, Anttonen A, Eriksson M, et al. Soluble syndecan-1 and serum basic fibroblast growth factor are new prognostic factors in lung cancer. *Cancer Res* 2002; 6218:5210–5217.
26. Nault JC, Guyot E, Laguillier C, et al. Serum proteoglycans as prognostic biomarkers of hepatocellular carcinoma in patients with alcoholic cirrhosis. *Cancer Epidemiol Biomarkers Prev* 2013; 228:1343–1352.
27. Szarvas T, Reis H, Kramer G, et al. Enhanced stromal syndecan-1 expression is an independent risk factor for poor survival in bladder cancer. *Hum Pathol* 2014; 454:674–682.
28. Anttonen A, Leppa S, Ruotsalainen T, et al. Pretreatment serum syndecan-1 levels and outcome in small cell lung cancer patients treated with platinum-based chemotherapy. *Lung Cancer* 2003; 412:171–177.
29. Marisa L, de Reynies A, Duval A, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med* 2013; 105: e1001453
30. Richman SD, Seymour MT, Chambers P, et al. KRAS and BRAF mutations in advanced colorectal cancer are associated with poor prognosis but do not preclude benefit from oxaliplatin or irinotecan: results from the MRC FOCUS trial. *J Clin Oncol* 2009; 2735:5931–5937.
31. Passardi A, Scarpi E, Tamberi S, et al. Impact of Pre-Treatment Lactate Dehydrogenase Levels on Prognosis and Bevacizumab Efficacy in Patients with Metastatic Colorectal Cancer. *PLoS One* 2015; 108:e0134732
32. Yin C, Jiang C, Liao F, et al. Initial LDH level can predict the survival benefit from bevacizumab in the first-line setting in Chinese patients with metastatic colorectal cancer. *Onco Targets Ther* 2014; 7:1415–1422.
33. Silvestris N, Scartozzi M, Graziano G, et al. Basal and bevacizumab-based therapy-induced changes of lactate dehydrogenases and fibrinogen levels and clinical outcome of previously untreated metastatic colorectal cancer patients: a multicentric retrospective analysis. *Expert Opin Biol Ther* 2015; 152:155–162.
34. Mundt F, Heidari-Hamedani G, Nilsonne G, et al. Diagnostic and prognostic value of soluble syndecan-1 in pleural malignancies. *Biomed Res Int* 2014; 2014:419853
35. Chen ZY, He WZ, Peng LX, et al. A prognostic classifier consisting of 17 circulating cytokines is a novel predictor of overall survival for metastatic colorectal cancer patients. *Int J Cancer* 2015; 136: 584–92.
36. Granchi C, Bertini S, Macchia M, et al. Inhibitors of lactate dehydrogenase isoforms and their therapeutic potentials. *Curr Med Chem* 2010; 177: 672–697.
37. Purushothaman A, Uyama T, Kobayashi F, et al. Heparanase-enhanced shedding of syndecan-1 by myeloma cells promotes endothelial invasion and angiogenesis. *Blood* 2010; 11512:2449–2457.
38. Beauvais DM, Rapraeger AC. Syndecan-1 couples the insulin-like growth factor-1 receptor to inside-out integrin activation. *J Cell Sci* 2010; 123Pt 21:3796–3807.
39. Reiland J, Sanderson RD, Waguespack M, et al. Heparanase degrades syndecan-1 and perlecan heparan sulfate: functional implications for tumor cell invasion. *J Biol Chem* 2004; 2799:8047–8055.
40. Fujiya M, Watari J, Ashida T, et al. Reduced expression of syndecan-1 affects metastatic potential and clinical outcome in patients with colorectal cancer. *Jpn J Cancer Res* 2001; 9210:1074–1081.

IV.1.3 Développement et validation d'un nomogramme et d'un score pronostique pour prédire la survie globale sous seconde ligne de chimiothérapie chez des patients atteints d'un cancer pancréatique avancé.

Vienot A, Beinse G, Louvet C, de Mestier L, Meurisse A, Fein F, Heyd B, Cleau D, d' Engremont C, Dupont-Gossart A-C, Lakkis Z, Tournigand C, Bouché O, Rousseau B, Neuzillet C, Bonnetain F, Borg C**,** Vernerey D
Development and validation of a prognostic nomogram and score to predict overall survival and usefulness of second-line chemotherapy in advanced pancreatic cancer. **Journal of the National Cancer Institute in revision**.

IV.1.3.1 Positionnement du problème

Le cancer pancréatique (PC) est la quatrième cause de décès par cancer dans les pays industrialisés et est en mesure de devenir la seconde en 2030 si les tendances actuelles se confirment. Le PC avancé (non résécable : localement avancé ou métastatique ; aPC) représente 80% des cas au diagnostic. Ces patients ont un pronostic extrêmement sombre puisqu'ils ont une médiane de survie estimée entre 9 et 15 mois en situation localement avancée et entre 6 et 9 mois en situation métastatique. La stratégie de traitement de ces patients repose sur la chimiothérapie. Leur prise en charge en première ligne est relativement bien établie et standardisée avec deux combinaisons thérapeutiques qui ont montrées un bénéfice de survie globale : le FOLFIRINOX et l'association gemcitabine plus *nab*-paclitaxel, en comparaison à la gemcitabine seule. Cette dernière est restée pendant longtemps le traitement de référence dans cette pathologie. Pour simplifier, ces deux nouvelles chimiothérapies sont administrées chez des patients en relativement bon état général puisqu'elles sont associées à une augmentation des toxicités. Les patients atteints d'aPC qui ne peuvent pas recevoir ces

traitements sont classiquement traités par chimiothérapie à base de gemcitabine.

Au-delà de la première ligne de chimiothérapie, environ la moitié des patients atteints d'un aPC restent en bon état général, et peuvent donc a priori recevoir des lignes ultérieures de chimiothérapie.
Néanmoins, tous les patients traités en seconde ligne par chimiothérapie (L2) ne semblent pas en tirer un bénéfice. Ainsi l'administration d'une L2 doit donc être discutée de manière individuelle. À noter, qu'aujourd'hui les éléments pour guider la sélection des patients pour une L2 après progression sous FOLFIRINOX ou gemcitabine plus *nab*-paclitaxel sont rares et il n'y a pas de stratégie consensuelle validée à ce jour.
L'âge, le PS, et la durée de première ligne (L1) sont des «paramètres pragmatiques», qui sont fréquemment utilisés en réunion de concertation pluridisciplinaire afin d'estimer le bénéfice potentiel d'une L2, mais le niveau de preuve est faible. Ainsi, l'identification des facteurs pronostiques d'intérêt pour améliorer la stratification du risque de décès pour ces patients sous L2 est un enjeu majeur afin d'améliorer les éléments disponibles pour la prise de décision thérapeutique. Dans ce contexte, il est urgent de développer des outils pronostiques utiles qui peuvent aider à estimer la survie globale (SG) des patients afin de guider les cliniciens dans leur prise de décision d'administrer d'une L2 et d'optimiser la conception des futurs essais cliniques.

Cette étude a pour objectif, dans un premier temps, de proposer et valider un nouveau modèle pronostique pour la SG chez des patients sous L2 atteints d'un aPC, sur la base d'un spectre de paramètres très important dans une population extrêmement bien « phénotypée ». Dans un second temps, disposant des patients n'ayant pas bénéficié d'un traitement de chimiothérapie sous L2 et ayant été inclus dans l'étude, cela nous a permis d'investiguer le bénéfice de recevoir la chimiothérapie à travers les différents groupes de pronostic de survie globale identifiés

IV.1.3.2 Population de l'étude

La cohorte de développement utilisée dans cette étude est constituée de tous les patients traités entre 2003 et 2013 au Centre Hospitalier Universitaire de Besançon pour un aPC.
Les patients ont été considérés comme éligibles à l'évaluation médicale posant l'indication d'une L2 s'ils avaient déjà reçu une ligne de chimiothérapie, y compris la chimiothérapie adjuvante en cas de rechute tumorale post-opératoire précoce (dans les 6 mois après la dernière administration de ce traitement), et s'ils n'étaient pas morts pendant la L1.
La cohorte de développement comporte donc tous les patients qui ont reçu une L2. L'information concernant les patients ayant reçu une première ligne et toujours en vie à la fin de cette L1 mais n'ayant pas reçu de traitement de L2 a aussi été recueillie.
La cohorte de validation externe inclus tous les patients atteints d'un aPC traités à l'Hôpital Henri Mondor, au Centre Hospitalier Universitaire de Reims ainsi qu'à l'Institut Mutualiste Montsouris entre 2010 et 2016 dans le cadre d'une L2.
Les patients des deux cohortes ont été pris en charge selon les protocoles cliniques standards en vigueur pour l'administration des L1 et L2.
Les données démographiques, liées aux antécédents de cancer, clinico-pathologiques, biologiques et radiologiques ont été collectées au début de la L2 (à la fin de la L1 pour ceux qui n'ont pas reçu de L2).

IV.1.3.3 Approche pronostique méthodologique

La SG sous L2 a été calculée comme le délai entre la date de progression sous L1 (ou la date de première administration de la L2 si aucune progression n'a été observée sous la L1) et le décès toutes causes confondues. Les patients en vie à la fin du suivi ont été censurés et renseignés comme étant vivants à cette date. La SSP sous L2 a été calculée comme le délai entre la date de progression sous L1 (ou la date de première administration de la L2 si

aucune progression n'a été observée sous L1) et la date de progression ou de mort toutes causes confondues ou la date de dernières nouvelles étant la date de censure si aucun évènement n'est renseigné avant cette date. La SG et la SSP ont été estimées par la méthode de Kaplan-Meier et décrites à l'aide de la médiane de survie ou de taux de survie aux temps d'intérêt accompagnés de leurs intervalles de confiances à 95%. Le suivi de l'étude a été estimé selon la méthode du Kaplan-Meier inverse.

La construction du modèle de Cox multivarié final pour prédire la SG a d'abord impliqué une étape univariée afin d'identifier les paramètres associés avec la SG puis, après avoir examiné la colinéarité potentielle entre les variables avec une matrice de corrélation, une étape multivariée dans laquelle tous les facteurs identifiés avec une p-value inférieure à 0.05 en univarié ont été considérés pour être inclus dans un même modèle de Cox. Afin de garder le maximum d'information toutes les variables quantitatives ont été utilisées dans leur forme continue. Les transformations réalisées lorsque nécessaires ont été validées par deux approches statistiques : les splines restreintes cubiques et les polynômes fractionnaires. En l'absence d'hypothèse clinique et ou de documentation dans la littérature aucune interaction n'a été considérée dans la construction du modelé multivarié final.

La validité et les performances du modèle multivarié ont été évaluées selon deux paramètres : la discrimination et la calibration. La valeur prédictive du modèle et sa capacité à discriminer la SG et donc à séparer les patients avec un pronostic différent ont été évalués avec l'indice de concordance de Harrell (C de Harrell). Mille échantillons aléatoires de la population de développement ont été obtenus par procédure de ré-échantillonnage (bootstrap) et ont été ensuite utilisés pour construire un intervalle de confiance à 95% autour du C de Harrell.

La calibration du modèle, quant à elle, se réfère à sa capacité à fournir des prédictions non biaisées dans des groupes de patients considérés comme similaires. Elle a été analysée graphiquement par une courbe de calibration.

Une validation interne du modèle a été proposée à l'aide d'une procédure de ré-échantillonnage en produisant des intervalles de confiance autour des estimations du modèle final.

L'amélioration de la capacité à discriminer la SG de notre modèle multivarié final par rapport à une approche dite classique, qui repose sur des paramètres pragmatiques tels que l'âge, le PS et la durée de L1, a été évaluée par l'indice de concordance de Harrell (C-index). Nous avons aussi utilisé les indices de reclassification tels que le NRI continu et l'IDI pour quantifier la performance dans la reclassification des probabilités de risque de décès à 6 mois après le début de L2, apportée par l'ajout des autres paramètres d'intérêt comparés à l'approche dite classique.

Le modèle multivarié final a été ensuite utilisé pour produire un nomogramme permettant l'estimation, pour un patient donné, de sa médiane et de ses probabilités individuelles de survie à 6, 12, 18 et 24 mois au début de la L2.

Une approche populationnelle a aussi été proposée en dérivant un score pronostique à partie des points du nomogramme. L'identification de trois groupes de pronostics différents a été investigué en définissant des valeurs seuils déterminées à l'aide de la méthode dite de Cox (1957). La capacité discriminante de cette approche a été évaluée par le C de Harrell en considérant la classification des patients dans les différents groupes à risque.
Le score pronostique développé chez les patients ayant reçu une L2 a été appliqué aux patients qui ne l'ont pas reçu afin d'évaluer le bénéfice du traitement dans chacun des groupes identifiés.

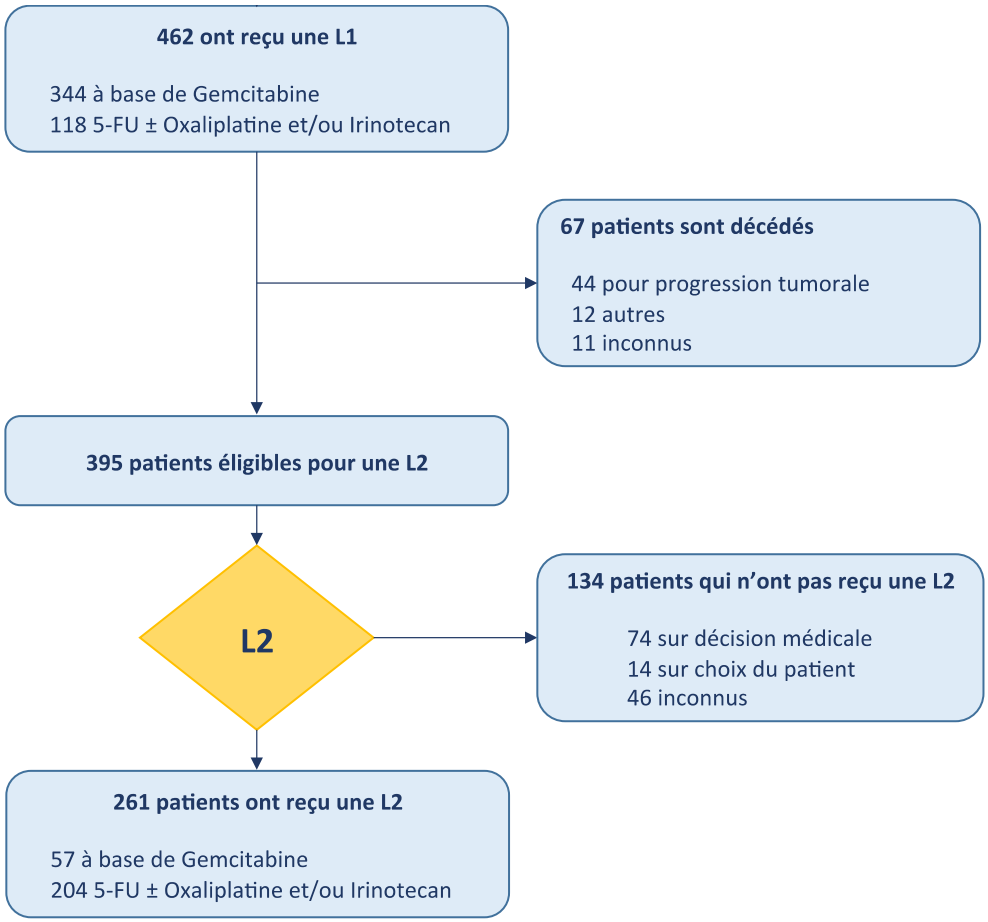
Enfin, en guise de validation externe, le même algorithme pour l'identification des groupes issus de la cohorte de développement a été appliqué dans la cohorte de validation. La capacité discriminante du modèle multivarié final et du score pronostique a été confirmée dans la cohorte de validation externe par le C de Harrell en considérant respectivement la prédiction de survie issue du modèle et la classification des patients dans un groupe à risque.

L'évaluation du bénéfice clinique apporté par la considération des prédictions issues du modèle final afin d'orienter la prise de décisions thérapeutiques a été évaluée par une analyse de courbe de décision dans les deux cohortes (développement et validation).

IV.1.3.4 Résultats

Parmi les 462 patients inclus dans l'étude qui avaient reçu une L1, 395 étaient éligibles pour une évaluation médicale posant l'indication de L2; 261 (66%) d'entre eux ont effectivement reçu une L2 et ont été inclus dans la cohorte de développement, tandis que 134 (34%) patients n'ont pas été traités (Figure 1).

Figure 1: « Flow chart » de l'étude



Comme attendu les caractéristiques des patients à la fin de la L1 dans les groupes traités et non traités par une L2 sont significativement différentes pour l'âge, l'ECOG PS, et la durée de L1 (tableau 1) puisque ce sont les paramètres pragmatiques actuellement utilisés en réunion de concertation pluridisciplinaire, pour discuter de l'état général du patient et donc de la possibilité de lui administrer réellement une L2.

Tableau 1: Caractéristiques des patients de la cohorte de développement selon l'administration ou non d'une L2.

		Population éligible à une L2 N = 395	Population éligible ayant reçu une L2 N = 261	Population éligible n'ayant pas reçu une L2 N = 134	p[a]
Paramètres démographiques	**Age, médiane [IQR], années**	67.2 [60.8 - 74.9]	65.8 [60.3 - 72.9]	71.5 [63.2 - 76.9]	**<.0001**
	Sexe, N (%)				.3824
	Homme	233 (59)	158 (61)	75 (56)	
	Femme	162 (41)	103 (39)	59 (44)	
	Statut tabagique, N (%)				.3022
	Non-fumeur ou ancien fumeur	296 (75)	200 (77)	96 (72)	
	Fumeur actif	97 (25)	60 (23)	37 (28)	
	Données manquantes	2	1	1	
Paramètres pathologiques au diagnostic de cancer	**Localisation tumorale primitive, N (%)**				.2540
	Tête	237 (61)	152 (59)	85 (65)	
	Corps et/ou queue	152 (39)	106 (41)	46 (35)	
	Données manquantes	6	3	3	
	Stade tumoral, N (%)				.4153
	Résecable	92 (23)	66 (25)	26 (19)	
	Localement avancé	96 (24)	61 (24)	35 (26)	
	Métastatique	207 (53)	134 (51)	73 (55)	
Extension tumorale au début de L2	**Nombre de sites métastatiques, N (%)**				.8655
	0 - 1	244 (62)	162 (62)	82 (61)	
	≥ 2	151 (38)	99 (38)	52 (39)	
	Métastases hépatiques, N (%)				.3497
	Non	167 (42)	106 (41)	61 (46)	
	Oui	228 (58)	155 (59)	73 (54)	
	Carcinose péritonéale, N (%)				.6475
	Non	242 (61)	162 (62)	80 (60)	
	Oui	153 (39)	99 (38)	54 (40)	
	Métastases pulmonaires, N (%)				.4297
	Non	304 (77)	204 (78)	100 (75)	
	Oui	91 (23)	57 (22)	34 (25	
	Métastases osseuses, N (%)				.1403
	Non	377 (95)	252 (97)	125 (93)	
	Oui	18 (5)	9 (3)	9 (7)	
Paramètres cliniques au début de L2	**PS (OMS), N (%)**				**<.0001**
	0	52 (15)	47 (19)	5 (6)	
	1	126 (38)	118 (47)	8 (9)	
	≥ 2	157 (47)	86 (34)	71 (85)	
	Données manquantes	60	10	50	
	Indice de masse corporelle, N (%),kg/m²				**.0201**
	Normal [18,5-25[	225 (60)	167 (64)	58 (49)	
	Maigreur <18,5	66 (17)	39 (15)	27 (23)	
	Surpoids [25-30[et obésité ≥30	87 (23)	54 (21)	33 (28)	
	Données manquantes	17	1	16	
	Douleur[b], N (%)				.5078

	Non	165 (49)	119 (48)	46 (52)	
	Oui	174 (51)	131 (52)	43 (48)	
	Données manquantes	56	11	45	
	Ictère, N (%)				.0728
	Non	280 (83)	212 (85)	68 (76)	
	Oui	59 (17)	38 (15)	21 (24)	
	Données manquantes	56	11	45	
	Ascite, N (%)				.9220
	Non	274 (80)	202 (80)	72 (80)	
	Oui	67 (20)	49 (20)	18 (20)	
	Données manquantes	54	10	45	
Traitement avant L2	**Résection tumorale primitive, N (%)**				**.0224**
	Non	111 (28)	83 (32)	28 (21)	
	Oui	284 (72)	178 (68)	106 (79)	
L1	**Protocole en L1, N (%)**				**.0094**
	5-FU ± Oxaliplatine and/or Irinotecan	99 (25)	76 (29)	23 (17)	
	À base de Gemcitabine	296 (75)	185 (71)	111 (83)	
	Meilleur réponse selon RECIST, N (%)				.1428
	Réponse complète ou partielle ou stabilité	223 (65)	170 (67)	53 (59)	
	Progression tumorale	119 (35)	82 (33)	37 (41)	
	Données manquantes	53	9	44	
	Durée de L1, médiane [IQR], mois	4.7 [2.3 - 8.1]	5.7 [2.9 - 8.7]	2.5 [0.8 - 5.8]	**<.0001**
	Cause d'arrêt de L1, N (%)				.9353
	Progression tumorale	265 (67)	176 (67)	89 (67)	
	Toxicité	22 (6)	15 (6)	7 (5)	
	Autres	108 (27)	70 (27)	38 (28)	

Abréviations: IQR=Interquartile Range, OMS=Organisation Mondiale de la Santé, L1= chimiothérapie de 1ere ligne, L2= chimiothérapie de 2nd ligne

[a] χ^2 ou test exact de Fisher ont été utilisés pour comparer les proportions, et le test de Wilcoxon pour comparer les variables quantitatives entre les groupes ayant reçu ou non une L2.
[b] correspond à la prescription d'antalgiques de palier 3

Parmi les 50 paramètres collectés et analysés au début de L2, 26 ont été trouvés comme étant significativement associés à la SG dans l'analyse de Cox univarié.

Le modèle multivarié final de Cox a identifié 9 facteurs indépendamment associés à la SG sous L2 : l'âge, le statut tabagique, la présence de métastases hépatiques, le PS, la douleur, l'ictère, l'ascite, la durée de L1, et le protocole en L2 (tableau 2)

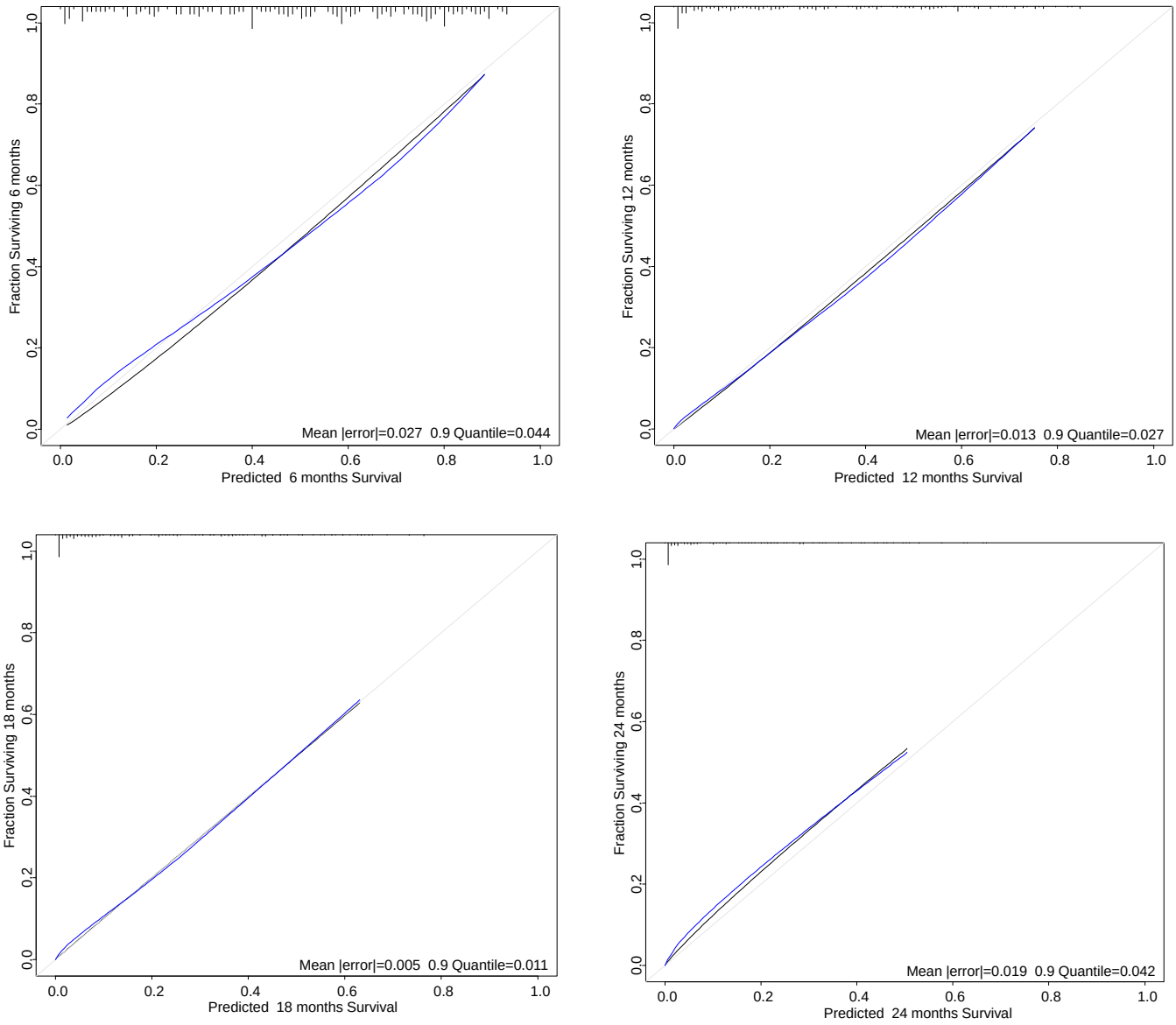
Tableau 2: Facteurs pronostiques associés avec la survie globale dans l'analyse de Cox multivariée (N = 248)

	Nombre de patients	Nombre de décès	Rapport de risque	IC à 95% du rapport de risque	p	Validation interne IC à 95% du rapport de risque par bootstrap	β	Score maximal dans le nomogramme
PS (OMS)								
0	47	40	1					
1	117	114	1.614	1.067 - 2.374		1.00897 - 2.39352	0.47879	25
≥ 2	84	81	3.055	1.935 - 4.822	**<.0001**	1.74622 - 4.72267	1.11671	58
Durée de L1 (transformation log), mois	248	235	0.478	0.325 - 0.702	**.0002**	0.33244 - 0.81427	-0.73868	100
Age, années	248	235	1.019	1.003 - 1.036	**.0168**	1.00237 - 1.03901	0.01921	55
Statut tabagique								
Non-fumeur ou ancien fumeur	190	178	1					
Fumeur actif	58	57	1.491	1.086 - 2.046	**.0134**	1.01718 - 2.13902	0.39946	21
Métastases hépatiques								
Non	99	93	1					
Oui	149	142	2.149	1.620 - 2.850	**<.0001**	1.48348 - 2.93914	0.76502	40
Douleur								
Non	118	108	1					
Oui	130	127	1.411	1.059 - 1.880	**.0186**	0.98883 - 1.85176	0.34442	18
Ictère								
Non	210	198	1					
Oui	38	37	1.958	1.327 - 2.890	**.0007**	1.18897 - 2.98240	0.67192	35
Ascite								
Non	200	188	1					
Oui	48	47	1.792	1.264 - 2.540	**.0011**	1.20746 - 2.76271	0.58335	30
Protocole en L2								
5-FU ± Oxaliplatine and/or Irinotecan	195	184	1					
À base de Gemcitabine	53	51	1.765	1.268 - 2.457	**.0008**	1.18119 - 2.49164	0.56827	30

Le modèle multivarié final est caractérisé par une excellente capacité discriminante (C-index = 0.75; 95% IC par les percentiles issue d'une procédure de bootstrap, 0.72 - 0.78) et une bonne calibration aux temps d'intérêt (6, 12, 18, et 24 mois) comme reporté sur la figure 1.

En guise de validation interne, les intervalles de confiance à 95% produits autour des rapports de risque du modèle multivarié reflètent sa robustesse.

Figure 1: Courbes de calibration à 6, 12, 18, et 24 mois pour le modèle multivarié final



Black: observed **Grey**: ideal calibrated model

Blue: optimism corrected B=20 based on observed-predicted

Vertical axis is the observed proportion of patients surviving at time of interest.

Nous nous sommes ensuite interrogés sur la valeur discriminante additionnelle des paramètres identifiés dans le modèle multivarié aux paramètres pragmatiques: l'âge, l'ECOG PS, et la durée de L1.

Nous avons donc défini deux modèles emboités :

- un modèle de référence avec uniquement les paramètres pragmatiques ;
- le modèle multivarié final comportant ces paramètres pragmatiques avec 6 autres variables.

Tableau 3: Détermination des deux modèles emboités à comparer

	Nombre de patients	Nombre de décès	Rapport de risque	IC à 95% du rapport de risque	p	Validation interne IC à 95% du rapport de risque par bootstrap	β	Score maximal dans le nomogramme
PS (OMS)								
0	47	40	1					
1	117	114	1.614	1.067 – 2.374		1.00897 – 2.39352	0.47879	25
≥ 2	84	81	3.055	1.935 – 4.822	**<.0001**	1.74622 – 4.72267	1.11671	58
Durée de L1 (transformation log), mois	248	235	0.478	0.325 – 0.702	**.0002**	0.33244 – 0.81427	(0.73868	100
Age, années	248	235	1.019	1.003 – 1.036	**.0168**	1.00237 – 1.03901	0.01921	55
Statut tabagique								
Non fumeur ou ancien fumeur	190	178	1					
Fumeur actif	58	57	1.491	1.086 – 2.046	**.0134**	1.01718 – 2.13902	0.39946	21
Métastases hépatiques								
Non	99	93	1					
Oui	149	142	2.149	1.620 – 2.850	**<.0001**	1.48348 – 2.93914	0.76502	40
Douleur								
Non	118	108	1					
Oui	130	127	1.411	1.059 – 1.880	**.0186**	0.98883 – 1.85176	0.34442	18
Ictere								
Non	210	198	1					
Oui	38	37	1.958	1.327 – 2.890	**.0007**	1.18897 – 2.98240	0.67192	35
Ascite								
Non	200	188	1					
Oui	48	47	1.792	1.264 – 2.540	**.0011**	1.20746 – 2.76271	0.58335	30
Protocole en L2								
5(FU ± Oxaliplatine and/or Irinotecan	195	184	1					
À base de Gemcitabine	53	51	1.765	1.268 – 2.457	**.0008**	1.18119 – 2.49164	0.56827	30

Modèle de référence

Modèle multivarié final

L'ajout de six paramètres identifiés dans le modèle multivarié (le statut tabagique, les métastases hépatiques, la douleur, l'ictère, l'ascite, et le protocole en L2), en supplément de ceux qui sont fréquemment utilisés en la réunion de concertation pluridisciplinaire (« paramètres pragmatiques »: âge, PS, et durée de L1) (tableau 3) entraine une augmentation significative de la C-statistique de 0,70 à 0,75 (différence moyenne par bootstrap : 0.048; 95% IC, 0.016 to 0.081)

De même, leur ajout reclasse de manière adéquate à 6 mois les patients à faible risque de décès et les personnes à risque plus élevé, tel que rapporté

par un NRI continu de 0,79 (IC 95%, 0,56 à 1,0; P < 0,0001) (figure 2), et un IDI de 0,14 (IC 95%, 0,10 à 0,18; P < 0,0001) (figure 3).

Figure 2: Valeur additionnelle des 6 autres paramètres pour la reclassification du risque de décès à 6 mois après le début de L2 (NRI continu).

Les lignes bleues chez les patients vivant à 6 mois indiquent un mouvement de la prédiction du risque par l'ajout des autres paramètres (vers le bas) (92/125=74%). Inversement, les lignes rouges chez les patients décédés à 6 mois indiquent une reclassification correcte, vers le haut, de l'évaluation du risque de décès à 6 mois lorsqu'on utilise l'ensemble des paramètres du modèle multivarié final (81/123=66%).

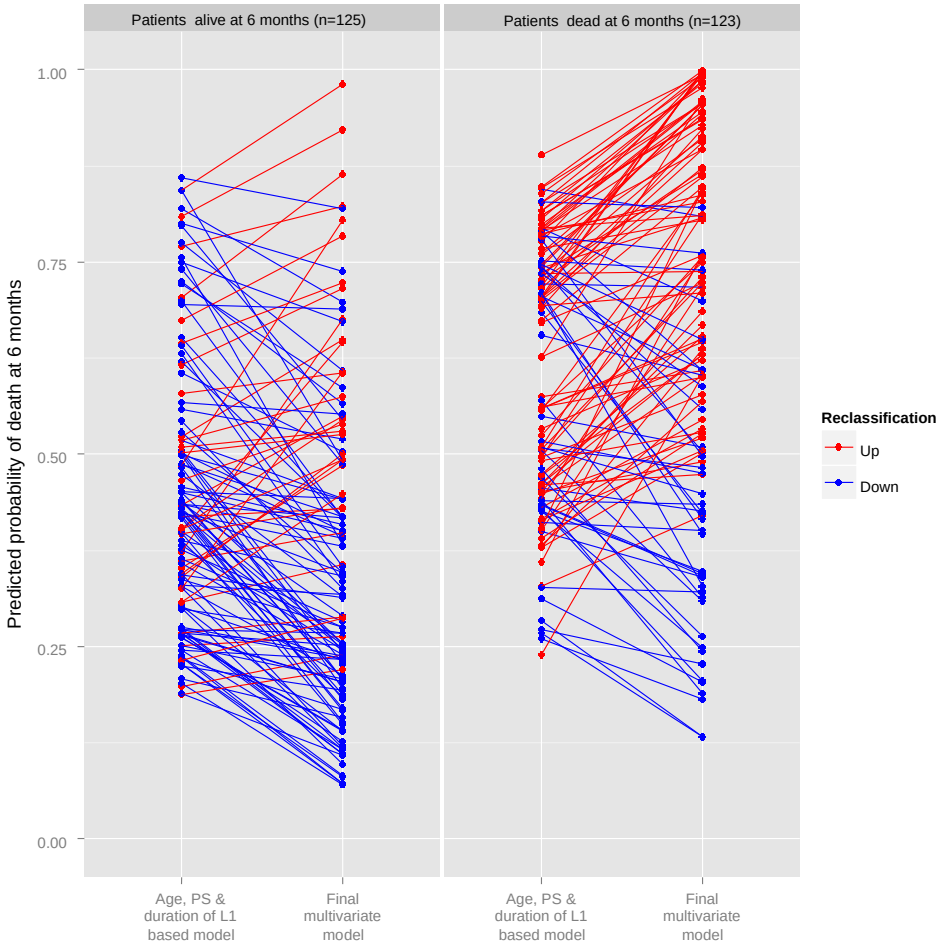
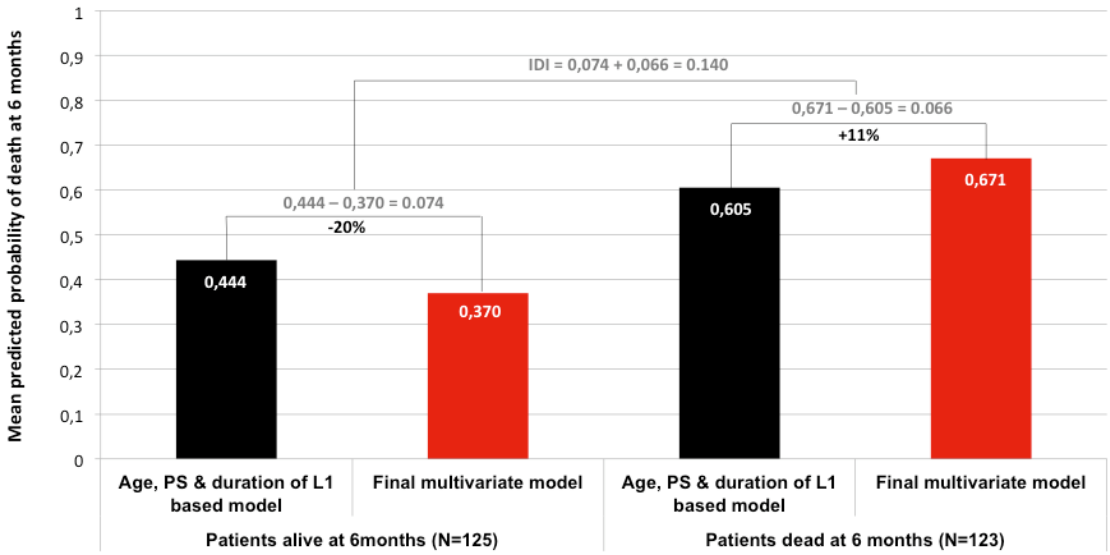
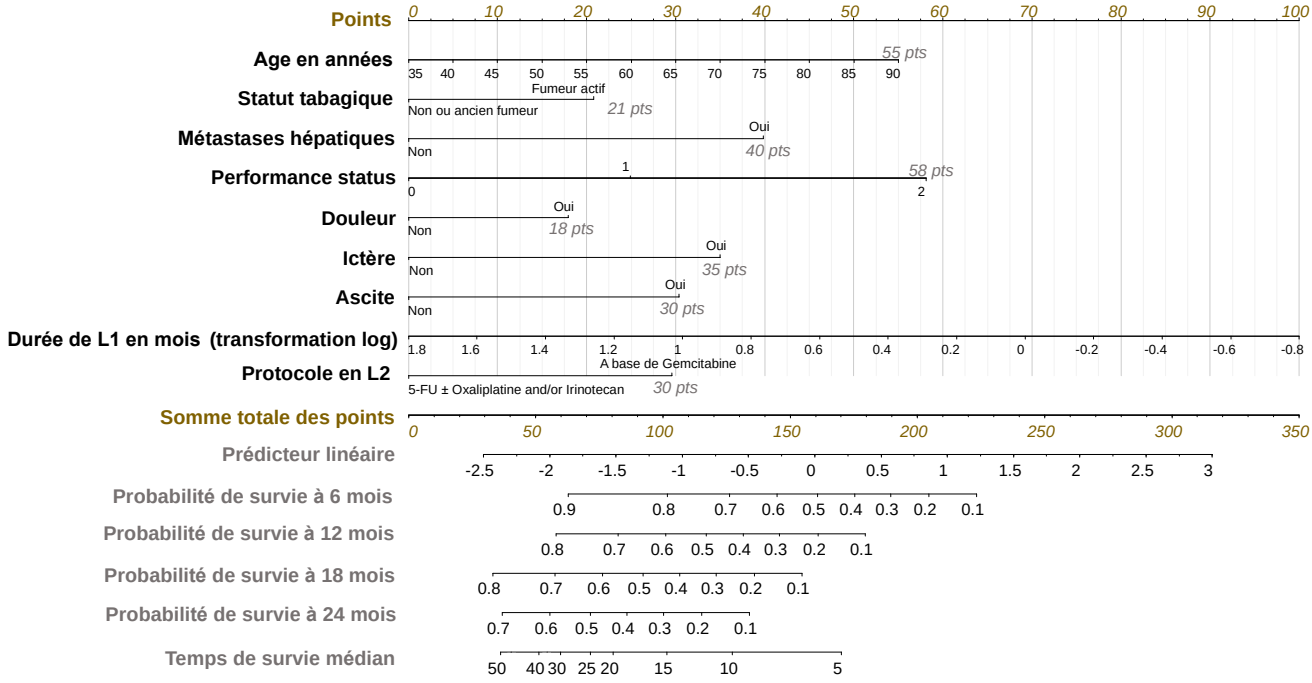


Figure 3: Amélioration de la prédiction du risque de décès à 6 mois en considérant les 6 paramètres ajoutés aux paramètres pragmatiques, évaluée par la méthode de l'IDI. L'IDI intègre séparément la variation des moyennes des probabilités de décès à 6 mois pour les patients toujours en vie et pour ceux décédés. Cette variation est adéquate si elle est à la hausse chez les patients décédés (augmentation de la probabilité de décès) et à la baisse chez les patients toujours en vie à 6 mois (diminution de la probabilité de décès).



Après avoir montré que les paramètres autres que ceux actuellement utilisés dans la pratique clinique courante apportaient de l'information intéressante et significative pour discriminer la survie globale sous L2, nous avons donc construit, sur la base du modèle multivarié final, un nomogramme permettant pour un patient donné de calculer sa médiane de survie et aussi ses probabilités individuelles de survie à 6, 12, 18 et 24 mois au début de L2 (Figure 4).

Figure 4: Nomogramme pour prédire la survie globale individuelle pour des patients avec un cancer pancréatique avancé au debut de L2 à 6,12,18 et 24 mois.



Les points sont assignés pour chaque facteur de risque en tracant une ligne verticale vers la ligne correspondant aux points. Le total des points pour les 9 facteurs est ensuite tracé sur la ligne correpondant à la somme totale des points. Une ligne verticale vers la bas permet de deduire les prédictions correspondantes aux temps d'intérêt (6, 8, 12 et 24 mois) ainsi que la mediane de survie du patient.

Une application web compatible avec les smartphones a été développée et fournit automatiquement les estimations de survie individuelles du nomogramme: http://www.umqvc.org/en/tool/proscap.html.

Pour les 248 patients impliqués dans le modèle multivarié, il a donc été possible d'obtenir un nombre total de points issu du nomogramme. Ce vecteur de 248 estimations a donc constitué notre score pronostique (Figure 5). En appliquant la méthode dite de Cox nous obtenons des seuils qui permettent d'isoler 3 groupes de pronostics différents (Figure 6 A). Les trois groupes à risque distincts ont été validés en externe (Figure 6 B).

Figure 5: Distribution du score pronostique dans la cohorte de développement

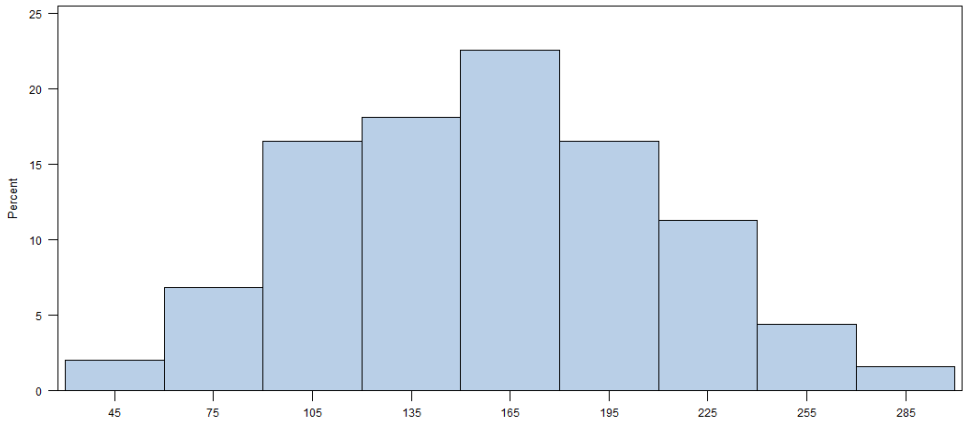
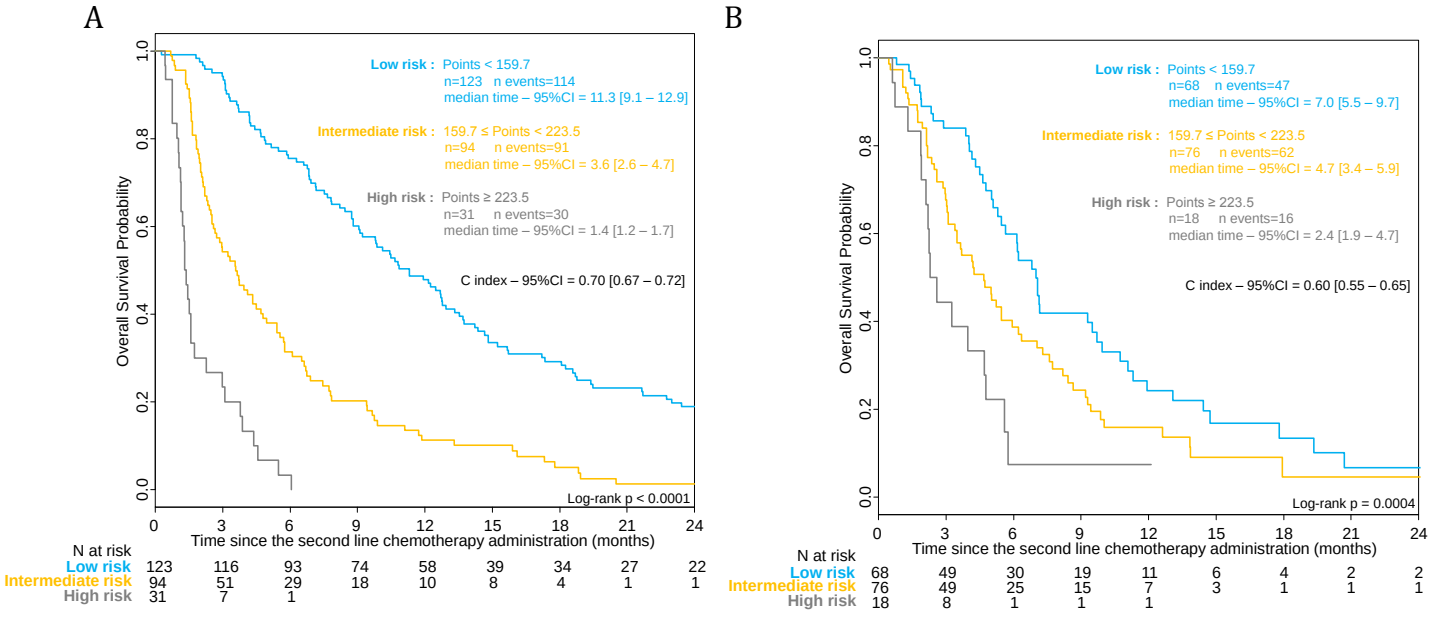
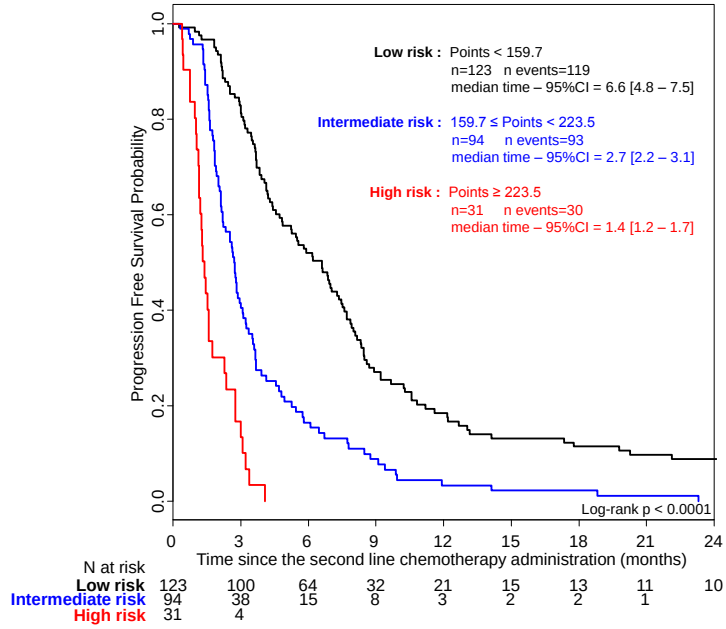


Figure 6 : Courbe de Kaplan-Meier de la survie globale pour A la cohorte de développement et B la cohorte de validation externe selon la catégorisation en 3 groupes avec la méthode de Cox.



De manière intéressante nous avons observé que la catégorisation des patients parmi ces 3 groupes permet aussi de stratifier convenablement la survie sans progression sous L2. Cette observation permet notamment d'affirmer que le résultat n'est pas attribuable aux lignes de chimiothérapie ultérieures à la L2.

Figure 7 : Courbe de Kaplan-Meier de la survie sans progression pour la cohorte de développement selon la catégorisation en 3 groupes avec la méthode de Cox.



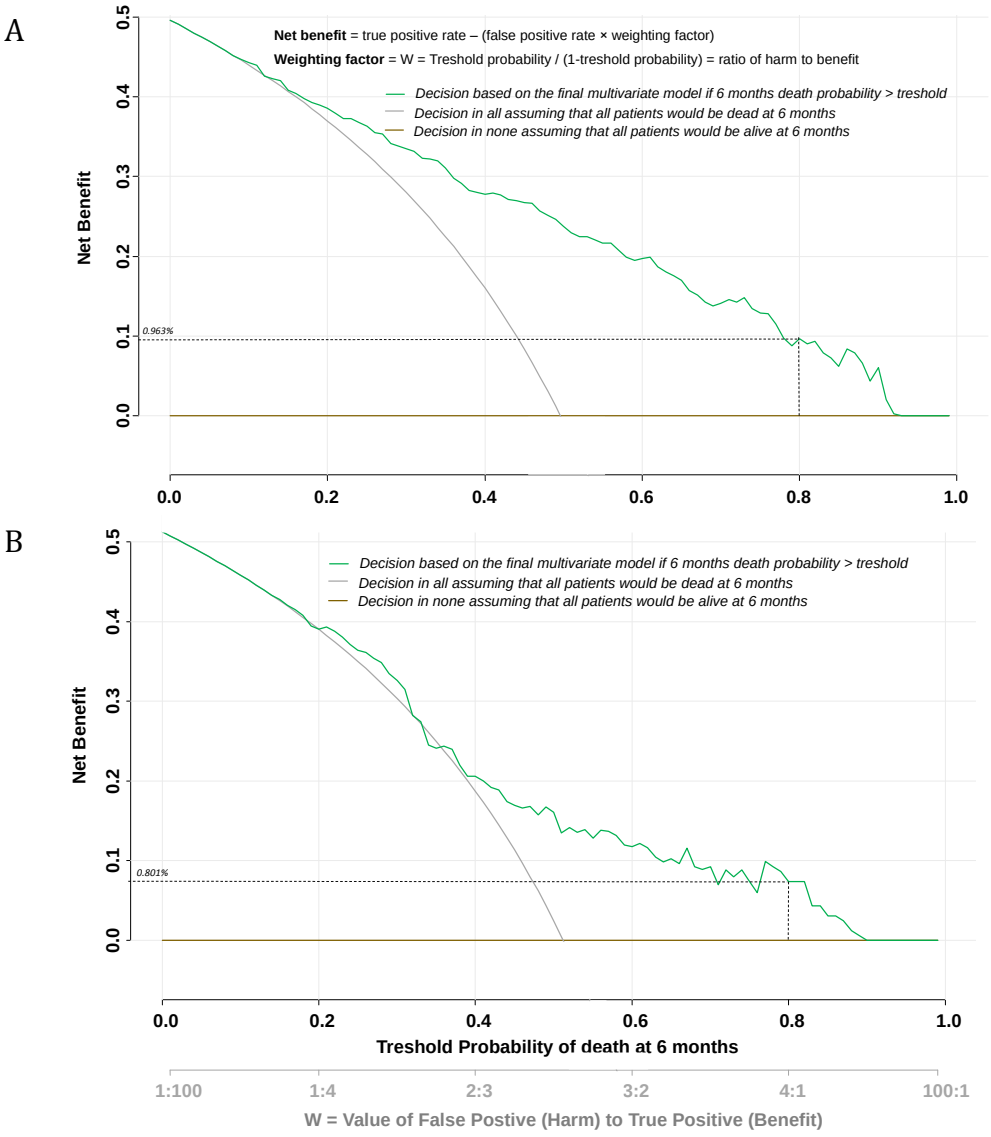
L'évaluation du bénéfice clinique apporté par la considération des prédictions issues du modèle final, afin d'orienter la prise de décision thérapeutique, a été confirmée par l'analyse de courbe de décision dans les deux cohortes (développement et validation) (figure 7 A et B).

Dans cette étude, compte tenu d'un pronostic de survie globale très sombre pour une partie de la population étudiée (groupe à haut risque) (Figure 6A) au début de L2 on peut se poser la question de l'utilité d'administrer la L2 chez des patients qui ont une très forte probabilité de décéder à 6 mois. La décision thérapeutique à l'étude est donc ici l'absence

d’administration une L2. Les seuils de probabilité eux correspondent à la probabilité de décéder à 6 mois.

Il est alors possible de représenter la courbe issue de la DCA comme suit :

Figure 8: Courbes de décision pour représenter le bénéfice net réalisé par la prise de décisions cliniques fondées sur la prédiction de décès à 6 mois issue du modèle multivarié final dans la cohorte de développement (A) et dans la cohorte de validation externe (B)



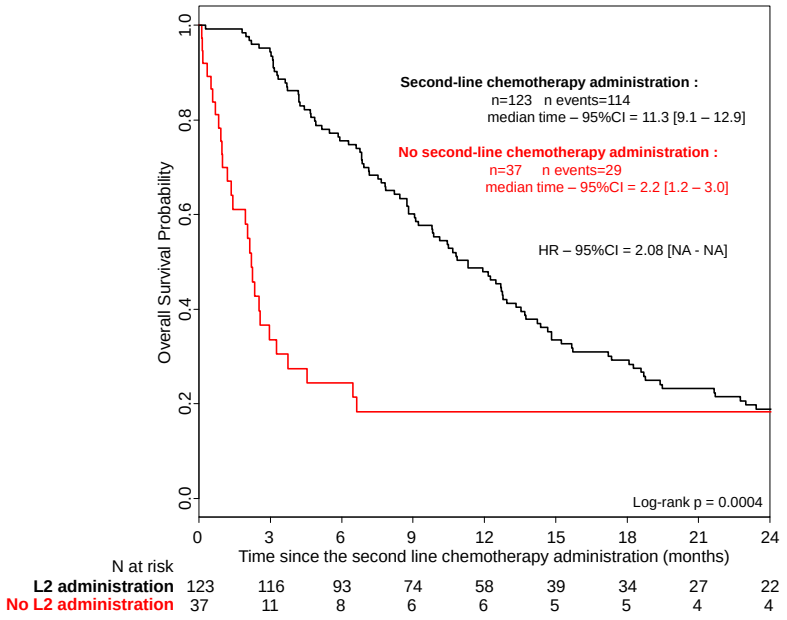
La DCA identifie donc un intervalle de seuil de probabilité de décès à 6 mois dans lequel les décisions entraîneront plus d'avantages pour les vrais positifs et la décision pour les faux positifs sera réduite. C'est en effet ce que l'on attend du modèle. Compte tenu que l'on souhaite éventuellement envisagé le non-traitement des patients de très mauvais pronostics, cette décision implique que nous souhaitons minimiser les faux positifs, c'est-à-dire les patients considérés de mauvais pronostics à tort, à savoir les patients encore en vie à 6 mois qui a priori devraient être traités. Sur la courbe, le bénéfice net des décisions basées sur les prédictions issues de notre modélisation multivariée finale est meilleur pour le seuil de probabilités supérieures à 20% (A) et 40% (B) comparé aux stratégies qui ont considéré l'ensemble de la population de la même manière (ne pas traiter personne ou ne pas traiter tout le monde). Dans l'ensemble, la courbe de décision montre que le bénéfice net de baser sa décision sur notre modèle multivarié final est bénéfique, en particulier pour les seuils > 0,5 ce qui peut être défini comme la zone d'intérêt dans le cancer pancréatique avancé.

Un des éléments les plus intéressants dans cette étude est la disponibilité des données de 134 patients qui n'ont pas reçu une L2. Ainsi, le « scoring » établi a pu être réalisé pour ces patients (le type de traitement qui n'a pas été attribué par définition a été a posteriori sélectionné sur la base des critères usuellement appliqués en réunion de concertation pluridisciplinaire). Ainsi pour chacun des groupes à risque identifiés nous disposions de patients traités et non traités par L2.

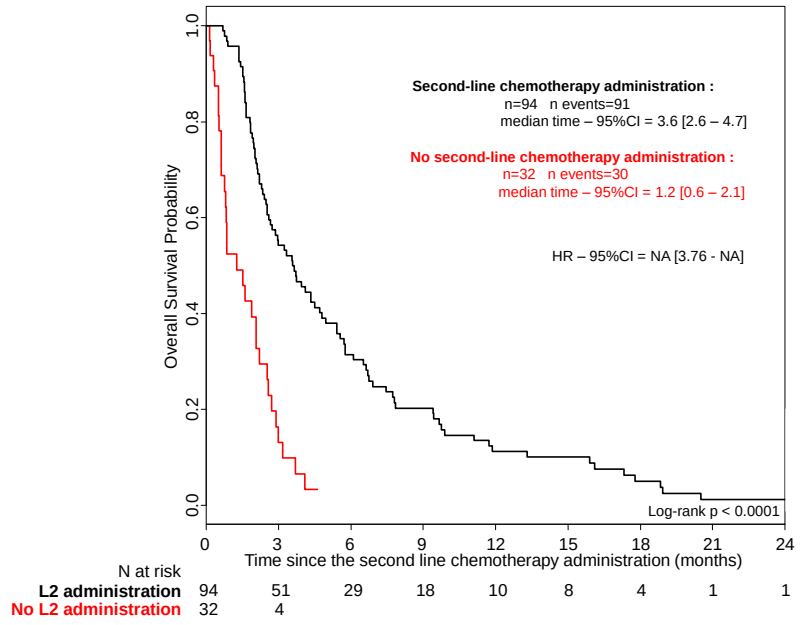
Voici ci-dessous les résultats qui quantifient le bénéfice en survie globale d'une L2 en fonction des groupes à risque:

Figure 9: Courbe de Kaplan-Meier de la survie globale en fonction de l'administration d'une L2 selon les trois groupes de pronostique distinct identifiés

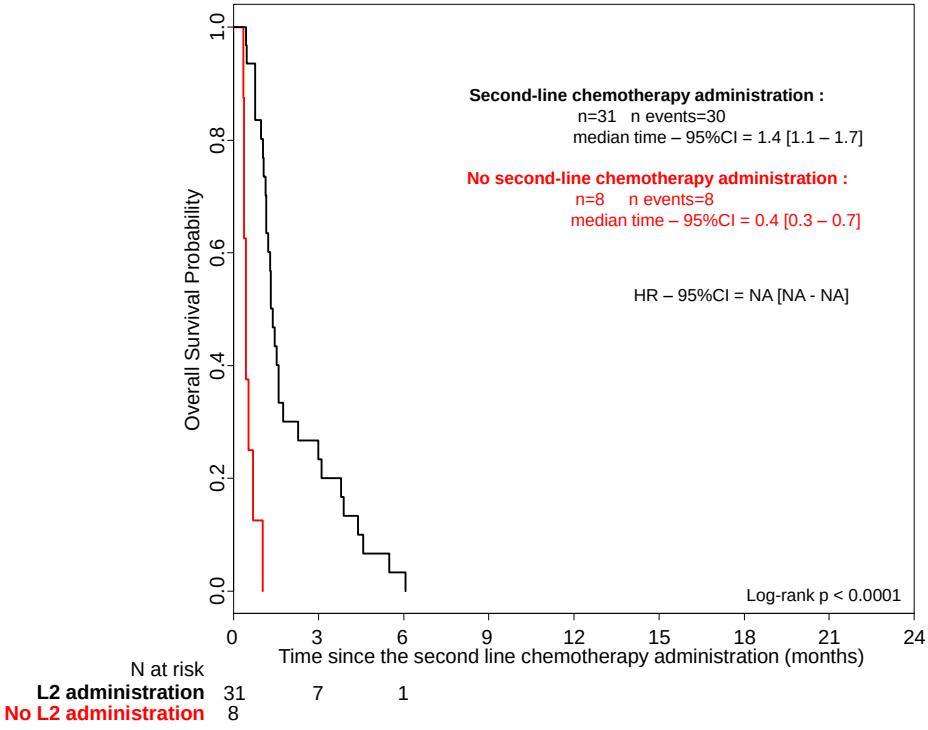
Groupe de bon pronostic



Groupe de pronostic intermédiaire



Groupe de mauvais pronostic



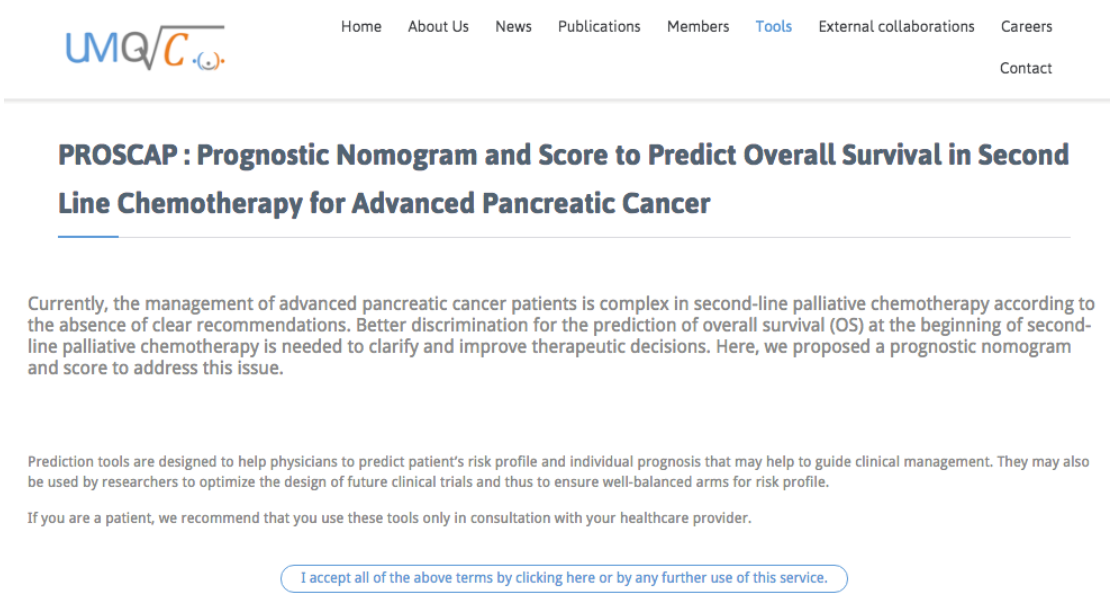
Un effet significatif d'un traitement de L2 est identifié dans chacun des groupes mais plus le pronostic de survie devient mauvais, plus l'intérêt d'une L2 (amplitude de l'effet) décroit.

Une interaction statistique significative (p = 0.0123) a pu être mise en évidence entre le groupe et l'administration d'une L2. Ceci signifie donc que dans un contexte de L2 pour les patients atteints d'un cancer pancréatique avancé, le pronostic de survie globale est prédictif du bénéfice attendu de l'administration d'une L2.

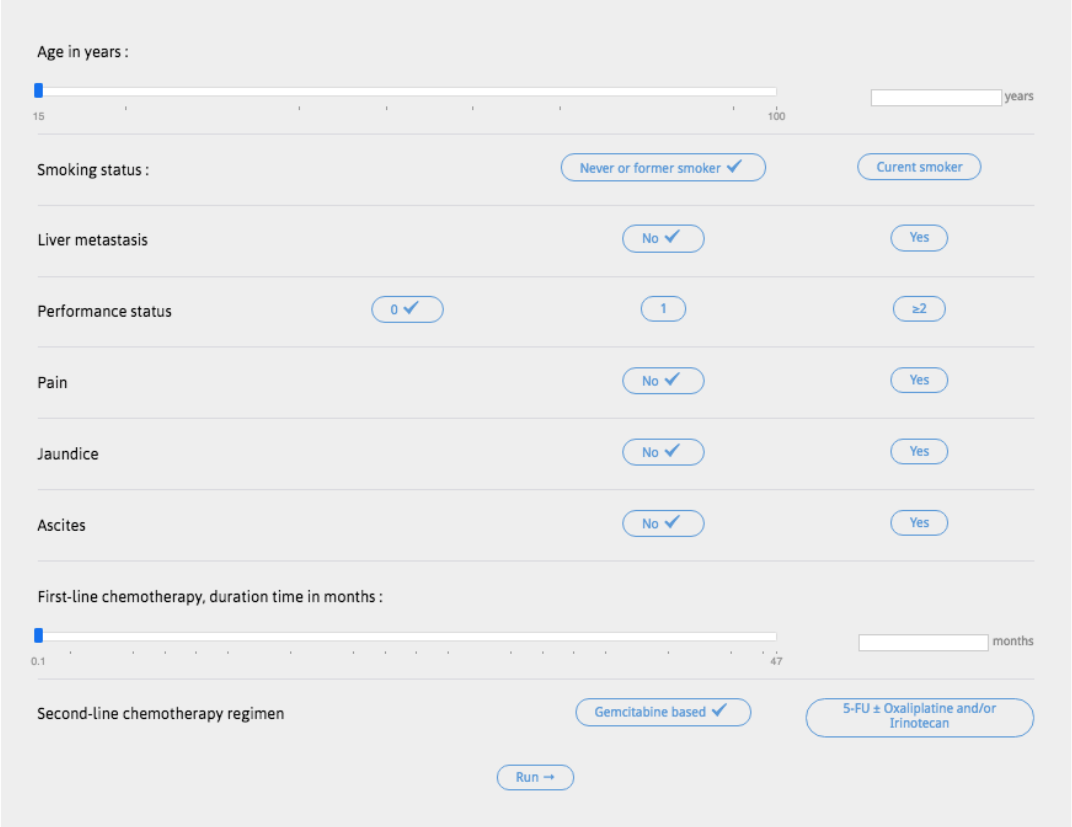
Enfin, une application en ligne sur le web (compatible avec les téléphones modernes) a été développée pour fournir des estimations de survie individualisées (nomogramme) et en population (score) pour les cliniciens et chercheurs.
L'adresse de cette application est la suivante : http://www.umqvc.org/en/tool/proscap.html .
Cette étape s'est avérée primordiale afin de rendre fonctionnels ces outils. En effet, l'utilisation des variables en continu ainsi que leur transformation pour garder le plus d'information possible rend l'utilisation des outils impossible sans automatisation des calculs.

Concrètement l'utilisateur rentre les données de son patient, et obtient un rapport qui lui fournit instantanément les estimations du nomogramme et du score.

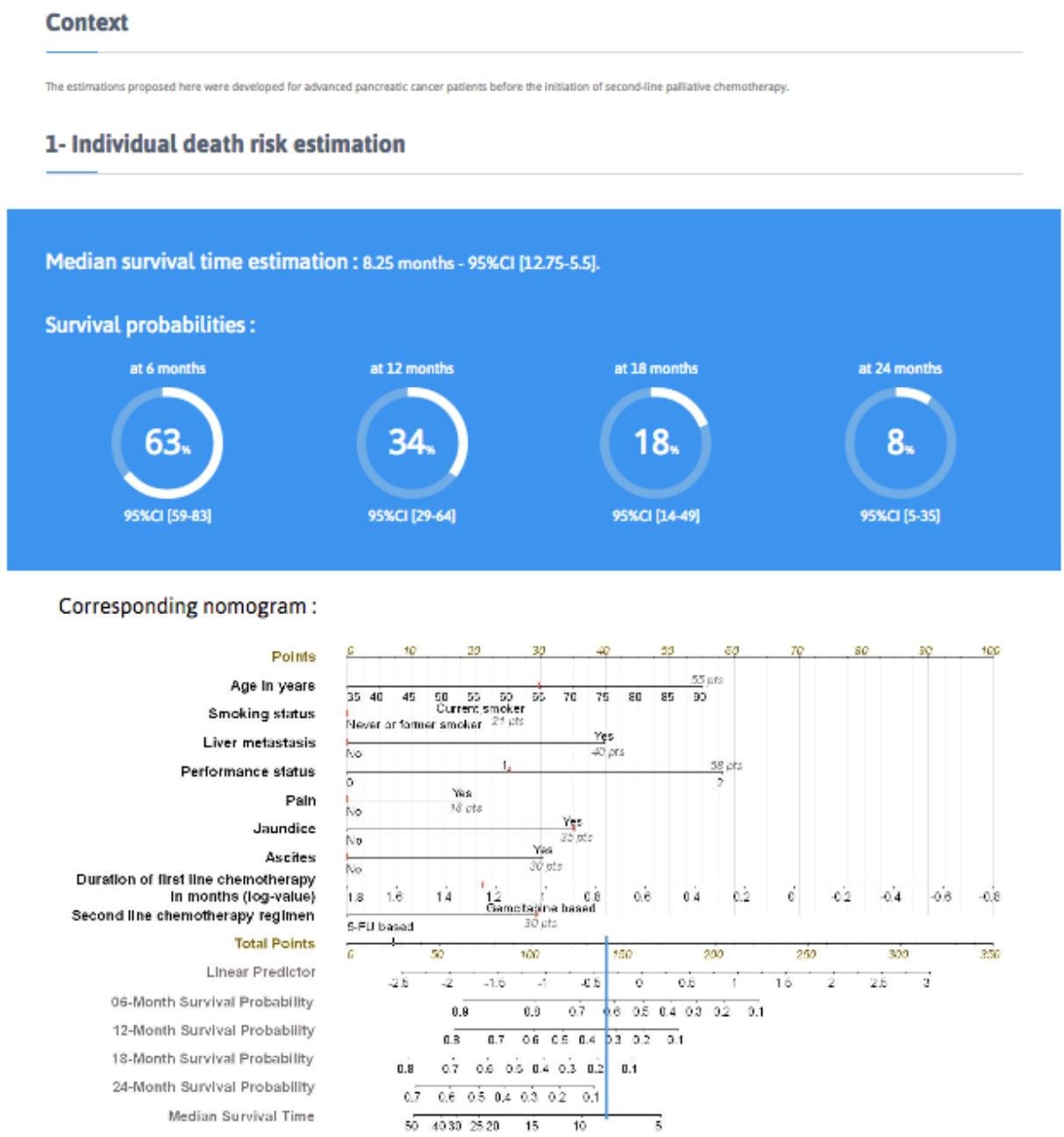
La page d'accueil de l'outil est une page d'information pour situer l'outil dans son contexte.



Ensuite les données du modèle multivarié propre au patient doivent être saisies.



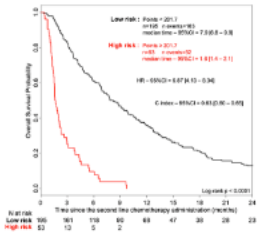
Les résultats issus du nomogramme et du score sont immédiatement visualisables comme ci-dessous :



2- Population death risk classification

2-A Two risk-groups approach

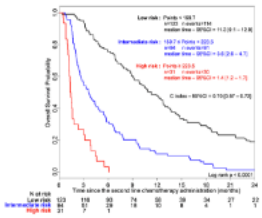
Corresponding risk population : Low risk (black curve)



Kaplan–Meier curves for overall according to the prognostic score group.

2-B Three risk-groups approach

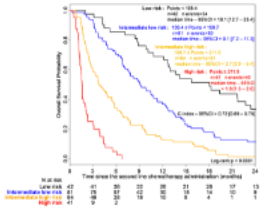
Corresponding risk population : Low risk (black curve)



Kaplan–Meier curves for overall according to the prognostic score group.

2-C Four risk-groups approach

Corresponding risk population : Intermediate low risk (blue curve)



Kaplan–Meier curves for overall according to the prognostic score group.

3- For the following data

- Used tool : Clinical tool
- Age in years : 64.57 / So 29.576651282329 points
- Smoking status : Never or former smoker / So 0 points
- Liver metastasis : No / So 0 points
- Performance status : 1 / So 24.92960415877 points
- Pain : 0 / So 0 points
- Jaundice : 1 / So 34.98548346 points
- Ascite : 0 / So 0 points
- First-line chemotherapy, duration time in months : 18.05 months (Log10 value = 1.2564772062417) / So 20.904722836859 points
- Second-line chemotherapy regimen : Gemcitabine based / So 29.588642526586 points

Total points : 139.98510426454

Results extraction in PDF file :

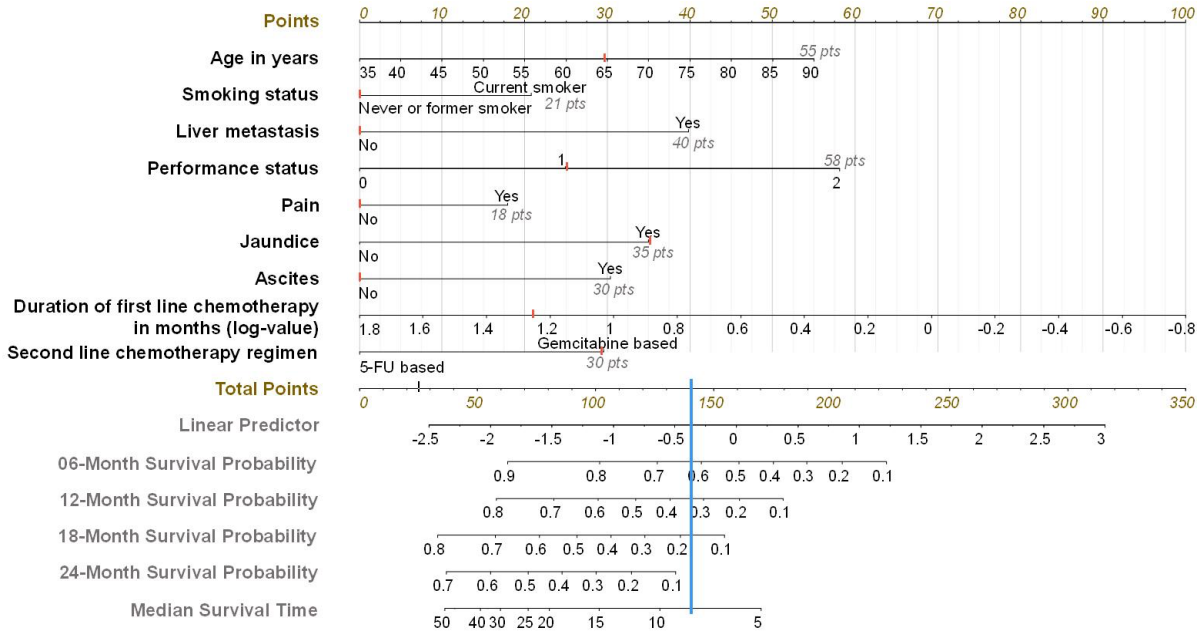
Une extraction de ces résultats peut être demandée pour les incorporer dans le dossier médical du patient par exemple.

L'extraction en question se présente sous la forme suivante :

1- Individual death risk estimation

Median survival time estimation : 8.25 months - 95%CI [12.75-5.5].
Survival probabilities at 6 months : 63% // 95%CI [59-83]
Survival probabilities at 12 months : 34% // 95%CI [29-64]
Survival probabilities at 18 months : 18% // 95%CI [14-49]
Survival probabilities at 24 months : 8% // 95%CI [5-35]

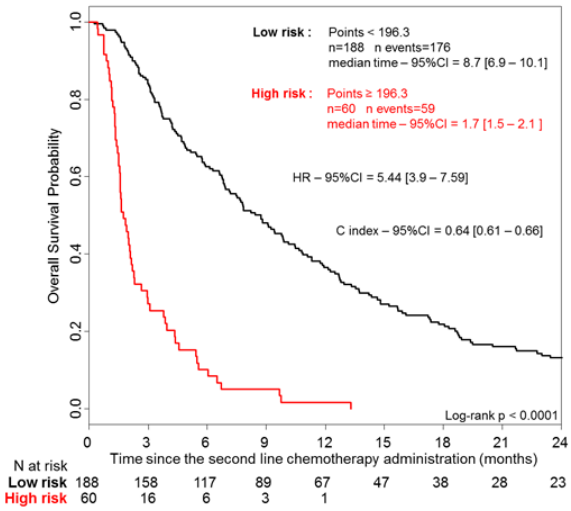
Corresponding nomogram :



2- Population death risk classification

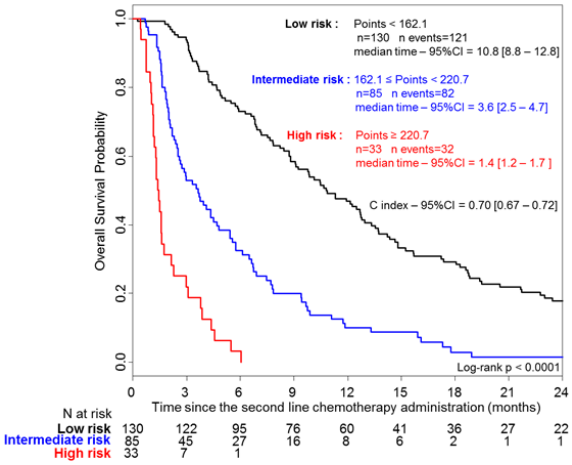
2-A Two risk-groups approach

Corresponding risk population :
Low risk (black curve)



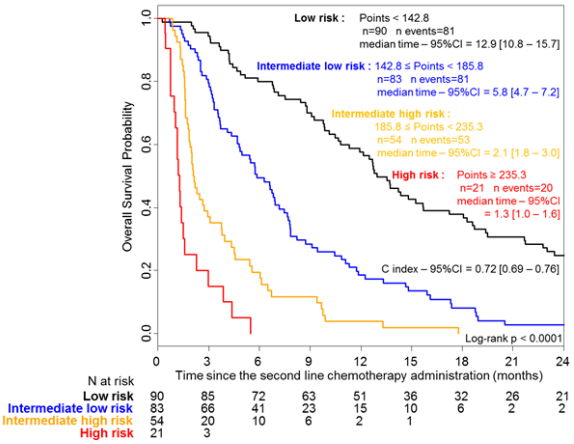
2-B Three risk-groups approach

Corresponding risk population :
Low risk (black curve)



2-C Four risk-groups approach

Corresponding risk population :
Intermediate low risk (blue curve)



3- For the following data

Used tool : Clinical tool
Age in years : 64.57 / So 29.576651282329 points
Smoking status : Never or former smoker / So 0 points
Liver metastasis : No / So 0 points
Performance status : 1 / So 24.92960415877 points
Pain : 0 / So 0 points
Jaundice : 1 / So 34.98548346 points
Ascite : 0 / So 0 points
First-line chemotherapy, duration time in months : 18.05 months
Second-line chemotherapy regimen : Gemcitabine based / So 29.588642526586 points

Total points : 139.98510426454

IV.1.3.5 Conclusion de l'étude

Nous proposons donc dans cette étude un nouveau modèle de prédiction de la SG au début de la L2 pour des patients atteints d'un cancer pancréatique avancé. Nous avons, sur la base de ce nouveau modèle, développé des outils pronostiques fonctionnels à l'échelle individuelle (nomogramme) et populationnelle (score) dont l'utilisation a été facilitée avec le développement d'une interface web qui automatise les calculs et synthétise les résultats (http://www.umqvc.org/en/tool/proscap.html).

Au-delà des outils proposés cette étude met en évidence une considérable hétérogénéité de pronostic pour la survie globale sous L2 chez des patients atteints d'un cancer pancréatique avancé.

De manière intéressante, nous avons observé une interaction significative entre la réponse au traitement sous L2 en termes de survie globale et le pronostic du patient. Il est donc essentiel dans cette situation de modéliser au mieux la survie globale afin d'offrir aux cliniciens des informations précises pour leur prise de décision.

Cette étude fournit des éléments qui permettre d'optimiser le design des essais cliniques futurs et offre l'opportunité de définir des stratégies de management adaptées aux différents profils de risque.

IV.1.3.6 Manuscrit

Manuscript
Click here to view linked References

Development and validation of a prognostic nomogram and score to predict overall survival and usefulness of second-line chemotherapy in advanced pancreatic adenocarcinoma

Angélique Vienot, MD[1], Guillaume Beinse, MD[2]*, Christophe Louvet, PhD[3]*, Louis de Mestier, MD[4]*, Aurélia Meurisse[5,6], Francine Fein, MD[1], Bruno Heyd, PhD[7], Denis Cleau, MD[8], Christelle d'Engremont, MD[1], Anne-Claire Dupont-Gossart, MD[1], Zaher Lakkis, MD[7], Christophe Tournigand, PhD[2], Olivier Bouché, PhD[4], Benoît Rousseau, MD[2], Cindy Neuzillet, MD[2], Franck Bonnetain, PhD[5,6], Christophe Borg, PhD[6,9,10] and Dewi Vernerey, PhD[5,6]

[1]Department of Gastroenterology, Besançon University Hospital, 3 Boulevard Alexandre Fleming, 25030, Besançon, France;

[2]Department of Medical Oncology, Henri Mondor University Hospital, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris Est Créteil University (UPEC), 51 Avenue du Maréchal de Lattre de Tassigny, 94010, Créteil, France;

[3]Department of Medical Oncology, Institute Mutualiste Montsouris, 42 Boulevard Jourdan, 75014, Paris, France;

[4]Department of Hepato-Gastroenterology, Reims University Hospital, Rue du Général Koenig, 51100, Reims, France;

[5]Methodological and Quality of Life in Oncology Unit, EA 3181, Besançon University Hospital, 3 Boulevard Alexandre Fleming, 25030, Besançon, France;

[6]INSERM, Unit 1098, University of Bourgogne – Franche-Comté, 1 Boulevard Fleming, 25020, Besançon, France;

[7]Department of Digestive Surgery and Liver Transplantation, Besançon University Hospital, 3 Boulevard Alexandre Fleming, 25030, Besançon, France;

[8]Department of Gastroenterology, Vesoul Hospital, 2 Avenue René Heymes, 70000, Vesoul, France;

[9]Department of Medical Oncology, Besançon University Hospital, 3 Boulevard Alexandre Fleming, 25030, Besançon, France;

[10]Clinical Investigation Center 1431, EFS Bourgogne – Franche-Comté, 1 Boulevard Fleming, 25020, Besançon, France.

*These authors should be considered as joint second authors

Corresponding author:

Angélique Vienot, Department of Gastroenterology, Besançon University Hospital, 3 Boulevard Alexandre Fleming, 25030, Besançon, France;

Tel.: +33 381 668 254, Fax: +33 381 668 417, E-mail: angelique.vienot@gmail.com.

Abstract word count: 300

Manuscript word count: 2990

Running title: Second-line chemotherapy in advanced pancreatic adenocarcinoma

ABSTRACT

Background

In advanced pancreatic ductal adenocarcinoma (aPDAC), there is no consensual strategy for second-line chemotherapy (L2). Better discrimination of overall survival (OS) may help clinical decision-making in this setting. We aimed to predict OS from the beginning of L2 and to assess the benefit from chemotherapy among the risk groups identified.

Methods

Analyses were derived from all consecutive aPDAC patients treated at Besancon hospital, France, during 11 years. The association of 50 parameters with OS was evaluated using univariate and multivariate Cox analyses. Based on the final model, a prognostic nomogram and score were developed, and externally validated in patients who received L2 in three other French centers.

Findings

The development cohort included 462 patients. 395 patients (85%) were eligible for L2, of which 261 (66%) were treated. Age, smoking status, liver metastases, performance status, pain, jaundice, ascites, duration of first-line chemotherapy, and type of L2 regimen were identified as independent prognostic factors for OS in L2. The final model exhibited excellent calibration and discrimination (C-index = 0.75, 95% confidence interval 0.72 to 0.78), and robust internal validity. The score determined three groups with median OS of 11.3, 3.6 and 1.4 months ($P < 0.0001$). By applying the score in the population eligible for L2 but untreated, the chemotherapy benefit was significant across all groups but with a magnitude of the effect decreased from low to high-risk groups ($P = 0.0123$ for treatment and risk groups interaction term). The clinical benefit centered accuracy of the final model was confirmed by a decision curve analysis. The score ability to discriminate OS was externally confirmed in 162 patients, by identifying the same risk groups.

Interpretation

This prognostic nomogram and score in patients with aPDAC can accurately predict OS before administration of L2 and can help clinicians in their therapeutic decisions.

Funding

None.

Key words

Advanced pancreatic adenocarcinoma, second-line chemotherapy, overall survival, prognostic factor, nomogram, score, decision curve analysis.

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is the fourth cause of cancer-related death in developed countries,[1] and it is expected to become the second leading cause of cancer death in 2030.[2] PDAC has the poorest prognostic among digestive malignancies with a 5-year survival rate of 5%-7%, with no significant change in death rate between 1997 and 2013.[3] Surgical resection of localized PDAC is the only treatment that can provide prolonged survival. However, diagnosis is made at an advanced stage in the vast majority of cases (> 80%). The median overall survival (OS) is about 9-15 months in patients with locally advanced PDAC, and 6-9 months in those with metastatic disease.[4,5]

Advanced PDAC (aPDAC) remains a challenging, non-curable disease. Gemcitabine chemotherapy has been the only standard of care in this setting for more than a decade, until 2010. Significant progress has been achieved in the landscape of aPDAC management during the 5 past years with the approval of two active cytotoxic combinations: the FOLFIRINOX (5-fluorouracil [5-FU], irinotecan, and oxaliplatin) and the gemcitabine plus *nab*-paclitaxel regimens. These regimens were shown to be superior to standard gemcitabine as first-line chemotherapy (L1) in patients with metastatic PDAC, yielding median OS of 11.1 and 8.5 months, respectively.[6,7] Currently, clinical practice guidelines for L1 are well established.[8,9]

Beyond L1, about half of patients with aPDAC remain in good general condition, and thus may receive subsequent line(s) of chemotherapy.[10] Combinations of 5-FU with oxaliplatin[11] or with nanoliposomal irinotecan[12] have shown interesting activity after failure of gemcitabine and can be proposed as second-line chemotherapy (L2).[8] Nevertheless, not all patients seem to benefit from L2 treatment, and L2 has to be discussed on an individual basis in terms of risk benefit for each patient. Of note, evidence to guide patient selection for L2 after progression under FOLFIRINOX or gemcitabine plus *nab*-paclitaxel regimen is scarce[13] and there is no validated strategy to date.

Age, Eastern Cooperative Oncology Group performance status (ECOG PS), and duration of L1 are so-called “pragmatic parameters”, which are frequently used in multidisciplinary meeting to estimate the potential benefit of L2, but the level of evidence is low.[14,15] Thus, identification of reliable factors for patient prognostic stratification is warranted to improve therapeutic decision-making in this setting. However, there

is no well-validated and widely accepted prognostic model for application in routine practice or clinical trial for L2 in aPDAC. In this context, it is urgent to develop useful prognostic tools that may help to estimate patient OS in order to guide clinicians decision for L2 administration and to optimize future clinical trial design.

In this prospective population-based cohort study, we aimed to develop and validate a prognostic nomogram and score to predict OS in patients with aPDAC who received L2 in routine clinical practice, using a broad spectrum of parameters. Then, we assessed the benefit from chemotherapy across risk groups identified by this score.

METHODS

Patients

All consecutive patients with histologically proven, aPDAC (i.e. metastatic, locally advanced, or recurrent after surgery), who were treated at Besancon University Hospital, France, between January 2003 and December 2013, were involved in the development cohort. Patients were considered to be eligible for medical evaluation of L2 indication if they had received one previous line of chemotherapy, including adjuvant chemotherapy in case of early post-operative tumour relapse (i.e. within six months after the last administration of this treatment), and if they were not dead during L1. The external validation cohort included consecutive patients with aPDAC who received L2 with the same inclusion criteria in three other French institutions (Institute Mutualiste Montsouris, Henri Mondor University Hospital, and Reims University Hospital) between January 2010 and April 2016. Main procedures for both cohorts are specified in Appendix A1.

Demographics, cancer history, pathological, clinical, biological, and radiological (tumour response according to RECIST v1.1 criteria) parameters at the beginning of L2 (or at the end of L1 for patients of the development cohort who did not received L2), as well as treatment outcomes were retrospectively collected from medical records.

Statistical analysis

Median value (interquartile range) and frequency (percentage) were provided for the description of continuous and categorical variables, respectively. Medians and proportions were compared using Student's t-test and Chi-square test (or Fisher's exact test, if appropriate), respectively.

OS was calculated from the date of progression under L1 (or the date of first administration of L2, if no progression was observed under L1) to the date of death from any cause. Survival data were censored at the last follow-up. Progression free survival (PFS) was calculated from the date of progression under L1 (or the date of first administration of L2, if no progression was observed under L1) to the date of progression or

death from any cause, or the date of the last follow-up, at which point data were censored. OS and PFS were estimated using the Kaplan-Meier method and described using median or rate at specific time points with 95% confidence intervals (95% CI). Follow-up duration was calculated using a reverse Kaplan-Meier estimation when feasible.[16]

Cox-proportional-hazard models were performed to estimate hazard ratio (HR) and 95% CI for factors associated with OS. The association of baseline parameters with OS was first assessed using univariate Cox analyses and then parameters with $P < 0.05$ were entered into a final multivariate Cox regression model, after considering collinearity among variables with a correlation matrix. When used in continuous in the Cox modelization, a potential non-linear relationship between predictors and OS was first investigated using restricted cubic splines modelization. Subsequently, a fractional polynomial method was applied to determine the best transformation for continuous variables. Hazard proportionality was checked by plotting log-minus-log survival curves and by the cumulative martingale process plots.

Sensitivity analyses to explore the reliability of the univariate analysis and the final model were performed with a stratified but also with a frailty approach by using a random component for the hazard function based on the L2 regimen.

Accuracy of the final model was verified regarding two parameters: discrimination and calibration. The predictive value and the discrimination ability of the final model were assessed with the Harrell's concordance index (C-index).[17] Random samples of the population were used to derive 95% CI for the C-statistic. Calibration was assessed by visual examination of calibration plot. Internal validation of the final model was performed with a bootstrap sample procedure.

Improvement in discrimination ability of our final model compared to the classical approach ("pragmatic parameters": age, ECOG PS and duration of L1) was evaluated with the Harrell's concordance index (C-index). We used continuous net reclassification improvement (cNRI) and integrated discrimination improvement (IDI) to quantify the performance in risk reclassification of adding others parameters of the final model to the classical approach at 6 months after the beginning of L2.

The final model was used to establish a nomogram allowing the estimation of median and individual post-L1 OS probabilities at 6, 12, 18, and 24 months. At a population level, a prognostic score was constructed with nomogram total points. To give a reasonable spread of risk we chose to distinguish three prognostic groups, according to their risk score level, which were identified based on cut points determined following the Cox's method.[18] Patients' baseline key characteristics were compared between prognostic risk groups using Fisher-exact test and Kruskal-Wallis test for categorical and quantitative parameters, respectively.

The prognostic score developed in patients who received L2 was applied to patients who did not receive L2 in order to assess the benefit from treatment in each groups identified.

Another multivariate analysis was performed by including relevant biological data recorded, but not initially selected in the multivariate analysis process due to their high rate of missing data.

The prognostic score discrimination ability was confirmed in an external validation cohort and evaluated with the Harrell's concordance index (C-index). To identify risk groups and to determine their survival profile, the same development cohort-derived risk predictive algorithm was applied.

A clinical benefit centered accuracy of the final model was evaluated by a decision curve analysis[19] for both cohorts.

All analyses were performed using SAS version 9.4 (SAS Institute) and R software version 2.15.2 (R Development Core Team; http://www.r-project.org). Values of $P < 0.05$ were considered statistically significant and all tests were two-sided. Details on the interpretation of important statistical concepts are given in Appendix A2.

RESULTS

Population-based prospective cohort

Among 462 patients who had received L1 for aPDAC, 395 (85%) were eligible for medical evaluation of L2 indication; 261 (66%) of them actually received L2 and were included in the development cohort, while 134 (34%) patients were not treated (Figure 1). Patient characteristics in the L2 treated and untreated groups were significantly different for age, ECOG PS, and duration of L1 (Table 1).

In the external validation cohort, 163 patients with aPDAC who received L2 treatment were analysed. The two cohorts displayed similar patient characteristics, except for the type of L1 and L2 regimens (Table S1).

Determinants of OS in patients receiving L2

In univariate Cox analysis, we identified 26 parameters as prognostics factors for OS with a $P < 0.05$ (Table 2). The main analyses did not include biological parameters because of the high-rate of missing data in some of them. A correlation matrix was used to detect significant correlations (coefficient ≥ 0.4) between investigated parameters (Table S2). The multivariate Cox analysis showed nine independent risk factors for OS: age, smoking status, liver metastases, ECOG PS, pain, jaundice, ascites, duration of L1, and type of L2 regimen (Table 3). The transformations used for continuous variables are summarized in Figure S1.

In the sensitivity analyses, associations remained unchanged in the univariate and multivariate analyses (Table S3 and S4).

Performance assessment and internal validation of the final model

The multivariate model exhibited good discrimination ability (C-index = 0.75; 95% bootstrap percentile CI, 0.72 to 0.78). The calibration plots showed an optimal agreement between model prediction and actual observation for predicting OS probability at 6, 12, 18, and 24 months (Figure S2). In the internal

validation, uncertainties around HR measured with a bootstrapping procedure reflected the robustness of the final model (Table 3).

Improvement in OS discrimination ability compared to the classical approach

The input of six parameters identified in the final model (smoking status, liver metastases, pain, jaundice, ascites, and type of L2 regimen) in addition to those frequently used in multidisciplinary meeting ("pragmatic parameters": age, ECOG PS, and duration of L1) (Table S5) significantly improved its discrimination capacity since the C-statistic increased from 0.70 to 0.75 (bootstrap mean difference, 0.048; 95% CI, 0.016 to 0.081).

Similarly, their addition adequately reclassify patients at lower risk of death and those at higher risk, as reported by a cNRI of 0.79 (95% CI, 0.56 to 1.0; $P < 0.0001$) at 6 months (Figure S3), with an IDI of 0.14 (95% CI, 0.10 to 0.18; $P < 0.0001$) (Figure S4).

Prognostic nomogram and score for OS

A nomogram integrating all significant independent factors for OS was built (Figure 2). The prognostic score was based on the total number of points obtained from the nomogram. This prognostic score assumed a normal distribution (Figure S5). Based on the Cox method, patients were categorized into three risk groups with median OS of 11.3, 3.6, and 1.4 months ($P < 0.0001$) (Figure 3-A). Approaches with two risk groups are presented in Figure S6. Patient characteristics in each risk group are described in Table S6.

This prognostic score was applied to patients who did not receive a L2. Similarly, patients were categorized into the same three risk groups previously identified with significantly different prognostic profiles (Figure S7). Then, the risk groups were compared according to L2 administration or not (Figure S8) and described in Table S7. The benefit of L2 was higher in better prognostic groups with a significant interaction term between risk groups and L2 administration ($P = 0.0123$) (Table S8).

PFS analysis

PFS analysis was investigated in the OS risk groups identified. Interestingly, the discriminative ability of the three-group model was confirmed in L2 PFS analysis (Figure S9). Approaches for two risk groups are presented in Figure S10.

Clinico-biological model

The multivariate model with biological parameters was established based on patients for whom the eleven parameters were available including two independent biological factors: neutrophil-to-lymphocyte ratio (NLR), and carbohydrate antigen 19-9 (CA19-9) (Table S9). Supplementary analyses are specified in Appendix A3, and provided in Tables S10 (stratified and random approaches) and S11 (multiple imputations).

External validation of the prognostic score

Information for the nine baseline parameters that were required for the score calculation was available for 162 (99%) patients from the external validation cohort.

The discrimination ability of the risk score algorithm developed in the main analysis was externally confirmed, with a good discrimination ability (C-index = 0.63; 95% bootstrap percentile CI 0.58 to 0.69). We identified the same risk groups with OS applying to Kaplan-Meier curves (Figure 3-B).

Clinical benefit analysis

The clinical benefit centered accuracy of the final model was confirmed by a decision curve analysis in both cohorts (Figure 4).

DISCUSSION

PDAC is a highly aggressive cancer and chemotherapy remains the cornerstone of advanced disease therapy. While the clinical indications for L1 are well-described and consensual,[8] decision-making regarding L2 is much more controversial issue. OS improvement with L2 has been reported in patients with gemcitabine-refractory PDAC.[20,21] Nevertheless, the analysis of 34 clinical trials comparing L2 versus best supportive care showed that the benefit remains modest and limited to selected patients.[15] Moreover, due to lack of high-level evidence, no guidelines are available for L2 after FOLFIRINOX[6] or gemcitabine plus *nab*-paclitaxel L1 regimen.[7]

A prognostic model is a useful tool for clinical management by predicting patient life expectancy. The purpose of implementing such a scoring system, using easily measurable clinical parameters, is to guide clinical decision-making in the routine practice and to propose patient risk stratification for clinical trials. Currently, no validated prognostic model is available to estimate OS in patients with aPDAC at the beginning of L2. Some scores have been proposed but failed to complete validation process due to small sample sizes with insufficient statistical power.[22] More recently, Sinn et al[14] developed a score based on prognostic factors including ECOG-PS, CA19-9, and duration of L1, but the performance and internal validation of the final model were not assessed.

From a large population-based study, we established a new straightforward prognostic model. We identified prognostic factors for survival of patients with aPDAC who received L2 and derived a prognostic model based on nine easily available clinical parameters. Smoking status is a factor associated with an increased risk of PDAC,[23] and here we demonstrate that it is also a negative prognostic factor.[24] In the clinico-biological model, two parameters were identified, CA19-9 and NLR, whose relevance was previously demonstrated in the L1 setting.[25] We developed a robust prognostic nomogram for individual estimation of patient survival. In addition, the prognostic score allows classifying patients into three risk groups. Both tools are available in an open access version on a smartphone-compatible website (http://www.umqvc.org/en/tool/proscap.html).

Our results highlight considerable heterogeneity in survival in patients with PDAC who receive L2. Our prognostic score may help to refine patient selection for L2; indeed, a benefit from L2 was observed mostly in the low-risk group. Even if "pragmatic parameters" are already frequently used for decision-making in multidisciplinary meeting, our final model improves the discrimination ability compared to this classical approach. Interestingly, the risk groups identified based on OS matched with the risk groups for PFS. Therefore, through the PFS analysis, the prognostic value of our scoring system was confirmed independently from the "background noise" of the eventual following lines of chemotherapy, and for both patient-related (OS) and tumour-related (PFS) outcomes. Overall, we confirmed the survival benefit of L2 for all patients with aPDAC, although a decreasing clinical benefit was observed from low to high-risk groups.

Our study has several strengths. We evaluated a large population including patients treated with FOLFIRINOX regimen in both cohorts, and patients treated with *nab*-paclitaxel plus gemcitabine in the external validation cohort. Parameters used in the model are clinically relevant, simple to collect for clinicians, and consistent with previous studies. We built our model in a rigorous methodological framework respecting a recent recommended checklist for nomograms[26] and providing a transparent reporting of the multivariate model as suggested in the TRIPOD statement.[27] Moreover, discrimination, calibration, and internal validation demonstrated satisfactory performance and validity of the model. Our prognostic score was externally replicated in routine clinical practice in three other institutions. The clinical benefit centered accuracy of the final model was confirmed by the decision curve analysis.

The inclusion period in our study was long (11 years); meanwhile, clinical practices have changed with the approval of the FOLFIRINOX and gemcitabine plus *nab*-paclitaxel chemotherapy, resulting in heterogeneity in L1 treatments. This highlights the reliability of our model in the era of these new treatment regimens and not only in gemcitabine-refractory patients. Noticeably, it would be of interest to assess similar prognostic tools in homogenous patient groups for L1. Additional variables, particularly albumin or serum C-reactive protein levels, could not be evaluated in our study due to the retrospective nature of the data collection with high rate of missing data, and may be relevant to assess in further studies.[28,29] The

additional prognostic value for OS of health related quality of life (HRQoL) in metastatic disease suggests that HRQoL should also be evaluated in the future prognostic model for PDAC.[30]

In conclusion, we propose a novel OS prediction model based on nine key independent clinical prognostic factors at the beginning of L2 in patients with aPDAC. Our study highlights considerable heterogeneity in the survival in patients who receive L2. We identify three distinct risk groups in which chemotherapy was found to be significantly beneficial but with unequal benefit magnitude across risk groups. This study allowed us to develop and externally validate a prognostic model for OS and provide functional tools at individual (nomogram) and population (score) levels. These tools can accurately predict OS before administration of L2 and thereby improve risk stratification and patient selection for routine clinical decision-making and future clinical trials.

Acknowledgements

We thank Florian Limousin (graphic designer and independent web developer in Pearlweb society) for the web development application support.

Author Contributions

Conception and design: AV, AM, FB, CB, DV.

Collection and assembly of data: All authors.

Data analysis and interpretation: AV, CN, CB, DV.

Manuscript writing: All authors.

Final approval of manuscript: All authors.

AV had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests

LM: Consulting (Novartis, Ipsen Pharma, Pfizer), Research funding (Novartis), Travel grants (Ipsen Pharma, Novartis, Amgen).

CL: Consulting (Celgene, Roche, Nucana, Merck), Research funding (Roche), Travel grants (Celgene, Roche).

CT: Honoria (Roche, Sanofi, Bayer, Lilly), Consulting (Lilly, Sandoz), Research Funding (Roche), Travel grants (Roche, Sanofi.

OB: Honoria (Amgen, Lilly, Pierre-Fabre, Novartis), Consulting (Roche, Merck-Serono).

CN: Consulting (Celgene), Research Funding (Celgene).

FB: Honoria (Roche), Consulting (Roche), Resarch Funding (Roche).

CB: Honoria (Roche, BMS, MSD companies), Consulting (Lilly, Servier, Bayer), Research Funding (Roche)

DV: Consulting (Celgene, Janseen).

The remaining authors declare no conflict of interest.

Funding

None.

Ethics committee approval

CNIL declaration

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016: Cancer Statistics, 2016. CA Cancer J Clin. 2016 Jan;66(1):7–30.

2. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting Cancer Incidence and Deaths to 2030: The Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States. Cancer Res. 2014 Jun 1;74(11):2913–21.

3. SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. http://seer.cancer.gov/csr/1975_2013/.

4. Ryan DP, Hong TS, Bardeesy N. Pancreatic Adenocarcinoma. N Engl J Med. 2014 Sep 11;371(11):1039–49.

5. Neuzillet C, Tijeras-Raballand A, Bourget P, Cros J, Couvelard A, Sauvanet A, et al. State of the art and future directions of pancreatic ductal adenocarcinoma therapy. Pharmacol Ther. 2015 Nov;155:80–104.

6. Conroy T, Desseigne F, Ychou M, Bouché O, Guimbaud R, Bécouarn Y, et al. FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. N Engl J Med. 2011 May 12;364(19):1817–25.

7. Von Hoff DD, Ervin T, Arena FP, Chiorean EG, Infante J, Moore M, et al. Increased Survival in Pancreatic Cancer with nab-Paclitaxel plus Gemcitabine. N Engl J Med. 2013 Oct 31;369(18):1691–703.

8. Ducreux M, Cuhna AS, Caramella C, Hollebecque A, Burtin P, Goéré D, et al. Cancer of the pancreas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015 Sep;26(suppl 5):v56–68.

9. Sohal DPS, Mangu PB, Khorana AA, Shah MA, Philip PA, O'Reilly EM, et al. Metastatic Pancreatic Cancer: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol Off J Am Soc Clin Oncol. 2016 May 31;

10. Walker EJ. Beyond first-line chemotherapy for advanced pancreatic cancer: An expanding array of therapeutic options? World J Gastroenterol. 2014;20(9):2224.

11. Oettle H, Riess H, Stieler JM, Heil G, Schwaner I, Seraphin J, et al. Second-Line Oxaliplatin, Folinic Acid, and Fluorouracil Versus Folinic Acid and Fluorouracil Alone for Gemcitabine-Refractory Pancreatic Cancer: Outcomes From the CONKO-003 Trial. J Clin Oncol. 2014 Aug 10;32(23):2423–9.

12. Wang-Gillam A, Li C-P, Bodoky G, Dean A, Shan Y-S, Jameson G, et al. Nanoliposomal irinotecan with fluorouracil and folinic acid in metastatic pancreatic cancer after previous gemcitabine-based therapy (NAPOLI-1): a global, randomised, open-label, phase 3 trial. The Lancet. 2016 Feb;387(10018):545–57.

13. Portal A, Pernot S, Tougeron D, Arbaud C, Bidault AT, de la Fouchardière C, et al. Nab-paclitaxel plus gemcitabine for metastatic pancreatic adenocarcinoma after Folfirinox failure: an AGEO prospective multicentre cohort. Br J Cancer. 2015 Sep 29;113(7):989–95.

14. Sinn M, Dälken L, Striefler JK, Bischoff S, Schweitzer N, Pelzer U, et al. Second-Line Treatment in Pancreatic Cancer Patients: Who Profits?—Results From the CONKO Study Group. Pancreas. 2016 Apr;45(4):601–5.

15. Rahma OE, Duffy A, Liewehr DJ, Steinberg SM, Greten TF. Second-line treatment in advanced pancreatic cancer: a comprehensive analysis of published clinical trials. Ann Oncol. 2013 Aug 1;24(8):1972–9.

16. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Control Clin Trials. 1996 Aug;17(4):343–6.

17. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and mesuaring and reducing errors. Stat Med. 1996 Feb 29;15(4):361–87.

18. Cox DR. Note on Grouping. J Am Stat Assoc. 1957 Dec;52(280):543–7.

19. Fitzgerald M, Saville BR, Lewis RJ. Decision Curve Analysis. JAMA. 2015 Jan 27;313(4):409.

20. Pelzer U, Schwaner I, Stieler J, Adler M, Seraphin J, Dörken B, et al. Best supportive care (BSC) versus oxaliplatin, folinic acid and 5-fluorouracil (OFF) plus BSC in patients for second-line advanced pancreatic cancer: A phase III-study from the German CONKO-study group. Eur J Cancer. 2011 Jul;47(11):1676–81.

21. Chiorean EG, Von Hoff DD, Tabernero J, El-Maraghi R, Ma WW, Reni M, et al. Second-line therapy after nab-paclitaxel plus gemcitabine or after gemcitabine for patients with metastatic pancreatic cancer. Br J Cancer [Internet]. 2016 Jun 28 [cited 2016 Jul 2]; Available from: http://www.nature.com/doifinder/10.1038/bjc.2016.185

22. Kim ST, Choi YJ, Park KH, Oh SC, Seo JH, Shin SW, et al. A Prognostic Model to Identify Patients with Advanced Pancreas Adenocarcinoma Who Could Benefit from Second-line Chemotherapy. Clin Oncol. 2012 Mar;24(2):105–11.

23. Bosetti C, Lucenteforte E, Silverman DT, Petersen G, Bracci PM, Ji BT, et al. Cigarette smoking and pancreatic cancer: an analysis from the International Pancreatic Cancer Case-Control Consortium (Panc4). Ann Oncol. 2012 Jul 1;23(7):1880–8.

24. Kawaguchi C, Sho M, Tanaka T, Akahori T, Kinoshita S, Nagai M, et al. Impact of Smoking on Pancreatic Cancer Patients Receiving Current Chemotherapy: Pancreas. 2015 Oct;44(7):1155–60.

25. Goldstein D, El-Maraghi RH, Hammel P, Heinemann V, Kunzmann V, Sastre J, et al. nab-Paclitaxel Plus Gemcitabine for Metastatic Pancreatic Cancer: Long-Term Survival From a Phase III Trial. JNCI J Natl Cancer Inst. 2015 Jan 31;107(2):dju413–dju413.

26. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. Lancet Oncol. 2015 Apr;16(4):e173–80.

27. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. Br J Cancer. 2015 Jan 20;112(2):251–9.

28. Vernerey D, Huguet F, Vienot A, Goldstein D, Paget-Bailly S, Van Laethem J-L, et al. Prognostic nomogram and score to predict overall survival in locally advanced untreated pancreatic cancer (PROLAP). Br J Cancer [Internet]. 2016 Jul 12 [cited 2016 Jul 16]; Available from: http://www.nature.com/doifinder/10.1038/bjc.2016.212

29. Hurwitz HI, Uppal N, Wagner SA, Bendell JC, Beck JT, Wade SM, et al. Randomized, Double-Blind, Phase II Study of Ruxolitinib or Placebo in Combination With Capecitabine in Patients With Metastatic Pancreatic Cancer for Whom Therapy With Gemcitabine Has Failed. J Clin Oncol. 2015 Dec 1;33(34):4039–47.

30. Gourgou-Bourgade S, Bascoul-Mollevi C, Desseigne F, Ychou M, Bouche O, Guimbaud R, et al. Impact of FOLFIRINOX Compared With Gemcitabine on Quality of Life in Patients With Metastatic

Pancreatic Cancer: Results From the PRODIGE 4/ACCORD 11 Randomized Trial. J Clin Oncol. 2013 Jan 1;31(1):23–9.

Tables and Figures Titles and Legends

Figure 1: Flow chart

*Almost half of patients (43%) in L2 received a third-line or more (up to seven lines).

Abbreviations: 5-FU=5-Fluorouracil, L1=first-line chemotherapy, L2=second-line chemotherapy

Table 1: Patients characteristics according to second-line chemotherapy administration in the development cohort

Table 2: Prognostics factors associated with overall survival in univariate analysis

Table 3: Prognostics factors associated with overall survival in multivariate analysis (N = 248)

Figure 2: Prognostic nomogram to predict individual overall survival probability at the beginning of second-line chemotherapy in patients with advanced pancreatic ductal adenocarcinoma

Points are assigned for each prognostic factor by drawing a line upward from the corresponding values to the "point" line. The total sum of points for nine prognostic factors is plotted on the "total points" line. A line is drawn down to read the corresponding predictions of 6, 8, 12 and 24 month survival probability and median survival time.

An online web-based application smartphone-compatible was developed that provides individualized survival estimates from the nomogram: http://www.umqvc.org/en/tool/proscap.html.

Abbreviations: 5-FU=5-Fluorouracil

Figure 3: Kaplan-Meier curves of overall survival for three risk groups for the development cohort (A) and the external validation cohort (B)

An online web-based smartphone-compatible application was developed that provides risk group classification for OS from the prognostic score: http://www.umqvc.org/en/tool/proscap.html.

Abbreviations: CI=confidence interval

Figure 4: Decision curves to plot the net benefit achieved by making clinical decisions based on the final multivariate model risk of death predictions at 6 months in the development cohort (A) and in the external validation cohort (B)

Decision curve analysis shows a threshold of risk of death at 6 months at which decisions will cause greater benefit for true positives and decision of false positives will be reduced. The net benefit for decisions based on our final multivariate modelization for threshold values greater than 20% (A) and 40% (B) is better than considering patients on the same level of risk. Here the decision is defined by the possibility of not treating a patient by L2. In this context, the clinician must ensure that the patient is at high-risk of death at 6 months and minimize false positives, i.e. patients who are still alive at 6 months which must be treated. Overall, the decision curve shows that the net benefit for decision based on our final multivariate model is of interest, especially for thresholds > 0.5, the area of interest.

Table 1: Patients characteristics according to second-line chemotherapy administration in the development cohort

		Population eligible for second-line chemotherapy N = 395	Second-line chemotherapy administration N = 261	No second-line chemotherapy administration N = 134	p[a]
Demographic parameters	**Age, median [IQR], years**	67.2 [60.8 - 74.9]	65.8 [60.3 - 72.9]	71.5 [63.2 - 76.9]	**<0.0001**
	Sex, N (%)				0.3824
	Male	233 (59)	158 (61)	75 (56)	
	Female	162 (41)	103 (39)	59 (44)	
	Smoking status, N (%)				0.3022
	Never smoker or former smoker	296 (75)	200 (77)	96 (72)	
	Current smoker	97 (25)	60 (23)	37 (28)	
	Missing	2	1	1	
Pathologic parameters at diagnosis of cancer	**Primary tumour site, N (%)**				0.2540
	Head	237 (61)	152 (59)	85 (65)	
	Body and/or Tail	152 (39)	106 (41)	46 (35)	
	Missing	6	3	3	
	Tumour stage, N (%)				0.4153
	Localized	92 (23)	66 (25)	26 (19)	
	Locally advanced	96 (24)	61 (24)	35 (26)	
	Metastatic	207 (53)	134 (51)	73 (55)	
Tumour extension at the beginning of L2	**Number of metastatic sites, N (%)**				0.8655
	0 - 1	244 (62)	162 (62)	82 (61)	
	≥ 2	151 (38)	99 (38)	52 (39)	
	Liver metastases, N (%)				0.3497
	No	167 (42)	106 (41)	61 (46)	
	Yes	228 (58)	155 (59)	73 (54)	
	Peritoneal metastases, N (%)				0.6475
	No	242 (61)	162 (62)	80 (60)	
	Yes	153 (39)	99 (38)	54 (40)	
	Lung metastases, N (%)				0.4297
	No	304 (77)	204 (78)	100 (75)	
	Yes	91 (23)	57 (22)	34 (25	
	Bone metastases, N (%)				0.1403
	No	377 (95)	252 (97)	125 (93)	
	Yes	18 (5)	9 (3)	9 (7)	
Clinical parameters at the beginning of L2	**Performance status (WHO), N (%)**				**<0.0001**
	0	52 (15)	47 (19)	5 (6)	
	1	126 (38)	118 (47)	8 (9)	
	≥ 2	157 (47)	86 (34)	71 (85)	
	Missing	60	10	50	
	Body mass index, N (%),kg/m^2				**0.0201**
	Normal weight [18,5-25[	225 (60)	167 (64)	58 (49)	
	Underweight <18,5	66 (17)	39 (15)	27 (23)	
	Overweight [25-30[and Obesity ≥30	87 (23)	54 (21)	33 (28)	
	Missing	17	1	16	
	Pain[b], N (%)				0.5078
	No	165 (49)	119 (48)	46 (52)	
	Yes	174 (51)	131 (52)	43 (48)	
	Missing	56	11	45	
	Jaundice, N (%)				0.0728
	No	280 (83)	212 (85)	68 (76)	
	Yes	59 (17)	38 (15)	21 (24)	
	Missing	56	11	45	
	Ascites, N (%)				0.9220
	No	274 (80)	202 (80)	72 (80)	
	Yes	67 (20)	49 (20)	18 (20)	
	Missing	54	10	45	
Previous treatment at the beginning of L2	**Primary tumour resection, N (%)**				**0.0224**
	Yes	111 (28)	83 (32)	28 (21)	
	No	284 (72)	178 (68)	106 (79)	
First-line chemotherapy	**Type of L1 regimen, N (%)**				**0.0094**
	5-FU ± Oxaliplatin and/or Irinotecan	99 (25)	76 (29)	23 (17)	
	Gemcitabine based	296 (75)	185 (71)	111 (83)	
	RECIST Best response, N (%)				0.1428
	Complete or partial response or stability	223 (65)	170 (67)	53 (59)	
	Progression disease	119 (35)	82 (33)	37 (41)	
	Missing	53	9	44	

Table 1: Patients characteristics according to second-line chemotherapy administration in the development cohort (continued)

		Population eligible for second-line chemotherapy N = 395	Second-line chemotherapy administration N = 261	No second-line chemotherapy administration N = 134	p[a]
First-line chemotherapy	**Duration of L1, median [IQR], months**	4.7 [2.3 - 8.1]	5.7 [2.9 - 8.7]	2.5 [0.8 - 5.8]	**<0.0001**
	Reason for discontinuation, N (%)				0.9353
	Progression disease	265 (67)	176 (67)	89 (67)	
	Toxicity	22 (6)	15 (6)	7 (5)	
	Other	108 (27)	70 (27)	38 (28)	
Median follow-up time (95% CI), months		All patients were followed until death (maximum time observed = 126.1) except 28 censored patients with a median follow-up equal to 11.7	All patients were followed until death (maximum time observed = 126.1) except 14 censored patients with a median follow-up equal to 14.1	25.6 [14.6 – 101.0]	

Abbreviations : IQR=Interquartile Range, WHO=World Health Organization, 5-FU=5-Fluorouracil, L1=first-line chemotherapy, L2=second-line chemotherapy

[a] χ^2 tests or Fisher's exact tests used to compare proportions, and Wilcoxon tests used to compare continuous variables between the groups with or without second-line chemotherapy administration.

[b] corresponding to prescription of morphine

Table 2: Prognostics factors associated with overall survival in univariate analysis

		Number of Patients	Number of Events	HR	95% CI	p
Demographic parameters	**Age, years**	261	247	1.015	1.001 - 1.030	**0.0400**
	Sex					
	Male	158	152	1		
	Female	103	95	0.800	0.618 - 1.036	0.0690
	Smoking status					
	Never smoker or former smoker	200	187	1		
	Current smoker	60	59	1.532	1.136 - 2.066	**0.0052**
	Missing	1	1			
	Personal history of cancer					
	No	210	197	1		
	Yes	49	48	1.157	0.843 - 1.588	0.3661
	Missing	2	2			
	Family history of cancer					
	No	146	138	1		
	Yes	113	107	0.854	0.662 - 1.101	0.2230
	Missing	2	0			
	Family history of pancreatic cancer					
	No	240	227	1		
	Yes	19	18	1.178	0.728 - 1.906	0.5055
	Missing	2	2			
Pathologic parameters at diagnosis of cancer	**Primary tumour site**					
	Head	152	142	1		
	Body and/or Tail	106	102	1.249	0.966 - 1.615	0.0900
	Missing	3	3			
	Primary tumour size, mm	242	229	1.005	0.997 - 1.013	0.2097
	Missing	19	18			
	Histological grade					
	Well or moderately differentiated	100	92	1		
	Poorly differentiated or undifferentiated	28	26	1.330	0.856 - 2.065	0.2049
	Missing	133	129			
	Tumour stage					
	Localized	66	61	1		
	Locally advanced	61	57	1.148	0.798 - 1.652	
	Metastatic	134	129	1.798	1.319 - 2.450	**0.0002**
Tumour extension at the beginning of L2	**Number of metastatic sites**					
	0 - 1	162	152	1		
	□ 🗎	99	95	1.692	1.305 - 2.194	**<0.0001**
	Lymph node metastases					
	No	240	230	1		
	Yes	21	17	0.798	0.487 - 1.308	0.3708
	Liver metastases					
	No	106	100	1		
	Yes	155	147	1.742	1.345 - 2.255	**<0.0001**
	Peritoneal metastases					
	No	162	148	1		
	Yes	99	99	1.606	1.243 - 2.077	**0.0003**
	Lung metastases					
	No	204	193	1		
	Yes	57	54	0.928	0.685 - 1.257	0.6288
	Bone metastases					
	No	252	239	1		
	Yes	9	8	0.857	0.423 - 1.735	0.6685
	Other metastases					
	No	257	243	1		
	Yes	4	4	1.193	0.442 - 3.216	0.7279
	Isolated lung metastases					
	No	248	235	1		
	Yes	13	12	0.521	0.291 - 0.934	**0.0285**

Table 2: Prognostics factors associated with overall survival in univariate analysis (continued)

		Number of Patients	Number of Events	HR	95% CI	p
Clinical parameters at the beginning of L2	**Performance status (WHO)**					
	0	47	40	1		
	1	118	115	1.971	1.366 - 2.846	
	≥ 2	86	83	4.620	3.114 - 6.854	**<0.0001**
	Missing	10	9			
	Body mass index, kg/m²	260	247	0.984	0.950 - 1.019	0.3564
	Missing	1	0			
	Body mass index, kg/m²					
	Normal weight [18,5-25[	167	160	1		
	Underweight <18,5	39	36	1.181	0.821 - 1.699	
	Overweight [25-30[and Obesity ≥30	54	51	0.938	0.684 - 1.287	0.5594
	Missing	1	0			
	Pain[a]					
	No	119	109	1		
	Yes	131	128	1.887	1.454 - 2.451	**<0.0001**
	Missing	11	10			
	Jaundice					
	No	212	200	1		
	Yes	38	37	2.683	1.871 - 3.847	**<0.0001**
	Missing	11	10			
	Ascites					
	No	202	190	1		
	Yes	49	48	2.137	1.549 - 2.948	**<0.0001**
	Missing	10	9			
Biological parameters at the beginning of L2	**Hemoglobin, g/dL**	233	223	0.822	0.753 - 0.896	**<0.0001**
	Missing	28	24			
	Neutrophils, mm³ (square root value)	219	210	1.016	1.011 - 1.022	**<0.0001**
	Missing	42	37			
	Lymphocytes, mm³ (inverse transformation value)	212	204	2.28E134	1.336E30 - 3.88E238	**0.0115**
	Missing	49	43			
	Neutrophil-to-lymphocyte ratio (square root value)	210	202	1.666	1.420 - 1.956	**<0.0001**
	Missing	51	45			
	Platelets, mm³ (log value)	230	220	1.528	0.868 - 2.689	0.1420
	Missing	31	27			
	Creatinine, µmol/L (log value)	247	234	0.786	0.261 - 2.363	0.6681
	Missing	14	13			
	Total bilirubin, µmol/L (square root value)	218	209	1.154	1.090 - 1.221	**<0.0001**
	Missing	43	38			
	Albumin, g/L	161	155	0.935	0.913 - 0.958	**<0.0001**
	Missing	100	92			
	CA19-9, UI/mL (square root value)	204	195	1.007	1.005 - 1.009	**<0.0001**
	Missing	57	52			
	CEA, ng/mL (log value)	159	150	1.855	1.468 - 2.343	**<0.0001**
	Missing	102	97			
Previous treatment at the beginning of L2	**Primary tumour resection**					
	Yes	83	76	1		
	No	178	171	1.575	1.198 - 2.072	**0.0012**
	Adjuvant chemotherapy					
	Yes	67	61	1		
	No	194	186	1.586	1.184 - 2.125	**0.0020**
	Radiotherapy					
	Yes	17	15	1		
	No	244	232	1.112	0.659 - 1.876	0.6902
	Neo-adjuvant chemotherapy					
	Yes	3	3	1		
	No	258	244	1.005	0.321 - 3.147	0.9935
	Biliary stent					
	Yes	83	78	1		
	No	177	168	0.843	0.644 - 1.103	0.2131
	Missing	1	1			

Table 2: Prognostics factors associated with overall survival in univariate analysis (continued)

		Number of Patients	Number of Events	HR	95% CI	p
Previous treatment at the beginning of L2	**Duodenal stent**					
	Yes	19	18	1		
	No	241	228	0.816	0.503 - 1.322	0.4087
	Missing	1	1			
	Alcohol celiac plexus					
	Yes	22	21	1		
	No	238	225	0.746	0.475 - 1.172	0.2037
	Missing	1	1			
First-line chemotherapy	**Type of L1 regimen**					
	5-FU ± Oxaliplatin and/or Irinotecan	76	69	1		
	Gemcitabine based	185	178	1.112	0.841 - 1.470	0.4565
	Number of cures	261	247	0.987	0.976 - 0.997	**0.0106**
	RECIST Best response					
	Complete or partial response or stability	170	161	1		
	Progression disease	82	77	1.379	1.049 - 1.813	**0.0211**
	Missing	9	9			
	Duration of L1 (log value), months	261	247	0.554	0.392 - 0.784	**0.0008**
	Toxicity of grade 3 or 4					
	No	198	189	1		
	Yes	63	58	0.830	0.617 - 1.115	0.2150
	Type of toxicity					
	Digestive	8	7	1		
	Neurological	17	14	0.959	0.385 - 2.388	
	Skin	4	4	2.291	0.652 - 8.053	
	Hematology	20	20	1.727	0.724 - 4.117	
	Other	14	13	1.702	0.670 - 4.323	0.2984
	Missing	198	189			
	Reason for discontinuation					
	Other	70	61	1		
	Toxicity	15	15	1.271	0.721 - 2.240	
	Progression disease	176	171	1.576	1.174 - 2.114	**0.0095**
	Locoregional progression					
	No	207	196	1		
	Yes	54	51	0.938	0.687 - 1.281	0.6856
	Metastatic progression					
	No	67	61	1		
	Yes	194	186	1.438	1.073 - 1.927	**0.0152**
Second-line chemotherapy	**Type of L2 regimen**					
	5-FU ± Oxaliplatin and/or Irinotecan	204	192	1		
	Gemcitabine based	57	55	1.489	1.099 - 2.018	**0.0102**

Abbreviations : HR=Hazard Ratio, CI=confidence interval, WHO=World Health Organization, CA 19-9=Carbohydrate Antigen 19-9, CEA=Carcinoembryonic Antigen, 5-FU=5-Fluorouracil, L1=first-line chemotherapy, L2=second-line chemotherapy

[a]corresponding to prescription of morphine

Table 3: Prognostics factors associated with overall survival in multivariate analysis (N = 248)

		Number of Patients	Number of Events	HR	95% CI	p	Internal validation BCA HR 95%	β	Max scores in nomogram
Demographic parameters	**Age, years**	248	235	1.019	1.003 - 1.036	**0.0168**	1.00237 - 1.03901	0.01921	55
	Smoking status								
	Never smoker or former smoker	190	178	1					
	Current smoker	58	57	1.491	1.086 - 2.046	**0.0134**	1.01718 - 2.13902	0.39946	21
Tumour extension at the beginning of L2	**Liver metastases**								
	No	99	93	1					
	Yes	149	142	2.149	1.620 - 2.850	**<0.0001**	1.48348 - 2.93914	0.76502	40
Clinical parameters at the beginning of L2	**Performance Status (WHO)**								
	0	47	40	1					
	1	117	114	1.614	1.067 - 2.374		1.00897 - 2.39352	0.47879	25
	□	84	81	3.055	1.935 - 4.822	**<0.0001**	1.74622 - 4.72267	1.11671	58
	Pain[a]								
	No	118	108	1					
	Yes	130	127	1.411	1.059 - 1.880	**0.0186**	0.98883 - 1.85176	0.34442	18
	Jaundice								
	No	210	198	1					
	Yes	38	37	1.958	1.327 - 2.890	**0.0007**	1.18897 - 2.98240	0.67192	35
	Ascites								
	No	200	188	1					
	Yes	48	47	1.792	1.264 - 2.540	**0.0011**	1.20746 - 2.76271	0.58335	30
Treatment	**Duration of L1 (log value), months**	248	235	0.478	0.325 - 0.702	**0.0002**	0.33244 - 0.81427	-0.73868	100
	Type of L2 regimen								
	5-FU ± Oxaliplatin and/or Irinotecan	195	184	1					
	Gemcitabine based	53	51	1.765	1.268 - 2.457	**0.0008**	1.18119 - 2.49164	0.56827	30

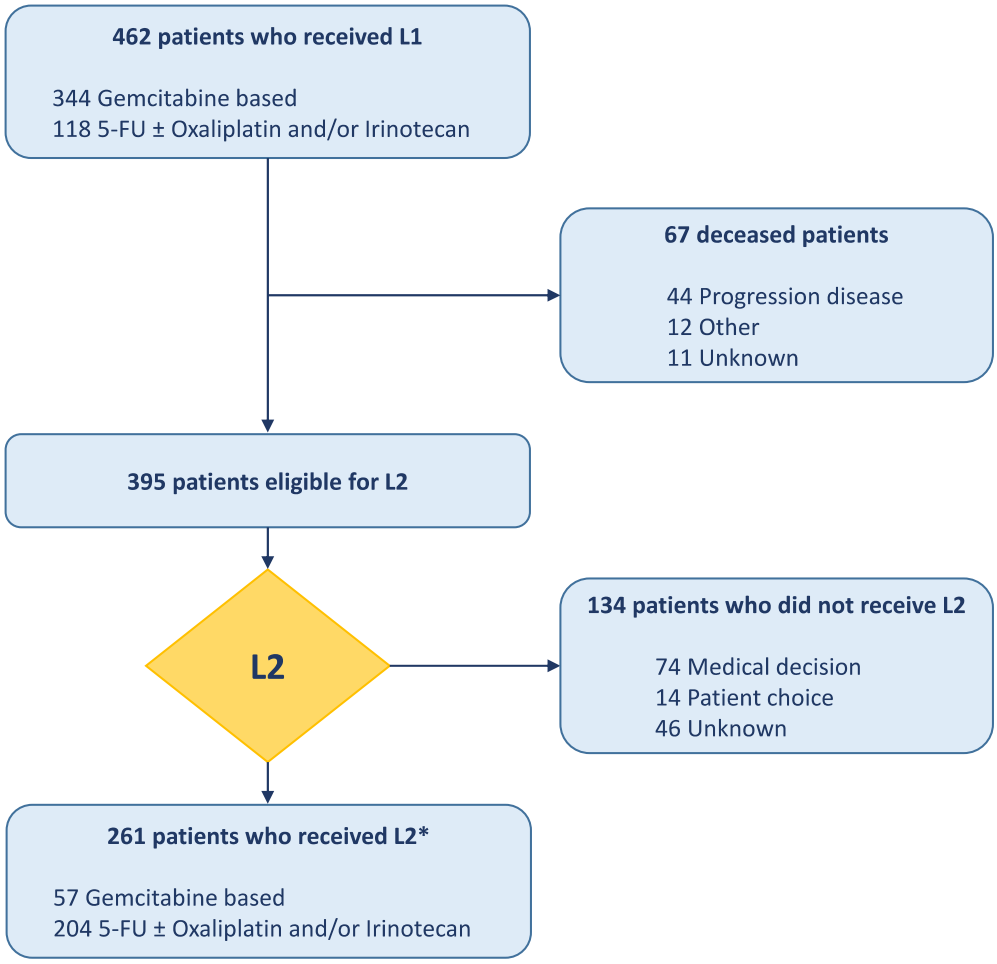
Abbreviations : HR=Hazard Ratio, CI=confidence interval, BCA=bias corrected and accelerated bootstrap, WHO=World Health Organization, 5-FU=5-Fluorouracil, L1=first-line chemotherapy, L2=second-line chemotherapy

The final multivariate Cox model was obtaine[illegible] proportional hazards model.

[a]corresponding to prescription of morphine

Figure 1

Figure 1: Flow chart

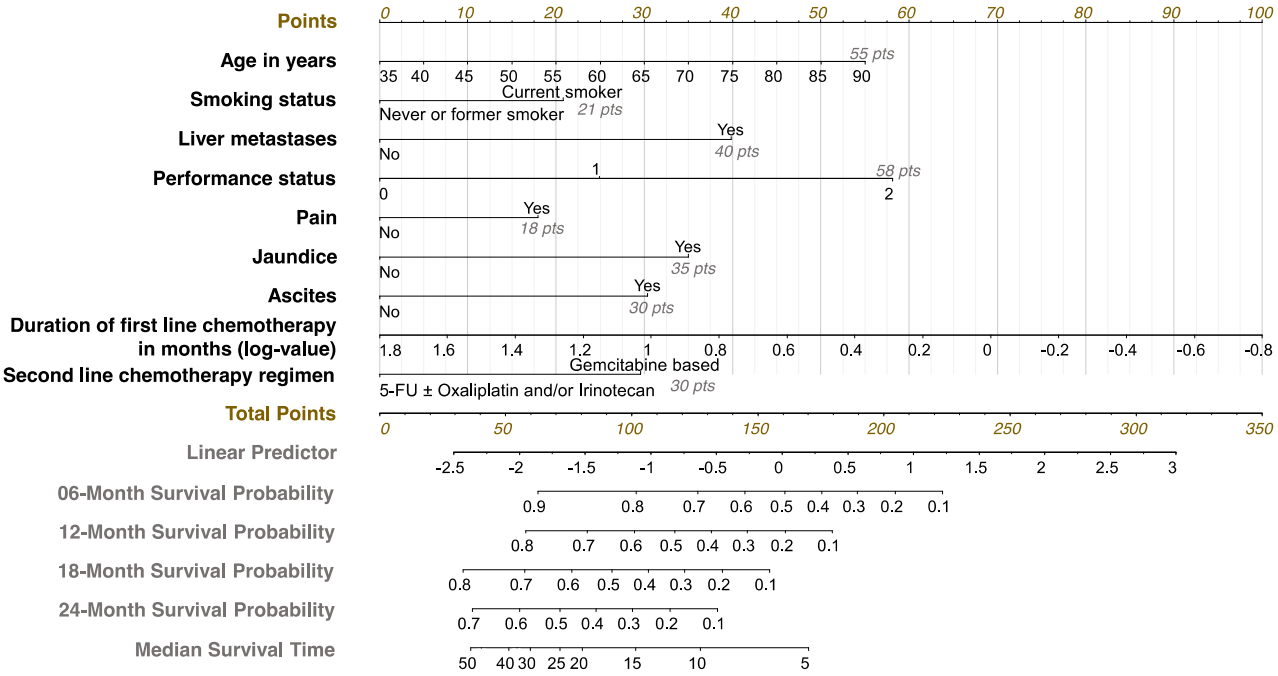


*Almost half of patients (43%) in L2 received a third-line or more (up to seven lines).

Abbreviations: 5-FU=5-Fluorouracil, L1=first-line chemotherapy, L2=second-line chemotherapy

Figure 2

Figure 2: Prognostic nomogram to predict individual overall survival probability at the beginning of second-line chemotherapy in patients with advanced pancreatic ductal adenocarcinoma



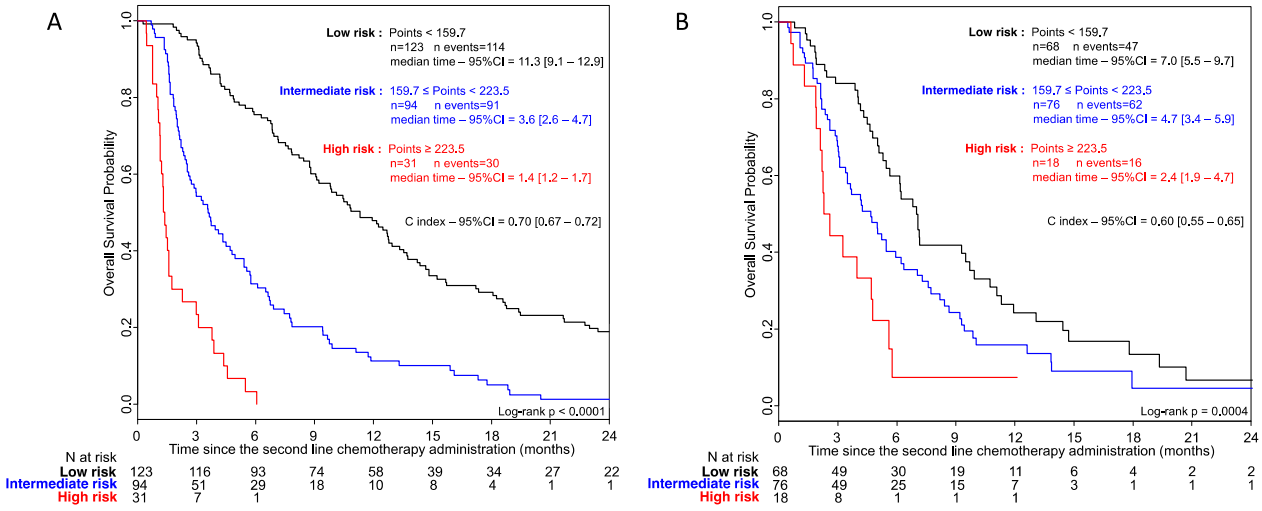
Abbreviations: 5-FU=5-Fluorouracil

Points are assigned for each prognostic factor by drawing a line upward from the corresponding values to the "point" line. The total sum of points for nine prognostic factors is plotted on the "total points" line. A line is drawn down to read the corresponding predictions of 6, 8, 12 and 24 month survival probability and median survival time.

An online web-based application smartphone-compatible was developed that provides individualized survival estimates from the nomogram: http://www.umqvc.org/en/tool/proscap.html.

Figure 3 Parts A-B

Figure 3: Kaplan-Meier curves of overall survival for three risk groups for the development cohort (A) and the external validation cohort (B)

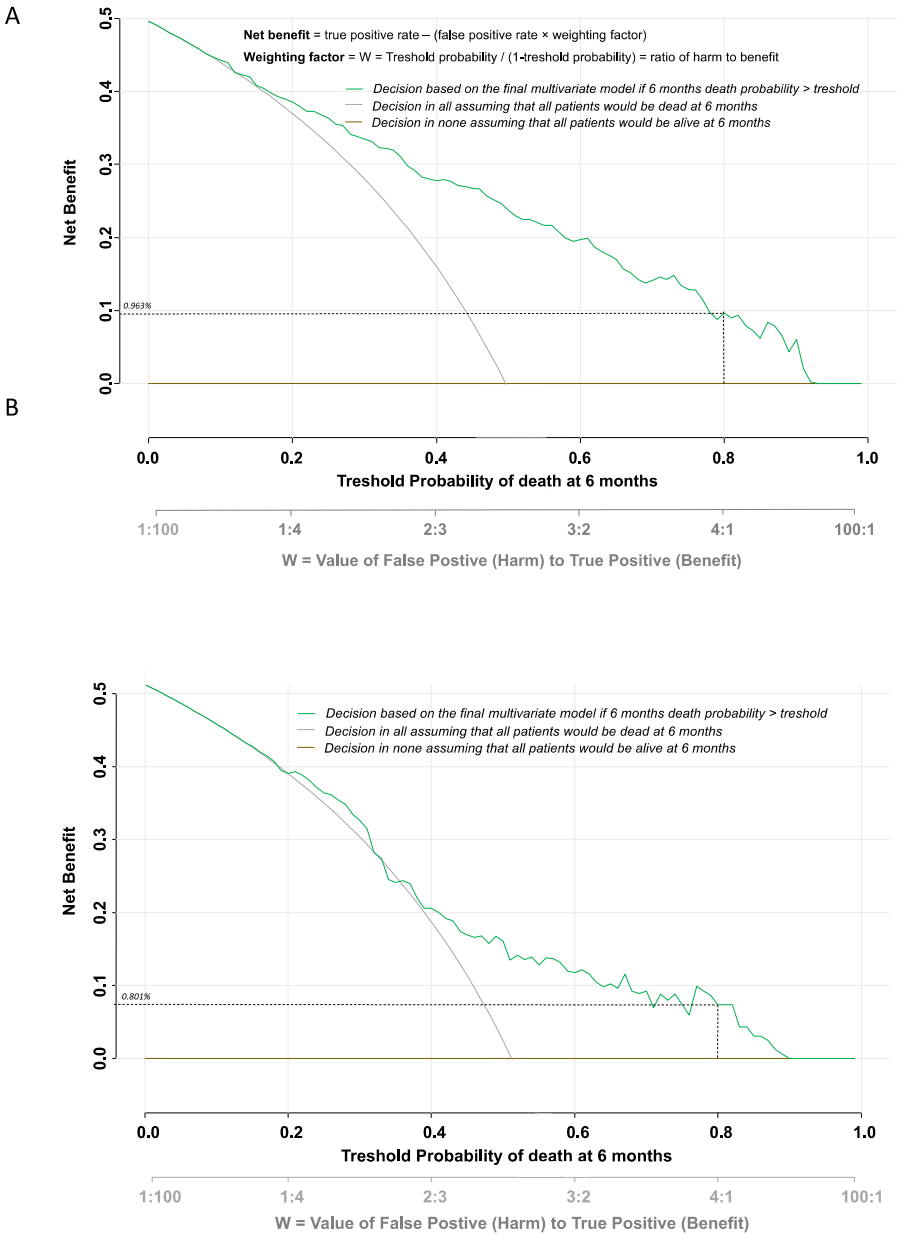


Abbreviations: CI=confidence interval

An online web-based smartphone-compatible application was developed that provides risk group classification for OS from the prognostic score: http://www.umqvc.org/en/tool/proscap.html.

Figure 4 Parts A-B

Figure 4: Decision curves to plot the net benefit achieved by making clinical decisions based on the final multivariate model risk of death predictions at 6 months in the development cohort (A) and in the external validation cohort (B)



Decision curve analysis shows a threshold of risk of death at 6 months at which decisions will cause greater benefit for true positives and decision of false positives will be reduced. The net benefit for decisions based on our final multivariate modelization for threshold values greater than 20% (A) and 40% (B) is better than considering patients on the same level of risk. Here the decision is defined by the possibility of not treating a patient by L2. In this context, the clinician must ensure that the patient is at high-risk of death at 6 months and minimize false positives, i.e. patients who are still alive at 6 months which must be treated. Overall, the decision curve shows that the net benefit for decision based on our final multivariate model is of interest, especially for thresholds > 0.5, the area of interest.

IV.2 Axe transplantation

IV.2.1 Pronostic des transplantations rénales réalisées à partir de donneurs à critères élargis.

Aubert O*, Kamar N*, Vernerey D*, Viglietti D, Martinez F, Duong-Van-Huyen J-P, Eladari D, Empana J-P, Rabant M, Verine J, Rostaing L, Congy N, Guilbeau-Frugier C, Mourad G, Garrigue V, Morelon E, Giral M, Kessler M, Ladrière M, Delahousse M, Glotz D, Legendre C, Jouven X, Lefaucheur C, Loupy A

Long term outcomes of transplantation using kidneys from expanded criteria donors: prospective, population based cohort study. **BMJ** 351: h3557. (2015)

*Co first authorship

Impact factor: 20.0

IV.2.1.1 Positionnement du problème

Le vieillissement de la population mondiale a modifié le paysage des maladies chroniques avec notamment une prévalence accrue de patients âgés en phase terminale d'insuffisance rénale. Pour ces patients, la transplantation rénale est la meilleure approche dépassant la dialyse en termes de qualité et de quantité de vie.

Dans le contexte des donneurs décédés (environ 85% des donneurs de reins en 2014 en France), ce vieillissement de la population marque une augmentation constante et spectaculaire de l'âge et aussi des comorbidités cardiovasculaires des donneurs. Ceci a conduit à la définition de donneurs à critères dits élargis. En effet, pour pallier à la pénurie d'organes, la sélection des reins issus de donneurs décédés a été étendue à partir des années 90 et surtout depuis les années 2000 en France. Les reins initialement nommés « marginaux » ont été rebaptisés reins de « donneurs à critères élargis » en 1997 par Kaufman et al. C'est en 2002 que des critères plus précis de donneur à critères élargis (ECD) ont été établis et publiés.

Il s'agissait des critères suivants :

- Donneur âgé de plus de 60 ans ;
- Ou donneur âgé de 50 à 59 ans avec 2 des 3 comorbidités suivantes :
    - Antécédent d'HTA,
    - Décès d'un accident vasculaire cérébral,
    - Créatinine supérieure à 1,5 mg/dL (132,6 µmol/L).

L'utilisation des greffons ECD répond donc à une double contrainte : la nécessité d'une part d'augmenter le nombre de greffons disponibles et d'autre part de prendre en compte l'évolution de l'épidémiologie des donneurs notamment de leur âge et comorbidités.

Aujourd'hui, environ 30% en Europe et 47% en France des reins issus de donneurs décédés sont des reins ECD. Dans les années futures les reins ECD seront la principale source d'organe pour la transplantation.

En dépit de leur potentiel, l'utilisation de reins ECD a évolué de façon inégale de telle sorte que, par exemple aux Etats-Unis, ils ne concernent que 16,6% des transplantations d'organes provenant de donneurs décédés. Ces disparités sont le reflet d'incertitudes persistantes concernant le pronostic et les déterminants de la survie de ces greffons. En effet, bien que les transplantations avec ces reins produisent un avantage en termes de survie du patient comparé à la dialyse elles sont malgré tout associées à un risque accru de perte de greffon par rapport aux transplantations avec des reins de donneurs à critères standards (SCD). Plus de 10 ans après la définition des donneurs à critères élargis, aucune étude en population n'a abordé la recherche des principaux déterminants pronostiques de la survie du greffon ECD en intégrant un spectre exhaustif de paramètres regroupant les paramètres du donneur, du receveur et le phénotype des anticorps anti-HLA dirigés contre le donneur.

Augmenter le nombre de transplantations en diminuant le nombre de greffons rénaux refusés par excès a été pointé comme une problématique cruciale par le comité international de Banff réuni en 2013 au Brésil. La nécessité de cette étude s'inscrit dans cette logique pour améliorer et guider l'utilisation des reins ECD grâce à une meilleure compréhension de leur pronostic à long terme.

L'objet de cette étude est donc d'identifier les facteurs pronostiques des reins ECD dans le but de caractériser des groupes de patients transplantés avec ces reins dont le pronostic des greffons est distinct et ainsi optimiser leur attribution.

IV.2.1.2 Population de l'étude

L'ensemble des patients transplantés rénaux entre le 1er janvier 2004 et le 1er janvier 2011 à l'Hôpital Necker (Pais), à l'Hôpital Saint-Louis (Paris), à l'Hôpital Foch (Suresnes) et à l'hôpital de Toulouse ont été inclus dans la cohorte de développement de cette étude en population (n=2763).

La cohorte de validation comporte 4128 patients transplantés dans les hôpitaux de Lyon, Montpellier, Nancy et Nantes entre le 1er janvier 2002 et le 31 décembre 2011.

Le système d'attribution des greffons était identique pour les 3 centres et suivait les règles de l'agence nationale française pour l'attribution des organes (Agence de la Biomédecine).

IV.2.1.3 Données

Pour le recueil des données nous avons utilisé les informations issues de deux bases de données informatisées : DIVAT (Données informatiques Validées en Transplantation) ainsi que les données issues du registre CRISTAL de l'Agence de la biomédecine.

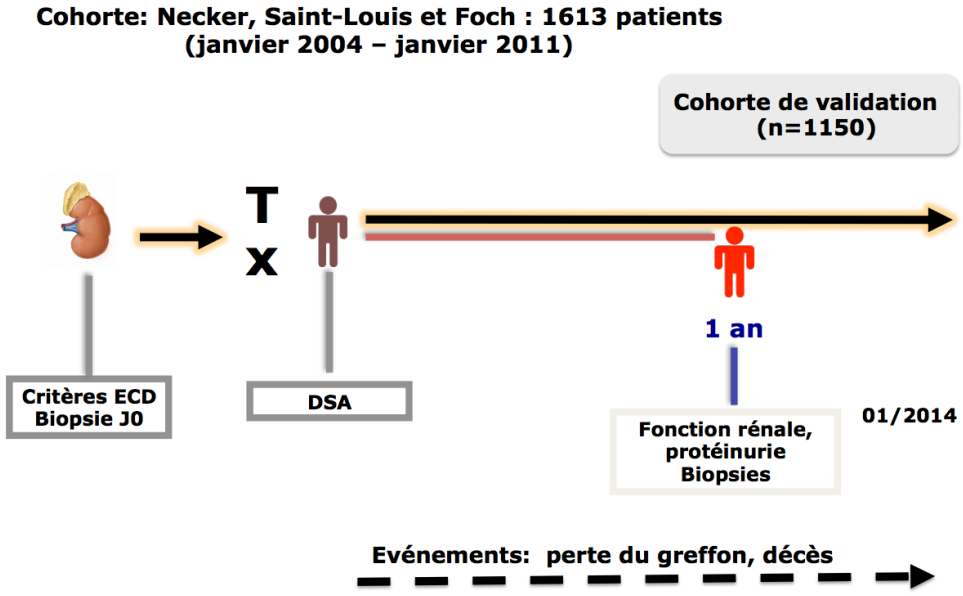
Par ailleurs, l'exhaustivité des bases de données a été complétée par un recueil d'informations sur d'autres supports, notamment pour obtenir un classement précis des donneurs selon les critères standard ou ECD, ainsi que l'obtention des données immunologiques, histologiques et de suivi.

La détection et l'identification des anticorps anti-HLA le jour de la transplantation ont été effectuées par technique de Luminex.

Pour la cohorte de validation, le DSA n'étant pas disponible le jour de la transplantation, la quantification de l'immunisation a été faite en considérant le cytotoxic PRA (panel reactive antibody) pour les patients transplantés avant

juillet 2009 et le cPRA (calculated PRA) pour les patients transplantés après juillet 2009.
Les biopsies ont été gradées de 0 à 3 selon la dernière classification de Banff mise à jour en 2013.
Tous les patients ont reçu un traitement standard d'induction le jour de la transplantation. Le traitement de maintenance est décrit précisément dans la publication associée.

Ci dessous voici graphiquement le design de l'etude:



IV.2.1.4 Approche méthodologique

La durée de survie du greffon a été calculée en utilisant la méthode de Kaplan-Meier. La survie a été analysée depuis la transplantation jusqu'à un maximum de 7 ans après la transplantation, avec comme évènement d'intérêt la perte du greffon définie par le retour en dialyse. Les données de survie du greffon ont été censurées au moment du décès du receveur. La survie des greffons rénaux a été représentée sur des courbes de Kaplan-Meier stratifiées selon le type de donneur (ECD ou SCD) et la présence ou non de DSA le jour de la greffe et comparées selon le test du log-rank.

Un modèle semi-paramétrique des risques proportionnels a été utilisé afin de quantifier les rapports de risque et les intervalles de confiances à 95% pour la perte du greffon. L'association des caractéristiques du donneur, du receveur, les données immunologiques et histologiques avec la perte du greffon a été évaluée dans des modèles de Cox univariés puis multivariés. Toutes les variables avec une p-value inférieure ou égale à 0.10 ont été incluses dans le modèle multivarié. Les facteurs identifiés étaient ensuite inclus dans un modèle final multivarié avec élimination pas à pas ascendante (backward stepwise). La validation interne du modèle a été réalisée par une méthode de ré-échantillonnage avec tirage aléatoire avec remise (bootstrap). Une validation externe de tous les résultats obtenus a été réalisée.

IV.2.1.5 Résultats

Population

Au total, 2 763 patients ont été inclus dans l'analyse principale. La Table 1 résume les caractéristiques des donneurs et receveurs dans l'ensemble de la population. En ce qui concerne la validation externe, celle-ci comprenait 4128 patients.

Table 1: Caractéristiques des receveurs et donneurs de la cohorte de validation (N=1150) et de la cohorte de développement (N=1613)

	Cohorte de validation externe (n=4128)		Cohorte de développement (n=2763)		P
	n		n		
Caractéristiques du receveur					
Age, moyenne (erreur standard), années	4128	49.36 (13.6)	2763	49.23 (13.5)	0.6913
Sexe masculin, No. (%)	4128	2548 (61.72)	2763	1643 (59.46)	0.0600
Causes d'insuffisance rénale	4128		2763		
Glomérulaires, No. (%)		1101 (26.67)		767 (27.76)	
Diabètes, No. (%)		622 (15.07)		264 (9.56)	
Hypertension, No. (%)		237 (5.74)		212 (7.67)	
Autre, No. (%)		2168 (52.52)		1520 (55.01)	**<0.0001**
Durée en dialyse, moyenne (erreur standard), années	4128	2.79 (3.44)	2763	3.90 (4.41)	**<0.0001**
Caractéristiques du donneur					
Age, moyenne (erreur standard), années	4127	48.16 (16.05)	2763	50.31 (15.98)	**<0.0001**
Age en catégorie, No. (%), années	4127		2763		
0-49		2043 (49.50)		1222 (44.23)	
50-59		1005 (24.35)		733 (26.53)	
>60		1079 (26.15)		808 (29.24)	**<0.0001**
Sexe masculin, No. (%)	4128	2457 (59.52)	2763	1509 (54.61)	**<0.0001**
Hypertension, No. (%)	3414	802 (23.49)	2709	661 (24.40)	0.4080
Diabètes, No. (%)	2271	47 (2.07)	2604	142 (5.45)	**<0.0001**
Créatinine > 1.5 mg/dL, No. (%)	4008	495 (12.35)	2736	270 (9.87)	**<0.0001**
Type du donneur					
Donneur décédé, No. (%)	4128	3755 (90.96)	2763	2370 (85.78)	**<0.0001**
Décès d'origine cardiovasculaire, No. (%)	3755	1963 (52.28)	2370	1323 (55.82)	0.0070
ECD, No. (%)	4015	1245 (31.01)	2763	916 (33.15)	0.0630
ECD/donneurs décédés, No. (%)	3646	1245 (34.15)	2370	916 (38.65)	**<0.0001**
Caractéristiques de base de la greffe					
Rang de greffe > 1, No. (%)	4128	611 (14.80)	2763	413 (14.95)	0.8670
Ischémie froide, No. (%)	4126		2753		
< 12 heures		803 (19.46)		674 (24.48)	
[12-24[heures		2319 (56.21)		1560 (56.67)	
≥ 24 heures		1004 (24.33)		519 (18.85)	**<0.0001**
HLA A/B/DR mismatch, moyenne (erreur standard),	4128	3.30 (1.37)	2700	3.12 (1.43)	**<0.0001**
Reprise retardée de fonction, No. (%)	4009	1151 (28.71)	2603	707 (27.16)	0.1710
Suivi (années), mediane, 95%[IC]	4128	6.08 (5.95 to 6.23)	2763	5.54 (5.42 to 5.68)	**<0.0001**
Survenues de décès, No. (%)	4128	366 (8.87)	2763	263 (9.52)	0.3570
Perte du greffon, No. (%)	4128	595 (14.41)	2763	333 (12.05)	**0.0050**

Abbreviations: ECD, expanded criteria donor, ESRD, end-stage renal disease, HLA, human leucocyte antigen, IC, confidence interval

Déterminants de la perte du greffon dans la cohorte de developpement

En analyse de Cox univariée, les paramètres initiaux de la greffe significativement associés à la perte du greffon étaient : l'âge du donneur, le type de donneur, l'antécédent d'hypertension du donneur, le critère ECD, la retransplantation, l'ischémie froide, le nombre d'incompatibilités HLA A/B/DR,

la présence d'un DSA le jour de la transplantation. Le diabète et la créatinine du donneur n'étaient pas significativement associés avec la perte du greffon (p = 0,5212 et p = 0,9004 respectivement).

L'analyse multivariée a identifié 5 facteurs majeurs indépendamment associés à la perte du greffon dont le critère donneur ECD et la présence de DSA circulant le jour de la transplantation (table 2).

Table 2: Facteurs associés avec la perte du greffon: Analyse multivariée (N=1613) [a]

		Nombre de patients	Nombre de perte de greffon	HR	95% IC	P	Validation interne HR 95% IC bootstrap BCA
ECD	**Non**	1835	187	1	-		-
	Oui	855	138	1.842	(1.467 to 2.311)	**<0.0001**	(1.463 to 2.275)
Temps d'ischémie froide	**<12 heures**	670	44	1	-		-
	12-24 heures	1514	198	1.457	(1.042 to 2.039)		(1.042 to 2.093)
	≥24 heures	506	83	1.727	(1.185 to 2.517)	**0.0173**	(1.195 to 2.506)
Rang de greffe	**1**	2278	241	1	-		-
	>1	412	84	1.544	(1.168 to 2.042)	**0.0023**	(1.129 to 2.046)
No. of HLA A/B/DR mismatches		2690	325	1.095	(1.013 to 1.184)	**0.0224**	(1.013 to 1.182)
Anti-HLA DSA à la transplantation	**Non**	2364	241	1	-		-
	Oui	326	84	2.988	(2.265 to 3.941)	**<0.0001**	(2.198 to 3.940)

Validation interne et externe du modèle

Validation interne

Pour la validation interne, nous avons mis en œuvre une technique de ré-échantillonnage par tirage au sort avec remise (Bootstrapping). Cette méthode a permis de ré-échantillonner la cohorte principale en générant 1 000 tirages au sort (1 000 bases de données générées à partir de la base initiale et 1 000 modèles multivariés ou les risques relatifs des variables d'intérêt étaient recalculés). L'incertitude autour des rapports de risque,

mesurés avec la procédure de tirage au sort avec remise, reflète le caractère robuste du modèle final (Tableau 2).

Validation externe

En ce qui concerne la validation externe, celle-ci comprenait 1 150 patients. L'âge médian des donneurs était de 50 ans (Interquartile: 38 – 59). Les caractéristiques cliniques de la cohorte de validation externe sont résumées dans la Table 1

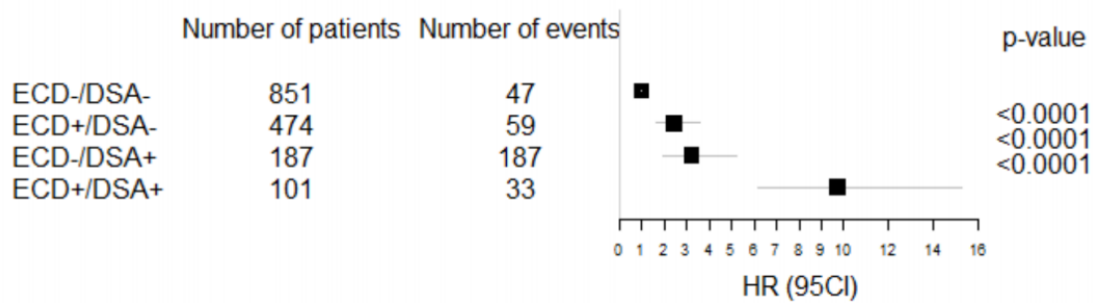
Le modèle multivarié a pu être répliqué dans sa totalité comme ci-dessous:

Table 3: Facteurs indépendamment associés à la perte du greffon dans la cohorte de validation externe en analyse multivarié.

		Nombre de patients	Nombre de perte de greffon	HR	95% CI	P
ECD	**Non**	2375	290	1	-	
	Oui	1235	235	1.692	(1.418 to 2.020)	**<0.0001**
Temps d'ischémie froide	**< 12 heures**	694	57	1	-	
	12-24 heures	2006	303	1.442	(1.078 to 1.928)	
	≥ 24 heures	910	165	1.687	(1.244 to 2.287)	**0.0033**
Rang de greffe	**1**	3030	414	1	-	
	> 1	580	111	1.251	(0.964 to 1.624)	**0.0920**
No. de mismatches HLA A/B/DR		3610	525	1.074	(1.004 to 1.149)	**0.0370**
Présence d'anticorps Anti-HLA le jour de la transplantation	**PRA < 5%**	2724	364	1	-	
	PRA ≥ 5%	886	161	1.380	(1.099 to 1.731)	**0.0060**

En résumé, l'analyse multivariée a donc permis d'identifier 5 facteurs majeurs indépendamment associés à la perte du greffon dont les principaux sont le critère du donneur ECD et la présence de DSA circulant le jour de la transplantation. Sur la base de ce constat on peut se demander si cumuler les risques associés aux deux caractéristiques est problématique.

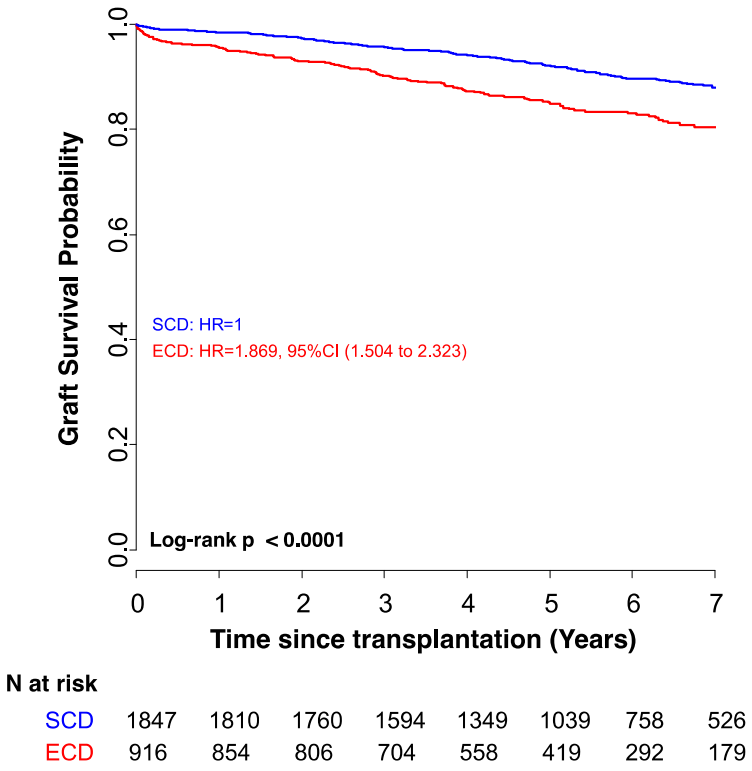
Figure 1: Rapports de risque de perte du greffon en fonction de la considération conjointe du type de donneur (ECD/SCD) et du risque immunologique initial à J0 de la greffe (DSA+/DSA-)



Survie des greffons rénaux dans la cohorte de développement

La survie post-transplantation des greffons issue de donneurs SCD était significativement meilleure que celle des greffons issus d'un donneur ECD ($p < 0,0001$, Figure 2).

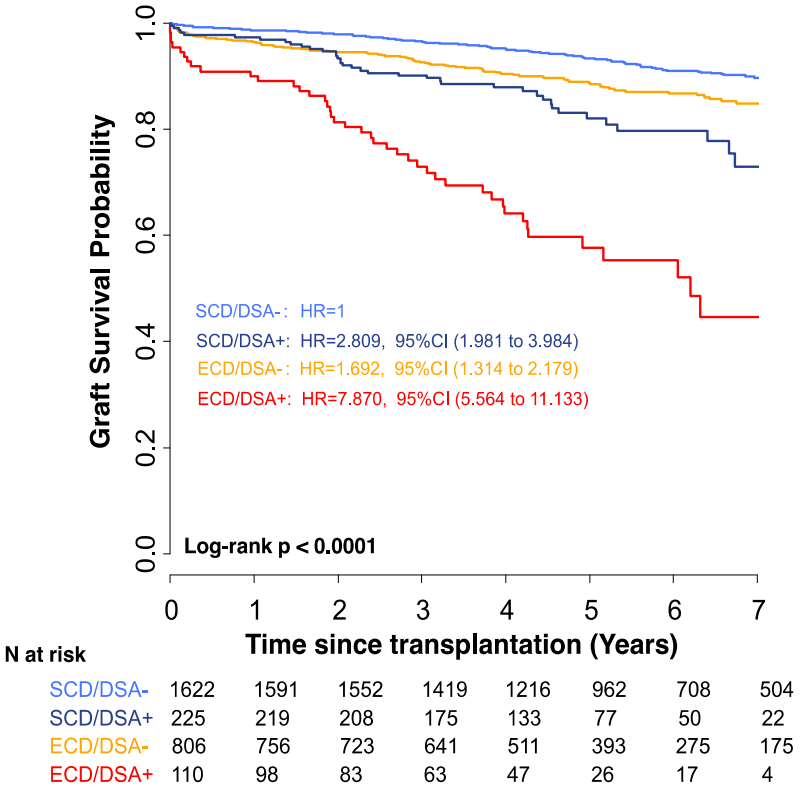
Figure 2 : Survie à long terme du greffon par la méthode de Kaplan-Meier selon le type de donneur : SCD (Donneurs à critères standard), ECD (Donneurs à critères élargis).



N at risk

SCD	1847	1810	1760	1594	1349	1039	758	526
ECD	916	854	806	704	558	419	292	179

Figure 3: Survie à long terme du greffon par la méthode de Kaplan-Meier selon la combinaison du type de donneur : SCD (Donneurs à critères standard), ECD (Donneurs à critères élargis) et de la présence de DSA circulant le jour de la transplantation.

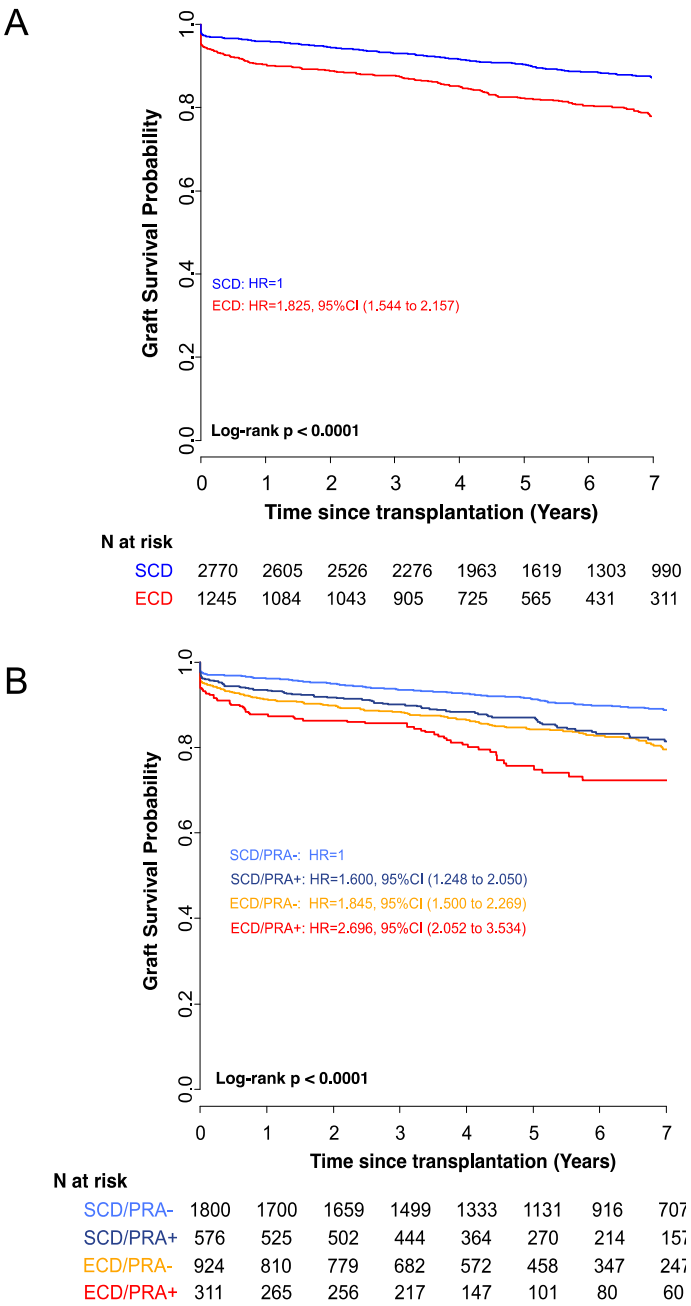


Le risque de perte du greffon en fonction du type de donneur (ECD/SCD) et du risque immunologique initial à J0 de la greffe (DSA+/DSA-) est représenté Figure 3. La survie post-greffe à 7 ans des greffons était respectivement de 93,3% dans le groupe SCD/DSA-, 75,7% dans le groupe SCD/DSA+, 86,3% dans le groupe ECD/DSA- et 47,2% dans le groupe ECD/DSA+ ($p < 0,0001$, Figure 3). L'absence de DSA circulant le jour de la greffe chez les receveurs d'un greffon issu d'un donneur ECD était donc associée à un gain de survie à 7 ans de 39,1% comparativement aux greffons issus de donneur SCD avec DSA circulant le jour de la transplantation ($p < 0,0001$).

Validation externe de la stratification de survie des greffons selon les deux caractéristiques DSA et ECD

La survie à long terme du greffon par la méthode de Kaplan-Meier selon la combinaison du type de donneur: SCD (Donneurs à critères standard), ECD (Donneurs à critères élargis) et de la présence de DSA circulant le jour de la transplantation est aussi confirmée (Figure 4).

Figure 4: Courbes de Kaplan-Meier pour la survie des greffons selon le type de donneur (A) et sa stratification sur le cytotoxic PRA (panel reactive antibody) (B).



IV.2.1.6 Conclusions

En conclusion, dans cette étude multicentrique (Hôpitaux Necker, Saint-Louis, Foch et Toulouse) comportant 2 763 patients très bien phénotypés, nous avons étudié les déterminants de la survie du greffon des patients transplantés à partir des donneurs à critères élargis et déterminé l'impact de la présence de DSA circulant le jour de la greffe. La politique de greffe, en réponse à la pénurie d'organes, a consisté ces dernières années en une augmentation de l'âge des donneurs. En 2013, 47,3% des greffons rénaux prélevés chez des donneurs décédés étaient issus de donneurs à critères élargis. Actuellement, avec la proportion croissante de patients immunisés dans le système anti-HLA sur la liste d'attente, certains greffons ECD sont attribués à des patients ayant des DSA le jour de la greffe sans que le risque cumulé n'ait été étudié sur un grand nombre de patients. Au terme de cette étude, nous démontrons que les receveurs d'un rein ECD présentent une survie acceptable du greffon à 7 ans (80,8% contre 90,9% pour les reins SCD). L'absence de DSA circulant le jour de la greffe chez les receveurs d'un greffon issu d'un donneur ECD était associée à une augmentation de 39,1% de la survie du greffon à 7 ans. Ainsi, l'utilisation des reins issus de donneurs ECD devrait être limitée aux patients non porteurs de DSA. En l'absence de cette prise de risque, l'écart entre la survie du greffon à 7 ans entre les reins ECD et SCD est minime (86,3% pour les reins ECD sans DSA et 93,3% dans le groupe SCD sans DSA à J0). La politique de greffe actuelle en France semble être une solution acceptable en réponse à la pénurie d'organes si celle-ci ne s'adresse pas à des patients immunisés avec des DSA le jour de la greffe.

IV.2.1.7 Manuscrit

OPEN ACCESS

Long term outcomes of transplantation using kidneys from expanded criteria donors: prospective, population based cohort study

Olivier Aubert,[1] Nassim Kamar,[2,3,4,5] Dewi Vernerey,[1] Denis Viglietti,[1,6] Frank Martinez,[7] Jean-Paul Duong-Van-Huyen,[1,8] Dominique Eladari,[1,9] Jean-Philippe Empana,[1] Marion Rabant,[8] Jerome Verine,[10] Lionel Rostaing,[2,3,4,5] Nicolas Congy,[4,11,12] Céline Guilbeau-Frugier,[4,13] Georges Mourad,[5,14] Valérie Garrigue,[5,14] Emmanuel Morelon,[5,15,16] Magali Giral,[5,16,17] Michèle Kessler,[5,16,18] Marc Ladrière,[5,16,18] Michel Delahousse,[1,19] Denis Glotz,[1,5,6] Christophe Legendre,[1,5,7,16] Xavier Jouven,[1,20] Carmen Lefaucheur,[1,5,6] Alexandre Loupy[1,5,7]

For numbered affiliations see end of article.

Correspondence to: A Loupy alexandreloupy@gmail.com

Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmj.h3557)

Cite this as: *BMJ* 2015;351:h3557
doi: 10.1136/bmj.h3557

Accepted: 17 June 2015

ABSTRACT

OBJECTIVES

To assess the long term outcomes of transplantation using expanded criteria donors (ECD; donors aged ≥60 years or aged 50-59 years with vascular comorbidities) and assess the main determinants of its prognosis.

DESIGN

Prospective, population based cohort study.

SETTING

Four French referral centres.

PARTICIPANTS

Consecutive patients who underwent kidney transplantation between January 2004 and January 2011, and were followed up to May 2014. A validation cohort included patients from another four referral centres in France who underwent kidney transplantation between January 2002 and December 2011.

MAIN OUTCOME MEASURES

Long term kidney allograft survival, based on systematic assessment of donor, recipient, and transplant clinical characteristics; preimplantation biopsy; and circulating levels of donor specific anti-HLA (human leucocyte antigen) antibody (DSA) at baseline.

RESULTS

The study included 6891 patients (2763 in the principal cohort, 4128 in the validation cohort). Of 2763 transplantations performed, 916 (33.2%) used ECD kidneys. Overall, patients receiving ECD transplants had lower allograft survival after seven years than patients receiving transplants from standard criteria donors (SCD; 80% *v* 88%, P<0.001). Patients receiving ECD transplants who presented with circulating DSA at the time of transplantation had worse allograft survival after seven years than patients receiving ECD kidneys without circulating DSA at transplantation (44% *v* 85%, P<0.001). After adjusting for donor, recipient, and transplant characteristics, as well as preimplantation biopsy findings and baseline immunological parameters, the main independent determinants of long term allograft loss were identified as allocation of ECDs (hazard ratio 1.84 (95% confidence interval 1.5 to 2.3); P<0.001), presence of circulating DSA on the day of transplantation (3.00 (2.3 to 3.9); P<0.001), and longer cold ischaemia time (>12 h; 1.53 (1.1 to 2.1); P=0.011). Recipients of ECD kidneys with circulating DSA showed a 5.6-fold increased risk of graft loss compared with all other transplant therapies (P<0.001). ECD allograft survival at seven years significantly improved with screening and transplantation in the absence of circulating DSA (P<0.001) and with shorter (<12 h) cold ischaemia time (P=0.030), respectively. This strategy achieved ECD graft survival comparable to that of patients receiving an SCD transplant overall, translating to a 544.6 allograft life years saved during the nine years of study inclusion time.

CONCLUSIONS

Circulating DSA and cold ischaemia time are the main independent determinants of outcome from ECD transplantation. Allocation policies to avoid DSA and reduction of cold ischaemia time to increase efficacy could promote wider implement of ECD transplantation in the context of organ shortage.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Organ transplants from expanded criteria donors (donors aged ≥60 years, or aged 50-59 years with vascular comorbidities) are increasingly becoming a main resource for treating end stage renal disease

Use of ECD kidney transplants has evolved unequally worldwide, and a high proportion of these kidneys are discarded

Organ shortages urgently need the optimisation and increased efficacy of ECD transplantation

WHAT THIS STUDY ADDS

In this large scale prospective study, the main independent determinants of improved ECD outcome were the absence of donor specific anti-human leucocyte antigen antibodies in transplant recipients, and reduction of cold ischaemia time

After correcting these two factors, ECD transplantation can yield satisfactory long term results, and achieve a graft survival comparable to that of transplants from standard criteria donor

Dedicated ECD specific allocation policies are mandatory to extend the life span of aged organs

Introduction

Global population ageing has modified the landscape of chronic diseases with an increased prevalence of older patients with end stage renal failure.[1] For such patients, kidney transplantation is the best approach, surpassing dialysis in terms of health related quality

and quantity of life as well as cost effectiveness.[2-4] Ageing of the general population reflects a constant and dramatic increase of the deceased donor age with accompanying burden of cardiovascular comorbidities, leading to the definition of an expanded criteria donor (ECD; that is, those aged ≥60 years, or aged 50-59 years with vascular comorbidities).[5] About 30% and 47% of deceased kidney donors are ECDs in Europe and France, respectively.[6 7] In the coming years, this dynamic will render ECD transplantation the main source for kidney transplants as well as other solid organ (heart, lung, and liver) transplants.[6 7]

But despite its potential, the use of ECD kidneys has evolved unevenly; only 16.6% of deceased donor transplants come from an ECD in the United States, and many harvested ECD organs are ultimately refused by transplant teams and discarded.[8] These discrepancies reflect uncertainties regarding the prognosis and determinants associated with ECD transplants. The discrepancies also contribute to a major public health issue (with a high proportion (10-20%) of viable organs being discarded worldwide) and to the deaths of patients with end stage renal disease waiting for a kidney transplant.[9-11]

Although ECD transplantation produces a benefit in recipient survival compared with dialysis, it is associated with an increased risk of graft loss compared with that associated with transplants from standard criteria donors (SCDs).[5 12 13] Currently, no comprehensive studies has addressed the primary independent determinants of ECD transplant outcomes by integrating deep donor and recipient phenotyping and large scale circulating anti-HLA (human leucocyte antigen) antibody characterisation. The need for such studies was identified by the International Banff Allograft Committee in 2013[14] and public healthcare agencies[15] to improve the use and allocation of ECD kidneys through a better understanding of long term outcomes. Thus, we conducted a large, prospective, population based study of extensively phenotyped kidney recipients to assess the distinct prognosis of ECD transplants and the factors that affect long term allograft survival.

Methods

Study population

We enrolled all 2763 consecutive patients who underwent kidney transplantation at four referral centres in France, including Necker Hospital in Paris (n=920), Saint-Louis Hospital in Paris (n=391), Foch Hospital in Suresnes (n=597), and Toulouse University Hospital (n=855), between 1 January 2004 and 1 January 2011. Patients were followed up to 1 May 2014. We also included an additional external validation cohort comprising 4128 patients who underwent kidney transplantation at centres in Lyon, Montpellier, Nancy, and Nantes (France) between 1 January 2002 and 31 December 2011 (web appendix, supplementary methods).

The transplantation allocation system was identical for all centres, and it followed the rules of the French national agency for organ procurement (Agence de la Biomédecine). All transplants were compatible based on ABO blood group. Cytotoxicity cross matching of negative immunoglobulin G T cell and B cell complements was required for all the recipients. Post-transplant immunosuppression therapy and treatment of allograft rejection episodes were standardised between centres (web appendix, supplementary methods).

The institutional review boards of Necker, Saint-Louis, Foch, and Toulouse Hospitals approved the study. The use of data from the external validation cohort was based on agreements between the centres that participate in the national database system (web appendix). All the recipients provided written informed consent.

Patient involvement

There was no patient involvement in this study. The selection of outcome measures was not informed by patients' priorities and experiences. No patients, service users, or lay people were involved in developing plans for participant recruitment and study conduct. The results of the study will be disseminated to participants through the physicians of each department. The participants are thanked in the acknowledgments.

Clinical data

We obtained clinical data on the donors and recipients in the development cohort (Necker, Saint-Louis, Foch, and Toulouse Hospitals) and the validation cohort (Lyon, Montpellier, Nancy, and Nantes) from two national registries: Données Informatiques Validées en Transplantation[16] and Agence de la Biomédecine.[17] Anonymised data from these registries were prospectively entered at a specific time point for each patient (on day 0, and six months and one year after transplantation), and were updated annually thereafter (supplementary methods in the web appendix shows the clinical and biological parameters assessed). Data from Necker, Saint-Louis, Foch, and Toulouse Hospitals were retrieved from the database on 1 May 2014. Data from the validation cohort were retrieved on 20 June 2014.

Definition of expanded criteria donor (ECD)

Expanded criteria donors included all deceased donors aged 60 years and older, and those aged over 50-59 years with at least two of the following three conditions: cerebrovascular cause of death, serum creatinine greater than 1.5 mg/dL (132.6 µmol/L), or a history of hypertension.

Detection and characterisation of donor specific anti-HLA antibodies

All the patients were tested for the presence of circulating donor specific anti-HLA antibodies (DSA) in banked serum samples (at the Jean Dausset Histocompatibility Laboratory) from Necker, Saint Louis, and Foch Hospitals and Toulouse Histocompatibility Laboratory at the time of transplantation (day 0 DSA). The presence of circulating DSA against HLA-A, HLA-B, HLA-Cw, HLA-DR, HLA-DQ, and HLA-DP was retrospectively determined using single antigen flow bead assays (One Lambda) on a Luminex platform. Beads

doi: 10.1136/bmj.h3557 | BMJ 2015;351:h3557 | the bmj

with a normalised mean fluorescence intensity higher than 500 arbitrary units were judged positive, as previously described.[18]

HLA typing of the transplant recipients was performed using the Innolipa HLA typing kit (Innogenetics). Donor HLA typing for HLA-Cw and HLA-DP was only performed if recipients had circulating anti-HLA-Cw or anti-HLA-DP (or both). Tissue typing was done using the microlymphocytotoxicity technique with One Lambda and tissue typing trays; traditional controls in molecular biology were included.

Kidney allograft phenotypes

In the principal cohort, kidney allograft biopsies were performed on the day of transplantation (preimplantation biopsies, n=1011) and at one year after transplantation (n=1743); biopsies were also taken from patients with acute allograft rejection during the first year after transplantation (n=438). Renal biopsies were fixed in a solution of alcohol, formalin, and acetic acid, and subsequently embedded in paraffin. The biopsy sections (4 µm thick) were stained with periodic acid-Schiff, Masson's trichrome, and haematoxylin and eosin. The graft biopsies were scored and graded by trained pathologists (JPD-V-H, MR, JV, and CG-F) according to the international Banff criteria for kidney allograft transplantation (web appendix).[14]

Statistical analysis

Continuous variables are described using means and standard deviations or medians and the interquartile range. We compared means and proportions between groups using Student's t test, analysis of variance (ANOVA; Mann-Whitney test for mean fluorescence intensity) or the χ^2 test (or Fisher's exact test if appropriate). The kidney survival analysis was performed from the time of transplantation to a maximum follow-up of seven years with kidney graft loss as the event of interest, defined as the patient's return to dialysis. For the 204 (7.4%) patients who died with a functioning graft, graft survival was censored at the time of death.[19] Kidney allograft survival according to ECD and DSA status was plotted using Kaplan-Meier curves and compared using the log rank test.

Cox proportional hazards models were applied to quantify the hazard ratios and the 95% confidence intervals for kidney graft loss. We compared the graft life years between recipients of ECD allografts with and without baseline levels of DSA, and estimated the graft life years saved. To this end, we compared the area under the Kaplan-Meier curves between the two groups starting from a common time point, namely, the time of transplantation.[20]

The associations of donor, recipient, and transplant parameters as well as immunological factors with graft loss were first assessed in univariate regression analyses. The factors identified in these analyses were then included in a final multivariate model using stepwise backward selection. In the sensitivity analysis, a competing risks approach was applied to consider the potential competition of patient death with kidney allograft failure. We estimated cumulative incidence functions from competing risks data and compared the subdistribution for each cause across groups. We then assessed the effect of predictive factors (that is, ECD and DSA status) on the subdistribution of graft loss in a competing risks setting with death, by fitting the proportional subdistribution hazards regression model described in the Fine and Gray method.[21] Propensity scores were used to test whether older recipient age affected allograft outcome and the associations identified in the primary analyses (web appendix).[22]

We confirmed the internal validity of the final model using a bootstrap procedure. A thousand datasets were generated from the resampling of the original dataset to estimate the biased corrected 95% confidence intervals and the accelerated bootstrap hazard ratios.[23]

We used an additional external validation cohort comprising 4128 kidney transplant recipients. We first determined kidney allograft survival according to ECD and DSA status. We then tested in the validation cohort, the set of risk factors identified in the final multivariate Cox model of the development cohort. We used SAS 9.2 (SAS) and R (version 2.10.1, R Foundation for Statistical Computing) for the descriptive and survival analyses. All statistical tests were two sided, and $P<0.05$ was considered significant.

Results

Baseline characteristics of recipients and donors

The main analysis included 2763 recipient patients. Median donor age was 52 years (interquartile range 40-61). A total of 1509 (54.6%) donors were male, and 2370 (85.8%) kidneys were from deceased donors, of whom 1323 (55.8%) had died of cerebrovascular causes. All the deceased donors were brain dead. A total of 661 (24.4%) donors had hypertension, and 142 (5.5%) donors presented with diabetes mellitus. A total of 916 (33.2%) kidneys fulfilled the criteria for ECD, whereas 1847 (66.8%) were classified as SCD. Web figure 1 details the donor age distribution for the entire cohort. Cold ischaemia time is the period starting from retrieval and cold perfusion of an organ until reperfusion with warm blood in the recipient. The median cold ischaemia time was 16.5 hours (interquartile range 12-21). Of the 2763 recipients, 335 (12.1%) had circulating DSA at the time of transplantation. The highest ranked DSA was HLA class II in 189 (56.4%) patients and HLA class I in 146 patients (43.6%). Mean fluorescence intensity of the immunodominant DSA had a median value of 2928 (interquartile range 1390-6835). Table 1 presents the baseline and immunological characteristics of the overall cohort according to ECD status.

Four distinct populations were identified on the basis of donor characteristics (SCD or ECD) and the presence or absence of circulating anti-HLA antibodies at the time of transplantation. These groups included: patients receiving SCD transplants without DSA (SCD/DSA–, n=1622); patients receiving SCD transplants with DSA (SCD/DSA+, n=225); patients receiving ECD transplants without DSA (ECD/DSA–, n=806); patients receiving ECD transplants with DSA

Table 1 | Baseline characteristics according to donor status

	All patients (n=2763)	Standard criteria donors (n=1847)	Expanded criteria donors (n=916)	P*
Recipient characteristics				
Age (years, mean (SD))	49.23 (13.51)	44.26 (12.31)	59.23 (10.21)	<0.001
Male sex (no (%))	1643 (59.5)	1105 (59.8)	538 (58.7)	0.582
Causes of end stage renal disease (no (%))				
Glomerulonephritis	767 (27.8)	561 (30.4)	206 (22.5)	<0.001
Diabetes	264 (9.5)	171 (9.2)	93 (10.1)	
Hypertension	212 (7.7)	129 (7.0)	83 (9.1)	
Other	1520 (55.0)	986 (53.4)	534 (58.3)	
Time since onset of dialysis (years, mean (SD))	3.90 (4.41)	3.85 (4.67)	4.02 (3.84)	0.3451
Donor characteristics				
Age (years, mean (SD))	50.31 (15.98)	42.34 (12.51)	66.37 (8.29)	<0.001
Donor age (years, no (%))				
0-50	1222 (44.2)	1222 (66.2)	0	<0.001
50-60	733 (26.6)	565 (30.6)	168 (18.3)	
≥60	808 (29.2)	60 (3.2)	748 (81.7)	
Male sex (no (%))	1509 (54.6)	1056 (57.2)	453 (49.5)	<0.001
Hypertension (no (%))†	661 (24.4)	167 (9.2)	494 (55.0)	<0.001
Diabetes mellitus (no (%))‡	142 (5.5)	40 (2.3)	102 (11.8)	<0.001
Creatinine >1.5 mg/dL (no (%))§	270 (9.9)	169 (9.3)	101 (11.0)	0.150
Donor type (no (%))				
Deceased donor	2370 (85.8)	1454 (78.7)	916 (100.0)	<0.001
Death from cerebrovascular disease¶	1323 (55.8)	627 (43.1)	696 (76.0)	<0.001
Transplant baseline characteristics				
Graft rank >1 (no (%))	413 (15.0)	288 (15.6)	125 (13.7)	0.177
Cold ischaemia time (no (%))**				
<12 h	674 (24.5)	594 (32.3)	80 (8.8)	<0.001
12-24 h	1560 (56.7)	910 (49.5)	650 (71.0)	
≥24 h	519 (18.8)	334 (18.2)	185 (20.2)	
Delayed graft function (no (%))††, ‡‡	707 (27.2)	410 (23.5)	297 (34.7)	<0.001
HLA A/B/DR mismatch (no, mean (SD))§§	3.12 (1.43)	3.07 (1.47)	3.24 (1.33)	0.005
Donor specific anti-HLA antibodies on day 0 (no (%))	335 (12.1)	225 (12.2)	110 (12.0)	0.896
HLA class of donor specific anti-HLA antibodies (no)				
HLA class I	191	123	68	0.214
HLA class II	245	166	79	0.704
HLA class I and II	114	76	38	0.889
Donor specific anti-HLA antibodies (mean fluorescence intensity (median (IQR))	2928 (1390-6835)	2755 (1360-6480)	3005 (1450-8369)	0.099
Follow-up (years, median (95% CI))	5.54 (5.42 to 5.68)	5.72 (5.54 to 5.86)	5.24 (5.02 to 5.46)	<0.001
Acute rejection in first year				
Antibody mediated rejection (no (%))	205 (7.4)	140 (7.6)	65 (7.1)	0.648
T cell mediated rejection (no (%))	233 (8.4)	156 (8.5)	77 (8.4)	0.972
Death events (no (%))	263 (9.5)	128 (6.9)	135 (14.7)	<0.001
Graft loss (no (%))	333 (12.1)	187 (10.1)	146 (15.4)	<0.001

Data are no (%) of patients or events, mean (standard deviation), or median (interquartile range). SD=standard deviation; IQR=interquartile range; HLA=human leucocyte antigen.
*χ^2 tests used to compare proportions, and unpaired tests used to compare continuous variables.
†Total group numbers were: all patients (n=2709), standard criteria donors (n=1811), and expanded criteria donors (n=898).
‡Total group numbers were: all patients (n=2604), standard criteria donors (n=1741), and expanded criteria donors (n=863).
§Total group numbers were: all patients (n=2736), standard criteria donors (n=1820), and expanded criteria donors (n=916).
¶Total group numbers were: all patients (n=2370), standard criteria donors (n=1454), and expanded criteria donors (n=916).
**Total group numbers were: all patients (n=2753), standard criteria donors (n=1838), and expanded criteria donors (n=915).
††Delayed graft function defined as use of dialysis in the first postoperative week.
‡‡Total group numbers were: all patients (n=2603), standard criteria donors (n=1746), and expanded criteria donors (n=857).
§§Total group numbers were: all patients (n=2700), standard criteria donors (n=1844), and expanded criteria donors (n=856).

(ECD/DSA+, n=110, 12% of patients receiving ECD transplants). Table 2 details the baseline clinical and immunological characteristics of these four patient groups. Web figure 2 shows the histological characteristics of donor biopsies before implantation in the four patient groups.

Long term kidney allograft survival, injury phenotypes, and function

Median follow-up after transplantation was 5.54 years (interquartile range 4.01-7.42). Median follow-up times were 5.24 years (3.64-6.97) for ECD recipients and 5.72 years (4.16-7.82) for SCD recipients (P<0.001).

Figure 1A depicts kidney allograft survival according to ECD or SCD status. Overall, patients receiving ECD transplants had a higher rate of graft loss than those receiving SCD transplants (hazard ratio 1.87 (95% confidence interval 1.50 to 2.32); log rank test P<0.001). After dividing patients into groups based on donor characteristics (SCD or ECD) and the presence of circulating DSA on day 0, ECD/DSA+ recipients showed worse

doi: 10.1136/bmj.h3557 | BMJ 2015;351:h3557 | the bmj

Table 2 | Baseline characteristics according to donor status and presence of DSA on the day of transplantation

	SCD/DSA− (n=1622)	SCD/DSA+ (n=225)	ECD/DSA− (n=806)	ECD/DSA+ (n=110)	P*
Recipient characteristics					
Age (years, mean (SD))	44.20 (12.20)	44.76 (11.65)	59.60 (9.98)	56.53 (11.41)	<0.001
Male sex (no (%))	1001 (61.7)	104 (46.2)	489 (60.7)	49 (44.6)	<0.001
Time since onset of dialysis (years, mean (SD))	3.56 (4.28)	5.93 (6.45)	3.74 (3.58)	6.01 (4.91)	<0.001
Donor characteristics					
Age (years, mean (SD))	42.25 (12.50)	43.01 (12.61)	66.43 (8.24)	65.95 (8.67)	<0.001
Donor age (years, no (%))					
0-50	1073 (66.1)	149 (66.2)	0	0	<0.001
50-60	496 (30.6)	69 (30.7)	147 (18.2)	21 (19.1)	
≥60	53 (3.3)	7 (3.1)	659 (81.8)	89 (80.9)	
Male sex (no (%))	934 (57.6)	122 (54.2)	393 (48.8)	60 (54.6)	0.001
Hypertension (no (%))†	142 (8.9)	25 (11.3)	427 (54.0)	67 (62.6)	<0.001
Diabetes mellitus (no (%))‡	34 (2.2)	6 (2.9)	94 (12.3)	8 (8.1)	<0.001
Creatinine >1.5 mg/dL (no (%))§	155 (9.7)	14 (6.3)	94 (11.7)	7 (6.4)	0.054
Donor type (no (%))					
Deceased donor	1273 (78.5)	181 (80.4)	806 (100)	110 (100)	<0.001
Death from cerebrovascular disease¶	552 (43.4)	75 (41.4)	612 (75.9)	84 (76.4)	<0.001
Transplant baseline characteristics					
Graft rank >1 (no (%))	193 (11.9)	95 (42.2)	85 (10.6)	40 (36.4)	<0.001
Cold ischaemia time (no (%))**					
<12 h	544 (33.7)	50 (22.5)	70 (8.7)	10 (9.1)	<0.001
12-24 h	803 (49.7)	107 (48.2)	575 (71.4)	75 (68.2)	
≥24 h	269 (16.6)	65 (29.3)	160 (19.9)	25 (22.7)	
Delayed graft function (no (%))††	343 (22.4)	67 (30.9)	256 (33.9)	41 (40.2)	<0.001
HLA A/B/DR mismatch (no, mean (SD))‡‡	3.07 (1.47)	3.09 (1.48)	3.21 (1.32)	3.44 (1.35)	0.015
HLA class of donor specific anti-HLA antibodies (no) on day 0					
HLA class I	—	123	—	68	0.214§§
HLA class II	—	166	—	79	0.704§§
HLA class I and II	—	76	—	38	0.889§§
Donor specific anti-HLA antibodies (mean fluorescence intensity (median (IQR))	—	2755 (1360-6480)	—	3005 (1450-8369)	<0.001
Follow-up (years, median (95% CI))	5.86 (5.70 to 6.12)	4.66 (4.34 to 4.92)	5.31 (5.08 to 5.54)	4.73 (4.24 to 5.19)	<0.001
Acute rejection in the first year					
Antibody mediated rejection (no (%))	78 (4.8)	62 (27.6)	31 (3.9)	34 (30.9)	<0.001
T cell mediated rejection (no (%))	142 (8.8)	14 (6.2)	66 (8.2)	11 (10.0)	0.561
Death events (no (%))	108 (6.7)	20 (8.9)	118 (14.6)	17 (15.5)	<0.001
Graft loss (no (%))	146 (9.0)	41 (18.2)	103 (12.8)	43 (39.1)	<0.001

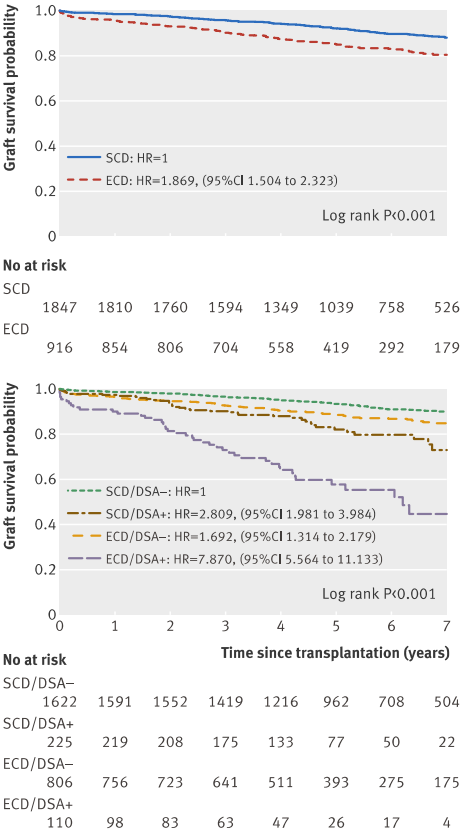
Data are no (%) of patients or events, mean (standard deviation), or median (interquartile range). SD=standard deviation; IQR=interquartile range; HLA=human leucocyte antigen; SCD=standard criteria donors; ECD=expanded criteria donors; DSA=donor specific anti-HLA antigen.
*χ^2 tests or Fisher's exact tests used to compare proportions, and ANOVA or Kruskal-Wallis test used to compare continuous variables between the four groups.
†Total group numbers were: SCD/DSA− (n=1589), SCD/DSA+ (n=222), ECD/DSA− (n=791), and ECD/DSA+ (n=107).
‡Total group numbers were: SCD/DSA− (n=1533), SCD/DSA+ (n=208), ECD/DSA− (n=764), and ECD/DSA+ (n=99).
§Total group numbers were: SCD/DSA− (n=1598), SCD/DSA+ (n=222), ECD/DSA− (n=806), and ECD/DSA+ (n=110).
¶Total group numbers were: SCD/DSA− (n=1273), SCD/DSA+ (n=181), ECD/DSA− (n=806), and ECD/DSA+ (n=110).
**Total group numbers were: SCD/DSA− (n=1616), SCD/DSA+ (n=222), ECD/DSA− (n=805), and ECD/DSA+ (n=110).
†† Delayed graft function defined as use of dialysis in the first postoperative week. Total group numbers were: SCD/DSA− (n=1529), SCD/DSA+ (n=217), ECD/DSA− (n=755), and ECD/DSA+ (n=102).
‡‡Total group numbers were: SCD/DSA− (n=1619), SCD/DSA+ (n=225), ECD/DSA− (n=752), and ECD/DSA+ (n=104).
§§P values represent the comparison between the SCD/DSA+ and ECD/DSA+ groups.

graft survival (44%, seven years post-transplant) than ECD/DSA− (85%), SCD/DSA+ (73%), and SCD/DSA− (90%) recipients (log rank test, P<0.001). Antibody mediated rejection was the main cause of graft loss in the ECD/DSA+ group (58.1% *v* 14.6% in the ECD/DSA- group; P<0.001).

Patients receiving ECD transplants without circulating DSA at the time of transplantation showed a 41% improvement in graft survival at seven years compared with those with circulating DSA at the time of transplantation (P<0.001, fig 1B).

Patients receiving ECD transplants with circulating DSA had a 4.4-fold increased risk of graft loss compared with those receiving ECD transplants without circulating DSA (P<0.001). ECD recipients with circulating DSA also had a 5.6-fold increased risk of graft loss compared with all other transplant therapies (P<0.001; fig 1B).

Mean allograft life years for ECD/DSA+ recipients was 4.58 (standard deviation 0.22) compared with 9.53 (0.12) for ECD/DSA− recipients (mean gain survival of 4.95 years). Overall, we calculated that exclusively allocating ECD kidneys to patients without circulating DSA would have translated to a 544.6 allograft life years saved during the nine years of study inclusion time. Competing risks regression models confirmed that the differences in graft survival between the four groups were not affected by competition with patient death (web fig 3).



Kaplan-Meier curves of kidney allograft survival by (A) donor type and (B) presence of DSA on the day of transplantation. Panel A shows the classic approach to determine kidney allograft survival based on donor status (that is, SCD or ECD). Panel B integrates donor status with the presence or absence of DSA on the day of transplantation. SCD=standard criteria donor; ECD=expanded criteria donor; DSA=donor specific anti-HLA antibodies; SCD/DSA–=patients receiving SCD transplants without DSA; SCD/DSA+=patients receiving SCD transplants with DSA; ECD/DSA–=patients receiving ECD transplants without DSA; ECD/DSA+=patients receiving ECD transplants with DSA; HR=hazard ratio

An allograft evaluation at one year revealed that the patients receiving ECD transplants with circulating DSA had decreased kidney function compared with the other groups (P<0.001, web fig 4). Allograft biopsies at one year after transplantation showed that ECD/DSA+ recipients had increased microcirculation inflammation (P<0.001), arteriosclerosis (P=0.014), and transplant glomerulopathy lesions (P<0.001) compared with patients receiving ECD transplants without circulating DSA. ECD/DSA+ recipients also showed increased chronic allograft injury lesions (arteriosclerosis, arterial hyalinosis, and atrophy scarring) compared with the two groups of patients receiving SCD transplants (P<0.002 for all comparisons, web fig 5).

Determinants of kidney allograft loss

Table 3 shows the associations of recipient and donor characteristics, transplant characteristics, and immunological parameters with graft loss. Table 4 shows the identified baseline independent predictors of graft loss. These predictors included: ECD status (hazard ratio 1.84 (95% confidence interval 1.47 to 2.31); P<0.001), presence of circulating DSA on day 0 (3.00 (2.27 to 3.94); P<0.001), cold ischaemia time between 12 and 24 hours (1.46 (1.04 to 2.04); P=0.028); cold ischaemia time longer than 24 hours (1.73 (1.19 to 2.52); P=0.004); graft rank more than 1 (1.54 (1.17 to 2.04); P=0.002), and number of HLA A/B/DR mismatches (1.10 (1.01 to 1.18); P=0.022). The set of risk factors identified in the primary analysis remained unchanged when we included continuous cold ischaemia time in the final multivariate Cox model (web table 1).

When we performed the Cox analysis in the ECD group, the major determinants independently associated with graft failure were cold ischaemia between 12 and 24 hours (hazard ratio 2.49 (1.02 to 6.13); P=0.046) and longer than 24 hours (3.77 (1.49 to 9.55); P=0.005) and the presence of circulating DSA at day 0 (4.59 (3.18 to 6.61); P<0.001). Among patients receiving ECD transplants, those with a high level of DSA mean fluorescence intensity (≥3000) were at a higher risk of graft failure than those with a low level of DSA mean fluorescence intensity (<3000) and those without circulating DSA (P<0.001, web fig 6). We saw no difference between HLA class I and HLA class II groups regarding allograft survival (web fig 7). For the 1011 patients with an available biopsy from the day of transplantation, baseline histology parameters were not independently associated with the risk of long term kidney allograft loss (web table 2).

Sensitivity analysis

The robustness of our study was assessed using sensitivity analyses. The independent associations of DSA and cold ischaemia time with graft outcome were consistent after excluding living donors (n=393, 14.2%) from the final multivariate model (web table 3). DSA and cold ischaemia time remained independently associated with graft loss when allograft quality for age and cardiovascular comorbidities was assessed using the kidney donor risk index (KDRI) score[24] instead of ECD status (web fig 8; KDRI adjusted DSA (hazard ratio 3.44 (95% confidence interval 2.60 to 4.55), P<0.001); KDRI adjusted cold ischaemia time (1.81 (1.12 to 2.92), P=0.015).

The use of non-a priori donor parameters (age, deceased donor, sex, comorbidities, and baseline renal function) instead of ECD criteria did not change the significant independent predictors of graft loss identified in the principal analyses (web table 4). When the final model was stratified by transplant centre, the set of risk factors identified in the principal analyses did not change (web table 5). Finally, in a matched sample with similar SCD and ECD recipient ages using the propensity score (web appendix, supplementary methods), the survival analysis produced similar results to those obtained in the primary analyses (web fig 9).

doi: 10.1136/bmj.h3557 | *BMJ* 2015;351:h3557 | the bmj

Table 3 | Factors associated with kidney allograft loss in univariate analyses

	No of patients/ events	Hazard ratio (95% CI)	P
Baseline recipient characteristics			
Age (per 1 year increment)	2763/333	1.000 (0.992 to 1.008)	0.973
Sex			
Female	1120/131	1	0.648
Male	1643/202	1.053 (0.845 to 1.311)	
Baseline donor characteristics			
Age (per 1 year increment)	2763/333	1.014 (1.007 to 1.021)	<0.001
Sex			
Female	1254/163	1	0.164
Male	1509/170	0.858 (0.692 to 1.064)	
Donor type			
Living related	393/25	1	0.001
Death from cardiovascular disease	1323/184	2.136 (1.406 to 3.245)	
Other cause of death	1047/124	1.716 (1.116 to 2.638)	
Deceased donor			
No	393/25	1	0.001
Yes	2370/308	1.945 (1.294 to 2.925)	
Hypertension			
No	2048/225	1	<0.001
Yes	661/102	1.647 (1.302 to 2.082)	
Diabetes mellitus			
No	2462/292	1	0.050
Yes	142/22	1.543 (1.000 to 2.380)	
Creatinine			
<1.5 mg/dL	2466/291	1	0.900
≥1.5 mg/dL	270/40	0.964 (0.547 to 1.700)	
Expanded criteria donor			
No	1847/187	1	<0.001
Yes	916/146	1.869 (1.504 to 2.323)	
Baseline transplant characteristics			
Graft rank			
1	2350/249	1	<0.001
>1	413/84	2.174 (1.697 to 2.785)	
Cold ischaemia time			
<12 h	674/44	1	
12-24 h	1560/203	1.877 (1.355 to 2.601)	<0.001
≥24 h	519/85	2.392 (1.662 to 3.443)	
Baseline immunological characteristics			
No of HLA A/B/DR mismatches	2700/326	1.080 (1.001 to 1.116)	0.047
Anti-HLA DSA on day 0			
No	2428/249	1	<0.001
Yes	335/84	3.440 (2.676 to 4.421)	

DSA=donor specific anti-HLA antibodies; HLA=human leucocyte antigen.

Internal and external validation

A bootstrapping procedure with 1000 samples from the original dataset confirmed the internal validity and robustness of the final model (biased corrected 95% confidence intervals and accelerated bootstrap hazard ratios, table 4). Baseline characteristics of the external validation cohort (4128 patients) were compared with those of the principal cohort in web table 6. In this cohort, the HLA sensitisation level at the time of transplantation was estimated using cytotoxic panel reactive antibodies (PRAs) or calculated PRAs without identifying specific anti-HLA antibodies. The Kaplan-Meier estimate of graft survival confirmed that patients receiving ECD transplants with circulating DSA on the day of transplantation had a higher risk of graft loss at seven years than those without circulating DSA and the two groups of patients receiving SCD transplants (web fig 10). The multivariate Cox model of the external cohort confirmed the independent predictors of allograft failure identified in the principal analyses (web table 7).

Discussion

Principal findings

In a large, prospective, population based study of extensively phenotyped recipients of kidney transplants, we assessed the long term allograft survival and determinants of ECD prognosis using an integrative epidemiological approach that combined clinical donor and recipient data, histopathology, and immunological characteristics. We identified a group of people with poor prognosis—that is, patients receiving ECD transplants with circulating DSA. This group had a 4.4-fold increased risk of graft loss compared with patients receiving ECD transplants without circulating DSA, and had a 5.6-fold increased risk of graft loss compared with all other transplant therapies. When patients receiving ECD transplants were extensively screened and transplanted without DSA, graft survival increased by 41%, reaching a seven year rate of allograft survival of 85% (versus 90% in patients receiving SCD transplants without circulating DSA).

Implication of results

In the past decade, the constant rise in donor age in the organ pool has generated new challenges for optimising organ use, especially when considering the characteristics of an older organ in terms of response to injury after transplantation. This problem has raised concerns beyond kidney transplantation—extending to lung, liver, and heart transplants—and has become a top priority for allocation agencies.

ECD kidney transplantation has evolved unequally between countries. ECD programmes have been largely promoted and implemented in Europe and France and represent about half of deceased donor kidney transplantations.[6] However, in the USA, a high proportion of ECD organs are discarded, and ECD organs represent only 16.6% of deceased donor allografts.[8] Data from US scientific registries indicate that over a six year period, 5139 (41%) of 12536 ECD kidneys were not transplanted.

ECD therapy has demonstrated benefits over dialysis in terms of patient survival[13] but has diminished long term allograft survival.[5 25] Therefore, the improvement of ECD outcomes and specific allocation is critical for the transplant community and allocation agencies as well as for general practitioners dealing with population ageing and the burden of chronic renal disease.

Currently, an important gap exists regarding the independent determinants of ECD outcome and methods for improving ECD allograft survival in the context of allograft assessments, which include contemporary tools for evaluating circulating anti-HLA antibodies. Although circulating DSA are known to impair allograft outcomes, their specific effect, the amplitude of their effect, and their independency from other relevant

Table 4 | Factors associated with kidney allograft loss in the multivariate analysis*

	No of patient/events	Hazard ratio (95% CI)	P	Internal validation hazard ratio 95% CI, BCA
Expanded criteria donor				
No	1835/187	1	<0.001	—
Yes	855/138	1.842 (1.467 to 2.311)		(1.463 to 2.275)
Cold ischaemia time				
<12 h	670/44	1	0.017	—
12-24 h	1514/198	1.457 (1.042 to 2.039)		(1.042 to 2.093)
≥24 h	506/83	1.727 (1.185 to 2.517)		(1.195 to 2.506)
Graft rank				
1	2278/241	1	0.002	—
>1	412/84	1.544 (1.168 to 2.042)		(1.129 to 2.046)
No of HLA A/B/DR mismatches	2690/325	1.095 (1.013 to 1.184)	0.022	(1.013 to 1.182)
Anti-HLA DSA on day 0				
No	2364/241	1	<0.001	—
Yes	326/84	2.988 (2.265 to 3.941)		(2.198 to 3.940)

*The final multivariate Cox model was obtained by entering risk factors from the univariate model that achieved P≤0.10 as the thresholds in a single multivariate proportional hazards model. The final multivariate model was adjusted for the following parameters: deceased donor, donor diabetes, expanded criteria donor, graft rank, cold ischaemia time, number of HLA A/B/DR mismatches, and circulating DSA. The test for potential interactions between graft rank, ECD status, cold ischaemia time, HLA mismatch, and circulating DSA generated P>0.10 for all the interactions.
DSA=donor specific anti-HLA antibodies; BCA=bias corrected and accelerated bootstrap; HLA=human leucocyte antigen.

predictors have not been specifically analysed in the recipients of ECD transplants.[18 26-29] The structure of the existing registries—the United Network of Organ Sharing in the USA,[30] Eurotransplant,[31 32] the Organ Procurement Agency in Spain,[33] and the French National Agency for Organ Procurement [17]—does not specifically integrate the whole spectrum of important determinants that might affect ECD prognosis.

In the present study, we used a prospective, extensively phenotyped cohort with a detailed and integrative strategy of allograft parameter analysis, and demonstrated that ECD kidney recipients with circulating DSA show a dramatic acceleration in antibody mediated injury and a poor outcome. Another important independent and strong contributor of ECD survival was longer cold ischaemia time (>12 h), which was associated with a 2.6-fold decrease in seven year allograft survival. Increased cold ischaemia time predisposes to delayed graft function, but its independent role in long term allograft outcome is still debated.[34 35] Our data suggest that ECD kidneys might be vulnerable to the detrimental effects of cold ischaemia time on allograft survival.

Finally, we demonstrated that preimplantation biopsy assessments did not have independent and additional predictive ability for long term ECD allograft outcome when considering cold ischaemia time and immunological parameters (that is, circulating DSA). Given this result and the fact that preimplantation biopsies increase cold ischaemia time, the current practice of discarding ECD kidneys based on preimplantation biopsy results might not be the best approach for decision making in the context of ECD transplants, as recently suggested in small preliminary studies.[9 36]

Our results support the implementation of active ECD specific allocation policies for avoiding DSA, decreasing cold ischaemia time, and performing adequate recipient matching. In the present study, exclusive allocation of ECD kidneys to patients without circulating DSA would have translated to a 544.6 allograft life years saved during the nine years of study inclusion time.

AUTHOR AFFILIATIONS

[1]Paris Translational Research Centre for Organ Transplantation, INSERM, UMR-S970, Paris Descartes University, 75015 Paris, France
[2]Department of Nephrology and Organ Transplantation, Rangueil, Toulouse, France
[3]INSERM U1043, Purpan, Toulouse
[4]Paul Sabatier University, Toulouse
[5]DIVAT (Données Informatiques VAlidées en Transplantation Network), France
[6]Department of Nephrology and Kidney Transplantation, Saint-Louis Hospital, Assistance Publique Hôpitaux de Paris
[7]Department of Kidney Transplantation, Necker Hospital, Assistance Publique Hôpitaux de Paris, Sorbonne Paris
[8]Department of Pathology, Necker Hospital, Paris
[9]Department of Renal Physiology, Georges Pompidou European Hospital, Paris
[10]Department of Pathology, Saint-Louis Hospital, Paris
[11]Molecular Immunogenetics Laboratory, EA 3034, Faculty of Medicine Purpan, IFR150 (INSERM), Toulouse, France
[12]Department of Immunology, Rangueil Hospital, Toulouse
[13]Department of Pathology, Rangueil, Toulouse
[14]Department of Nephrology and Kidney Transplantation, Montpellier, France
[15]Department of Nephrology and Kidney Transplantation, Groupement Hospitalier, Hôpital Edouard Herriot, Lyon, France
[16]Centaure Network, France
[17]Department of Nephrology and Kidney Transplantation, Nantes, France
[18]Department of Nephrology and Kidney Transplantation, Nancy, France
[19]Department of Nephrology and Kidney Transplantation, Foch Hospital, Suresnes, Paris
[20]Department of Cardiology, Georges Pompidou European Hospital, Paris

We thank all patients included in the present study.

Contributors: OA, NK, DV, XJ, CaL, and AL have shared first authorship. OA, NK, CaL, DG, FM, ChL, XJ, and AL designed the study. OA, AL, CaL, DV, JPD-V-H, NK, LR, NC, MR, JV, FM, MD, GM, EM, MG, and CG-F collected the data. Data analysis was performed by OA, DW, AL, CaL,

NK, J-PE, and XJ. OA, NK, DW, DE, ChL, XJ, CaL, and AL wrote the final report. All authors contributed to both protocol and final report. AL is guarantor. All the authors had full access to all of the data and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding: There was no specific funding for this study.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work, no financial relationships with any organisations that might have an interest in the submitted work in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: The research protocol was reviewed and approved by the local institutional review boards.

Patient consent: Codes were used to ensure the strict donor and recipient anonymity and blind assay. Informed consent was obtained from participants at time of transplantation.

Data sharing: No additional data available.

The guarantor affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

1 Réseau Epidémiologie et Information en Néphrologie. Annual report. 2013. www.socnephrologie.org/PDF/enephro/registres/rapport_2013/rapport_rein2013.pdf.
2 Evans RW, Manninen DL, Garrison LP Jr, et al. The quality of life of patients with end-stage renal disease. *N Engl J Med* 1985;312:553-9.
3 Port FK, Wolfe RA, Mauger EA, et al. Comparison of survival probabilities for dialysis patients vs cadaveric renal transplant recipients. *JAMA* 1993;270:1339-43.
4 Rana A, Gruessner A, Agopian VG, et al. Survival benefit of solid-organ transplant in the United States. *JAMA Surg* 2015;150:252-9.
5 Port FK, Bragg-Gresham JL, Metzger RA, et al. Donor characteristics associated with reduced graft survival: an approach to expanding the pool of kidney donors. *Transplantation* 2002;74:1281-6.
6 Eurotransplant. Annual report. 2013. https://www.eurotransplant.org/cms/mediaobject.php?file=AR20135.pdf.
7 L'Agence de la Biomédecine. Annual report. 2013. www.agence-biomedecine.fr/rapport-annuel-2013.
8 Organ Procurement and Transplantation Network/Scientific Registry of Transplant Recipients. Annual report. 2012. srtr.transplant.hrsa.gov/annual_reports/2012/pdf/2012_SRTR_ADR.pdf.
9 Sung RS, Christensen LL, Leichtman AB, et al. Determinants of discard of expanded criteria donor kidneys: impact of biopsy and machine perfusion. *Am J Transplant* 2008;8:783-92.
10 Cecka JM, Gritsch HA. Why are nearly half of expanded criteria donor (ECD) kidneys not transplanted? *Am J Transplant* 2008;8:735-6.
11 Tanriover B, Mohan S, Cohen DJ, et al. Kidneys at higher risk of discard: expanding the role of dual kidney transplantation. *Am J Transplant* 2014;14:404-15.
12 Metzger RA, Delmonico FL, Feng S, et al. Expanded criteria donors for kidney transplantation. *Am J Transplant* 2003;3 Suppl 4:114-25.
13 Merion RM, Ashby VB, Wolfe RA, et al. Deceased-donor characteristics and the survival benefit of kidney transplantation. *JAMA* 2005;294:2726-33.
14 Haas M, Sis B, Racusen LC, et al. Banff 2013 meeting report: inclusion of c4d-negative antibody-mediated rejection and antibody-associated arterial lesions. *Am J Transplant* 2014;14:272-83.
15 Sack K. In discarding of kidneys, system reveals its flaws. *New York Times* 2012 Sept 19.
16 Données Informatiques Validées en Transplantation (DIVAT) software. Homepage. www.divat.fr/.
17 Agence de la Biomédecine. Homepage. www.sipg.sante.fr/portail/.
18 Lefaucheur C, Loupy A, Hill GS, et al. Preexisting donor-specific HLA antibodies predict outcome in kidney transplantation. *J Am Soc Nephrol* 2010;21:1398-406.
19 Lamb KE, Lodhi S, Meier-Kriesche HU. Long-term renal allograft survival in the United States: a critical reappraisal. *Am J Transplant* 2011;11:450-62.
20 Lee ET. Statistical methods for survival data analysis/ 2nd edition, John Wiley & Sons, 1992.
21 Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing Risk. *J Am Stat Assoc* 1999;94:496-509.
22 D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998;17:2265-81.
23 Efron B. Bootstrap methods: another look at the jackknife. *Ann Stats* 1979;7:1-26.
24 Rao PS, Schaubel DE, Guidinger MK, et al. A comprehensive risk quantification score for deceased donor kidneys: the kidney donor risk index. *Transplantation* 2009;88:231-6.
25 Ojo AO, Hanson JA, Meier-Kriesche H, et al. Survival in recipients of marginal cadaveric donor kidneys compared with other recipients and wait-listed transplant candidates. *J Am Soc Nephrol* 2001;12:589-97.
26 Nankivell BJ, Alexander SI. Rejection of the kidney allograft. *New Engl J Med* 2010;363:1451-62.
27 Sellares J, de Freitas DG, Mengel M, et al. Understanding the causes of kidney transplant failure: the dominant role of antibody-mediated rejection and nonadherence. *Am J Transplant* 2012;12:388-99.
28 Nankivell BJ, Kuypers DR. Diagnosis and prevention of chronic kidney allograft loss. *Lancet* 2011;378:1428-37.
29 Gaston RS, Cecka JM, Kasiske BL, et al. Evidence for antibody-mediated injury as a major determinant of late kidney allograft failure. *Transplantation* 2010;90:68-74.
30 Tuttle-Newhall JE, Krishnan SM, Levy MF, et al. Organ donation and utilization in the United States: 1998-2007. *Am J Transplant* 2009;9:879-93.
31 Frei U, Noeldeke J, Machold-Fabrizii V, et al. Prospective age-matching in elderly kidney transplant recipients—a 5-year analysis of the Eurotransplant Senior Program. *Am J Transplant* 2008;8:50-7.
32 Smits JM, Persijn GG, van Houwelingen HC, et al. Evaluation of the Eurotransplant Senior Program. The results of the first year. *Am J Transplant* 2002;2:664-70.
33 Lledo-Garcia E, Riera L, Passas J, et al. Spanish consensus document for acceptance and rejection of kidneys from expanded criteria donors. *Clin Transplant* 2014;28:1155-66.
34 Kayler LK, Srinivas TR, Schold JD. Influence of CIT-induced DGF on kidney transplant outcomes. *Am J Transplant* 2011;11:2657-64.
35 Debout A, Foucher Y, Trebern-Launay K, et al. Each additional hour of cold ischemia time significantly increases the risk of graft failure and mortality following renal transplantation. *Kidney Int* 2015;87:343-9.
36 Azancot MA, Moreso F, Salcedo M, et al. The reproducibility and predictive value on outcome of renal biopsies from expanded criteria donors. *Kidney Int* 2014;85:1161-8.

© BMJ Publishing Group Ltd 2015

Web appendix: Supplementary material

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

V Discussion

V.1 Discussion des travaux de la thèse

V.1.1 Objectif méthodologique

Le premier objectif de ce travail était d'identifier les étapes clefs des recherches à visée pronostique et de proposer une approche méthodologique et une stratégie d'analyse adéquate et rigoureuse basées sur les techniques statistiques disponibles les plus pertinentes ainsi que sur des développements personnels. Quatre étapes pouvant être complémentaires ont été identifiées dans cette thèse.

Très schématiquement la première étape consiste en l'identification et la caractérisation du phénotype de la pathologie à étudier. En pratique, cette étape est simple et ne requiert pas de processus analytique si la pathologie étudiée possède une hiérarchisation ou classification et donc des phénotypes de patients déjà bien reconnus, pour lesquels la prise en charge proposée et/ou le pronostic de survie sont clairement différenciés. En revanche lorsqu'on veut s'affranchir des classifications connues, l'identification de phénotypes de patients en oncologie ou de rejets en transplantation n'est pas quelque chose d'aisé et doit être réalisée de la façon la moins arbitraire possible. A cet effet, nous avons pu voire qu'une approche non supervisée par des outils provenant de la statistique descriptive multidimensionnelle est possible notamment avec l'utilisation successive des méthodes de classification hiérarchique et de l'analyse en composantes principales (ACP).

Dans un second temps s'en suit naturellement la recherche de facteurs pronostiques de l'évènement d'intérêt dans la population initialement considérée. Ainsi, le choix de la variable à expliquer ou critère de jugement du type « temps jusqu'à événement » censuré à droite nécessite une définition précise de l'évènement d'intérêt orienté sur le patient ou sur la maladie. Le

plus souvent, dans le domaine médical, l'analyse de données de survie se réalise à l'aide d'un modèle semi-paramétrique dit de Cox. En effet, ce modèle permet de prendre en compte l'effet de plusieurs variables explicatives pour modéliser la survenue d'un événement (partie paramétrique) sans faire d'hypothèse paramétrique sur les distributions de survie et la fonction de risque instantané de base (partie non-paramétrique). Tous les modèles de survie, y compris le modèle de Cox, reposent sur des hypothèses précises. Ce dernier suppose que le rapport des risques instantanés correspondants à deux populations de caractéristiques différentes ne dépend pas du temps, c'est ce qu'on appel l'hypothèse de proportionnalité des risques. Lorsqu'on a écrit que le risque instantané de faire l'évènement dépend des covariables de façon log-linéaire, le modèle est encore très loin d'être précisé. En effet, le codage avec lequel les covariables sont introduites dans le modèle repose aussi sur des hypothèses. Une perte d'information dans une étude dont l'objet est de fournir des estimations précises du risque de faire l'évènement n'est pas acceptable. Les travaux menés dans cette thèse s'inscrivent pour la plupart dans cette volonté de précision, et notamment ceux dont le but est de fournir des outils d'estimation du risque. Ainsi, malgré des habitudes solidement ancrées de catégorisation des variables quantitatives, il paraît clair que les inconvénients décrits dans la partie méthodologie l'emportent et que les solutions méthodologiques telles que la méthode des splines cubiques restreintes ou celle des polynômes fractionnaires sont des alternatives à privilégier.

Ensuite, la troisième étape a pour objectif de déterminer la valeur additionnelle apportée par un nouveau facteur pronostique identifié aux facteurs conventionnels déjà connus et fonctionnels dans la pratique clinique. Un intérêt majeur de la recherche médicale repose sur la quantification de l'information additionnelle apportée par un nouveau marqueur (moléculaire, génétique, radiologique, etc.) aux spectres de paramètres conventionnels. L'analyse de données de survie par une modélisation de Cox permet donc, comme cela a été évoqué précédemment, de tester l'association d'un facteur avec la distribution de survie et aussi de tester des associations indépendantes dans le cas de l'évaluation conjointe dans un même modèle

de plusieurs covariables. Cependant, pour qu'un nouveau facteur pronostique soit qualifié de bon prédicteur, une association indépendante avec l'évènement est nécessaire mais doit être complétée par d'autres analyses orientées, notamment sur la stratification du risque de faire l'évènement. Autrement dit, considérer qu'un facteur apporte une information additionnelle pour la prédiction de survenue de l'évènement requiert une association indépendante significative avec la distribution de survie, mais il doit aussi apporter une amélioration de la stratification du risque et donc être discriminant. Dans le cas des données censurées, les indices de discrimination ou indices de concordance constituent une partie de la réponse à cette problématique. Certains auteurs ont proposé de nouvelles méthodes complémentaires, le NRI et l'IDI, qui apportent une information additionnelle à l'index de concordance et qui sont basées sur la quantité et la direction de la reclassification des sujets en termes de probabilité d'évènements, après l'ajout d'un nouveau marqueur de risque à un modèle de prédiction. Ces informations C-index, NRI et IDI constituent un spectre d'information permettant de valider et d'illustrer la valeur additionnelle apportée par l'ajout d'un facteur à un modèle dit de référence constitué de facteurs pronostiques de l'évènement dit conventionnels. Le C-index permet d'apprécier la capacité discriminante du facteur tandis que le NRI et l'IDI permettent d'apprécier la nature de l'information apportée par ce facteur (marqueur de mauvais pronostic s'il reclasse le risque de faire l'évènement à la hausse chez les cas ou marqueur de bon pronostic s'il reclasse à la baisse cette probabilité chez les non-cas ou les deux).

En pratique, la présentation des statistiques du NRI et de l'IDI dans les publications sont faites sous la forme d'une estimation d'un intervalle de confiance et d'une p-value ce qui rend le résultat de l'analyse difficile d'accès et peu didactique pour les acteurs de la recherche et du soin non familiarisés avec les notions statistiques impliquées dans ces analyses. J'ai ainsi pu proposer dans cette thèse des développements personnels, notamment par des représentations graphiques, des techniques dites de reclassification du risque, afin de rendre les résultats plus didactiques et informatifs. Il est important de rappeler que ces statistiques de reclassification doivent être

employées dans un contexte de modèle bien calibré sans quoi elles peuvent véhiculer une information erronée (Leening *et al*, 2014).

Enfin, la quatrième étape identifiée consiste à considérer conjointement le spectre de paramètres pronostiques de l'occurrence de l'évènement étudié et développer des outils de prédiction adéquats permettant une stratification précise du risque d'évènement, à l'échelle individuelle et populationnelle. Une recherche visant à développer un nouveau modèle de prédiction doit absolument proposer une évaluation approfondie de ce dernier pour quantifier ses performances, dans l'objectif de fournir des prédictions dans un cadre méthodologique clairement établi. La validité, justesse ou encore précision d'un modèle de prédiction doit être vérifiée selon deux paramètres : la discrimination et la calibration. De plus, un modèle pronostique déterminé sur un échantillon de sujets est, par essence, aussi bon que possible pour cet échantillon. Ainsi, il ne faut pas s'attendre à ce qu'il soit automatiquement aussi performant pour de nouveaux sujets. Il est donc important d'étudier la valeur pronostique de ce modèle sur un autre échantillon que celui qui a servi à le définir. En effet, Il n'est pas acceptable de publier le développement d'un modèle pronostique sans validation, en tout cas au moins sans validation interne. Prétendre qu'un modèle est cliniquement pertinent, est admissible seulement avec une validation externe indépendante. Par ailleurs, un modèle est qualifié d'utile s'il fournit des prédictions précises qui informent les sujets et les cliniciens. Un autre rôle important des modèles de prédiction est d'apporter des éléments fiables pour une décision clinique éclairée afin d'améliorer les résultats concernant l'évènement d'intérêt. Une des méthodes qui permet de répondre à cette problématique est l'analyse par courbe de décision (Fitzgerald *et al*, 2015). Une fois les performances du modèle rigoureusement évaluées il est intéressant de proposer des outils pronostiques pouvant être exploités en pratique clinique. L'utilisation de l'information contenue dans la modélisation de Cox offre la possibilité de proposer aux patients, aux cliniciens et aux chercheurs, une information pertinente, objective et surtout personnalisée en termes de prédiction de l'évènement étudié à l'échelle individuelle

(nomogramme) et populationnelle (score). Dans un contexte contemporain où l'outil informatique est prédominant et avec un accès illimité à internet via les Smartphones, il est inévitable de fournir une interface web où les calculs qui correspondent aux outils sont automatisés. J'ai ainsi dans cette thèse proposé des développements via des applications web qui ont permis de rendre accessible et automatiser des outils de prédiction et de stratification du risque de décès en particulier pour les patients atteints d'un cancer pancréatique: http://www.umqvc.org/en/tool/prolap.html & http://www.umqvc.org/en/tool/proscap.html.

Le processus méthodologique proposé dans ce travail constitue une base qui garantit, si toutes les recommandations sont respectées, une bonne qualité analytique pour une recherche à visée pronostique. Naturellement, ce processus devra être enrichi à l'avenir notamment en tenant compte des nombreux développements proposés continuellement par la communauté des chercheurs en épidémiologie.

V.1.2 Objectif clinique

Un second objectif de ce travail est plus appliqué avec des réponses concrètes à des problématiques importantes posées dans le contexte de la transplantation rénale et de l'oncologie digestive.

V.1.2.1 Travaux d'intérêt de la thèse en oncologie digestive

Les travaux de ce manuscrit de thèse qui relèvent du champ de l'oncologie digestive apportent des éléments importants en termes de pronostic de survie des patients atteints d'un cancer pancréatique ou d'un cancer colorectal métastatique.

Tout d'abord nous avons développé et validé un nomogramme et un score pronostique pour prédire, au diagnostic, avant l'initiation du traitement d'induction la survie globale de patients atteints d'un cancer pancréatique localement avancé. Des outils de prédiction individuelle (nomogramme) et

populationnelle (score) ont été proposés et automatisés: http://www.umqvc.org/en/tool/prolap.html. Le modèle multivarié final de Cox sur lequel a reposé la construction de ces outils se compose de cinq facteurs indépendamment associés à la survie globale (SG) : l'âge, la douleur, l'albumine, la taille de la tumeur et le marqueur tumoral CA 19-9. Ces paramètres sont des facteurs aujourd'hui bien reconnus dans cette pathologie et illustrent le bon sens clinique des outils proposés. Les outils fournis dans cette étude ont été développés à l'aide de la plus importante population de patients atteints d'un cancer pancréatique localement avancé, provenant d'un essai thérapeutique tout en intégrant un spectre important de paramètres recueillis au diagnostic. Les résultats ont été validés en externe sur une population de patients rencontrés dans la pratique clinique courante et par conséquent non sélectionnée. Aussi, nous avons pu mettre en évidence l'existence d'une importante hétérogénéité de pronostic de survie au sein de cette population de patients qui doit être prise en compte dans le futur. En effet, les développements issus de cette étude vont permettre d'optimiser le design des essais cliniques à deux niveaux. Le premier consiste en la meilleure sélection des patients dans le cas de critères d'éligibilité basée sur le pronostic de survie et l'autre repose sur la bonne maitrise de la répartition des patients de pronostics différents à travers les groupes lors du processus de randomisation, ce qui permettra in fine une meilleure estimation de l'effet traitement ou intervention testé. Nous recommandons aussi des analyses en sous-groupes dans les essais futurs sur la base des différents groupes à risque identifiés. Par ailleurs, le nomogramme et le score vont peut-être offrir l'opportunité de définir, dans le futur, des stratégies de management adaptées aux différents profils de risque. Dans cette étude nous n'avons pas réussi à conférer aux outils pronostiques développés une valeur prédictive de la réponse aux traitements, notamment celle de la radiothérapie associée à la chimiothérapie. En effet, la cohorte de développement analysée se compose de la population de patients inclus dans l'essai thérapeutique de phase 3 internationale LAP07 dont l'objectif principal était d'évaluer le bénéfice de la radio-chimiothérapie comparée à la chimiothérapie seule en termes de survie globale chez des patients atteints d'un cancer pancréatique localement

avancé, contrôlé après une chimiothérapie d'induction à base de gemcitabine de 4 mois. Globalement, cet essai n'a pas reporté de bénéfice de survie aux patients traités par radio-chimiothérapie, en comparaison aux patients traités par chimiothérapie seule. Le rationnel de cet essai était basé sur plusieurs observations rétrospectives sérieuses d'un bénéfice en faveur de la radio-chimiothérapie pour des patients qualifiés de bon pronostic. Ces observations ont guidé la construction de l'essai LAP07 en proposant la randomisation chimiothérapie versus radio-chimiothérapie uniquement aux patients contrôlés après 4 mois de chimiothérapie d'induction, donc a priori de bon pronostic. Devant le résultat négatif observé dans l'essai, un des enjeux du travail proposé avec la construction du nomogramme et du score pronostique était de pouvoir identifier un effet différentiel de la radio-chimiothérapie à travers les groupes de pronostics différents identifiés. Une tendance favorable a été suggérée dans le groupe de bon pronostic mais n'a pas été trouvée comme statistiquement significative. Le travail et les outils proposés constituent ainsi une base de travail pour le futur. En effet, le modèle final ne propose qu'un indice de discrimination de 0.6 ce qui est acceptable mais très largement améliorable. Des paramètres importants, par exemple le ratio neutrophile – lymphocyte et aussi la qualité de vie relative à la santé, devront dans le futur être intégrés dans le but d'améliorer le pouvoir discriminant du modèle et ainsi peut-être lui conférer un pouvoir prédictif de la réponse aux traitements.

Un autre travail de cette thèse propose aussi le développement et la validation d'un nomogramme et d'un score pronostique pour prédire la survie globale, mais cette fois-ci pour des patients atteints d'un cancer pancréatique avancé donc non résécable sous seconde ligne de chimiothérapie : http://www.umqvc.org/en/tool/proscap.html. Cette étude répond à une question clinique importante. En effet, la prise en charge en première ligne des patients atteints d'un cancer pancréatique avancé est relativement bien établie et standardisée avec deux combinaisons thérapeutiques : le FOLFIRINOX et l'association gemcitabine plus *nab*-paclitaxel, qui ont montré un bénéfice de survie globale par rapport à l'administration de gemcitabine seule, qui est restée longtemps le traitement de référence dans cette

pathologie. En revanche au-delà de la première ligne de chimiothérapie, il n'y a pas de stratégie consensuelle validée à ce jour pour guider la sélection des patients pouvant tirer un bénéfice cliniquement pertinent d'une seconde ligne de chimiothérapie. Or, la moitié des patients atteints d'un cancer pancréatique avancé restent en bon état général à la fin de la première ligne de chimiothérapie et peut donc a priori recevoir dans des conditions favorables des lignes ultérieures de chimiothérapie. Cependant, tous les patients traités en seconde ligne de chimiothérapie ne semblent pas en tirer un bénéfice important en termes de survie globale. Ainsi, l'identification de l'ensemble des facteurs pronostiques de la survie globale et le développement d'outils de prédiction pour améliorer la stratification du risque de décès pour ces patients sous seconde ligne de chimiothérapie est un enjeu majeur afin d'améliorer les éléments disponibles pour la prise de décision thérapeutique. Sur la base d'une large population très consciencieusement phénotypée (plus de 50 paramètres analysés tous très bien renseignés) de patients traités sur une période de 11 ans pour un cancer pancréatique avancé au Centre Hospitalier Universitaire de Besançon, nous avons donc proposé un nouveau modèle de prédiction de la survie globale à l'initiation de la seconde ligne de chimiothérapie basé sur neuf paramètres cliniquement pertinents. Nous avons, sur la base de ce nouveau modèle, qui possède une excellente capacité discriminante de la survie globale, développé des outils pronostiques fonctionnels à l'échelle individuelle (nomogramme) et populationnelle (score) dont l'utilisation a été facilitée par le développement d'une interface web qui automatise les calculs et synthétise les résultats (http://www.umqvc.org/en/tool/proscap.html). Au-delà des outils proposés, cette étude met en évidence comme attendu une considérable hétérogénéité de pronostic pour la survie globale sous seconde ligne de chimiothérapie chez des patients atteints d'un cancer pancréatique avancé. Un des éléments clefs de cette étude est la disponibilité de patients très bien phénotypés qui n'ont pas reçu un traitement de seconde ligne à l'issue de l'échec de la première ligne. Aussi, pour chacun des groupes à risque distincts identifiés, nous disposons de patients traités et non traités par seconde ligne de chimiothérapie et nous observons une interaction significative entre la

réponse au traitement de seconde ligne en termes de survie globale et le pronostic du patient à son initiation. Un effet significatif de l'administration d'un traitement par chimiothérapie de seconde ligne est identifié dans chacun des groupes mais plus le pronostic de survie devient mauvais, plus l'intérêt d'un traitement (amplitude de l'effet) décroit. Ceci signifie donc que dans un contexte de seconde ligne pour des patients atteints d'un cancer pancréatique avancé, le pronostic de survie globale est prédictif du bénéfice attendu de l'administration de cette dernière. Les outils développés dans cette étude offrent l'opportunité de définir des stratégies de management adaptées aux différents profils de risque. Un autre élément clef de l'étude repose sur la démonstration de l'intérêt de prendre en compte l'intégralité de l'information délivrée par le modèle et donc par les outils proposés. En effet, aujourd'hui les trois paramètres pragmatiques utilisés par les cliniciens pour la prise de décision de l'administration de la seconde ligne sont l'âge, le performance status, et la durée de la première ligne qui sont d'ailleurs présents dans le modèle final. Nous nous sommes donc interrogés sur la valeur discriminante additionnelle apportée par les autres paramètres identifiés dans le modèle multivarié et avec une procédure méthodologique rigoureuse d'étude de la discrimination et de la reclassification du risque. Nous avons pu démontrer que le modèle est plus intéressant lorsque nous le considérons dans sa totalité puisque beaucoup plus discriminant. A la fin de la première ligne, la question clinique qui se pose pour ces patients est la possibilité de ne pas administrer une seconde ligne aux patients qui ont une très forte probabilité de ne pas survivre à six mois après l'initiation de cette dernière. Il est alors fondamental d'utiliser le meilleur modèle de prédiction possible afin de limiter les faux positifs et ainsi ne pas priver à tort un patient qui peut tirer un bénéfice cliniquement pertinent de cette thérapie. Avec ces considérations il apparaît naturel d'utiliser les informations issues du modèle proposé.

Le troisième travail proposé s'inscrivant dans le champ de l'oncologie digestive se consacre à l'étude de la valeur pronostique pour la survie globale du Syndecan-1 (CD138) au diagnostic chez des patients atteints d'un cancer du colorectal métastatiques (mCRC) initialement non résécable. Aujourd'hui, diverses options thérapeutiques sont disponibles en première ligne pour les

patients atteints d'un mCRC. Dans ce contexte, l'identification au diagnostic de biomarqueurs permettant une stratification du risque de décès est une problématique cruciale afin de mieux connaître le pronostic de la maladie, d'individualiser plus précisément les différentes stratégies de traitement et par conséquent d'améliorer le rapport coût-efficacité des interventions. Les deux scores pronostiques de la SG actuellement les plus utilisés pour les patients atteints d'un mCRC sont le score de Kohne et celui du Groupe Coopérateur Multidisciplinaire en Oncologie GERCOR. Ces scores sont, tous deux, basés sur le performance status (PS) qui est un paramètre ayant pour objet d'évaluer l'état fonctionnel d'un patient qui peut être influencé par de multiples facteurs liés aux patients (anorexie, état psychologique) et/ou liés à la maladie tumorale (inflammation, taille tumorale, anémie, …). La substitution du PS par un paramètre biologique plus reproductible et aussi plus spécifique et représentatif de la tumeur ou de l'état métabolique du patient pourrait être intéressant pour mieux individualiser les différentes options thérapeutiques au regard de leur mode d'action. L'invasion, l'angiogenèse et le contrôle du microenvironnement tumoral sont des caractéristiques communes de la progression du cancer et peu de biomarqueurs spécifiques de ces caractéristiques sont disponibles au moment du diagnostic. Le Syndecan-1 (SDC-1, aussi appelé CD138) est un protéoglycane transmembranaire de type héparane sulfate qui constitue un biomarqueur d'intérêt. Dans ce travail nous avons démontré que le CD138 soluble est un facteur pronostique indépendant de la SG chez les patients atteints de mCRC qui apporte une valeur additionnelle aux autres paramètres conventionnels en termes de discrimination du risque de décès. L'évaluation de la combinaison biologique CD138-LDH au diagnostic offre une amélioration de la capacité de stratification du risque de décès et ainsi peut servir de base pour adapter des interventions thérapeutiques. Nous avons mis en évidence une population conséquente de très mauvais pronostic pour laquelle une intensification de traitement pourrait être proposée. Par conséquent, les résultats de ce travail sont en train d'être mis à profit dans le cadre d'un projet interventionnel visant à étudier l'impact de l'intensification thérapeutique dans une population de patients de mauvais pronostic, sur la base du score pronostique biologique

basé sur le dosage sérique du Syndecan-1 (CD138) et LDH. En effet comme précisé précédemment la multiplicité des thérapeutiques en première ligne de traitement des cancers colorectaux (CRC) invite aujourd'hui à déterminer des stratégies de prise en charge guidées par des biomarqueurs prédictifs ou pronostiques. Cette individualisation thérapeutique est un axe de recherche important. L'intensification thérapeutique est une stratégie thérapeutique qui peut améliorer la survie des patients pris en charge pour un CRC. Des études randomisées ont déjà montré l'intérêt d'une stratégie d'intensification par FOLFOXIRI seul puis en combinaison au bevacizumab (Falcone *et al*, 2007; Loupakis *et al*, 2014a). Malgré l'objectivation d'une amélioration de la survie globale, l'utilisation en pratique clinique de cette stratégie reste faible. En effet, le surcroit de toxicités observé chez les patients traités par FOLFOXIRI pourrait faire craindre un risque en termes de coût/toxicité/efficacité. De plus, l'influence de ces traitements sur la qualité de vie relative à la santé n'a pas été décrite. Les patients pouvant bénéficier de cette intensification ne sont pas clairement définis. Il semble que les trithérapies pourraient être particulièrement efficaces pour un sous-groupe de patients dont les tumeurs portent une mutation du gène BRAF (5 à 8% des CRC), qui confère aux CRC un pronostic particulièrement péjoratif (Venderbosch *et al*, 2014). Dans ce sous-groupe de patients, l'intensification thérapeutique a un bénéfice relatif majoré (46% d'amélioration de la survie globale et 43% d'amélioration de la survie sans progression) (Loupakis *et al*, 2014). Ces résultats posent donc la problématique de notre capacité à déterminer une population dont l'évolution naturelle de la maladie serait péjorative et justifierait d'une intensification des chimiothérapies. Comme démontré dans notre étude le dosage au diagnostic de la maladie métastatique du CD138 et du LDH, avant traitement, dans le sérum, permet d'identifier des maladies dont l'évolution est défavorable (médiane de survie globale de 19,8 mois contre 38,9 mois lorsque les taux de LDH et de Syndecan-1 sont élevés ou faibles, respectivement). Les patients de mauvais pronostic identifiés par ce test sérique représentent 30% de la cohorte, ce qui permet d'envisager la proposition d'une stratégie d'intensification thérapeutique spécifique pour cette population. Sur la base de ce rationnel l'étude CLavSyn a pour objectif d'évaluer l'intérêt d'une

intensification thérapeutique par FOLFOXIRI plus bevacizumab dans une population de mauvais pronostic de survie globale définie par des taux élevés de LDH et de Syndecan-1. L'amélioration de la qualité de vie et l'optimisation des coûts en matière de santé seront aussi évaluées. Il s'agira d'une phase II randomisée qui si les résultats intermédiaires sont positifs, conduira à une phase III. L'investigateur principal de cette étude est le Dr Marine Jary de l'équipe d'oncologie médicale du Centre Hospitalier Régional Universitaire de Besançon.

V.1.2.2 Travaux complémentaires à la thèse en transplantation rénale

Les travaux de ce manuscrit de thèse qui relèvent du champ de la transplantation rénale apportent des réponses concrètes en termes de pronostic du greffon, dans le contexte de la réponse allo-immune induite par la transplantation à partir d'un donneur génétiquement différent, représentant aujourd'hui un obstacle majeur au succès de la greffe. Nous apportons ainsi une vision éclairée du rejet aigu d'allogreffe en transplantation rénale grâce à une approche innovante qui intègre des outils mathématiques, immunologiques et histologiques appliqués à de larges cohortes de patients transplantés très bien phénotypés.

Tout d'abord nous avons individualisé quatre profils distincts de rejets : le rejet médié par les lymphocytes T sans endartérite, le rejet médié par anticorps sans endartérite, le rejet vasculaire médié par anticorps et le rejet vasculaire médié par les lymphocytes T. Ces deux dernières entités étaient absentes de la classification internationale puisque jusqu'à présent, le diagnostic de rejet opposait le rejet médié par des lymphocytes T au rejet médié par anticorps. L'étude réalisée met en lumière un nouveau profil de rejet de greffe, appelé rejet vasculaire qui se caractérise par une inflammation de la paroi des artères des greffons en réponse à la présence d'anticorps anti-HLA dirigés contre le donneur (rejet vasculaire médié par anticorps) ou à l'agression par les lymphocytes T (rejet vasculaire médié par des lymphocytes T sans endartérite). Nous avons mis en lumière l'intérêt de l'identification

précise de ces nouvelles entités de rejet en démontrant qu'elles possèdent des profils pronostiques de survie du greffon distincts. Le rejet vasculaire médié par anticorps a en effet un pronostic extrêmement péjoratif dans l'évolution de la greffe, justifiant clairement de l'individualiser des autres rejets aigus d'allogreffe. Par conséquent, ces lésions ne doivent donc plus être interprétées comme des arguments directs en faveur du diagnostic de rejet médié par les lymphocytes T, comme c'est actuellement le cas dans la classification internationale. Elles font partie du spectre général de modifications inflammatoires de l'endothélium en réponse aux anticorps anti-HLA dirigés contre le donneur. Nos résultats ont donc des implications thérapeutiques majeures. Le rejet vasculaire est traditionnellement considéré comme une forme clinique sévère de rejet ne répondant pas au traitement classique. Il nécessite des thérapeutiques immunosuppressives puissantes avec néanmoins une efficacité aléatoire. L'individualisation précise de ces formes de rejet et leur caractérisation physiopathologique ouvre la voie à des probables futurs progrès thérapeutiques. En effet, le choix de la stratégie thérapeutique est déterminant pour l'évolution des rejets vasculaires médiés par anticorps. Les patients traités exclusivement par une thérapie de type rejet cellulaire ont un risque significativement plus important de perte de greffon que les patients recevant une thérapie intensive ciblant les anticorps anti-HLA de type rejet humoral. L'identification précise des rejets vasculaires et de leurs mécanismes devrait ainsi permettre la mise en place de stratégies thérapeutiques spécifiques permettant d'améliorer leur pronostic. Ensuite, nous avons étudié le rôle pronostic de la propriété des anticorps anti-HLA à fixer le complément. Pour ce faire nous avons utilisé la même approche qui consiste à considérer les données immunologiques, histologiques, cliniques et épidémiologiques sur de larges cohortes de patients. Déterminer la capacité des DSA à fixer le C1q qui caractérise la première étape de l'activation du complément permet d'identifier précocement le potentiel cytotoxique des anticorps anti-HLA avant qu'ils ne déclenchent la cascade du complément. Jusqu'à présent, seules quelques données suggèrent que l'existence d'une liaison spécifique du C1q aux DSA pourrait être plus précise pour définir le rejet humoral que la simple détection des immunoglobulines G.

Ainsi, la détection des DSA qui fixent le complément peut avoir un rôle majeur dans la stratification du risque de perte des greffons et donc dans l'évaluation du pronostic de la greffe. Dans notre étude, parmi les patients présentant des DSA détectables dans la première année post-transplantation, seulement un petit quart ont des DSA fixant le complément. Nous démontrons une association robuste entre la capacité des DSA à fixer le complément et le risque de perte de greffon. Grâce à cette capacité des DSA, nous isolons dans cette étude un groupe de patients qui présente une augmentation dramatique du risque de perte de greffons dans les cinq premières années de greffe Ce risque est 11,2 fois plus important que chez les patients avec des DSA ne fixant pas le complément. Ces patients ne sont pas identifiés par l'approche actuelle qui consiste à détecter les DSA et éventuellement leur MFI en post-greffe. Nous proposons par ailleurs un modèle de stratification du risque de perte de greffons dans la première année de greffe dans lequel cette propriété des DSA augmente significativement sa performance discriminative. Avec ce résultat nous ouvrons, pour cette population à haut risque, la voie à de futures études thérapeutiques, telles que l'inhibition de la cascade du complément ou encore sa modulation. Dans un autre travail, nous avons étudié l'apport du microscope moléculaire dans la stratification du risque de perte du greffon. Ainsi dans une population très bien phénotypée de patients transplantés rénaux, pour lesquels un diagnostic de rejet humoral aigu prouvé par une biopsie rénale a été réalisée, nous avons démontré l'utilité de la mesure des transcrits par étude moléculaire des biopsies en complément des informations traditionnellement utilisées dans la stratification du risque de perte du greffon. Avec une approche méthodologique très rigoureuse autour de cette notion de reclassification du risque, nous avons quantifié le bénéfice apporté par le microscope moléculaire aux paramètres classiques (immunologie, histologie, clinique). Ce score produit par ce nouvel outil qu'est le microscope moléculaire peut donc être utilisé comme un outil compagnon des paramètres conventionnels comme l'histologie et permet d'apporter de la précision à la prédiction du risque de perte du greffon.

V.1.2.1 Travail d'intérêt de la thèse en transplantation rénale

Enfin, la dernière étude présentée dans ce travail dans le cadre de la transplantation rénale, s'est concentrée sur l'identification de déterminants pronostiques de la survie du greffon des patients transplantés à partir des donneurs à critères élargis et la détermination de l'impact de la présence de DSA circulant le jour de la greffe. En effet la politique de greffe, en réponse à la pénurie d'organes, a consisté ces dernières années en une augmentation de l'âge des donneurs. Ceci a conduit à la définition de donneurs à critères dits élargis. Ces greffons rénaux représentent en France environ la moitié des reins prélevés chez des donneurs décédés et donc une étude approfondie de leur pronostic de survie présente un intérêt majeur. Dans cette étude, nous démontrons que les receveurs d'un rein ECD présentent une survie acceptable du greffon à 7 ans. Cependant un déterminant majeur de mauvais pronostic de la survie des greffons ECD est la présence de DSA le jour de la transplantation. Ainsi, cette caractéristique des patients étant connue avant la transplantation l'utilisation des reins issus de donneurs ECD devrait être limitée aux patients non porteurs de DSA. En effet, en l'absence de cette prise de risque, l'écart entre la survie du greffon entre les reins ECD et SCD est minime. Ce travail a donc permis de valider *a posteriori* la politique de greffe actuelle en France avec l'encouragement de l'utilisation de rein ECD ce qui semble être une solution acceptable en réponse à la pénurie d'organes si celle-ci ne s'adresse pas à des patients immunisés avec des DSA le jour de la greffe.

V.1.2.2 Travaux de la thèse présentés en annexes

Deux travaux importants pouvant s'inscrire dans le champ méthodologique de ce travail de thèse sont présentés en annexe. Il s'agit d'une part de la confirmation de la valeur pronostique additionnelle, pour la survie globale, apportée aux paramètres classiques des données de qualité

de vie relative à la santé, au moment du diagnostic, chez les patients âgés atteints de cancer du poumon avancé non à petites cellules. Considérer la qualité de vie au diagnostic permet d'identifier des groupes de patients à risque de décès différents. L'information véhiculée par la qualité de vie pourrait ainsi être utile notamment dans la sélection des patients pouvant être éligibles à une bi- chimiothérapie (cf annexe 2).

Fiteni F*, **Vernerey D***, Bonnetain F, Vaylet F, Sennélart H, Trédaniel J, Moro-Sibilot D, Herman D, Laizé H, Masson P, Derollez M, Clément-Duchêne C, Milleron B, Morin F, Zalcman G, Quoix E, Westeel V (2016) Prognostic value of health-related quality of life for overall survival in elderly non-small-cell lung cancer patients. Eur J Cancer Oxf Engl 1990 52: 120–128, doi:10.1016/j.ejca.2015.10.004.
*Co first authorship

D'autre part avec la transplantation rénale comme modèle, nous avons investigué les déterminants et les conséquences de l'artériosclérose accélérée. En effet, le rôle des anticorps circulants en plus des facteurs de risque cardiovasculaires traditionnels dans le développement de l'artériosclérose accélérée et leurs conséquences cliniques à long terme n'a pas encore été démontré à l'échelle populationnelle. Dans cette étude nous démontrons que les anticorps circulants sont un déterminant majeur pour le développement dans sa forme sévère de l'artériosclérose, observé sur les biopsies de greffon et aussi des événements cardiovasculaires majeurs des patients, ceci indépendamment des facteurs de risque cardiovasculaire traditionnels (cf annexe 3).

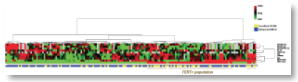
Loupy A*, **Vernerey D***, Viglietti D*, Aubert O, Duong Van Huyen J-P, Empana J-P, Bruneval P, Glotz D, Legendre C, Jouven X, Lefaucheur C (2015) Determinants and Outcomes of Accelerated Arteriosclerosis: Major Impact of Circulating Antibodies. Circ Res 117: 470–482, doi:10.1161/CIRCRESAHA.117.306340.
*Co first authorship

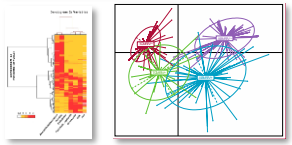
V.1.3 Mise en perspective de l'ensemble des travaux cliniques dans le processus méthodologique proposé

Les travaux cliniques réalisés dans cette thèse ont bénéficié du raisonnement méthodologique proposé dans la partie méthode de ce manuscrit.

Ci-après voici une mise en perspective des travaux présentés dans le schéma méthodologique d'analyse proposé :

Population globale de patients atteints d'une pathologie





Identification et caractérisation du phénotype de la pathologie à étudié

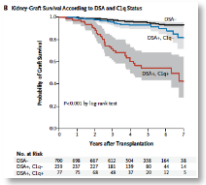
(Localisation tumorale: locale, loco régionale, métastatique
Maladie résécable ou non
Caractéristiques moléculaires, histologiques , …)

* *Lancet, 2013*

Population de patients avec un phénotype plus homogène de la pathologie

Recherche de facteurs pronostiques

(Paramètres cliniques, fonctionnels, histologiques, immunologiques, biologiques, de qualité de vie …)

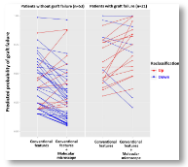


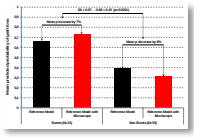
** *NEJM, 2013*
*** *JASN, 2014*
† *EJC, 2016*
†† *IJC,2016*

Evènement d'intérêt et variable à expliquer

(Décés / Survie Globale, Perte du greffon / Survie du greffon, …)

Détermination de la valeur additionnelle apportée par un nouveau facteur pronostique aux facteurs conventionnels

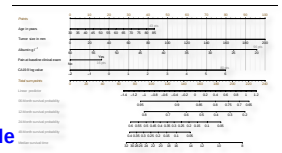


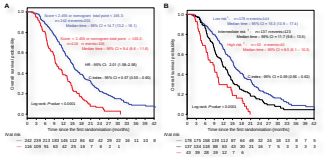


¶*BMJ, 2015*
††† *Circulation research, 2015*

Développement ou amélioration de modèles intégratifs à visée pronostique

¶¶ *BJC 2016 /*
¶¶¶ *JNCI in review*





Outils pronostiques de prédictions à l'échelle individuelle (nomogramme) & populationnelle (score)

- Amélioration de la prise en charge des patients au regard du pronostic ou/ et si une valeur prédictive à la réponse au traitements passe à travers le gradient pronostic: médecine de précision
- Optimisation des essais cliniques futurs (sélection, randomisation, analyse en sous groupe…)

***** Lefaucheur C, Loupy A, Vernerey D et al. Antibody-mediated vascular rejection of kidney allografts: a population-based study. **Lancet 2013**
****** Loupy A, Lefaucheur C, Vernerey D et al. Complement-binding anti-HLA antibodies and kidney-allograft survival. **NEJM 2013**
******* Loupy A, Lefaucheur C, Vernerey D et al. Molecular Microscope Strategy to Improve Risk Stratification in Early Antibody-Mediated Kidney Allograft Rejection. **JASN. 2014**
† Fiteni F*** and** Vernerey D+, Bonnetain F et al. Prognostic value of health-related quality of life for overall survival in elderly non-small-cell lung cancer patients. **EJC. 2016**
†† Jary M, …, Vernerey D Prognostic value of soluble Syndecan-1 in metastatic colorectal cancer patients treated with bevacizumab and chemotherapy **IJC 2016**
††† Loupy A, Vernerey D+, Viglietti D et al Determinants and Outcomes of Accelerated Arteriosclerosis: Major Impact of Circulating Antibodies**, Circulation research 2015**
¶ Aubert O and Kamar N and Vernerey D+ et al Long term outcomes of transplantation using kidneys from expanded criteria donors: prospective, population based cohort study, **BMJ 2015**
¶¶ Vernerey D, Goldstein D, Huguet F et al. Prognostic Score and Nomogram to Predict Overall Survival in Locally Advanced Pancreatic Cancer. **BJC 2016**
¶¶¶ Vienot A, …, Vernerey D Development and validation of a prognostic nomogram and score to predict overall survival and usefulness of second-line chemotherapy in advanced pancreatic adenocarcinoma**". JNCI in review**

+co-first authorship

V.2 Discussion générale

La médecine de précision, également appelée médecine personnalisée, a pour objectif de proposer au patient une prise en charge et un traitement adaptés à ses caractéristiques individuelles et aussi à celles de sa maladie ou tumeur. Autrement dit, cette approche souhaite offrir la possibilité de classer les individus en sous-groupes qui diffèrent dans la biologie et/ou le pronostic de la maladie afin d'individualiser les interventions et les traitements à ceux qui en tireront bénéfice, tout en évitant les effets secondaires aux autres. Aujourd'hui et en particulier en oncologie, les traitements qui s'inscrivent dans ce contexte sont les thérapies ciblées et l'immunothérapie. Le patient est bien évidement au centre de la médecine de précision, mais il ne faut pas être dupe, il y a aussi une volonté d'optimiser le rapport coût-efficacité des nouvelles thérapies très onéreuses.

Dans ce contexte, le président Barack Obama a lancé en 2015 un programme nommé PMI (Precision Medicine Initiative) et a exprimé avec conviction que la science peut, aujourd'hui, offrir un grand potentiel pour améliorer la santé (www.whitehouse.gov/precisionmedicine). L'objet de cette initiative est l'accélération des progrès par la constitution d'une cohorte d'au moins un million d'américains volontaires qui partageront leurs données de santé et seront suivis. Le projet présente des ambitions à court terme, en particulier sur la recherche en oncologie, ainsi qu'une vision à plus long terme, dont le but est de générer suffisamment de données et de connaissances pour déployer cette approche à tous les champs de la médecine. Le but avoué est de fonder une nouvelle ère dans la recherche biomédicale, mettant les patients au centre de la réflexion, et s'appuyant sur le libre partage de données issues de ces derniers i.e. l'e-medecine. Cette initiative a été relayée avec enthousiasme par Francis Collins et Harold Varmus dans le New England Journal of Medicine (Collins & Varmus, 2015). L'initiative semble arriver au moment opportun puisque nous possédons les technologies de biologie moléculaire, de génomique et de bioinformatique avec une facilité pour la constitution de larges bases de données épidémiologiques

indispensables pour analyser un spectre d'information aussi riche. Ce besoin de constitution de larges banques de données a aussi été relayé sous une autre forme par l'éditeur en chef du Lancet ; Richard Horton ; avec un appel à financer de larges cohortes de réplication, afin de valider les « découvertes » pas toujours obtenues dans des conditions optimales (Horton, 2015). Cet appel a été lancé suite à un symposium sur la reproductibilité et la fiabilité de la recherche biomédicale qui s'est tenu au Wellcome Trust à Londres. Ces initiatives vont certainement trouver écho à moyen terme puisqu'il va falloir les structurer et aussi prendre le temps de constituer de telles cohortes.

Cependant, il est important de préciser que de telles initiatives existent déjà et en particulier en cancérologie digestive. Sous l'impulsion de la Fondation ARCAD (http://www.fondationarcad.org/) et de son président le Pr Aimery de Grammont, le programme d'essais cliniques ARCAD a été lancé en 2008. Il s'agit d'une collaboration mondiale de cliniciens, de statisticiens et de scientifiques spécialisés dans les tumeurs malignes gastro-intestinales. Récemment, la première version de la base de données dédiée au cancer colorectal avancé, intégrant les données individuelles de la plupart des essais cliniques récents, a été proposée à la communauté. Un travail miroir est en cours de constitution pour le cancer pancréatique avancé. L'objectif du programme est de faciliter l'accès aux données pour le développement de futurs essais cliniques et aussi pour des recherches originales ou de validation.

Nous comprenons bien que dans ce contexte et encore plus dans un futur proche la recherche pronostique et les études de cohortes ont et auront une importante place dans la recherche biomédicale. A son échelle, cette thèse s'inscrit dans le sillage de ces initiatives, en particulier pour le domaine de la transplantation rénale et du cancer pancréatique, puisque les cohortes étudiées se situent parmi les mieux phénotypées et les plus importantes en termes de nombre de patients au monde. Ainsi, l'analyse de ces données nous a permis d'apporter des réponses à des problématiques pronostiques importantes : existence d'un hétérogénéité importante de survie globale chez les patients atteints de cancer pancréatique avancé en seconde ligne ou

localement avancé au diagnostic, démonstration de la pathogénicité des DSA post-transplantation rénale dans seulement 25% des cas lorsque leur capacité à fixer le complément est active, etc.

Enfin, la facilité d'accès aux données en temps réel dans le futur et même aujourd'hui avec les bilans biologiques, radiologiques, les examens cliniques réalisés de façon régulière lors des visites de suivis des patients et aussi par exemple avec le lancement à l'Hôpital de Besançon du monitoring de la qualité de vie relative à la santé pour les patients atteints de cancer, suggère fortement l'étude de l'évènement d'intérêt de manière dynamique. Le travail proposé dans cette thèse ne traite pas de cet aspect dynamique du pronostic. Il est naturel de penser que le pronostic du patient (par exemple ses prédictions de survie) peut évoluer au cours du temps. C'est même en réalité l'objectif recherché pour, par exemple, des patients dont les prédictions de survie sont extrêmement mauvaises au diagnostic. Les modèles pronostiques incluent des paramètres qui sont inamovibles tels que l'âge, le sexe et aussi des paramètres qui peuvent fluctuer comme la biologie ou encore la qualité de vie et qui peuvent être le reflet de l'efficacité de la prise en charge. Dans cette optique, il paraît dans un premier temps intéressant d'étudier les propriétés des outils développés au moment t à un moment t+x afin de constater si le pouvoir discriminant des outils est toujours présent. Nous pourrions même imaginer, si le scoring obtenu au temps t est aussi performant au temps t+x, que le critère de jugement d'un essai clinique puisse être, pour un patient classé comme à haut risque de faire l'évènement à t, le fait d'être reclassé dans une catégorie à moindre risque à t+X, ce qui serait une preuve d'efficacité de la prise en charge testée. Dans un second temps, il serait aussi intéressant de pouvoir modéliser l'évolution des variables qui peuvent fluctuer dans le temps comme la biologie à l'aide de variables dépendantes du temps introduites dans la modélisation de Cox. Cette modélisation ne permettrait pas de mettre à jour le pronostic du patient de manière dynamique mais peut être utile pour identifier les paramètres pour lesquels leur fluctuation s'avère intéressante sur la prédiction de l'évènement au temps t. Enfin, tout un champ de la statistique explore cette idée de prédiction dynamique afin de mettre à jour les prédictions d'évènements

d'intérêt et donc le pronostic des patients en fonction de l'occurrence d'autres évènements cliniques ou de l'évolution de biomarqueurs. Il est donc possible aujourd'hui d'obtenir des prédictions issues d'un modèle conjoint pour un événement récurrent (récidive ou rejet du greffon) et un événement terminal (décès ou perte du greffon). Les prédictions individuelles proposées sont dynamiques, dans le sens où le temps et la fenêtre de prédiction peuvent varier, afin de pouvoir mettre à jour la prédiction lors de la survenue de nouveaux événements. C'est l'idée de ce qui est proposé par les prédictions issues des modèles à fragilité partagés ou les modèles conjoints (Mauguen et al, 2013). Le landmarking est aussi une méthode intéressante qui permet de prendre en compte l'histoire clinique du patient dans un modèle tel qu'un modèle de Cox (Dynamic Prediction in Clinical Survival Analysis, 2011). Les analyses dites « landmark » correspondent à un modèle de survie marginal à des temps de prédiction fixes et connus. Cependant, dans la pratique clinique courante, les prédictions devraient se faire ou être mises à jour à chaque fois qu'une nouvelle mesure de biomarqeur est obtenue et pas seulement à quelques temps sélectionnés. En effet, les visites de suivi des patients et donc les mesures longitudinales des biomarqueurs sont souvent réalisées de façon irrégulière après le diagnostic. En conséquence, les praticiens ont besoin de modèles de prédiction dynamiques qui ont la caractéristique d'être applicables quelque soit le temps post diagnostic ou traitement considéré. Des chercheurs ont récemment proposé une approche intéressante pouvant effectuer des prédictions dynamiques de manière continue au fil du temps (Huang *et al*, 2016). Cette approche propose une procédure en deux étapes. Dans la première étape, une prédiction de survie est fournie classiquement sur l'information de covariables au temps d'origine en ignorant l'information de l'évolution des covariables en post-origine. La seconde étape consiste en une actualisation ou amélioration des prédictions obtenues dans l'étape précédente en prenant en compte cette fois ci l'information post origine des covariables. Dans ce travail des études par des techniques de simulation montrent que cette approche à deux étapes permet d'obtenir des estimations stables pour la modélisation de l'effet des biomarqueurs sur l'évènement d'intérêt considéré, et présente de bonnes performances prédictives. Il s'agit

d'un bon compromis entre les deux approches majeures, à savoir (i) la modélisation conjointe des données longitudinales et de survie et (ii) l'analyse dites « landmark ». Cette nouvelle approche compte tenu de sa compatibilité avec la pratique clinique pourrait offrir des perspectives intéressantes dans le futur.

L'idée de vouloir donner un aspect dynamique aux prédictions de survie des patients est élégante. En effet, une évolution différente des covariables chez des patients qui possèdent un même profil de risque au diagnostic ou à l'initiation du traitement offre la possibilité d'une prise en charge plus personnalisée. Cependant, une meilleure appréhension du pronostic du patient avec la considération d'une information longitudinale n'autorise pas pour autant de négliger la volonté d'améliorer les outils pronostiques construits à l'aide des seules informations disponibles au diagnostic de la maladie. C'est, en effet, à ce moment que le choix du parcours de soins du patient doit être le plus optimal pour lui proposer la prise en charge la plus adaptée afin d'optimiser sa survie dans une qualité de vie satisfaisante. Il est donc essentiel de continuer de rechercher les sources d'information non connues à ce jour pour que les modèles proposés puissent être améliorés et arborés une excellente capacité discriminante. La double application clinique proposée dans cette thèse montre que cela est possible avec des modèles qui en transplantation rénale sont caractérisés par un indice de discrimination excellent (C >0.8 pour la survie des greffons à 1 an post greffe) mais qui sont encore perfectibles. En revanche, au diagnostic pour les patients atteints d'un cancer du pancréas localement avancé la discrimination de la survie globale reste relativement faible (C = 0.6) et ce malgré l'étude de populations conséquentes avec un spectre de paramètres important. Ceci est la preuve que le chemin pour comprendre les mécanismes du cancer du pancréas, entre autre, est encore important et nécessite des efforts de recherche conséquents dans le futur.

VI Perspectives du travail de recherche

VI.1 Evaluation et enrichissement de la méthodologie

Ce travail illustre mon apprentissage et mes développements dans le contexte de la recherche pronostique. .Il est important de noter que la stratégie d'analyse proposée peut véhiculer des limites. Une d'entre elles réside dans le choix de réaliser une analyse univariée préliminaire dans la construction d'un modèle multivarié de Cox afin de mettre en avant les facteurs intéressants d'un point de vue statistique et éventuellement les enrichir de facteurs connus d'un point de vue clinique. Cette approche a été décrite comme n'étant pas toujours appropriée (Sun *et al*, 1996) car elle peut rejeter à tord des variables intéressantes pour la phase multivariée. Ainsi, des procédures dites « stepwise » peuvent constituer une information additionnelle dans des analyses de sensibilité. Une autre limite réside donc dans le choix de la modélisation par un modèle de Cox du lien existant entre la fonction de risque et les covariables étudiées. *A priori* il n'y pas de raison de remettre en question la modélisation par le modèle de Cox lorsque nous sommes dans des conditions d'application convenables mais il est intéressant de noter que des méthodes d'apprentissage automatique telles que les forets aléatoires ou « Random Forest » qui sont des méthodes non paramétriques peuvent offrir des alternatives intéressantes à la stratégie traditionnelle de Cox dans des conditions extrêmes (très peu d'évènements ou encore un nombre de facteurs à investiguer plus nombreux que le nombre de patients ce qui se retrouve typiquement en épidémiologie génétique) (Mogensen *et al*, 2012). Il serait intéressant d'évaluer la performance de ces alternatives dans les études proposées dans cette thèse afin de comparer ces approches en termes de capacité à discriminer dans des conditions qui ne contre indiquent pas l'utilisation du modèle de Cox. Ainsi, sur la base de ces résultats et des

nombreux développements proposés continuellement par la communauté des chercheurs en épidémiologie le processus méthodologique proposé dans ce travail pourra être enrichi à l'avenir.

VI.2 Score pronostique dans la première année post-transplantation rénale de la survie du greffon à long terme

Un travail miroir de ce qui a été fait en oncologie digestive avec le développement et la validation d'un nomogramme et d'un score pronostique, qui aurait permis d'illustrer la dernière étape méthodologique proposée dans le champ de la transplantation rénale, devait être initialement présenté dans cette thèse. Cette recherche a pour objectif de proposer des outils de prédiction individuel et populationnel de la perte du greffon à long terme sur la base du spectre complet des informations connus comme étant d'intérêt aujourd'hui et recueillies dans la première année post-transplantation. En effet, il est désormais connu que les informations histologiques et immunologiques recueillies dans la première année post greffe sont plus informatives dans le long terme que les mêmes données le jour de la transplantation. Le développement des outils dans une cohorte de 1 016 patients est à ce jour terminé. Une validation externe française a aussi été réalisée sur une cohorte importante de 1 257 patients. Nous sommes en attente de deux cohortes de validation américaines (Mayo Clinic ; Rochester et John Hopkins ; Baltimore) afin de garantir une adéquation internationale des outils proposés. Ce travail est donc en cours de finalisation.
Pour illustration et information voici le nomogramme (Figure 1) et le score (Figure 2) proposés ainsi que sa validation externe française (Figure 3) :

Figure 1: Nomogramme pour prédire la probabilité individuelle de perte du greffon à 3, 5 et 7 ans

Les points sont assignés pour chaque facteur de risque du modèle multivarié utilisé en traçant une ligne verticale vers la ligne correspondant aux points. Le total des points pour les 5 facteurs est ensuite tracé sur la ligne correspondant à la somme totale des points. Une ligne verticale vers le bas est alors tracée pour déduire les prédictions correspondantes aux temps d'intérêt ainsi que la médiane de survie du patient.

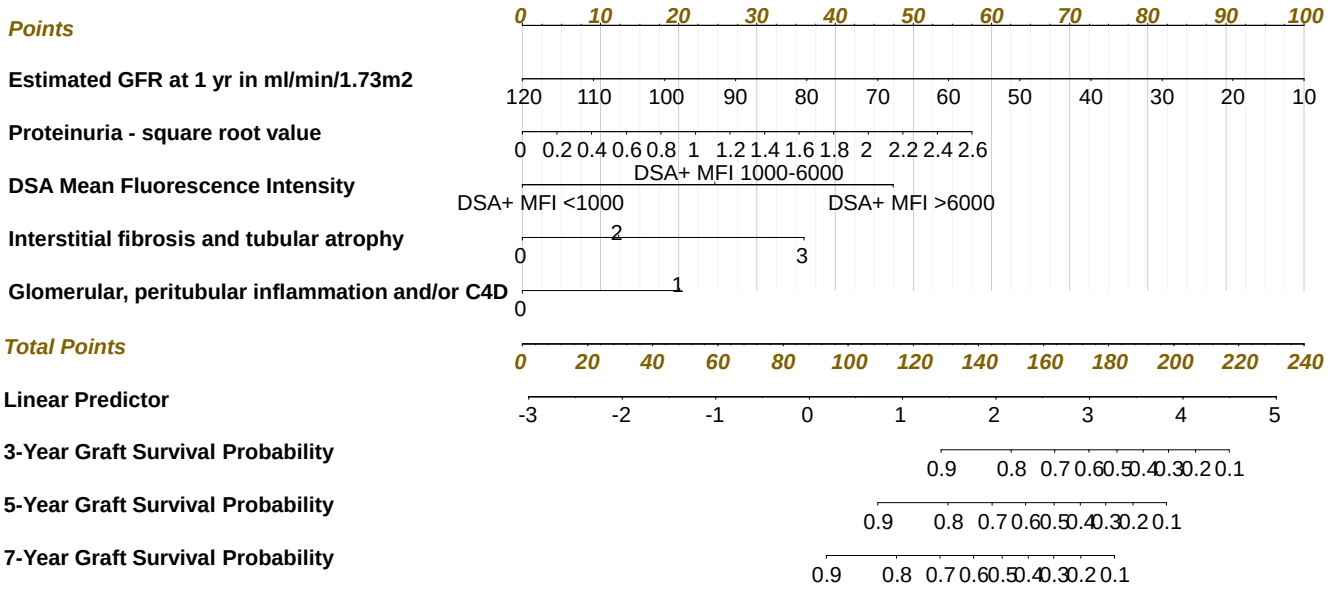
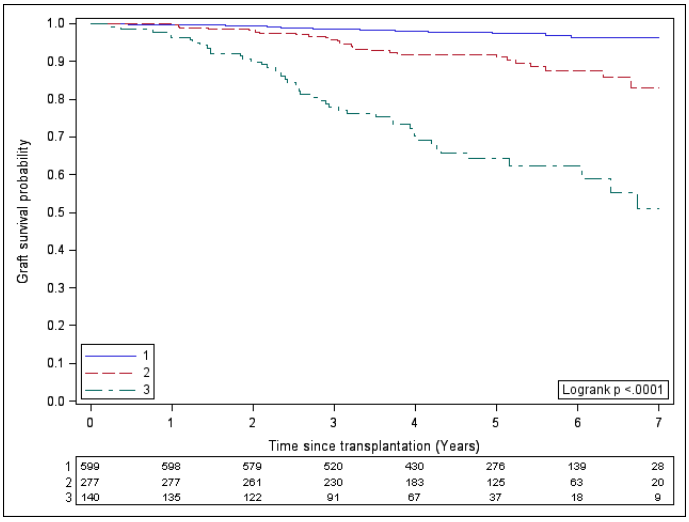
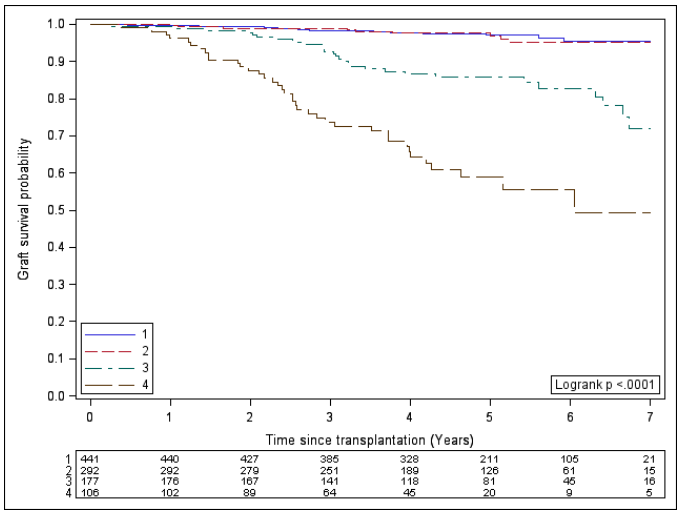


Figure 2: Courbes de Kaplan-Meier pour la survie du greffon pour le score construit avec les points du nomogramme en catégorisant selon la méthode de Cox les patients en (A) 3 groupes (B) 4 groupes et (C) 5 groupes.

A



B



C

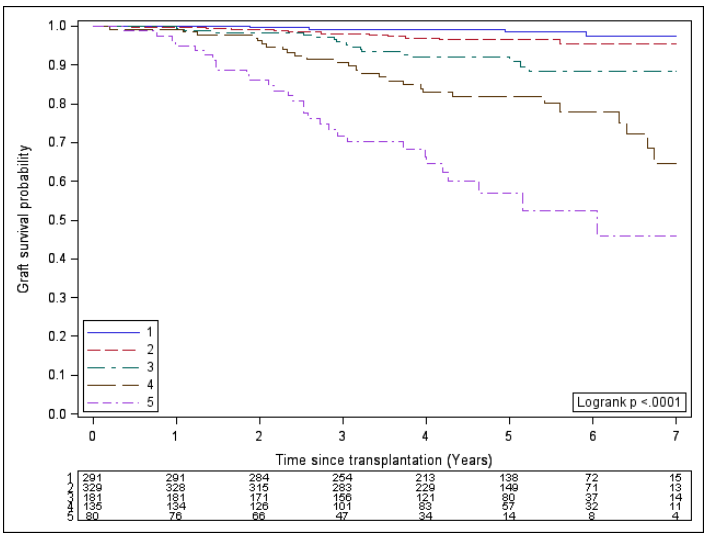
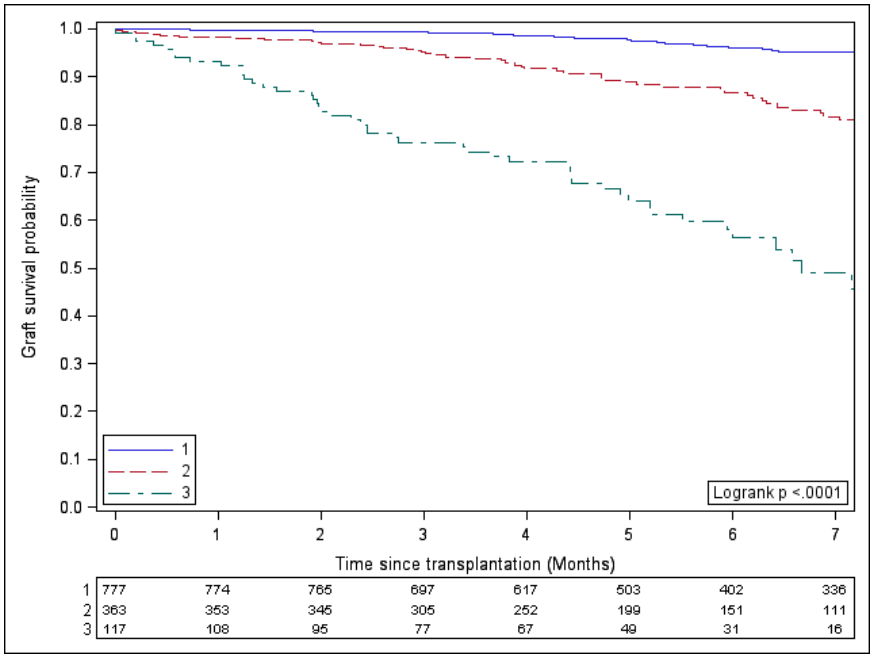
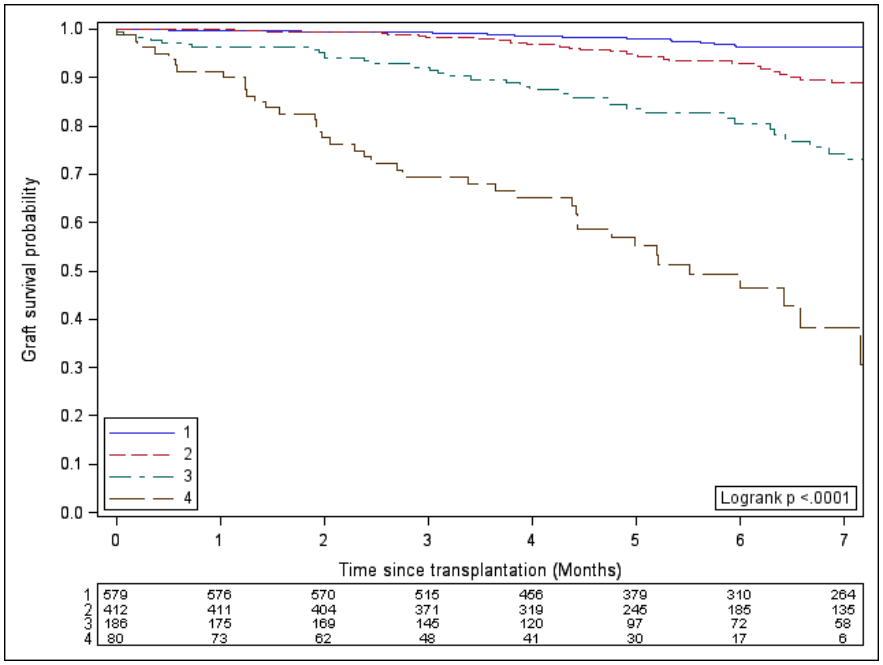


Figure 3: Courbes de Kaplan-Meier pour la survie du greffon en fonction des groupes identifiés dans la cohorte de développement dans la validation externe en catégorisant les patients en (A) 3 groupes (B) 4 groupes et (C) 5 groupes.

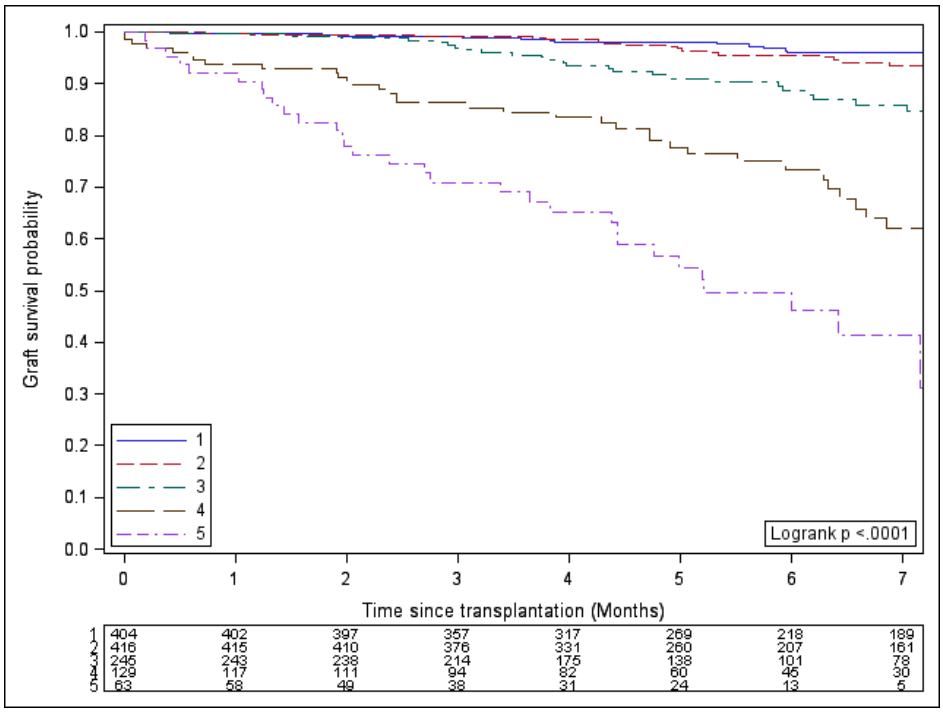
A



B



C



VI.3Score pronostique le jour de la transplantation rénale de la survie du greffon à long terme

Dans cette thèse l'étude du pronostic à long terme des greffons issus de donneurs dits à critères élargis a montré qu'il n'est pas opportun de cumuler les risques, c'est-à-dire de proposer les reins ECD à des receveurs possédant des DSA circulants le jour de la transplantation puisque le résultat en termes de survie du greffon n'est pas acceptable comparé aux greffons ECD transplantés chez des patients DSA.

Afin de préciser les résultats proposés, il serait très intéressant d'utiliser les cohortes disponibles de patients pour développer et valider des outils pronostiques prédictifs de la survie du greffon mais cette fois-ci avec les caractéristiques disponibles le jour de la greffe. Nous pouvons tout à fait envisager qu'un tel outil puisse aider et assister les médecins pour améliorer l'allocation des reins et ainsi optimiser les années de vie fonctionnelles des greffons transplantés.

VI.4Travaux en cours en oncologie digestives

Des travaux de recherche pronostique comme proposés dans ce manuscrit sont en cours pour :

- la survie globale au début de première ligne de chimiothérapie dans le cancer pancréatique avancé (collaboration A Vienot & C Borg)
- la survie globale au début de la seconde ligne de chimiothérapie pour le cholangiocarcinome (collaboration C Neuzillet & A Lièvre)
- la survie globale au début de la première ligne de chimiothérapie pour le cancer colorectal en situation adjudante (collaboration A De Grammont & T Andre)

VII Références de la thèse

Akl MR, Nagpal P, Ayoub NM, Prabhu SA, Gliksman M, Tai B, Hatipoglu A, Goy A, Suh KS (2015) Molecular and clinical profiles of syndecan-1 in solid and hematological cancer for prognosis and precision medicine. *Oncotarget* **6**: 28693–28715, doi:10.18632/oncotarget.4981.

Altman DG, Lausen B, Sauerbrei W, Schumacher M (1994) Dangers of using 'optimal' cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst* **86**: 829–835.

Altman DG, Royston P (2006) The cost of dichotomising continuous variables. *BMJ* **332**: 1080, doi:10.1136/bmj.332.7549.1080.

Anderberg MR (1973) Cluster analysis for applications / Michael R. Anderberg (New York: Academic Press).

André T, de Gramont A, Vernerey D, Chibaudel B, Bonnetain F, Tijeras-Raballand A, Scriva A, Hickish T, Tabernero J, Van Laethem JL, Banzi M, Maartense E, Shmueli E, Carlsson GU, Scheithauer W, Papamichael D, Möehler M, Landolfi S, Demetter P, Colote S, Tournigand C, Louvet C, Duval A, Fléjou J-F, de Gramont A (2015) Adjuvant Fluorouracil, Leucovorin, and Oxaliplatin in Stage II to III Colon Cancer: Updated 10-Year Survival and Outcomes According to BRAF Mutation and Mismatch Repair Status of the MOSAIC Study. *J Clin Oncol Off J Am Soc Clin Oncol* **33**: 4176–4187, doi:10.1200/JCO.2015.63.4238.

Baccini A (2010) Statistique Descriptive Multidimensionnelle.

Balachandran VP, Gonen M, Smith JJ, DeMatteo RP (2015) Nomograms in oncology: more than meets the eye. *Lancet Oncol* **16**: e173-180, doi:10.1016/S1470-2045(14)71116-7.

Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF (1996) Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* **276**: 637–639.

Bellera CA, Penel N, Ouali M, Bonvalot S, Casali PG, Nielsen OS, Delannes M, Litière S, Bonnetain F, Dabakuyo TS, Benjamin RS, Blay J-Y, Bui BN, Collin F, Delaney TF, Duffaud F, Filleron T, Fiore M, Gelderblom H, George S,

Grimer R, Grosclaude P, Gronchi A, Haas R, Hohenberger P, Issels R, Italiano A, Jooste V, Krarup-Hansen A, Le Péchoux C, Mussi C, Oberlin O, Patel S, Piperno-Neumann S, Raut C, Ray-Coquard I, Rutkowski P, Schuetze S, Sleijfer S, Stoeckle E, Van Glabbeke M, Woll P, Gourgou-Bourgade S, Mathoulin-Pélissier S, Definition for the Assessment of Time-to-event Endpoints in Cancer Trials Initiative (2015) Guidelines for time-to-event end point definitions in sarcomas and gastrointestinal stromal tumors (GIST) trials: results of the DATECAN initiative (Definition for the Assessment of Time-to-event Endpoints in CANcer trials)†. *Ann Oncol Off J Eur Soc Med Oncol ESMO* **26**: 865–872, doi:10.1093/annonc/mdu360.

Bonnetain F, Bonsing B, Conroy T, Dousseau A, Glimelius B, Haustermans K, Lacaine F, Van Laethem JL, Aparicio T, Aust D, Bassi C, Berger V, Chamorey E, Chibaudel B, Dahan L, De Gramont A, Delpero JR, Dervenis C, Ducreux M, Gal J, Gerber E, Ghaneh P, Hammel P, Hendlisz A, Jooste V, Labianca R, Latouche A, Lutz M, Macarulla T, Malka D, Mauer M, Mitry E, Neoptolemos J, Pessaux P, Sauvanet A, Tabernero J, Taieb J, van Tienhoven G, Gourgou-Bourgade S, Bellera C, Mathoulin-Pélissier S, Collette L (2014) Guidelines for time-to-event end-point definitions in trials for pancreatic cancer. Results of the DATECAN initiative (Definition for the Assessment of Time-to-event End-points in CANcer trials). *Eur J Cancer Oxf Engl 1990* **50**: 2983–2993, doi:10.1016/j.ejca.2014.07.011.

Bouyer J (2012) Régression logistique - Modélisation des variables quantitatives.

Cancer Statistics Review, 1975-2013 - SEER Statistics http://seer.cancer.gov/csr/1975_2013/ (accessed: 23/08/2016).

Chessel D, Thioulouse J, Dufour AB (2014) Introduction à la classification hiérarchique.

Chibaudel B, Bonnetain F, Tournigand C, Bengrine-Lefevre L, Teixeira L, Artru P, Desramé J, Larsen AK, André T, Louvet C, de Gramont A (2011) Simplified prognostic model in patients with oxaliplatin-based or irinotecan-based first-line chemotherapy for metastatic colorectal cancer: a GERCOR study. *The Oncologist* **16**: 1228–1238, doi:10.1634/theoncologist.2011-0039.

Collins FS, Varmus H (2015) A New Initiative on Precision Medicine. *N Engl J Med* **372**: 793–795, doi:10.1056/NEJMp1500523.

Collins GS, Ogundimu EO, Altman DG (2016) Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* **35**: 214–226, doi:10.1002/sim.6787.

Collins GS, Reitsma JB, Altman DG, Moons KGM (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Cancer* **112**: 251–259, doi:10.1038/bjc.2014.639.

Connor RJ (1972) Grouping for Testing Trends in Categorical Data. *J Am Stat Assoc* **67**: 601–604, doi:10.2307/2284446.

Conroy T, Desseigne F, Ychou M, Bouché O, Guimbaud R, Bécouarn Y, Adenis A, Raoul J-L, Gourgou-Bourgade S, de la Fouchardière C, Bennouna J, Bachet J-B, Khemissa-Akouz F, Péré-Vergé D, Delbaldo C, Assenat E, Chauffert B, Michel P, Montoto-Grillot C, Ducreux M, Groupe Tumeurs Digestives of Unicancer, PRODIGE Intergroup (2011) FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. *N Engl J Med* **364**: 1817–1825, doi:10.1056/NEJMoa1011923.

Cook NR, Ridker PM (2009) Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann Intern Med* **150**: 795–802.

Cormack RM (1971) A review of classification. *Journal of the Royal Statistical Society, A* 321–367.

Cox DR (1957) Note on Grouping. *J Am Stat Assoc* **52**: 543–547, doi:10.1080/01621459.1957.10501411.

Cox DR (1972) Regression Models and Life-Tables. *J R Stat Soc Ser B Methodol* **34**: 187–220.

D'Agostino RB, Grundy S, Sullivan LM, Wilson P, CHD Risk Prediction Group (2001) Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA* **286**: 180–187.

D'Agostino RB, Nam B-H (2003) Evaluation of the Performance of Survival Analysis Models: Discrimination and Calibration Measures. In Handbook of Statistics, (Elsevier), pp. 1–25.

Day RM, Hao X, Ilyas M, Daszak P, Talbot IC, Forbes A (1999) Changes in the expression of syndecan-1 in the colorectal adenoma-carcinoma sequence. *Virchows Arch Int J Pathol* **434**: 121–125.

Desot E, de Mestier L, Volet J, Delmas C, Garcia B, Geoffroy P, Abdelli N, Baule M, Dubroeucq O, Marquis E, Bouché O (2013) Prognostic factors in patients with non resectable metastatic colorectal cancer in the era of targeted biotherapies: relevance of Köhne's risk classification. *Dig Liver Dis Off J Ital*

Soc Gastroenterol Ital Assoc Study Liver **45**: 330–335, doi:10.1016/j.dld.2012.10.016.

Doward LC, McKenna SP (2004) Defining patient-reported outcomes. *Value Health J Int Soc Pharmacoeconomics Outcomes Res* **7 Suppl 1**: S4-8, doi:10.1111/j.1524-4733.2004.7s102.x.

Dray S, Dufour A. (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 1–20.

Ducreux M, Cuhna AS, Caramella C, Hollebecque A, Burtin P, Goéré D, Seufferlein T, Haustermans K, Van Laethem JL, Conroy T, Arnold D, ESMO Guidelines Committee (2015) Cancer of the pancreas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol Off J Eur Soc Med Oncol ESMO* **26 Suppl 5**: v56-68, doi:10.1093/annonc/mdv295.

Dynamic Prediction in Clinical Survival Analysis (2011) https://www.crcpress.com/Dynamic-Prediction-in-Clinical-Survival-Analysis/van-Houwelingen-Putter/p/book/9781439835333 (accessed: 29/09/2016).

Efron B (1979) Bootstrap Methods: Another Look at the Jackknife. *Ann Stat* **7**: 1–26.

Evans RW, Manninen DL, Garrison LP, Hart LG, Blagg CR, Gutman RA, Hull AR, Lowrie EG (1985) The quality of life of patients with end-stage renal disease. *N Engl J Med* **312**: 553–559, doi:10.1056/NEJM198502283120905.

Fakih MG (2015) Metastatic colorectal cancer: current state and future directions. *J Clin Oncol Off J Am Soc Clin Oncol* **33**: 1809–1824, doi:10.1200/JCO.2014.59.7633.

Falcone A, Ricci S, Brunetti I, Pfanner E, Allegrini G, Barbara C, Crinò L, Benedetti G, Evangelista W, Fanchini L, Cortesi E, Picone V, Vitello S, Chiara S, Granetto C, Porcile G, Fioretto L, Orlandini C, Andreuccetti M, Masi G, Gruppo Oncologico Nord Ovest (2007) Phase III trial of infusional fluorouracil, leucovorin, oxaliplatin, and irinotecan (FOLFOXIRI) compared with infusional fluorouracil, leucovorin, and irinotecan (FOLFIRI) as first-line treatment for metastatic colorectal cancer: the Gruppo Oncologico Nord Ovest. *J Clin Oncol Off J Am Soc Clin Oncol* **25**: 1670–1676, doi:10.1200/JCO.2006.09.0928.

Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* **136**: E359-386, doi:10.1002/ijc.29210.

Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, Forman D, Bray F (2013) Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer Oxf Engl 1990* **49**: 1374–1403, doi:10.1016/j.ejca.2012.12.027.

Fitzgerald M, Saville BR, Lewis RJ (2015) Decision curve analysis. *JAMA* **313**: 409–410, doi:10.1001/jama.2015.37.

Foucher Y, Combescure C, Ashton-Chess J, Giral M (2012) Prognostic markers: data misinterpretation often leads to overoptimistic conclusions. *Am J Transplant Off J Am Soc Transplant Am Soc Transpl Surg* **12**: 1060–1061, doi:10.1111/j.1600-6143.2011.03889.x.

Garcia GG, Harden P, Chapman J, World Kidney Day Steering Committee 2012 (2012) The global role of kidney transplantation. *Lancet Lond Engl* **379**: e36-38, doi:10.1016/S0140-6736(12)60202-5.

Gehan EA (1965) A GENERALIZED WILCOXON TEST FOR COMPARING ARBITRARILY SINGLY-CENSORED SAMPLES. *Biometrika* **52**: 203–223.

Gehlenborg N, Wong B (2012) Points of view: Heat maps. *Nat Methods* **9**: 213–213, doi:10.1038/nmeth.1902.

Gordon A (1999) Classification, 2nd Edition (Chapman & Hall/CRC Monographs on Statistics & Applied Probability).

Gourgou-Bourgade S, Cameron D, Poortmans P, Asselain B, Azria D, Cardoso F, A'Hern R, Bliss J, Bogaerts J, Bonnefoi H, Brain E, Cardoso MJ, Chibaudel B, Coleman R, Cufer T, Dal Lago L, Dalenc F, De Azambuja E, Debled M, Delaloge S, Filleron T, Gligorov J, Gutowski M, Jacot W, Kirkove C, MacGrogan G, Michiels S, Negreiros I, Offersen BV, Penault Llorca F, Pruneri G, Roche H, Russell NS, Schmitt F, Servent V, Thürlimann B, Untch M, van der Hage JA, van Tienhoven G, Wildiers H, Yarnold J, Bonnetain F, Mathoulin-Pélissier S, Bellera C, Dabakuyo-Yonli TS, Definition for the Assessment of Time-to-event Endpoints in Cancer Trials Initiative (2015) Guidelines for time-to-event end point definitions in breast cancer trials: results of the DATECAN initiative (Definition for the Assessment of Time-to-event Endpoints in CANcer trials)†. *Ann Oncol Off J Eur Soc Med Oncol ESMO* **26**: 873–879, doi:10.1093/annonc/mdv106.

Grambsch PM, Therneau TM (1994) Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81**: 515–526, doi:10.1093/biomet/81.3.515.

Greenland S (1995) Problems in the average-risk interpretation of categorical dose-response analyses. *Epidemiol Camb Mass* **6**: 563–565.

Grimes DA (2008) The Nomogram Epidemic: Resurgence of a Medical Relic. *Ann Intern Med* **149**: 273–275, doi:10.7326/0003-4819-149-4-200808190-00010.

Harrell F (2001) Regression Modeling Strategies. *Springer Series in Statistics*.

Harrell FE, Lee KL, Mark DB (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* **15**: 361–387, doi:10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4.

Hashimoto Y, Skacel M, Adams JC (2008) Association of loss of epithelial syndecan-1 with stage and local metastasis of colorectal adenocarcinomas: an immunohistochemical study of clinically annotated tumors. *BMC Cancer* **8**: 185, doi:10.1186/1471-2407-8-185.

Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, Briggs A, Udumyan R, Moons KGM, Steyerberg EW, Roberts I, Schroter S, Altman DG, Riley RD (2013) Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *BMJ* **346**: e5595, doi:10.1136/bmj.e5595.

Hill C (2009) Analyse statistique des données de survie. *LAVOISIER MÉDECINE SCIENCES*.

Hippocrate (0460 - 0377 av. J.-C.) Pronostic (Les Belles Lettres).

Horton R (2015) Offline: What is medicine's 5 sigma? *The Lancet* **385**: 1380, doi:10.1016/S0140-6736(15)60696-1.

Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S (1997) A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* **16**: 965–980.

Hosmer DW, Lemesbow S (1980) Goodness of fit tests for the multiple logistic regression model. *Commun Stat - Theory Methods* **9**: 1043–1069, doi:10.1080/03610928008827941.

Hothorn T, Lausen B (2003) On the exact distribution of maximally selected rank statistics. *Comput Stat Data Anal* **43**: 121–137, doi:10.1016/S0167-9473(02)00225-6.

Huang X, Yan F, Ning J, Feng Z, Choi S, Cortes J (2016) A two-stage approach for dynamic prediction of time-to-event distributions. *Stat Med* **35**: 2167–2182, doi:10.1002/sim.6860.

Ioannidis JPA (2008) Why most discovered true associations are inflated. *Epidemiol Camb Mass* **19**: 640–648, doi:10.1097/EDE.0b013e31818131e7.

Jameson JL, Longo DL (2015) Precision medicine--personalized, problematic, and promising. *N Engl J Med* **372**: 2229–2234, doi:10.1056/NEJMsb1503104.

Jary M, Vernerey D, Lecomte T, Dobi E, Ghiringhelli F, Monnien F, Godet Y, Kim S, Bouche O, Fratte S, Goncalves A, Leger J, Queiroz L, Adotevi O, Bonnetain F, Borg C (2015) Prognostic Value of Angiopoietin-2 for death risk stratification in Patients with Metastatic Colorectal Carcinoma. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* doi:10.1158/1055-9965.EPI-14-1059.

Jr DWH, Lemeshow S (2004) Applied Logistic Regression (John Wiley & Sons).

Kaplan EL, Meier P (1958) Nonparametric Estimation from Incomplete Observations. *J Am Stat Assoc* **53**: 457–481, doi:10.2307/2281868.

Kattan MW (2003) Judging New Markers by Their Ability to Improve Predictive Accuracy. *J Natl Cancer Inst* **95**: 634–635, doi:10.1093/jnci/95.9.634.

Kay R (1984) Goodness of fit methods for the proportional hazards regression model: a review. *Rev Dépidémiologie Santé Publique* **32**: 185–198.

Köhne CH, Cunningham D, Di Costanzo F, Glimelius B, Blijham G, Aranda E, Scheithauer W, Rougier P, Palmer M, Wils J, Baron B, Pignatti F, Schöffski P, Micheel S, Hecker H (2002) Clinical determinants of survival in patients with 5-fluorouracil-based treatment for metastatic colorectal cancer: results of a multivariate analysis of 3825 patients. *Ann Oncol Off J Eur Soc Med Oncol* **13**: 308–317.

Kramar A, Negrier S, Sylvester R, Joniau S, Mulders P, Powles T, Bex A, Bonnetain F, Bossi A, Bracarda S, Bukowski R, Catto J, Choueiri TK, Crabb S, Eisen T, El Demery M, Fitzpatrick J, Flamand V, Goebell PJ, Gravis G, Houédé N, Jacqmin D, Kaplan R, Malavaud B, Massard C, Melichar B, Mourey L, Nathan P, Pasquier D, Porta C, Pouessel D, Quinn D, Ravaud A, Rolland F, Schmidinger M, Tombal B, Tosi D, Vauleon E, Volpe A, Wolter P, Escudier B, Filleron T, DATECAN Renal Cancer group (2015) Guidelines for the definition of time-to-event end points in renal cell cancer clinical trials: results of the DATECAN project†. *Ann Oncol Off J Eur Soc Med Oncol ESMO* **26**: 2392–2398, doi:10.1093/annonc/mdv380.

Lawless JF (1982) Statistical models and methods for lifetime data (New York: Wiley).

Leening MJG, Vedder MM, Witteman JCM, Pencina MJ, Steyerberg EW (2014) Net Reclassification Improvement: Computation, Interpretation, and ControversiesA Literature Review and Clinician's Guide. *Ann Intern Med* **160**: 122–131, doi:10.7326/M13-1522.

Lefaucheur C, Loupy A, Vernerey D, Duong-Van-Huyen J-P, Suberbielle C, Anglicheau D, Vérine J, Beuscart T, Nochy D, Bruneval P, Charron D, Delahousse M, Empana J-P, Hill GS, Glotz D, Legendre C, Jouven X (2013) Antibody-mediated vascular rejection of kidney allografts: a population-based study. *Lancet* **381**: 313–319, doi:10.1016/S0140-6736(12)61265-3.

Leppä S, Vleminckx K, Van Roy F, Jalkanen M (1996) Syndecan-1 expression in mammary epithelial tumor cells is E-cadherin-dependent. *J Cell Sci* **109 (Pt 6)**: 1393–1403.

Loupakis F, Cremolini C, Masi G, Lonardi S, Zagonel V, Salvatore L, Cortesi E, Tomasello G, Ronzoni M, Spadi R, Zaniboni A, Tonini G, Buonadonna A, Amoroso D, Chiara S, Carlomagno C, Boni C, Allegrini G, Boni L, Falcone A (2014a) Initial therapy with FOLFOXIRI and bevacizumab for metastatic colorectal cancer. *N Engl J Med* **371**: 1609–1618, doi:10.1056/NEJMoa1403108.

Loupakis F, Cremolini C, Salvatore L, Masi G, Sensi E, Schirripa M, Michelucci A, Pfanner E, Brunetti I, Lupi C, Antoniotti C, Bergamo F, Lonardi S, Zagonel V, Simi P, Fontanini G, Falcone A (2014b) FOLFOXIRI plus bevacizumab as first-line treatment in BRAF mutant metastatic colorectal cancer. *Eur J Cancer Oxf Engl 1990* **50**: 57–63, doi:10.1016/j.ejca.2013.08.024.

Loupy A, Lefaucheur C, Vernerey D, Chang J, Hidalgo LG, Beuscart T, Verine J, Aubert O, Dubleumortier S, Duong van Huyen J-P, Jouven X, Glotz D, Legendre C, Halloran PF (2014) Molecular Microscope Strategy to Improve Risk Stratification in Early Antibody-Mediated Kidney Allograft Rejection. *J Am Soc Nephrol JASN* doi:10.1681/ASN.2013111149.

Loupy A, Lefaucheur C, Vernerey D, Prugger C, Duong van Huyen J-P, Mooney N, Suberbielle C, Frémeaux-Bacchi V, Méjean A, Desgrandchamps F, Anglicheau D, Nochy D, Charron D, Empana J-P, Delahousse M, Legendre C, Glotz D, Hill GS, Zeevi A, Jouven X (2013) Complement-binding anti-HLA antibodies and kidney-allograft survival. *N Engl J Med* **369**: 1215–1226, doi:10.1056/NEJMoa1302506.

Maeda T, Desouky J, Friedl A (2006) Syndecan-1 expression by stromal fibroblasts promotes breast carcinoma growth in vivo and stimulates tumor angiogenesis. *Oncogene* **25**: 1408–1412, doi:10.1038/sj.onc.1209168.

Malvezzi M, Bertuccio P, Levi F, La Vecchia C, Negri E (2014) European cancer mortality predictions for the year 2014. *Ann Oncol Off J Eur Soc Med Oncol ESMO* **25**: 1650–1656, doi:10.1093/annonc/mdu138.

Mantel N (1966) Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep* **50**: 163–170.

Marrugat J, Subirana I, Comín E, Cabezas C, Vila J, Elosua R, Nam B-H, Ramos R, Sala J, Solanas P, Cordón F, Gené-Badia J, D'Agostino RB, VERIFICA Investigators (2007) Validity of an adaptation of the Framingham cardiovascular risk function: the VERIFICA Study. *J Epidemiol Community Health* **61**: 40–47, doi:10.1136/jech.2005.038505.

Mauguen A, Rachet B, Mathoulin-Pélissier S, MacGrogan G, Laurent A, Rondeau V (2013) Dynamic prediction of risk of death using history of cancer recurrences in joint frailty models. *Stat Med* **32**: 5366–5380, doi:10.1002/sim.5980.

Mitselou A, Skoufi U, Tsimogiannis KE, Briasoulis E, Vougiouklakis T, Arvanitis D, Ioachim E (2012) Association of syndecan-1 with angiogenesis-related markers, extracellular matrix components, and clinicopathological features in colorectal carcinoma. *Anticancer Res* **32**: 3977–3985.

Mogensen UB, Ishwaran H, Gerds TA (2012) Evaluating Random Forests for Survival Analysis using Prediction Error Curves. *J Stat Softw* **50**: 1–23.

Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG (2010) CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* **340**: c869, doi:10.1136/bmj.c869.

Moher D, Schulz KF, Altman DG (2001) The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet Lond Engl* **357**: 1191–1194.

Molica S, Mauro FR, Callea V, Giannarelli D, Lauria F, Rotoli B, Cortelezzi A, Liso V, Foà R (2010) The utility of a prognostic index for predicting time to first treatment in early chronic lymphocytic leukemia: the GIMEMA experience. *Haematologica* **95**: 464–469, doi:10.3324/haematol.2009.011767.

Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS (2015) Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* **162**: W1-73, doi:10.7326/M14-0698.

Moser BK, Coombs LP (2004) Odds ratios for a continuous outcome variable without dichotomizing. *Stat Med* **23**: 1843–1860, doi:10.1002/sim.1776.

Nankivell BJ, Alexander SI (2010) Rejection of the kidney allograft. *N Engl J Med* **363**: 1451–1462, doi:10.1056/NEJMra0902927.

Neuzillet C, Tijeras-Raballand A, Bourget P, Cros J, Couvelard A, Sauvanet A, Vullierme M-P, Tournigand C, Hammel P (2015) State of the art and future directions of pancreatic ductal adenocarcinoma therapy. *Pharmacol Ther* **155**: 80–104, doi:10.1016/j.pharmthera.2015.08.006.

ONT-WHO Organ Donation and Transplantation Activities 2013. Global Observatory on Donation and Transplantation (GODT). http://www.transplant-observatory.org/ (accessed: 15/08/2016).

Pan JJ, Ng WT, Zong JF, Lee SWM, Choi HCW, Chan LLK, Lin SJ, Guo QJ, Sze HCK, Chen YB, Xiao YP, Kan WK, O'Sullivan B, Xu W, Le QT, Glastonbury CM, Colevas AD, Weber RS, Lydiatt W, Shah JP, Lee AWM (2016) Prognostic nomogram for refining the prognostication of the proposed 8th edition of the AJCC/UICC staging system for nasopharyngeal cancer in the era of intensity-modulated radiotherapy. *Cancer* doi:10.1002/cncr.30198.

Pencina MJ, D'Agostino RB, Sr (2015) Evaluating discrimination of risk prediction models: The c statistic. *JAMA* **314**: 1063–1064, doi:10.1001/jama.2015.11082.

Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS (2008) Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* **27**: 157-172-212, doi:10.1002/sim.2929.

Pencina MJ, D'Agostino RB Sr, Steyerberg EW (2011) Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* **30**: 11–21, doi:10.1002/sim.4085.

Pencina MJ, Steyerberg EW, D'Agostino RB (2016) Net reclassification index at event rate: properties and relationships. *Stat Med* doi:10.1002/sim.7041.

Peto R, Peto J (1972) Asymptotically Efficient Rank Invariant Test Procedures. *J R Stat Soc Ser Gen* **135**: 185–207, doi:10.2307/2344317.

Popovici V, Budinska E, Tejpar S, Weinrich S, Estrella H, Hodgson G, Van Cutsem E, Xie T, Bosman FT, Roth AD, Delorenzi M (2012) Identification of a poor-prognosis BRAF-mutant-like population of patients with colon cancer. *J Clin Oncol Off J Am Soc Clin Oncol* **30**: 1288–1295, doi:10.1200/JCO.2011.39.5814.

Prentice RL (1978) Linear rank tests with right censored data. *Biometrika* **65**: 167–179, doi:10.1093/biomet/65.1.167.

Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM (2014) Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* **74**: 2913–2921, doi:10.1158/0008-5472.CAN-14-0155.

Royston P, Altman DG (1994) Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. *J R Stat Soc Ser C Appl Stat* **43**: 429–467, doi:10.2307/2986270.

Royston P, Sauerbrei W (2008) Multivariable Model - Building: A Pragmatic Approach to Regression Anaylsis based on Fractional Polynomials for Modelling Continuous Variables.

Rubin DB (1987) Introduction. In Multiple Imputation for Nonresponse in Surveys, (John Wiley & Sons, Inc.), pp. 1–26.

Ryan DP, Hong TS, Bardeesy N (2014) Pancreatic adenocarcinoma. *N Engl J Med* **371**: 1039–1049, doi:10.1056/NEJMra1404198.

Schemper M (2003) Predictive accuracy and explained variation. *Stat Med* **22**: 2299–2308, doi:10.1002/sim.1486.

Schirripa M, Bergamo F, Cremolini C, Casagrande M, Lonardi S, Aprile G, Yang D, Marmorino F, Pasquini G, Sensi E, Lupi C, De Maglio G, Borrelli N, Pizzolitto S, Fasola G, Bertorelle R, Rugge M, Fontanini G, Zagonel V, Loupakis F, Falcone A (2015) BRAF and RAS mutations as prognostic factors in metastatic colorectal cancer patients undergoing liver resection. *Br J Cancer* **112**: 1921–1928, doi:10.1038/bjc.2015.142.

Seidel C, Sundan A, Hjorth M, Turesson I, Dahl IM, Abildgaard N, Waage A, Borset M (2000) Serum syndecan-1: a new independent prognostic marker in multiple myeloma. *Blood* **95**: 388–392.

Sellarés J, de Freitas DG, Mengel M, Reeve J, Einecke G, Sis B, Hidalgo LG, Famulski K, Matas A, Halloran PF (2012) Understanding the causes of kidney transplant failure: the dominant role of antibody-mediated rejection and nonadherence. *Am J Transplant Off J Am Soc Transplant Am Soc Transpl Surg* **12**: 388–399, doi:10.1111/j.1600-6143.2011.03840.x.

Siegel RL, Miller KD, Jemal A (2016) Cancer statistics, 2016. *CA Cancer J Clin* **66**: 7–30, doi:10.3322/caac.21332.

Sorbye H, Dragomir A, Sundström M, Pfeiffer P, Thunberg U, Bergfors M, Aasebø K, Eide GE, Ponten F, Qvortrup C, Glimelius B (2015) High BRAF Mutation Frequency and Marked Survival Differences in Subgroups According to KRAS/BRAF Mutation Status and Tumor Tissue Availability in a Prospective Population-Based Metastatic Colorectal Cancer Cohort. *PloS One* **10**: e0131046, doi:10.1371/journal.pone.0131046.

Steyerberg EW, Eijkemans MJC, Habbema JDF (2001) Application of Shrinkage Techniques in Logistic Regression Analysis: A Case Study. *Stat Neerlandica* **55**: 76–88, doi:10.1111/1467-9574.00157.

Strobl C, Malley J, Tutz G (2009) An Introduction to Recursive Partitioning: Rationale, Application and Characteristics of Classification and Regression Trees, Bagging and Random Forests. *Psychol Methods* **14**: 323–348, doi:10.1037/a0016973.

Sun G-W, Shook TL, Kay GL (1996) Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol* **49**: 907–916, doi:10.1016/0895-4356(96)00025-X.

Taieb J, Zaanan A, Le Malicot K, Julié C, Blons H, Mineur L, Bennouna J, Tabernero J, Mini E, Folprecht G, Van Laethem JL, Lepage C, Emile J-F, Laurent-Puig P (2016) Prognostic Effect of BRAF and KRAS Mutations in Patients With Stage III Colon Cancer Treated With Leucovorin, Fluorouracil, and Oxaliplatin With or Without Cetuximab: A Post Hoc Analysis of the PETACC-8 Trial. *JAMA Oncol* 1–11, doi:10.1001/jamaoncol.2015.5225.

Taylor JMG, Yu M (2002) Bias and Efficiency Loss Due to Categorizing an Explanatory Variable. *J Multivar Anal* **83**: 248–263, doi:10.1006/jmva.2001.2045.

Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ (2011) On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* **30**: 1105–1117, doi:10.1002/sim.4154.

Van Cutsem E, Cervantes A, Adam R, Sobrero A, Van Krieken JH, Aderka D, Aranda Aguilar E, Bardelli A, Benson A, Bodoky G, Ciardiello F, D'Hoore A, Diaz-Rubio E, Douillard J-Y, Ducreux M, Falcone A, Grothey A, Gruenberger T, Haustermans K, Heinemann V, Hoff P, Köhne C-H, Labianca R, Laurent-Puig P, Ma B, Maughan T, Muro K, Normanno N, Österlund P, Oyen WJG, Papamichael D, Pentheroudakis G, Pfeiffer P, Price TJ, Punt C, Ricke J, Roth A, Salazar R, Scheithauer W, Schmoll HJ, Tabernero J, Taïeb J, Tejpar S, Wasan H, Yoshino T, Zaanan A, Arnold D (2016) ESMO consensus guidelines for the management of patients with metastatic colorectal cancer.

Ann Oncol Off J Eur Soc Med Oncol ESMO **27**: 1386–1422, doi:10.1093/annonc/mdw235.

Venables WN, Ripley BD (2003) Modern Applied Statistics with S (New York: Springer).

Venderbosch S, Nagtegaal ID, Maughan TS, Smith CG, Cheadle JP, Fisher D, Kaplan R, Quirke P, Seymour MT, Richman SD, Meijer GA, Ylstra B, Heideman DAM, de Haan AFJ, Punt CJA, Koopman M (2014) Mismatch repair status and BRAF mutation status in metastatic colorectal cancer patients: a pooled analysis of the CAIRO, CAIRO2, COIN, and FOCUS studies. *Clin Cancer Res Off J Am Assoc Cancer Res* **20**: 5322–5330, doi:10.1158/1078-0432.CCR-14-0332.

Vickers AJ, Elkin EB (2006) Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak Int J Soc Med Decis Mak* **26**: 565–574, doi:10.1177/0272989X06295361.

Von Hoff DD, Ervin T, Arena FP, Chiorean EG, Infante J, Moore M, Seay T, Tjulandin SA, Ma WW, Saleh MN, Harris M, Reni M, Dowden S, Laheru D, Bahary N, Ramanathan RK, Tabernero J, Hidalgo M, Goldstein D, Van Cutsem E, Wei X, Iglesias J, Renschler MF (2013) Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *N Engl J Med* **369**: 1691–1703, doi:10.1056/NEJMoa1304369.

Wang X, Zuo D, Chen Y, Li W, Liu R, He Y, Ren L, Zhou L, Deng T, Wang X, Ying G, Ba Y (2014) Shed Syndecan-1 is involved in chemotherapy resistance via the EGFR pathway in colorectal cancer. *Br J Cancer* **111**: 1965–1976, doi:10.1038/bjc.2014.493.

Windeler J (2000) Prognosis - what does the clinician associate with this notion? *Stat Med* **19**: 425–430.

Yang N, Mosher R, Seo S, Beebe D, Friedl A (2011) Syndecan-1 in breast cancer stroma fibroblasts regulates extracellular matrix fiber organization and carcinoma cell motility. *Am J Pathol* **178**: 325–335, doi:10.1016/j.ajpath.2010.11.039.

VIII Annexes

VIII.1 Annexe 1: Publications complémentaires à la thèse servant de support pour la partie méthode :

VIII.1.1 Identification d'une nouvelle forme de rejet aigu en transplantation rénale

Lefaucheur C, Loupy A, Vernerey D, Duong-Van-Huyen J-P, Suberbielle C, Anglicheau D, Vérine J, Beuscart T, Nochy D, Bruneval P, Charron D, Delahousse M, Empana J-P, Hill GS, Glotz D, Legendre C, Jouven X
Antibody-mediated vascular rejection of kidney allografts: a population-based study. **Lancet** 381: 313–319, doi:10.1016/S0140-6736(12)61265-3. (2013)
Impact factor: 44.0

VIII.1.1.1 Contexte

En transplantation, le rejet aigu médié par anticorps (ABMR) est actuellement individualisé par la classification internationale de Banff et de plus en plus reconnu comme un évènement majeur dans l'évolution des greffes HLA-incompatibles, en étant responsable d'un nombre important de perte de greffons.
Cette étude est fondée sur l'analyse des corrélations entre la présence d'anticorps anti-HLA détectés dans le sérum, leur capacité à activer la cascade du complément, l'histologie et l'évolution clinique des greffons dans une cohorte de patients présentant un rejet aigu. Nous avons émis l'hypothèse que l'existence de phénotypes distincts de rejets aigus

cliniquement pertinents pouvait être individualisée sur la base de critères histologiques et immunologiques. Nous avons conçu cette étude comme une réévaluation critique des différents types de rejet d'allogreffe rénale en utilisant des outils modernes appliqués à une large cohorte historique, reflet de 10 ans de rejet aigu en transplantation rénale.

VIII.1.1.2 Population de l'étude

Cette étude inclut tous les patients ayant présentés un épisode de rejet aigu confirmé par histologie parmi l'ensemble des patients transplantés rénaux, entre janvier 1998 et décembre 2008, dans deux centres de transplantation parisiens, l'Hôpital Necker et l'Hôpital Saint-Louis. 2 079 transplantations rénales consécutives ont été réalisées dans les deux centres pendant cette période. Les patients ont été suivis jusqu'au 31 mars 2010. Nous avons validé nos résultats sur une cohorte externe indépendante constituée de 672 patients receveurs d'une transplantation rénale à l'Hôpital Foch, entre janvier 2004 et décembre 2010.

Les données cliniques des donneurs et des receveurs proviennent de 2 bases de données informatisées : Données Informatiques Validées en Transplantation (DIVAT) et CRISTAL de l'Agence de la Biomédecine.

Nous avons identifié l'ensemble des patients qui ont bénéficié d'une biopsie de greffon avec une indication clinique, c'est-à-dire dégradation aiguë de la fonction rénale ou protéinurie significative. Le diagnostic initial de rejet a été posé conformément à la classification internationale en vigueur au moment du diagnostic. Tous les patients inclus dans l'étude ont été analysés rétrospectivement afin de définir : 1) les paramètres histologiques et immunohistochimiques utilisés dans la classification actuelle de Banff ; 2) la présence des anticorps anti-HLA-DSA identifiés au moment de la biopsie par technique sensible.

VIII.1.1.3 Approche méthodologique

Pour redéfinir le spectre des rejets en s'affranchissant de la

classification internationale actuelle, nous avons utilisé une analyse par cluster hiérarchique (dendogramme) et une analyse en composantes principales basées sur les critères diagnostiques histologiques, le marquage de la fraction C4d du complément dans le greffon et la présence de DSA.

Pour l'étude de la survie des greffons, nous avons considéré le rejet comme temps d'origine et la perte de greffon comme événement d'intérêt. En cas de décès avec greffon fonctionnel, nous avons censuré la survie des greffons au moment du décès. La survie des greffons a été estimée par des courbes de survie selon la méthode de Kaplan-Meier. La comparaison entre les survies en fonction des différents types de rejet a été réalisée par le test du log-rank.

L'ensemble des résultats de cluster hiérarchique et de survie des greffons a été répliqué dans la cohorte de validation externe.

VIII.1.1.4 Résultats

Parmi les 2 079 receveurs d'une greffe rénale dans les 2 centres de transplantation (Hôpital Necker et Hôpital Saint-Louis), 302 (15%) ont connu un épisode de rejet aigu documenté histologiquement. La Table 1 détaille les caractéristiques des donneurs et des receveurs au moment de la transplantation chez les patients ayant subi un épisode de rejet et dans le reste de la population générale.

Table 1 : Caractéristiques des receveurs d'une transplantation rénale au moment de la greffe

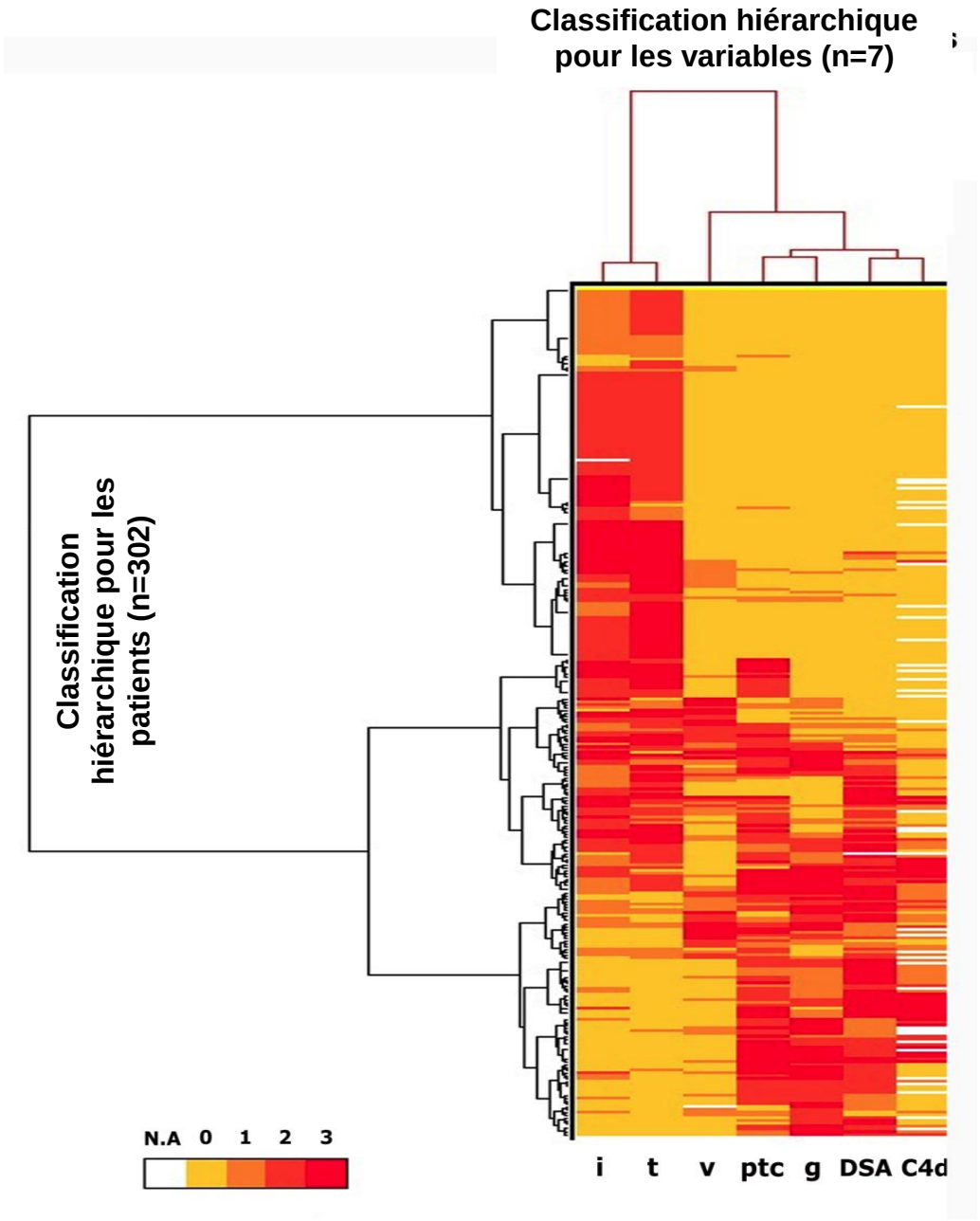
Caractéristiques	Patients sans rejet	Patients avec rejet	p *
Age du receveur — ans **	48,8 ± 13	45 ± 10	<0,001
Receveur de sexe masculin — no. (%)	1225 (69%)	148 (49%)	<0,001
Retransplantation — no. (%)	212 (27,8%)	122 (44,9%)	<0,0001
Durée début dialyse — mois **	51 ± 54	53 ± 62	NS
Age du donneur — (ans) **	51 ± 16	50 ± 12	NS
Donneur de sexe masculin — no. (%)	970 (54,6%)	169 (56%)	NS
Donneur cadavérique — no. (%)	1444 (8,2%)	239 (79%)	NS
Cause cardiovasculaire de décès du donneur —	622 (43,1%)	98 (41,0%)	NS
Ischémie froide — heures **	18,8 ± 10	17,8 ± 8	NS
Retard de prise de fonction — no. (%)	906 (50,9%)	177 (58,6%)	0,02
Etiologie de la néphropathie			
Diabètes — no. (%)	211 (12%)	18 (6%)	0,003
Vasculaire — no. (%)	164 (9,2%)	15 (5%)	0,02
Glomérulopathie — no. (%)	433 (24,4%)	125 (41,4%)	<0,001
Congénitale — no. (%)	291 (16,3%)	88 (29,1%)	<0,001
Autres — no. (%)	70 (3,9%)	14 (4,6%)	NS
Interstitielle — no. (%)	171 (9,6%)	35 (11,6%)	NS
Indéterminée — no. (%)	437 (24,6%)	7 (2,3%)	<0,0001
Immunologie			
HLA A+B mismatch **	2,05 ± 1,09	1.2 ± 0,4	NS
HLA DR mismatch **	0,8 ± 0,7	1.1 ± 0,3	NS
Groupe sanguin A/B/O/AB — no.	747/176/730/84 †	122/49/104/27	NS

* Test du Chi2 ; ** Valeurs moyennes ± SD ; † 40 patients avec données manquantes

En se basant sur l'analyse des 7 items utilisés pour le diagnostic de rejet aigu (DSA circulants, dépôt de C4d dans le greffon, glomérulite, capillarite péritubulaire, endartérite, tubulite et inflammation interstitielle), nous présentons dans la Figure 1 une photographie du rejet chez les 302 patients inclus dans notre étude. Le dendrogramme oriente vers l'existence de différentes signatures de rejet que nous allons identifier par analyse en composantes principales.

Figure 1: Présentation des caractéristiques morphologiques et immunologiques de chacun des patients transplantés rénaux au moment de l'épisode de rejet aigu (1 ligne=1 patient). Analyse des profils des rejets par cluster non supervisé.

La coloration de chacune des variables correspond à sa gradation de 0 à 3.



Nous avons ainsi identifié 4 types distincts de rejets : le rejet vasculaire médié par les lymphocytes T (26 patients, 9%), le rejet vasculaire médié par les anticorps (64 patients, 21%), le rejet cellulaire classique sans endartérite (139 patients, 46%) et le rejet médié par les anticorps sans lésions d'endartérite (73 patients, 24%). La Figure 2A illustre la projection des patients selon 2 axes qui permettent de séparer au mieux (plus de 75% de l'inertie de la base de données) les différents types de rejet. Le premier axe, l'axe horizontal, est

déterminé par l'opposition de l'inflammation interstitielle et de la tubulite aux éléments qui constituent la signature humorale d'un rejet (DSA, C4d, glomérulite, capillarite péritubulaire). L'axe vertical est représenté par les lésions d'endartérite (Figure 2B). Comme illustré par le cercle de corrélations, la présence de DSA est l'élément diagnostic le plus discriminant pour séparer les 4 types de rejet. L'analyse en composantes principales réalisée sur la cohorte de validation (n=672) confirme que le rejet vasculaire médié par anticorps a un phénotype histologique et immunologique distinct par rapport aux 3 autres types de rejets identifiés.

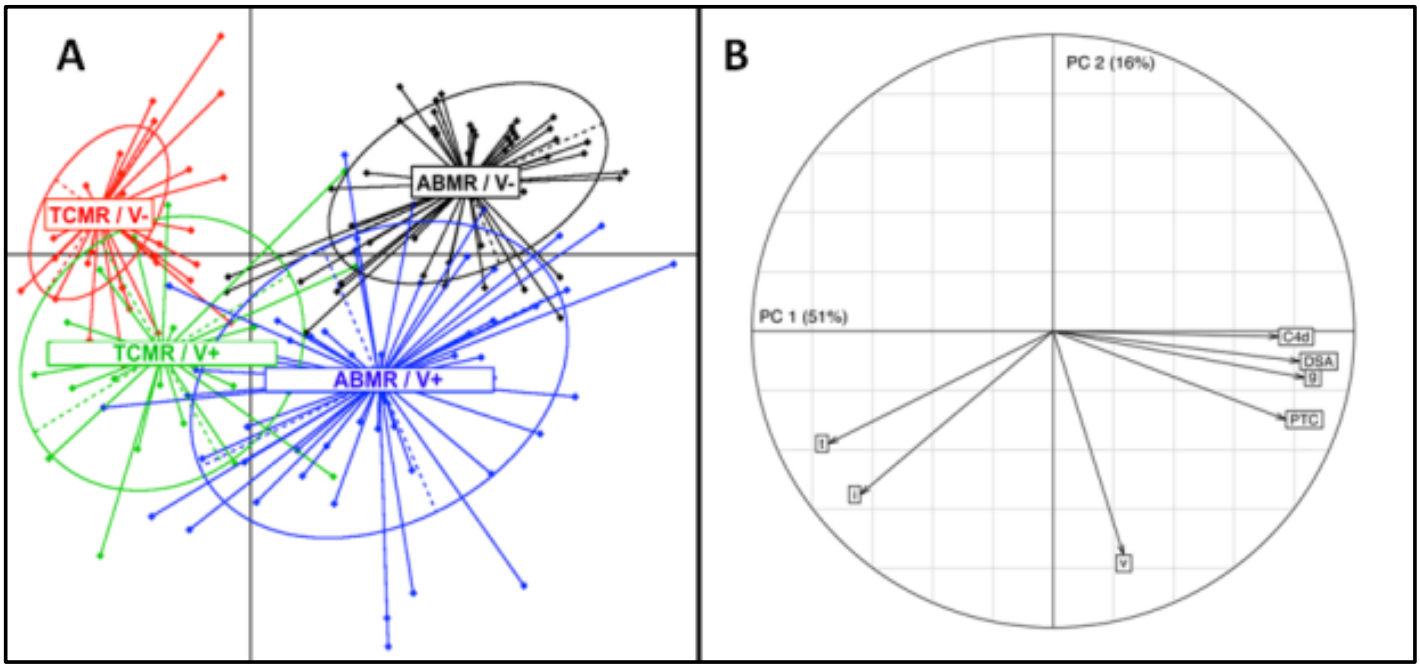
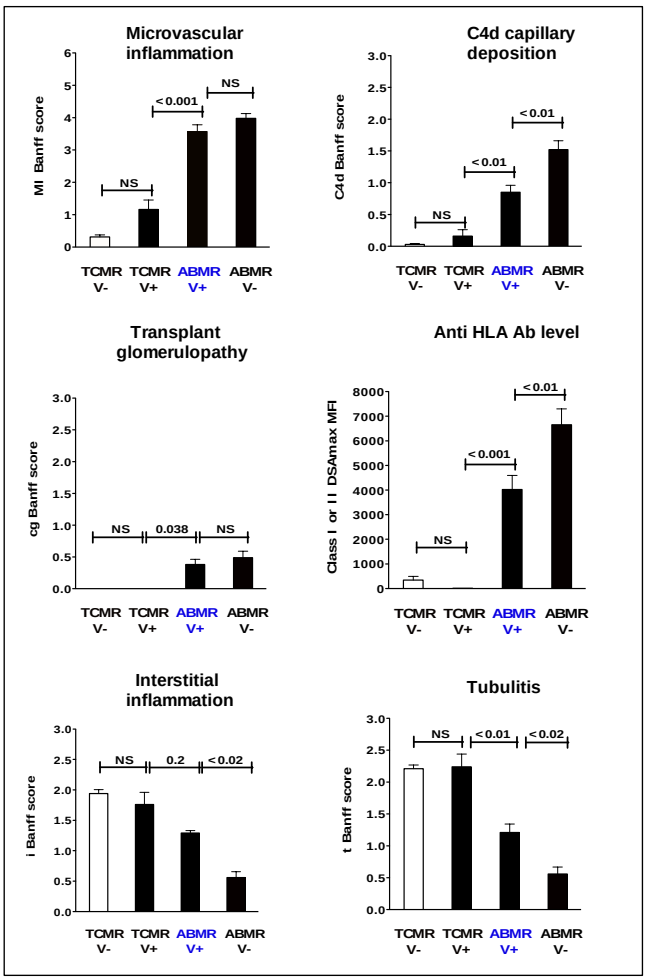


Figure 2: Analyse en composantes principales (ACP) des patients transplantés rénaux porteurs d'un diagnostic de rejet aigu d'allogreffe.
Panel A: Projection des 302 patients avec diagnostic de rejet aigu confirmé histologiquement sur les 2 premiers axes de l'ACP en utilisant 7 variables: DSA, glomérulite, inflammation interstitielle, tubulite, endartérite, capillarite péritubulaire et C4d.
TCMR/V- rejet médié par les lymphocytes T sans endartérite, ABMR/V- rejet médié par les anticorps sans endartérite, TCMR/V+ rejet vasculaire médié par les lymphocytes T et ABMR/V+ rejet vasculaire médié par les anticorps
Panel B: Le cercle de corrélations permet de caractériser les axes sur lesquels ont été projetés les patients. L'axe horizontal est défini par DSA/glomérulite/capillarite péritubulaire/C4d en opposition à l'inflammation interstitielle/tubulite. L'axe vertical est défini par la présence de lésions d'endartérite.

Nous avons identifié l'existence de 2 nouveaux types de rejets, absents de la classification internationale en vigueur : le rejet vasculaire médié par les anticorps et le rejet vasculaire médié par les lymphocytes T. Nous avons démontré que la présentation histologique et immunologique des 4 entités de rejets identifiées est aussi distincte (Figure 3). 20 (31%) des 64 patients avec rejet vasculaire médié par les anticorps et 26 (36%) des 73 patients avec rejet aigu médié par anticorps sans atteinte vasculaire présentaient des DSA de classe I au moment du rejet. 44 patients (69%) avec rejet vasculaire médié par anticorps et 47 (64%) avec rejet médié par anticorps classique présentaient des DSA de classe II.

Figure 3: Comparaison des paramètres morphologiques et immunologiques entre les 4 types de rejets aigus identifiés par l'analyse en composantes principales.

TCMR/Vrejet médié par lymphocytes T sans endartérite, ABMR/V- rejet médié par anticorps sans endartérite, TCMR/V+ rejet vasculaire médié par lymphocytes T et ABMR/V+ rejet vasculaire médié par anticorps



L'identification et donc la caractérisation précise des rejets ont des conséquences cliniques majeures. Notamment, certains patients avec rejet vasculaire médié par anticorps (45,3 %) ont été diagnostiqués au moment de la biopsie rénale de manière erronée comme des rejets médiés par des lymphocytes T (Table 3). En conséquence, ces 42 patients ont reçu un traitement de type « rejet cellulaire » (corticoïdes, sérum anti-lymphocytaire, OKT3) et cela a eu un impact délétère significatif en termes de survie des greffons. Nous avons ainsi déterminé que les patients avec rejet vasculaire médié par anticorps qui reçoivent une thérapie de type « rejet cellulaire » ont un risque significativement plus important de perdre leur greffon que ceux qui reçoivent une thérapie de type « humorale », basée sur l'association d'échanges plasmatiques, d'immunoglobulines intraveineuses et de Rituximab (log-rank p = 0,35).

	Nouvelle approche			
	TCMR/V- (n=139)	**TCMR/V+ (n=26)**	**ABMR/V+ (n=64)**	**ABMR/V- (n=73)**
Classification internationale				
Rejet cellulaire *(n, %)*	139 (100%)	24 (92,3%	**29 (45,3%)***	0
Rejet médié par anticorps *(n, %)*	0	2 (7,7%)	35 (54,7%)	73 (100%)
Stratégie thérapeutique				
st *(n, %)*	139 (100%)	19 (73%)	18 (28,1%)^{V}	0
st + [OKT3 / Thymo] *(n, %)*	0	5 (19,2%)	11 (17,2%)^{V}	0
st / IVIg (n, %)	0	2 (7,7%)	13 (20,3%)	29 (39,7%)
st / PP / IVIg / Rituximab *(n, %)*	0	0	22 (34,4%)	44 (60,3%)

Table 3: Comparaison entre le diagnostic de rejet basé sur la classification internationale et le diagnostic proposé par la nouvelle approche.

TCMR/V- rejet médié par lymphocytes T sans endartérite, ABMR/V- rejet médié par anticorps sans endartérite, TCMR/V+ rejet vasculaire médié par lymphocytes T et ABMR/V+ rejet vasculaire médié par anticorps.

st – stéroïdes. Thymo – thymoglobuline. IVIg – immunoglobulines intraveineuses. PP – échanges plasmatiques.

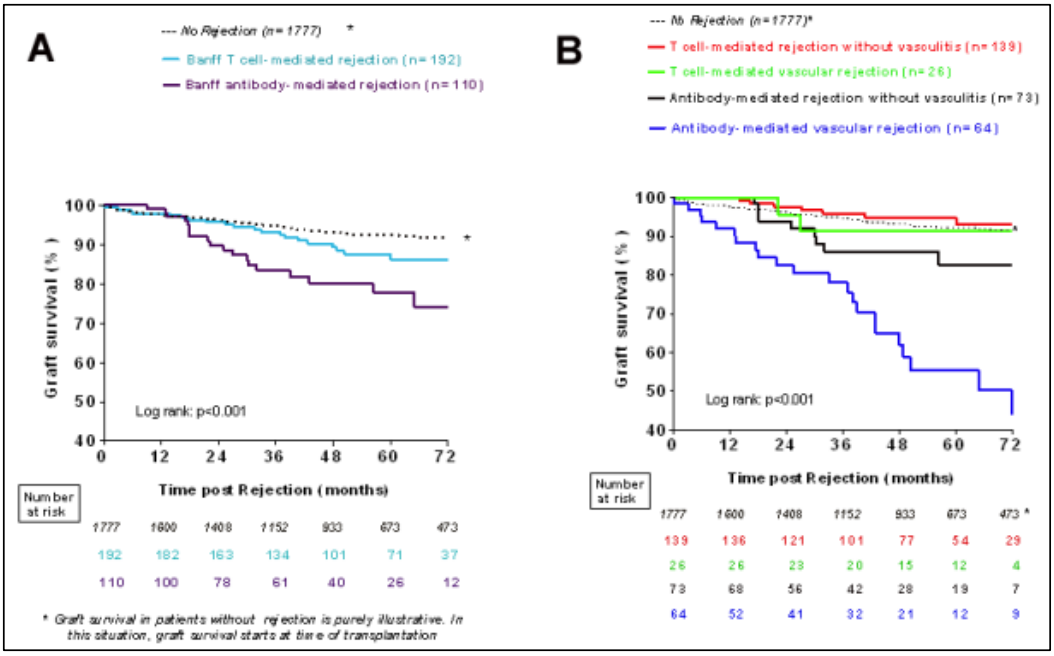
* Cas de rejets vasculaires médiés par anticorps diagnostiqués comme des rejets cellulaires utilisant la classification de Banff.

γ Cas de rejets vasculaires médiés par anticorps diagnostiqués comme des rejets cellulaires recevant une thérapie de type « rejet cellulaire ».

Les 4 entités de rejets que nous avons identifiées dans cette étude ont des profils évolutifs distincts. Nous avons déterminé que les patients porteurs d'un rejet vasculaire médié par anticorps ont la moins bonne survie de greffons (82,5% à 24 mois et 50,3% à 72 mois) par rapport aux patients avec un rejet médié par anticorps sans endartérite (93,8% à 24 mois et 82,6% à 72 mois) et aux patients avec un rejet cellulaire avec et sans endartérite (95,6% et 97,7% respectivement à 24 mois et 91,3% et 93,2% respectivement à 72 mois) ($p < 0,0001$). Cette population que nous avons identifiée comme ayant le risque le plus élevé de perte de greffons n'est actuellement pas reconnue par la classification internationale (Figure 4). Nous démontrons par analyse de Cox que le risque de perte de greffons est 9,07 fois plus élevé (95%CI 3,62-19,7) chez les patients avec rejet vasculaire médié par anticorps que chez les patients diagnostiqués avec rejet médié par des lymphocytes T ($p < 0,0001$).

En comparaison, le rejet médié par anticorps sans endartérite augmente ce risque de seulement 2,93 fois (1,1-7,9 ; p=0,0237).

Figure 4: Courbes de survie des greffons rénaux selon le phénotype de rejet : Panel A – diagnostics définis selon la classification internationale actuelle de Banff et Panel B – selon la nouvelle approche.



VIII.1.1.5 Conclusion de l'étude

Grâce à une approche en population, nous avons identifié 4 entités de rejets distinctes du point de vue clinique, histologique et immunologique qui ne sont pas reconnues comme telles par la classification internationale en vigueur.
Deux types de rejets sont associés à la présence de DSA : le rejet médié par anticorps « classique », sans lésions d'endartérite, et le rejet vasculaire médié par anticorps. Ces deux entités de rejet ont les survies de greffons les plus basses par rapport aux autres types de rejets et à la population générale des transplantés rénaux.
Nous apportons des données sur la performance des éléments utilisés pour le diagnostic des rejets. Parmi eux, la détection des DSA est l'élément le plus informatif. Cette étude démontre dans un modèle humain, que l'impact des

DSA dépasse l'atteinte endothéliale de la microcirculation et est responsable d'une atteinte de l'endothélium au niveau de la macrocirculation.

VIII.1.1.6 Manuscrit

LE QUOTIDIEN DU MEDECIN

www.lequotidiendumedecin.fr
N° 9198
JEUDI 29 NOVEMBRE 2012

ARNAQUES AUX ANNUAIRES PRO
Une nouvelle vague ruineuse
PAGE 3

LOCAUX ADAPTÉS AU HANDICAP
Le CNPS veut un aménagement
PAGE 4

MALADIE CŒLIAQUE
La piste d'un blé sans gluten
PAGE 6

ÉDITORIAL

Sus au AHC !

Dans le cadre de l'assemblée générale de l'International Longevity Center (ILC), le Pr V. G. Rodwin (université de New York) a présenté les bases d'une politique visant à réduire les hospitalisations évitables (AHC, pour Avoidable Hospital Condition) : une nécessité, quand on sait qu'on en a dénombré plus de 400 000 en 2004, soit un total de 2 millions de journées d'hospitalisation.

Certes, on peut se rassurer en constatant que les Anglais font plus mal que nous avec plus de 600 000 hospitalisations évitables, pour un total de 4 millions de journées. Mais les chiffres français ne sont pas glorieux, surtout à une époque où l'on veut faire des économies partout. Ils montrent, en tout cas, que la coordination ville-hôpital est loin d'être parfaite, en particulier pour la gestion des maladies chroniques.

Que faire pour mieux maîtriser ce problème ? Tout d'abord analyser les causes locales déterminant de fortes disparités régionales (les AHC sont deux fois plus fréquentes en Seine-Saint-Denis que dans les Pyrénées-Orientales) afin d'accomplir des efforts de *disease management* et d'améliorer les soins de santé primaire dans les régions les plus touchées. Une évaluation qui doit s'effectuer sur une période suffisante, en se servant des données administratives accessibles.

L'amélioration de la prise en charge ambulatoire des maladies chroniques (insuffisance cardiaque, diabète, BPCO…), mais aussi de certaines affections aiguës (pneumonies, infections bactériennes locales…).

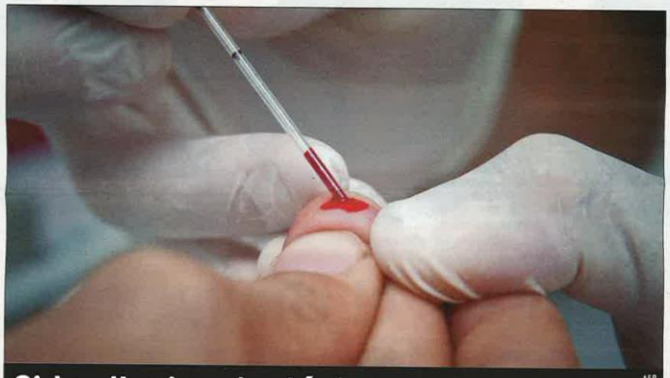
On peut ainsi espérer une économie de quelques milliards d'euros… Y-a-plus-qu'à…

> Dr ALAIN MARIÉ

Le rejet du 3e type

Des Français ont découvert une nouvelle forme de rejet de greffe rénale, qui multiplie par 9 le risque de perte du rein : le rejet vasculaire médié par les anticorps. Celui-ci s'ajoute aux deux types déjà connus, à savoir les réactions à médiation cellulaire ou à médiation humorale. Le pronostic de cette troisième forme peut être transformé grâce à un traitement spécifique par anticorps anti-HLA, déjà utilisé dans les trois centres parisiens où ce travail de recherche a été conduit.

PAGE 5



AFP

Sida : l'enjeu du dépistage

À la veille de la Journée mondiale de lutte contre le sida, l'association AIDES publie le premier bilan de son offre de dépistage rapide mise en place depuis 2012 dans 23 régions françaises. Les résultats sont encourageants. L'épidémie reste toujours dynamique avec 6 000 nouvelles infections par an et le dépistage, un enjeu. Prochains défis : l'élargissement du dépistage rapide aux hépatites B et C et l'introduction des autotests, actuellement à l'étude.

PAGE 8

La vie conventionnelle après l'avenant 8

Les signataires veulent dépasser les divisions

PAGE 3

Le PLFSS revient aujourd'hui au Sénat

L'encadrement du secteur privé à l'hôpital attendra une loi de 2013

PAGE 4

Une auto-transplantation hépatique à Rennes

Un cancer primitif inopérable totalement enlevé ex vivo

PAGE 6

42e ANNEE - 1, RUE AUGUSTINE-VARIOT - CS 80004 - 92245 MALAKOFF CEDEX - TEL. : 01. 73. 28. 12. 70 - ISSN 0399 - 2659 - CPPAP 0412 T 81257

Plus sérieusement :

12tl2776

Articles

AC

D-12-02776R3 6160

S0140-6736(12)61265-3

Embargo: Nov 23, 2012—00:01 (GMT)

Antibody-mediated vascular rejection of kidney allografts: a population-based study

Carmen Lefaucheur, Alexandre Loupy*, Dewi Vernerey, Jean-Paul Duong-Van-Huyen, Caroline Suberbielle, Dany Anglicheau, Jérôme Vérine, Thibaut Beuscart, Dominique Nochy, Patrick Bruneval, Dominique Charron, Michel Delahousse, Jean-Philippe Empana, Gary S Hill, Denis Glotz, Christophe Legendre, Xavier Jouven*

Summary

Background **Rejection of allografts has always been the major obstacle to transplantation success. We aimed to improve characterisation of different kidney-allograft rejection phenotypes, identify how each one is associated with anti-HLA antibodies, and investigate their distinct prognoses.**

Methods **Patients who underwent ABO-compatible kidney transplantations in Necker Hospital and Saint-Louis Hospital (Paris, France) between Jan 1, 1998, and Dec 31, 2008, were included in our population-based study. We assessed patients who provided biopsy samples for acute allograft rejection, which was defined as the association of deterioration in function and histopathological lesions. The main outcome was kidney allograft loss—ie, return to dialysis. To investigate distinct rejection patterns, we retrospectively assessed rejection episodes with review of graft histology, C4d in allograft biopsies, and donor-specific anti-HLA antibodies.**

Findings **2079 patients were included in the main analyses, of whom 302 (15%) had acute biopsy-proven rejection. We identified four distinct patterns of kidney allograft rejection: T cell-mediated vascular rejection (26 patients [9%]), antibody-mediated vascular rejection (64 [21%]), T cell-mediated rejection without vasculitis (139 [46%]), and antibody-mediated rejection without vasculitis (73 [24%]). Risk of graft loss was 9·07 times (95 CI 3·62–19·7) higher in antibody-mediated vascular rejection than in T cell-mediated rejection without vasculitis (p<0·0001), compared with an increase of 2·93 times (1·1–7·9; P=0·0237) in antibody-mediated rejection without vasculitis and no significant rise in T cell-mediated vascular rejection (hazard ratio [HR] 1·5, 95% CI 0·33–7·6; p=0·60).**

Interpretation **We have identified a type of kidney rejection not presently included in classifications: antibody-mediated vascular rejection. Recognition of this distinct phenotype could lead to the development of new treatment strategies that could salvage many kidney allografts.**

Funding **None.**

Introduction

Kidney transplantation is the treatment of choice for patients with end-stage renal disease, improving survival and quality of life and lowering costs compared with dialysis.[1–3] The alloimmune response induced by transplantation from a donor who differs genetically from the kidney recipient has always been the major obstacle to graft success. When not controlled, this response will destroy the graft,[4] leading to increased morbidity, mortality, and costs.

In the past decade, the view of acute allograft rejection has changed greatly. Although rejection was initially thought to be a T cell-mediated process, humoral mechanisms have progressively been recognised and are now acknowledged to lead to high rates of graft loss in HLA-incompatible kidney transplants.[5,6] International classification presently includes two acute rejection phenotypes: T cell-mediated rejection and antibody-mediated acute rejection.[7,8] However, rejection phenotypes and outcomes are more complex in clinical practice than in theory. Phenotypes might be difficult to identify or missed entirely, with direct negative effects on therapeutic management.

Vascular rejection of kidney allografts—defined by endarteritis, which is infiltration of immune cells beneath the endothelium—has been thought to be a T cell-mediated process.[4] Since the late 1990s, vascular rejection has been reported as a severe clinical disorder that does not respond to usual treatments directed at T cells.[4] More recently, clinical findings have suggested its association with alloantibodies,[9–12] challenging the notion of a unique T cell-mediated rejection process in vascular rejection. Importantly, these findings have been extended to transplantation of hearts,[13] composite tissues,[14] and small bowels,[15] and are reinforced by murine models showing that alloantibodies interact with vessels,[16,17] pointing towards a general idea of vascular damage and immune-mediated arteriosclerosis.[18]

We thus aimed to redefine rejection patterns by addressing their distinct clinical, histological, and immunological phenotypes, and their prognoses. In view of new immunosuppressive drugs available, such information would be particularly useful for adaptation of treatment strategies to rejection type, which might improve long-term prognoses of kidney allografts. We postulated that vascular rejection could be associated

Published Online November 23, 2012 http://dx.doi.org/10.1016/S0140-6736(12)61265-3

See Online/Comment http://dx.doi.org/10.1016/S0140-6736(12)61577-3

*These authors contributed equally

Nephrology and Kidney Transplantation (C Lefaucheur MD, Prof D Glotz MD), **Department of Immunology and Histocompatibility** (C Suberbielle MD, Prof D Charron MD), **and Department of Pathology** (J Vérine MD), **Saint-Louis Hospital, Assistance Publique Hôpitaux de Paris, INSERM U940, Paris, France; L'Université Paris Descartes, Sorbonne Paris Cité, and Department of Kidney Transplantation, Necker Hospital, Assistance Publique Hôpitaux de Paris, Paris, France** (A Loupy MD, D Anglicheau MD, Prof C Legendre MD); **Biostatistics Unit, Paris Cardiovascular Research Centre, L'Université Paris Descartes, UMR-S970, Paris, France** (A Loupy, D Vernerey MPH, Prof J P Duong-Van-Huyen MD, T Beuscart MPH, Prof P Bruneval MD, J-P Empana MD, Prof X Jouven MD); **Department of Pathology** (Prof J P Duong-Van-Huyen, T Beuscart, D Nochy MD, Prof P Bruneval, Prof G S Hill MD), **and Department of Cardiology** (Prof X Jouven), **Georges Pompidou European Hospital, Paris, France; and Department of Pathology** (Prof J P Duong-Van-Huyen) **and Department of Nephrology and Kidney Transplantation** (M Delahousse MD), **Hopital Foch. Suresnes, France**

Correspondence to: Dr Alexandre Loupy, Necker Hospital, 149 Rue de Sèvres, Paris 75015, France **alexandreloupy@gmail.com**

with anti-HLA antibodies, which might have important clinical implications for graft survival.

Methods

Participants

For **Données Informatiques Validées en Transplantation** see http://divat.fr

For **Agence de la Biomédecine** see http://www.agence-biomedecine.fr

See Online for appendix

Patients who underwent kidney transplantation in Necker Hospital and Saint-Louis Hospital (Paris, France) between Jan 1, 1998, and Dec 31, 2008, were included in our population-based study. Patients were followed up until March 31, 2010. We used an additional independent validation sample of patients who underwent kidney transplantation in Foch Hospital (Suresnes, France) between Jan 1, 2004, and Jan 31, 2010 (appendix).

All transplants were ABO compatible (ie, donors and recipients had the same blood type). Negative IgG T-cell and B-cell complement-dependent cytotoxicity cross-matching was necessary for all recipients. We excluded patients with primary non-functioning grafts due to surgical vascular or urological problems.

The study was approved by the institutional review boards of the two hospitals in Paris. Use of data from Foch Hospital is based on agreement between centres that participate in the national database system (appendix). All patients provided written informed consent.

	Kidney recipients without rejection (n=1777)	Kidney recipients with rejection (n=302)	p value*
Recipient age (years)	49 (13) [A1]	45 (10)	<0·0001
Men	1225 (69%)	148 (49%)	<0·0001
Retransplantation	212 (12%)	122 (40%)	<0·0001
Time since dialysis (months)	51 (54)	53 (62)	0·56
Donor age (years)	51 (16)	50 (12)	0·29
Male donor	970 (55%)	169 (56%)	0·66
Deceased donor	1444 (81%)	239 (79%)	0·38
Cardiovascular cause of donor death	622 (43%)	98 (41%)	0·39
Cold ischaemia time (h)	18·8 (10·3)	17·8 (8·1)	0·09
Delayed graft function	906 (51%)	177 (59%)	0·02
Cause of nephropathy			
Diabetes	211 (12%)	18 (6%)	0·003
Vascular	164 (9%)	15 (5%)	0·02
Glomerulopathy	433 (24%)	125 (41%)	<0·0001
Congenital	291 (16%)	88 (29%)	0·0009
Other	70 (4%)	14 (5%)	0·53
Interstitial nephropathy	171 (10%)	35 (12%)	0·29
Unknown	437 (25%)	7 (2%)	<0·0001
Immunology			
HLA A+B mismatch	2·1 (1·1) [A1]	1·2 (0·4)	<0·0001
HLA DR mismatch	0·8 (0·7)	1·1 0·3)	<0·0001
Blood group†			
Type A	747 (42%)	122 (40%)	0·41
Type B	176 (10%)	49 (16%)	0·0027
Type O	730 (41%)	104 (34%)	0·0134
Type AB	84 (5%)	27 (9%)	0·0058

Data are mean (SD) or n (%) NS=non-significant. *χ^2 tests for comparison of proportions and unpaired *t* test for comparison of continuous variables †Data missing for 40 patients without rejection. [A1: Rounded for consistency.]

Table 1: **Baseline characteristics**

Procedures

We took clinical data for donors and recipients from two validated computer databases (Données Informatiques Validées en Transplantation and Agence de la Biomédecine), in which data are prospectively entered at specific points for each patient (day 0, 6 months, and 1 year post-transplantation) and follow-up is updated annually thereafter (appendix). All data were retrieved from the database on March 31, 2010. Data from Foch Hospital were retrieved on Dec 22, 2011.

We assessed all patients who provided an indication biopsy sample (for deterioration in function, proteinuria, or impaired function) as part of standard care between Jan 1, 1998, and March 31, 2010, for acute clinical rejection episodes. Acute rejection was defined as the association of deterioration in function and histopathological lesions of allograft rejection, according to consensus rules represented by the international Banff classification criteria in use at the time.[7,19,20]

Patients with rejection episodes deemed to be T cell-mediated were given methylprednisolone pulses (500 mg/day for 3 days). Patients whose rejection was not controlled by this regimen received additional rabbit antithymocyte globulin (1·5 mg/kg/day for 5 days) or muromonab-CD3 (5 mg/day for 5 days). Patients who had episodes of antibody-mediated rejection were initially given methylprednisolone pulses (500 mg/day for 3 days) and intravenous immune globulin (2 g/kg, repeated every 3 weeks for three rounds). After January, 2004, all patients with antibody-mediated rejection received four plasmaphereses and two weekly doses of rituximab (375 mg/m²) as additional treatment.[21] The main outcome was kidney allograft loss, which was defined as the return of patient to dialysis.

Patients with rejection episodes were retrospectively reassessed between March, 2010, and December, 2010, with review of histology and immunochemistry for C4d in allograft biopsies and identification of circulating donor-specific anti-HLA antibodies on sera saved at time of biopsies. All graft biopsies of patients with biopsy-proven acute rejection were reviewed by two renal pathologists (DN, GSH) masked to clinical information. All biopsies were scored and graded from 0 to 3 according to the updated Banff criteria[7,8] for several histological factors: glomerular inflammation, tubulitis, interstitial inflammation, endarteritis (vasculitis), peritubular capillary inflammation, transplant glomerulopathy, interstitial fibrosis, tubular atrophy, arteriosclerosis, and hyaline arteriolar thickening. The microcirculation inflammation score was defined as the sum of glomerular and peritubular capillary inflammation, and the tubular and interstitial score by the sum of interstitial inflammation and tubulitis. C4d staining was done by immunochemistry on paraffin sections

with human C4d polyclonal antibody (Biomedica Gruppe, Vienna, Austria).

Presence of circulating donor-specific anti-HLA-A, anti-HLA-B, anti-HLA-DR, and anti-HLA-DQ antibodies at the time of biopsy was retrospectively analysed with single-antigen flow bead assays (One Lambda, Canoga Park, CA, USA) on the Luminex platform. All beads showing a normalised mean intensity of fluorescence higher than 500 arbitrary units were judged positive, as previously described.[6] For each patient, we recorded the number, specificities, and mean intensity of fluorescence of all donor-specific anti-HLA antibodies detected. We defined the maximum mean intensity of fluorescence of donor-specific anti-HLA antibodies as the highest ranked donor-specific bead. We categorised concentration of donor-specific antibody at time of rejection as absence (0), presence (1, 2, or 3) according to thresholds previously defined.[6]

HLA typing of transplant recipients was done by molecular biology (Innolipa HLA Typing Kit, Innogenetics, Ghent, Belgium). For all donors, tissue typing was done with the microlymphocytotoxicity technique with One Lambda INC tissue-typing trays and was controlled by molecular biology.

Statistical analysis

We provide mean (SD) values for descriptive analyses of continuous variables, with the exception of mean intensity of fluorescence, for which we use mean (SE) because of its wide distribution. We compared means and proportions of identified rejection phenotypes with Student's *t* test, ANOVA, or χ^2 test (Fisher's exact test when appropriate).

We addressed the hypothesis that vascular rejection with donor-specific antibodies is an independent event with unsupervised methods such as hierarchical cluster analysis and principal component analysis on the basis of combined histological lesions, C4d staining, and concentrations of donor-specific anti-HLA antibodies. We used Kaplan-Meier analysis to assess kidney allograft survival after rejection. We compared kidney allograft survival with the log-rank test across rejection phenotypes. A Bonferroni correction for several tests was used when we did two-by-two survival analysis comparison. The time of origin was time of acute rejection and the event of interest was a graft loss. In case of death with a functioning graft, we censored graft survival at time of death.[22]

To identify factors associated with risk of graft failure after antibody-mediated vascular rejection, we did univariate Cox analyses with variables such as baseline characteristics (donor or recipient age, cold ischaemia time, donor sex, donor's cause of death, type of nephropathy, and HLA mismatches) and characteristics of rejection (histopathology, immunology, and type of antirejection treatment used). All variables with a p value of 0·20 or less were then included in one multivariate Cox model. The proportionality assumption of the Cox model was verified with the log-graphic method.

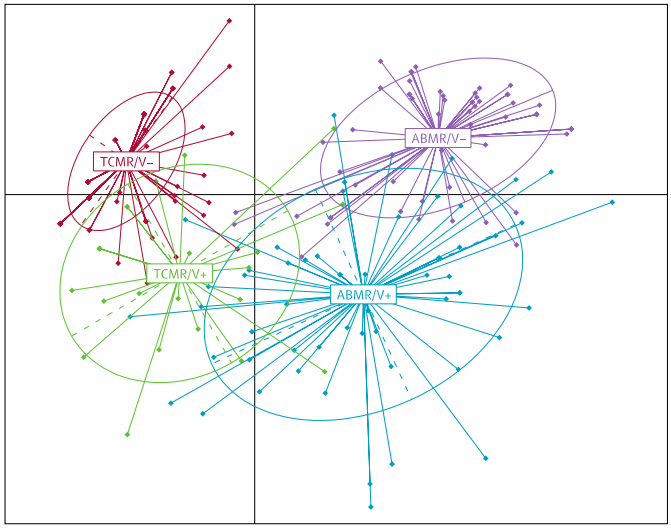


Figure 1: **Identification of four distinct rejection patterns according to clinical, histological, and immunological variables**
The unsupervised principal component analysis examined kidney recipients with acute biopsy-proven rejection with seven variables: glomerulitis, peritubular capillaritis, donor-specific anti-HLA antibodies, C4d deposition, interstitial inflammation, tubulitis, and endarteritis. The horizontal axis opposes cellular rejection (interstitial inflammation and tubulitis) and antibody-mediated rejection (donor-specific anti-HLA antibodies, glomerulitis, peritubular capillaritis and C4d), as recognised by the international Banff classification. The vertical axis defines the presence or absence of lesions of endarteritis (appendix).

All these statistical analyses were subsequently replicated and confirmed in the independent validation sample. We did hierarchical cluster analysis and dendrograms with the hcluster module of the amap package (version 0·8-7) of R (version 2.10.1). We did the principal components analysis with the dudi.pca module of the ade4 package (version 1.5-1) of R (version 2.10.1).[23] We used Stata (version 11.0) for descriptive and survival analyses. All tests were two-sided and we used an α of 0·05, unless otherwise stated.

Role of the funding source

There was no funding source for this study. CL and AL had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

2079 patients received a kidney allograft in Necker Hospital or Saint-Louis Hospital, of whom 302 (15%) had acute biopsy-proven rejection. Table 1 shows the characteristics of recipients at time of renal transplantation who subsequently did and did not have acute allograft rejection. Acute biopsy-proven rejection occurred at a median of 3·1 months post-transplant (IQR 1·0–11·3). 790 patients provided 952 indication biopsy samples. 147 samples (15%) showed acute tubular necrosis, 64 (7%) borderline lesions, 76 (8%) calcineurin inhibitor toxic effects, 68 (7%) recurrent disease, 40 (4%)

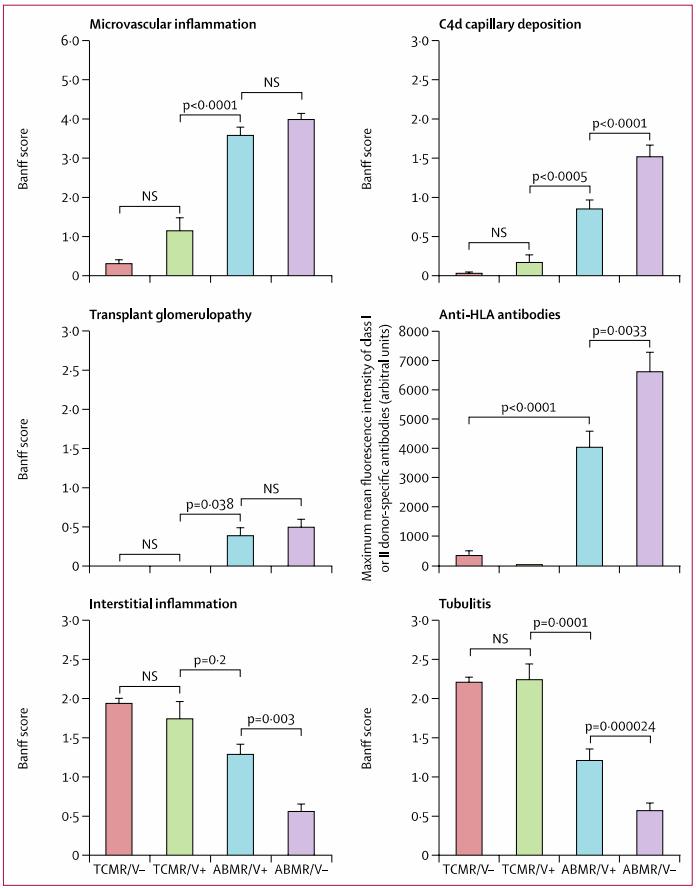


Figure 2: **Comparison of morphological and immunological variables in the four rejection patterns**
Bars represent SD. NS=not significant. TCMR/V–=T cell-mediated rejection without vasculitis. TCMR/V+=T cell-mediated vascular rejection. ABMR/V+=antibody-mediated vascular rejection. ABMR/V–=antibody-mediated rejection without vasculitis.

BK virus nephropathy, 124 (13%) interstitial fibrosis and tubular atrophy, and 44 (5%) transplant glomerulopathy, and 87 (9%) had other diagnoses. Median follow-up after transplantation was 43·7 months (27·2–64·1) in patients with acute rejection and 49·5 months (27·4–72·0) in those who did not experience rejection. 123 (7%) of the 1777 patients who did not experience rejection had graft loss, as did 40 (13%) of the 302 who had rejection.

We retrospectively identified four distinct rejection patterns: T cell-mediated vascular rejection (26 patients [9%]), antibody-mediated vascular rejection (64 [21%]), T cell-mediated rejection without vasculitis (139 [46%]), and antibody-mediated rejection without vasculitis (73 [24%]; figure 1). Median time between kidney transplantation and rejection was 1·6 months (IQR 0·2–4·4) in T cell-mediated vascular rejection, 1·1 months (0·4–4·4) in antibody-mediated vascular rejection, 3·3 months (3·0–12·0) in T cell-mediated rejection without vasculitis, and 7·0 months (1·1–18·0) in antibody-mediated rejection without vasculitis.

Antibody-mediated rejection both with and without vasculitis was associated with microvascular inflammation, transplant glomerulopathy, and anti-HLA antibodies (figure 2). 20 (31%) of the 64 patients with antibody-mediated vascular rejection and 26 (36%) of the 73 patients with antibody-mediated rejection without vasculitis had circulating class I immunodominant donor-specific anti-HLA antibodies. 44 patients (69%) with antibody-mediated vascular rejection and 47 (64%) with antibody-mediated rejection without vasculitis had circulating class II immunodominant donor-specific anti-HLA antibodies. Of patients with antibody-mediated vascular rejection, 33 patients (52%) had endarteritis graded as v1 with Banff criteria, 19 (30%) had v2, and 12 (19%) had v3. Of those with T cell-mediated vascular rejection, 16 (62%) had v1 endarteritis, 5 (19%) had v2, and 5 (19%) had v3. Additionally, 56 patients (88%) with antibody-mediated vascular rejection, 73 (100%) with antibody-mediated rejection without vasculitis, 16 (12%) with T cell-mediated rejection without vasculitis, and 4 (15%) with T cell-mediated vascular rejection had glomerular and peritubular capillary inflammation; 46 (72%), 27 (37%), 134 (96%), and 23 (88%) had interstitial inflammation; and 40 (63%), 21 (29%), 137 (99%) and 23 (88%) had tubulitis (appendix). C4d staining was positive in 36 patients (56%) with antibody-mediated vascular rejection, of whom 27 (75%) had intimal staining for C4d in arteries and arterioles (appendix). C4d staining was positive in 45 patients (62%) with antibody-mediated rejection without vasculitis, 6 (4%) with T cell-mediated rejection without vasculitis, and 2 (8%) with T cell-mediated vascular rejection.

Importantly, some patients with antibody-mediated vascular rejection were misclassified at time of biopsy as having T cell-mediated rejection (table 2), and received inappropriate treatment. Notably, the 42 patients with antibody-mediated vascular rejection treated with a T cell-mediated rejection strategy had a significantly greater risk of graft loss than did the 22 who received appropriate treatment (log rank p=0·035; appendix).

The Kaplan-Meier estimate of graft survival after T cell-mediated rejection without vasculitis (97·7% at 24 months; 93·2% at 72 months) was similar to that in patients after T cell-mediated vascular rejection (95·6% at 24 months; 91·3% at 72 months; figure 3). Graft survival after antibody-mediated rejection without vasculitis (93·8% at 24 months; 82·6% at 72 months) was significantly lower than after T cell-mediated rejection without vasculitis (p=0·0237). Patients with antibody-mediated vascular rejection had the poorest graft survival (82·5% at 24 months; and 50·3% to 72 months) when compared with those with T cell-mediated rejection without vasculitis (p<0·0001; figure 3). Cox univariate analysis showed that risk of graft loss was 9·07 times

(95% CI 3·62–19·7) higher in antibody-mediated vascular rejection than in T cell-mediated rejection without vasculitis (p<0·0001). By comparison, antibody-mediated rejection without vasculitis was associated with an increase of 2·93 times (1·1–7·9; p=0·0237). T cell-mediated vascular rejection was not associated with an increased risk of graft loss compared with T cell-mediated rejection without vasculitis (hazard ratio 1·5, 95% CI 0·33–7·6; p=0·60).

In univariate analysis, none of the clinical factors investigated were associated with graft loss after antibody-mediated vascular rejection (data not shown; all p>0·20). Transplant glomerulopathy (p=0·03), microcirculation inflammation (p=0·18), interstitial inflammation and tubulitis score (p=0·09), endarteritis scores (p=0·07), concentration of donor-specific anti-HLA antibodies at time of biopsy (p=0·06), and antibody-targeting regimen (p=0·10) were potential prognostic factors. When entered in one multivariate Cox regression analysis, a high interstitial inflammation and tubulitis score, high grade of endarteritis lesions, and the concentration of circulating donor-specific anti-HLA antibodies in the biopsy at time of rejection were independent predictors of graft loss (table 3). Conversely, the use of an antibody-targeting strategy was independently associated with a reduced risk of graft loss compared with non-antibody directed treatment strategies (table 3).

The principal component analysis of the independent validation cohort (n=672) confirmed that antibody-mediated vascular rejection with donor-specific anti-HLA antibodies had a distinct morphological and immunological phenotype compared with other rejection phenotypes (appendix). Additionally, antibody-mediated vascular rejection with donor-specific anti-HLA antibodies had a distinct phenotype characterised by the highest risk of graft loss (log-rank test p<0·0001; appendix).

Discussion

We have defined four relevant clinical phenotypes of rejection after kidney transplantation. With contemporary immunological and histopathological techniques, we have shown that the previously unrecognised phenotype of antibody-mediated vascular rejection is characterised by endarteritis, associated with circulating donor-specific anti-HLA antibodies, and has the poorest graft survival. Our data suggest that antibody-depleting strategies could be beneficial in the long term in patients with donor-specific anti-HLA antibody-associated vascular lesions.

Lesions of vascular rejection have long interested transplant groups because of their presentation and severity.[24,25] Vascular rejection has been considered to be a severe disorder that does not respond to traditional treatment for T cell-mediated rejection and so necessitates potent antilymphocyte-antibody treatment (muromonab-CD3 or antithymocyte globulin).[26] Endarteritic lesions form part of the Banff criteria only for diagnosis of T cell-mediated rejection. Only cases with transmural arteritis or arterial fibrinoid change, or both, and necrosis

	TCMR/V– (n=139)	TCMR/V+ (n=26)	ABMR/V+ (n=64)	ABMR/V– (n=73)
Traditional approach				
T cell-mediated rejection	139 (100%)	24 (92%)	29 (45%)	0
Antibody-mediated rejection	0	2 (8%)	35 (55%)	73 (100%)
Treatment strategy				
Steroids	139 (100%)	19 (73%)	18 (28%)	0
Steroids plus muromonab-CD3 or rabbit antithymocyte globulin	0	5 (19%)	11 (17%)	0
Steroids and intravenous immune globulin	0	2 (8%)	13 (20%)*	29 (40%)
Steroids, plasmapheresis, intravenous immune globulin, and rituximab	0	0	22 (34%)	44 (60%)

Data are n (%). TCMR/V–=T cell-mediated rejection without vasculitis. TCMR/V+=T cell-mediated vascular rejection. ABMR/V+=antibody-mediated vascular rejection. ABMR/V–=antibody-mediated rejection without vasculitis. *These patients were classified as having antibody-mediated rejection, but were not treated as such; steroids and intravenous immune globulin was practice at that time in the referring centre.

Table 2: **Comparison between initial diagnosis of rejection made by international classification and the diagnosis of the new approach**

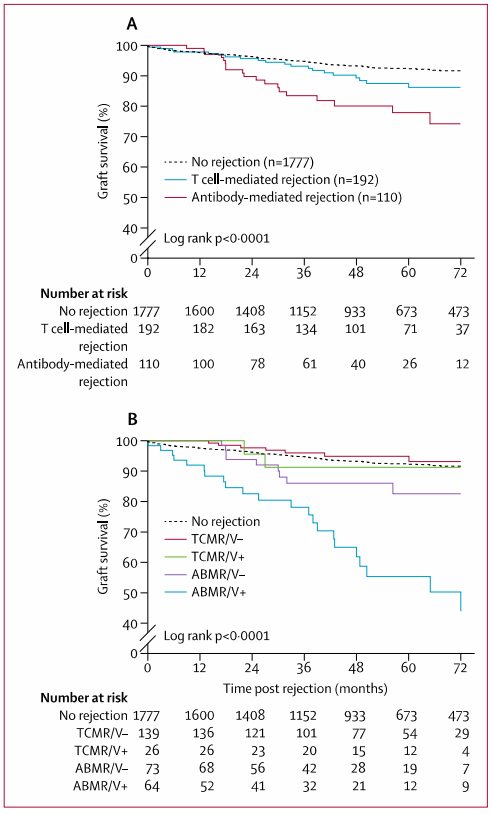


Figure 3: **Kaplan-Meier curves for kidney graft survival by acute rejection phenotype**
Initial diagnoses as per (A) Banff classifications and (B) our new approach. Graft survival in patients without rejection is purely illustrative; graft survival in these individuals starts at time of transplantation.

	Number of patients	Number of events	Hazard ratio (95% CI)	p value
Interstitial inflammation and tubulitis score*				
≤3	32	7	1	..
>3	32	14	4·33 (1·5–12·1)	0·005
Endarteritis score				
<3	52	15	1	..
3	12	6	5·17 (1·8–14·6)	0·002
DSA_{max} MFI				
<3000	41	9	1	..
>3000	23	12	3·88 (1·5–9·8)	0·004
Treatment strategy				
Steroids and intravenous immune globulin	13	7	1	..
Steroids plus muromonab-CD3 or rabbit antithymocyte globulin	29	11	0·4 (0·2–1·3)	0·1
Steroids, plasmapheresis, intravenous immune globulin, and rituximab	22	3	0·16 (0·04–0·66)	0·01

Hazard ratios were estimated in a single Cox proportional hazards model. DSA_{max} MFI=maximum mean intensity of fluorescence of donor-specific anti-HLA antibodies. *Interstitial inflammation and tubulitis score was defined as the sum of interstitial inflammation and tubulitis, and was graded from 0 to 6.

Table 3: **Multivariate analysis of factors associated with graft loss in patients with antibody-mediated vascular rejection**

Panel: Research in context

Systematic review

We searched PubMed for reports in any language published before April 30, 2012, with the search terms "kidney allograft rejection" and "antibody mediated rejection". We identified 691 reports overall, of which eight were results of the Banff working group and three were reviews. We did another search with the search terms "vascular rejection", "vasculitis/endarteritis", and "Banff allograft rejection classification". We identified 37 reports overall, of which eight were reviews. Previous studies focusing on vascular rejection of kidney allografts are limited by small samples and absence of contemporary techniques that are increasingly used worldwide for graft-rejection phenotyping. Additionally, little is known about the association of vascular rejection with donor-specific anti-HLA antibodies.

Interpretation

With a population-based approach, we have identified a type of allograft rejection that is not represented by the present classification of renal rejection. This type of rejection—antibody-mediated vascular rejection—is characterised by the association of circulating antibodies with arteritis lesions of allograft arteries and carries the highest risk of graft loss of all rejection phenotypes. If antibody-mediated vascular rejection becomes widely recognised, anti-HLA antibody-targeting strategies could reduce risk of graft loss. Overall, our results support the notion of immunological development of arteriosclerosis in people, which has been established in murine models.

of medial smooth muscle cells are deemed to be grade III T cell-mediated rejection or antibody-mediated rejection with the Banff criteria.[7,8] Patients with v1 or v2 lesions are judged to have T cell-mediated rejection. Our study now permits the division of all cases of endarteritis into two separate profiles of rejection on the basis of donor-specific anti-HLA antibody status.

Antibody-mediated vascular rejection is a separate form of acute rejection, with a distinct prognosis. Our approach has allowed this rejection type to be epidemiologically established. In the past decade, several reports have suggested that anti-HLA antibodies can be associated with endarteritis lesions.[9–12] Our study thus reinforces the idea that lesions of endarteritis form part of the range of general endothelial inflammatory changes mediated by anti-HLA antibodies. Importantly, our investigation provides the basis for the application of antibody-mediated transplant arteriopathy developed in mice[17] into a human model represented by kidney transplantation. This idea extends beyond the specialty of renal transplantation to reports of immunological damage to arteries in transplantation of hearts,[13] composite tissues,[14] and small bowels.[15] Generally, our findings reinforce those from previous studies,[18] supporting the notion of immunological development of arteriosclerosis and atherogenesis in people, which has already been established in murine models.[16]

Overall, our study presents a contemporary picture of acute allograft rejection on the basis of techniques that are becoming standard in organ transplantation worldwide. Our principal component analysis allowed us to hierarchically rank the variables that characterise the different rejection profiles. Thus, we could establish that microcirculation inflammation and anti-HLA donor-specific antibodies on the one hand, and the presence of endarteritis lesions on the other, allow effective discrimination. Not surprisingly, C4d is not a principal discriminating factor in the diagnosis of rejection. Although several groups agree that C4d staining is specific when it occurs, it is not a sensitive indicator of humoral rejection activity.[9,27–29]

The importance of precise identification of type of kidney allograft rejection is primarily therapeutic, because it enables assessment of treatment effectiveness. The resistance of antibody-mediated rejection to treatments traditionally given for T cell-mediated rejection has been recognised since the 1990s,[30] and the improvement in diagnostic criteria for antibody-mediated rejection has led to advances in treatment.[21,31–34] The guiding principle of treatment in antibody-mediated rejection is diminution of donor-specific anti-HLA antibody concentration and production.[35] Plasmapheresis, intravenous immune globulin, rituximab, and more recently the proteasome inhibitor bortezomib[36] are used in such regimens. Our findings support the view that interventions directed towards a decrease in donor-specific anti-HLA antibody concentration and production could be useful. Further studies will be necessary to define which treatments are most effective in reversal of antibody-mediated vascular rejection; the observational design of our study does not allow us to make any formal therapeutic conclusions about the superiority of one treatment.

Our study has some limitations. It was not designed to provide mechanistic insights or pathways linking anti-HLA antibodies with vascular cell biology. Additionally, few graft losses were reported in some subgroups, indicating that we might not have had precision in our point estimates.

In conclusion, advances in HLA testing techniques, together with the ability to better characterise kidney allograft rejection histologically, have allowed different profiles of allograft rejection to be elucidated. Our study has identified a clinically relevant profile of rejection linking vascular rejection lesions in the allograft with circulating donor-specific anti-HLA antibodies (panel). The prognosis and course of this type of rejection is substantially different from other types of rejection, with antibody-mediated vascular rejection having the poorest outcome. Because rejection is still the leading cause of kidney allograft loss, we hope that the recognition of this distinct type of rejection will lead to development of new treatment strategies that could salvage many kidney allografts.

Contributors

CL, AL, DA, MD, J-PE, DG, CL, and XJ designed the study. CL, AL, JPD-V-H, CS, JV, DN, PB, DC, MD, and GSH collected data. AL, DV, JPD-V-H, CS, DN, PB, DC, J-PE, and XJ analysed data. AL, DV, J-PE, and XJ prepared figures. CL, AL, DV, DA, TB, J-PE, GSH, DG, CL, and XJ interpreted data. CL, AL, DV, DA, PB, J-PE, GSH, DG, CL, and XJ wrote the report.

Conflicts of interest

We declare that we have no conflicts of interest.

Acknowledgments

We thank Corinne Antoine (French National Agency for Organ Procurement, Saint-Denis, France) for help with data collection.

References

1 Evans RW, Manninen DL, Garrison LP Jr, et al. The quality of life of patients with end-stage renal disease. *N Engl J Med* 1985; **312:** 553–59.
2 Port FK, Wolfe RA, Mauger EA, Berling DP, Jiang K. Comparison of survival probabilities for dialysis patients *vs* cadaveric renal transplant recipients. *JAMA* 1993; **270:** 1339–43.
3 Garcia GG, Harden P, Chapman J, for the World Kidney Day Steering Committee 2012. The global role of kidney transplantation. *Lancet* 2012; **379:** e36–38.
4 Nankivell BJ, Alexander SI. Rejection of the kidney allograft. *N Engl J Med* 2010; **363:** 1451–62.
5 Einecke G, Sis B, Reeve J, et al. Antibody-mediated microcirculation injury is the major cause of late kidney transplant failure. *Am J Transplant* 2009; **9:** 2520–31.
6 Lefaucheur C, Loupy A, Hill GS, et al. Preexisting donor-specific HLA antibodies predict outcome in kidney transplantation. *J Am Soc Nephrol* 2010; **21:** 1398–406.
7 Solez K, Colvin RB, Racusen LC, et al. Banff 07 classification of renal allograft pathology: updates and future directions. *Am J Transplant* 2008; **8:** 753–60.
8 Sis B, Mengel M, Haas M, et al. Banff 09 meeting report: antibody mediated graft deterioration and implementation of Banff working groups. *Am J Transplant* 2010; **10:** 464–71.
9 Sis B, Jhangri GS, Bunnag S, Allanach K, Kaplan B, Halloran PF. Endothelial gene expression in kidney transplants with alloantibody indicates antibody-mediated damage despite lack of C4d staining. *Am J Transplant* 2009; **9:** 2312–23.
10 Sis B, Einecke G, Chang J, et al. Cluster analysis of lesions in nonselected kidney transplant biopsies: microcirculation changes, tubulointerstitial inflammation and scarring. *Am J Transplant* 2010; **10:** 421–30.
11 Herzenberg AM, Gill JS, Djurdjev O, Magil AB. C4d deposition in acute rejection: an independent long-term prognostic factor. *J Am Soc Nephrol* 2002; **13:** 234–41.
12 Lefaucheur C, Nochy D, Hill GS, et al. Determinants of poor graft outcome in patients with antibody-mediated acute rejection. *Am J Transplant* 2007; **7:** 832–41.
13 Loupy A, Cazes A, Guillemain R, et al. Very late heart transplant rejection is associated with microvascular injury, complement deposition and progression to cardiac allograft vasculopathy. *Am J Transplant* 2011; **11:** 1478–87.
14 Kaufman CL, Ouseph R, Blair B, et al. Graft vasculopathy in clinical hand transplantation. *Am J Transplant* 2012; **12:** 1004–16.
15 de Serre NP, Canioni D, Lacaille F, et al. Evaluation of c4d deposition and circulating antibody in small bowel transplantation. *Am J Transplant* 2008; **8:** 1290–96.
16 Ait-Oufella H, Herbin O, Bouaziz JD, et al. B cell depletion reduces the development of atherosclerosis in mice. *J Exp Med* 2010; **207:** 1579–87.
17 Hirohashi T, Uehara S, Chase CM, et al. Complement independent antibody-mediated endarteritis and transplant arteriopathy in mice. *Am J Transplant* 2010; **10:** 510–17.
18 Hill GS, Nochy D, Bruneval P, et al. Donor-specific antibodies accelerate arteriosclerosis after kidney transplantation. *J Am Soc Nephrol* 2011; **22:** 975–83.
19 Racusen LC, Solez K, Colvin RB, et al. The Banff 97 working classification of renal allograft pathology. *Kidney Int* 1999; **55:** 713–23.
20 Racusen LC, Colvin RB, Solez K, et al. Antibody-mediated rejection criteria—an addition to the Banff 97 classification of renal allograft rejection. *Am J Transplant* 2003; **3:** 708–14.
21 Lefaucheur C, Nochy D, Andrade J, et al. Comparison of combination plasmapheresis/IVIg/anti-CD20 versus high-dose IVIg in the treatment of antibody-mediated rejection. *Am J Transplant* 2009; **9:** 1099–107.
22 Lamb KE, Lodhi S, Meier-Kriesche HU. Long-term renal allograft survival in the United States: a critical reappraisal. *Am J Transplant* 2011; **11:** 450–62.
23 Saldanha AJ. Java Treeview—extensible visualization of microarray data. *Bioinformatics* 2004; **20:** 3246–48.
24 Loucaidou M, McLean AG, Cairns TD, et al. Five-year results of kidney transplantation under tacrolimus-based regimes: the persisting significance of vascular rejection. *Transplantation* 2003; **76:** 1120–23.
25 Shimizu T, Ishida H, Shirakawa H, et al. Clinicopathological analysis of acute vascular rejection cases after renal transplantation. *Clin Transplant* 2010; **24** (suppl 22): 22–26.
26 Halloran PF. Immunosuppressive drugs for kidney transplantation. *N Engl J Med* 2004; **351:** 2715–29.
27 Mauiyyedi S, Crespo M, Collins AB, et al. Acute humoral rejection in kidney transplantation: II: morphology, immunopathology, and pathologic classification. *J Am Soc Nephrol* 2002; **13:** 779–87.
28 Bohmig GA, Exner M, Habicht A, et al. Capillary C4d deposition in kidney allografts: a specific marker of alloantibody-dependent graft injury. *J Am Soc Nephrol* 2002; **13:** 1091–99.
29 Loupy A, Hill GS, Suberbielle C, et al. Significance of C4d Banff scores in early protocol biopsies of kidney transplant recipients with preformed donor-specific antibodies (DSA). *Am J Transplant* 2011; **11:** 56–65.
30 Trpkov K, Campbell P, Pazderka F, Cockfield S, Solez K, Halloran PF. Pathologic features of acute renal allograft rejection associated with donor-specific antibody: analysis using the Banff grading schema. *Transplantation* 1996; **61:** 1586–92.
31 Montgomery RA, Zachary AA, Racusen LC, et al. Plasmapheresis and intravenous immune globulin provides effective rescue therapy for refractory humoral rejection and allows kidneys to be successfully transplanted into cross-match-positive recipients. *Transplantation* 2000; **70:** 887–95.
32 Rocha PN, Butterly DW, Greenberg A, et al. Beneficial effect of plasmapheresis and intravenous immunoglobulin on renal allograft survival of patients with acute humoral rejection. *Transplantation* 2003; **75:** 1490–95.
33 Faguer S, Kamar N, Guilbeaud-Frugier C, et al. Rituximab therapy for acute humoral rejection after kidney transplantation. *Transplantation* 2007; **83:** 1277–80.
34 Loupy A, Suberbielle-Boissel C, Zuber J, et al. Combined posttransplant prophylactic IVIg/anti-CD 20/plasmapheresis in kidney recipients with preformed donor-specific antibodies: a pilot study. *Transplantation* 2010; **89:** 1403–10.
35 Takemoto SK, Zeevi A, Feng S, et al. National conference to assess antibody-mediated rejection in solid organ transplantation. *Am J Transplant* 2004; **4:** 1033–41.
36 Everly MJ, Everly JJ, Susskind B, et al. Bortezomib provides effective therapy for antibody- and cell-mediated acute rejection. *Transplantation* 2008; **86:** 1754–61.

VIII.1.2 Etude du rôle pronostique des anticorps anti-HLA fixant le complément sur le risque de perte du greffon

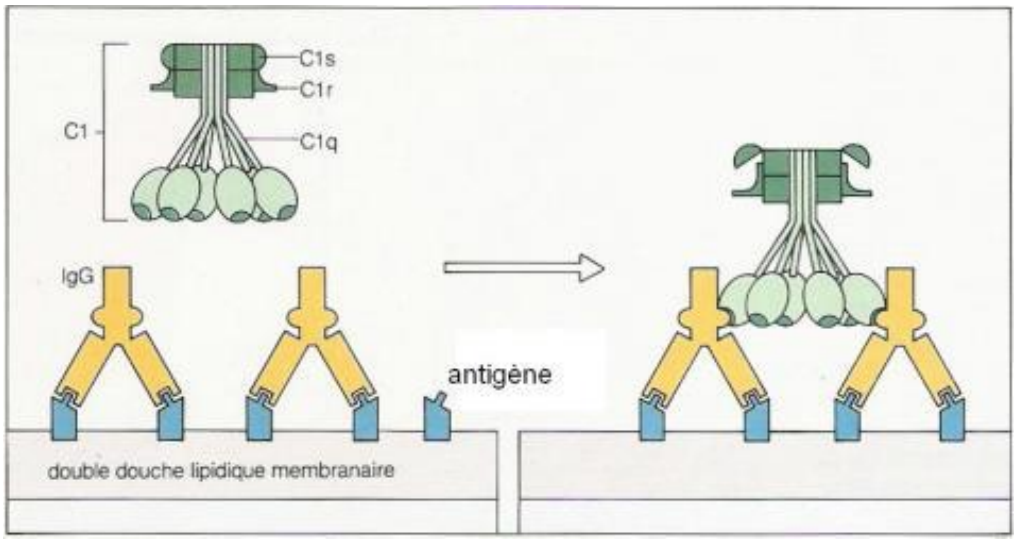
Loupy A, Lefaucheur C, Vernerey D, Prugger C, Duong van Huyen J-P, Mooney N, Suberbielle C, Frémeaux-Bacchi V, Méjean A, Desgrandchamps F, Anglicheau D, Nochy D, Charron D, Empana J-P, Delahousse M, Legendre C, Glotz D, Hill GS, Zeevi A, Jouven X
Complement-binding anti-HLA antibodies and kidney-allograft survival. **N Engl J Med** 369: 1215–1226, doi:10.1056/NEJMoa1302506. (2013)
Impact factor: 54.4

VIII.1.2.1 Positionnement du problème

Nous savons que la présence ou l'absence d'anticorps anti-HLA spécifiques du donneur (DSA) ont une valeur prédictive mais modérée pour la survenue d'un rejet humoral ainsi que pour la perte de greffon. Il existe un besoin évident de définir des profils distincts de DSA, en se basant sur leurs caractéristiques immunologiques et leur association à des phénotypes cliniques et histologiques distincts corrélés à l'évolution des greffons.
Dans le contexte d'émergence de nouvelles thérapies, incluant les inhibiteurs du complément, l'identification d'une population à risque élevé de perte de greffon pourrait être particulièrement utile pour définir des stratégies thérapeutiques adaptées et ainsi augmenter la survie à long terme des greffons.

Dans cette étude, nous nous sommes concentrés sur la première étape de l'activation du complément: la fixation de la fraction C1q du complément aux anticorps anti-HLA (Figure 1).

Figure 1: Le C1 est la première protéine de la voie classique du complément. C'est un complexe constitué du C1q, du C1r et du C1s. Les têtes globulaires peuvent se lier aux régions constantes des immunoglobulines, ce qui change la conformation du C1r, qui peut alors cliver et activer le zymogène C1s.



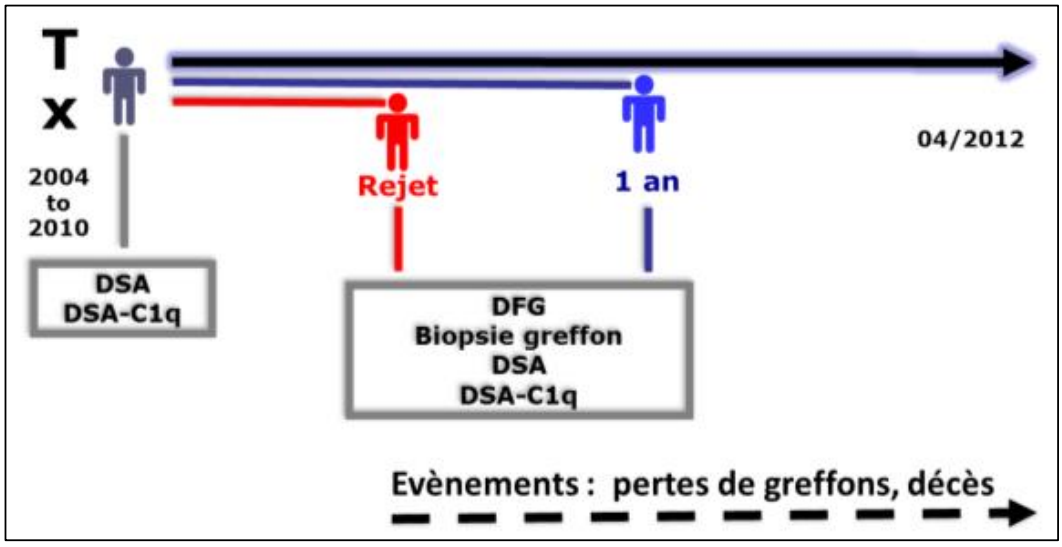
L'objectif de cette étude est d'investiguer le rôle pronostique de cette propriété des DSA à fixer le complément et de l'associer aux phénotypes histologiques, immunohistochimiques et cliniques pertinents. Notre hypothèse est que la détection de la capacité des DSA à fixer le complément peut augmenter les performances prédictives, en particulier de la perte du greffon, de la détection d'anticorps anti-HLA.

VIII.1.2.2 Population d'étude

Cette étude en population inclue l'ensemble des patients qui ont reçu une greffe rénale, entre janvier 2004 et janvier 2010, dans deux centres majeurs de transplantation en Ile-de-France (Hôpital Necker et Hôpital Saint-Louis). 1 016 patients ont été inclus : 695 patients transplantés à l'Hôpital Necker et 321 à l'Hôpital Saint-Louis. Les résultats sont validés sur une cohorte externe, représentée par la population de patients transplantés rénaux à l'Hôpital Foch pendant la même période.

La force de cette étude est de réaliser une analyse concomitante des caractéristiques des anticorps anti-HLA et de leurs conséquences dans le greffon. L'ensemble des patients a été évalué au moment de la transplantation pour la présence de DSA, leurs MFI et leur propriété à fixer le complément. Tous les patients inclus dans cette étude ont été évalués en post-transplantation pour la présence de DSA, leurs MFI, leur capacité à fixer le complément, l'histologie du greffon et la présence de C4d sur les biopsies. Cette évaluation a été faite au moment le plus informatif de la première année post-transplantation : au moment du rejet (pour les 171 patients présentant un épisode de rejet aigu dans la première année) et à 1 an pour le reste de la population (Figure 2).

Figure 2 : Design de l'étude visant à déterminer le rôle des anticorps anti-HLA spécifiques du donneur fixant le complément en transplantation rénale.



VIII.1.2.3 Approche méthodologique

La durée de survie du greffon a été calculée en utilisant la méthode de Kaplan-Meier. La survie a été analysée depuis la transplantation, avec comme évènement d'intérêt la perte du greffon définie par le retour en

dialyse. Les données de survie du greffon ont été censurées au moment du décès du receveur. La survie des greffons rénaux a été représentée sur des courbes de Kaplan-Meier stratifiées selon le statut C1q des DSA.

Un modèle semi-paramétrique des risques proportionnels a été utilisé afin de quantifier les rapports de risque et les intervalles de confiances à 95% pour la perte du greffon. L'association des caractéristiques du donneur, du receveur, les données immunologiques et histologiques avec la perte du greffon a été évaluée dans des modèles de Cox univariés puis multivariés. Pour déterminer les facteurs indépendants associés à la perte de greffons, nous avons construit un modèle multivarié par blocs de variables, incluant l'ensemble des caractéristiques cliniques, histologiques, fonctionnelles et immunologiques. Le bloc immunologique intègre la présence de DSA et leur statut C1q au moment de la greffe et en post-transplantation.

Enfin, nous avons évalué le gain de performance prédictive de survie des greffons ajouté par l'intégration du test du C1q aux autres paramètres du modèle multivarié final. Pour cela, nous avons comparé l'indice de concordance de Harrell (C-index) de deux modèles emboités : un modèle dit de « de référence » et un modèle intégrant le test du C1q dans le modèle de référence. Une procédure de ré-échantillonnage nous a permis de produire un intervalle de confiance autour de la différence des C des deux modèles. En l'absence de modèle validé de prédiction du risque de perte de greffon en transplantation, nous avons constitué un modèle « de référence » basé sur le résultat de l'étude multivariée des facteurs prédictifs classiques. Nous avons aussi utilisé les indices de reclassification tels que le NRI continu et l'IDI pour quantifier la performance dans la reclassification des probabilités de risque de perte du greffon apporté par l'ajout du statut C1q des DSA au modèle de référence.

VIII.1.2.4 Résultats

Nous avons identifié 3 groupes de patients selon le statut C1q des DSA détectés en post-transplantation : les patients sans DSA (DSA-, n=700, 69%),

les patients avec des DSA qui ne fixent pas le complément (DSA+/C1q-, n=239, 23%) et les patients avec des DSA qui fixent le complément (DSA+/C1q+, n=77, 8%). Les anticorps qui fixent le complément représentent 24% de l'ensemble des DSA détectés dans la première année post-transplantation.

La table 1 décrit l'ensemble des caractéristiques des 3 groupes de patients. Globalement, nous ne montrons pas de différences majeures entre les patients DSA+/C1q- et ceux DSA+/C1q+ en dehors de taux plus élevés de retransplantation et de receveurs immunisés dans le groupe DSA+/C1q+.

Table 1 : Caractéristiques de la population d'étude. Identification de 3 groupes de patients selon la détection d'anticorps anti-HLA spécifiques du donneur fixant le complément en post-transplantation**.**

DSA- — patients sans DSA post-transplantation. DSA+/C1q- — patients avec des DSA post-transplantation qui ne fixent pas le complément. DSA+/C1q+ — patients avec des DSA post-transplantation fixant le complément. J0 — jour de la transplantation.

p_Y — p résultant de la comparaison entre les patients DSA+/C1q- et DSA+/C1q+

** — p<0,01. ** — p<0,001. *** — p<0,0001 — p résultant de la comparaison entre les patients DSA- et DSA+.*

† — Moyennes ± SD

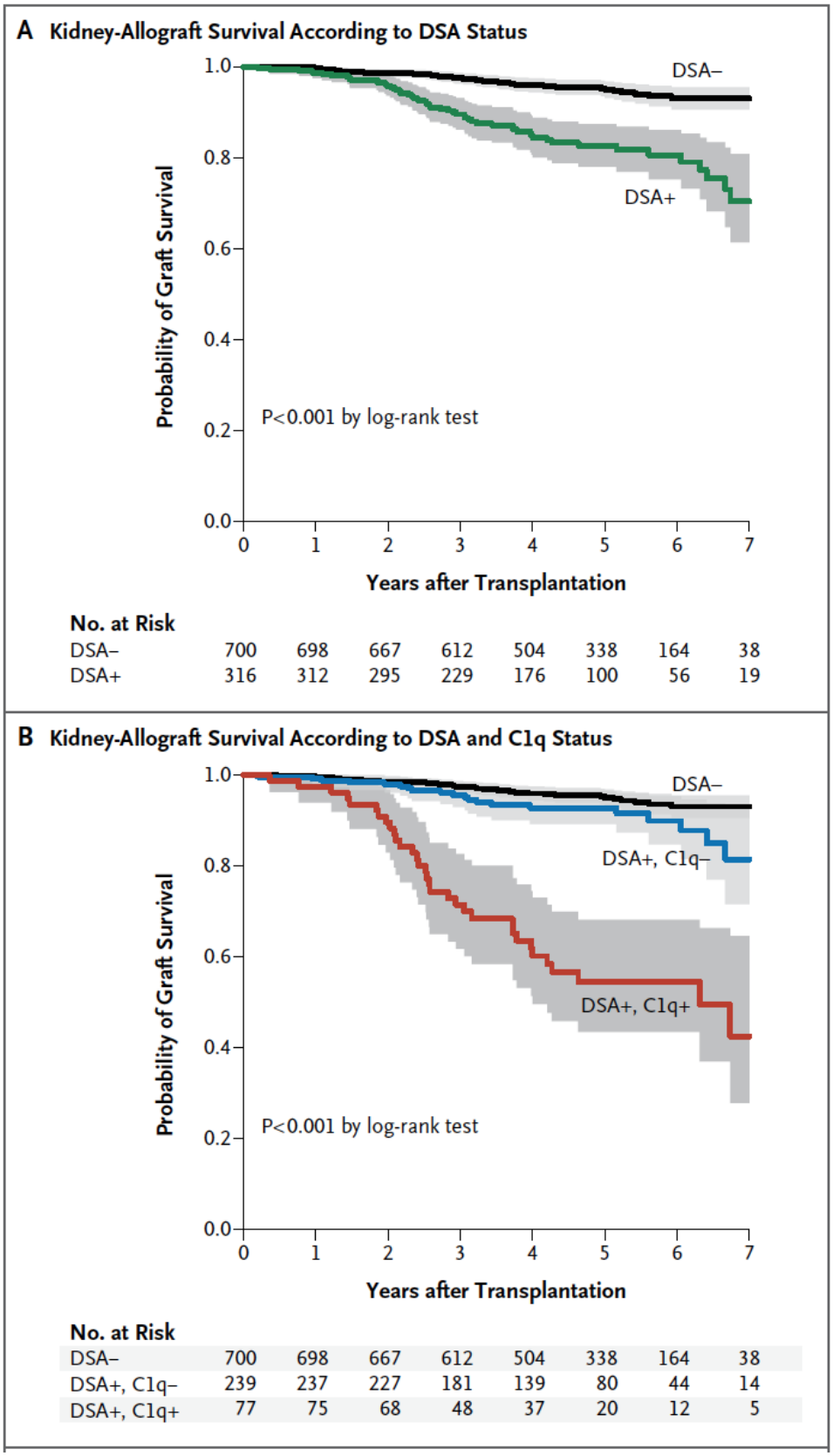
Patients	Population générale (N=1016)		DSA- (N=700)		DSA+ C1Q- (N=239)		DSA+ C1Q+ (N=77)		P [Y]
	N		N		N		N		
Age receveur — ans †	1016	47,6 ± 13	700	47,8 ± 13	239	46,5 ± 13	77	48,3 ± 13	0,2905
Receveur de sexe masculin — no. (%)	1016	599 (59%)	700	436 (62%) *	239	124 (52%)	77	39 (51%)	0,8506
Retransplantation — no. (%)	1016	181 (18%)	700	78 (11%)***	239	70 (29%)	77	33 (43%)	0,0272
Durée dialyse — ans †	874	4,9 ± 4,7	593	4,4 ± 4 **	208	5,6 ± 5	73	6,5 ± 6	0,2592
Age donneur (ans) †	1016	51 ± 16	700	51,3 ± 16	239	50,1 ± 17	77	51,8 ± 16	0,4419
Donneur de sexe masculin — no. (%)	1016	559 (55%)	700	387 (55%)	239	129 (54%)	77	43 (56%)	0,7746
Donneur cadavérique — no. (%)	1016	835 (82%)	700	550 (79%)***	239	213 (89%)	77	72 (94%)	0,2606
Ischémie froide (heures) †	1016	17 ± 9,6	700	16,6 ± 10 *	239	18,1 ± 9	77	20,3 ± 9	0,0598
Etiologie de la néphropathie	1016		700		239		77		
Diabète — no. (%)		101 (9,9%)	-	76 (10,9%)	-	19 (8,0%)	-	6 (7,8%)	
Vasculaire — no. (%)		82 (8,1%)	-	57 (%)	-	19 (8,0%)	-	6 (7,8%)	
Glomérulopathie — no. (%)		267 (26,3%)	-	180 (25,7%)	-	65 (27,2%)	-	22 (28,6%)	
Congénitale — no. (%)		185 (18,2%)	-	131 (18,7%)	-	43 (18,0%)	-	11 (14,3%)	
Autre — no. (%)		48 (4,7%)	-	27 (3,9%)	-	15 (6,3%)	-	6 (7,8%)	
Interstitielle no. (%)		120 (11,8%)	-	82 (11,7%)	-	29 (12,1%)	-	9 (11,7%)	
Indéterminée — no. (%)		213 (21,0%)	-	147 (21,0%)	-	49 (20,5%)	-	17 (22,1%)	0,9923
Immunologie									
HLA A/B/DR mis match †	1016	3,2 ± 1,5	700	3,1±1,5 **	239	3,5±1,4	77	3,3±1,4	0,1791
Groupe sanguin receveur A/B/O/AB — no.	1016	461/94/424/37	700	312/58/305/25	239	115/27/87/10	77	34/9/32/2	0,8015
Ac anti-HLA à J0 — no. (%)	1016	345 (34%)	700	132 (19%)***	239	153 (64%)	77	60 (78%)	<0,0001
DSA à J0 — no. (%)	1016	196 (19%)	700	10 (1%)***	239	128 (54%)	77	58 (75%)	<0,0001
Classe DSA à J0 Classe I/II/I+II — no	1016	125/246/82	700	8/6/5	239	81/98/57	77	36/42/20	0,8356

Parmi les patients DSA+/C1q-, 128 (54%) étaient immunisés au moment de la greffe, dont 22 (8%) présentaient des DSA qui fixaient le complément le jour de la transplantation. Dans le groupe des patients DSA+/C1q+, 58 (77%) étaient immunisés au moment de la transplantation, dont seulement 23 (30%) avaient des DSA avec la capacité à fixer le complément. La plupart des patients DSA+/C1q+ en post-transplantation (54 patients, 70%) présentaient des DSA détectables par technique Luminex au moment de la greffe, mais ces anticorps n'avaient pas la propriété de fixer le complément au moment de la greffe.

La présence de DSA/C1q+ circulant post transplantation s'associe à la présence de lésions humorales dans les greffons. Parmi les patients DSA+/C1q+, 67 (87%) ont des lésions d'inflammation au niveau de la microcirculation, 28 (36%) ont une inflammation interstitielle avec un score ≥ 2 et 17 (22%) ont une glomérulopathie du transplant. Nous observons un dépôt de la fraction C4d du complément chez 47 (61%) des patients DSA+/C1q+. Les patients DSA+/C1q+ ont des scores d'inflammation de la microcirculation, de glomérulopathie du transplant et de C4d supérieurs à ceux des patients DSA+/C1q- et des patients DSA-.

Les phénotypes immunologiques et histologiques s'associent aussi à des profils évolutifs distincts. La fonction rénale au moment des biopsies est significativement moins bonne chez les patients DSA+/C1q+ (DFG estimé à 42 ± 22 mL/min/1,73m2) que chez les patients DSA+/C1q- et chez les patients DSA- (DFG estimé à 51 ± 20 mL/min/1,73m2 et 54 ± 19 mL/min/1,73m2 respectivement, $p < 0,0001$).
Aussi, en termes de survie à 5 ans des greffons, nous démontrons que les patients DSA+/C1q+ ont des survies significativement moins bonnes (54%) que celles des patients DSA+/C1q– et des patients sans DSA (93% et 94% respectivement) (Figure 3, $p < 0,0001$) démontrant un rôle pronostique du C1q pour la survie des greffons.

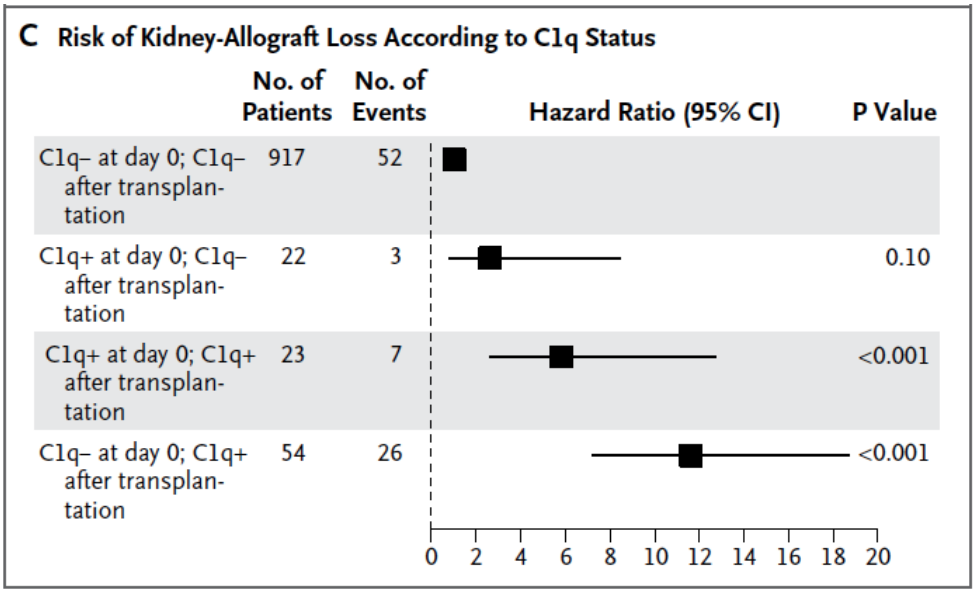
Figure 3: Courbes de survie des greffons selon le statut des DSA détectés en post-transplantation:
(Panel A) approche classique qui consiste à détecter les DSA par technique Luminex SA (Panel B) nouvelle approche basée sur la propriété des DSA à fixer le complément.


A Kidney-Allograft Survival According to DSA Status
1.0
0.8
0.6
0.4
0.2
0.0
Probability of Graft Survival
DSA–
DSA+
P<0.001 by log-rank test
0
1
2
3
4
5
6
7
Years after Transplantation
No. at Risk
DSA– 700 698 667 612 504 338 164 38
DSA+ 316 312 295 229 176 100 56 19
B Kidney-Allograft Survival According to DSA and C1q Status
1.0
0.8
0.6
0.4
0.2
0.0
Probability of Graft Survival
DSA–
DSA+, C1q–
DSA+, C1q+
P<0.001 by log-rank test
0
1
2
3
4
5
6
7
Years after Transplantation
No. at Risk
DSA– 700 698 667 612 504 338 164 38
DSA+, C1q– 239 237 227 181 139 80 44 14
DSA+, C1q+ 77 75 68 48 37 20 12 5

Le risque de perte de greffons est 5,9 fois plus important chez les patients DSA+/C1q+ par rapport aux patients DSA+/C1q- (95%CI 3,4-10,2). Ce risque reste significativement plus élevé chez les patients DSA+/C1q+ et à un niveau comparable, indifféremment du moment de la détection des DSA/C1q en post-transplantation : au moment du rejet (RR = 4,630, 95%CI 1,83-11,24, p = 0,0011) ou à 1 an post greffe (RR = 5,69, 95%CI, 1,67-19,40, p = 0,0054).
Pour 32% des patients avec des DSA dont les MFI sont supérieures à 6 000 (considérés à haut risque selon les résultats des études antérieures), les DSA n'ont pas la capacité à fixer le complément, et en conséquence leur risque de perte de greffons n'est pas plus élevé.

Nous observons une variabilité de la capacité des DSA à fixer le complément entre le moment de la transplantation et en post-transplantation. Les patients chez qui nous mettons en évidence l'apparition en post-transplantation de cette propriété (DSA+/C1q- à J0 devenant DSA+/C1q+ en post-transplantation) et les patients dont les DSA gardent cette capacité à fixer le complément en post-transplantation (DSA+/C1q+ à J0 restant DSA+/C1q+ en post-greffe) ont les risques les plus importants de perte de greffon (Figure 4).

Figure4 : Evaluation du risque relatif de perte de greffon selon l'évolution de la capacité des DSA à fixer le complément entre le moment de la greffe et en post-transplantation.



Le modèle multivarié final (backward stepwise) est présenté dans la Table 2

Table 2 : Modèle final de l'analyse multivariée des facteurs associés à la perte de greffons

Le modèle final est constitué de l'ensemble des paramètres significativement associés à la perte de greffons dans chacun des blocs.
DFG — débit de filtration glomérulaire. DSA — anticorps anti-HLA spécifiques du donneur.

Variable	No. of Patients	No. of Events	Hazard Ratio (95% CI)	P Value
Estimated GFR at 1 yr				
≥60 ml/min/1.73 m²	313	7	1.00	
≥30 and <60 ml/min/1.73 m²	579	36	2.45 (1.09–5.53)	
<30 ml/min/1.73 m²	111	42	12.49 (5.56–28.06)	<0.001
Interstitial fibrosis and tubular atrophy†				
Low score: 0 or 1	738	45	1.00	
High score: 2 or 3	265	40	2.22 (1.41–3.49)	0.005
Glomerular and peritubular inflammation and transplant glomerulopathy				
No	809	42	1.00	
Yes	194	43	2.26 (1.31–3.89)	0.003
C1q-binding donor-specific anti-HLA antibodies after transplantation				
No	928	52	1.00	
Yes	75	33	4.78 (2.69–8.49)	<0.001

* Risk factors were identified with the use of backward elimination, with a P value of 0.05 or lower for retention in the model.
† Banff scores range from 0 to 3, with higher scores indicating more severe abnormality.

En se basant sur cette association forte, nous avons construit un modèle de stratification du risque en transplantation rénale. Nous proposons cette stratification dans la première année après la transplantation : au moment du rejet ou à 1 an post-transplantation pour les patients sans rejet dans cette première année. Nous démontrons que l'addition de la propriété des DSA à fixer le complément augmente significativement les performances du modèle de référence, notamment sa capacité à discriminer les patients qui vont perdre ou non leur greffon (C-index 0,81 vs 0.85 ; différence moyenne obtenue par bootstrap 0.013 avec un IC à 95% de 0.012 à 0.015), la capacité de reclassification adéquate des patients (NRI = 0,75, 95% CI 0,54-0,97) et

leur discrimination (IDI = 0,03, p = 0,008).

VIII.1.2.5 Conclusion de l'étude

La propriété des DSA à fixer le complément augmente significativement le risque de perte de greffons dans les 5 ans post-transplantation et devient donc un facteur pronostique majeur en transplantation rénale. Nous avons identifié une population de patients ayant développé dans la 1ère année suivant la transplantation des DSA qui fixent le complément. Ils présentent un risque majeur de perte de greffons. Jusqu'à présent, l'impact de ces DSA ayant la propriété de fixer le complément, identifiés par des techniques sensibles, n'avait jamais été évalué en population.

En présence de DSA fixant le complément, le phénotype histologique observé dans les biopsies des greffons est dominé par la présence de lésions humorales.

Il n'existe pas de corrélation absolue entre la capacité des DSA à fixer le complément et le dépôt de la fraction C4d dans le greffon. Nous démontrons que la détection des DSA fixant le complément a une meilleure performance diagnostique que le C4d, marqueur classique d'activation du complément dans le greffon et aussi que la capacité des DSA à fixer le complément augmente la performance pronostique des DSA en post-transplantation.

VIII.1.2.6 Manuscrit

ORIGINAL ARTICLE

Complement-Binding Anti-HLA Antibodies and Kidney-Allograft Survival

Alexandre Loupy, M.D., Ph.D., Carmen Lefaucheur, M.D., Ph.D., Dewi Vernerey, M.P.H., Christof Prugger, M.D., Jean-Paul Duong van Huyen, M.D., Ph.D., Nuala Mooney, Ph.D., Caroline Suberbielle, M.D., Ph.D., Véronique Frémeaux-Bacchi, M.D., Ph.D., Arnaud Méjean, M.D., François Desgrandchamps, M.D., Dany Anglicheau, M.D., Ph.D., Dominique Nochy, M.D., Dominique Charron, M.D., Ph.D., Jean-Philippe Empana, M.D., Ph.D., Michel Delahousse, M.D., Christophe Legendre, M.D., Denis Glotz, M.D., Ph.D., Gary S. Hill, M.D.,* Adriana Zeevi, Ph.D., and Xavier Jouven, M.D., Ph.D.

ABSTRACT

BACKGROUND

Anti-HLA antibodies hamper successful transplantation, and activation of the complement cascade is involved in antibody-mediated rejection. We investigated whether the complement-binding capacity of anti-HLA antibodies plays a role in kidney-allograft failure.

METHODS

We enrolled patients who received kidney allografts at two transplantation centers in Paris between January 1, 2005, and January 1, 2011, in a population-based study. Patients were screened for the presence of circulating donor-specific anti-HLA antibodies and their complement-binding capacity. Graft injury phenotype and the time to kidney-allograft loss were assessed.

RESULTS

The primary analysis included 1016 patients. Patients with complement-binding donor-specific anti-HLA antibodies after transplantation had the lowest 5-year rate of graft survival (54%), as compared with patients with non–complement-binding donor-specific anti-HLA antibodies (93%) and patients without donor-specific anti-HLA antibodies (94%) (P<0.001 for both comparisons). The presence of complement-binding donor-specific anti-HLA antibodies after transplantation was associated with a risk of graft loss that was more than quadrupled (hazard ratio, 4.78; 95% confidence interval [CI], 2.69 to 8.49) when adjusted for clinical, functional, histologic, and immunologic factors. These antibodies were also associated with an increased rate of antibody-mediated rejection, a more severe graft injury phenotype with more extensive microvascular inflammation, and increased deposition of complement fraction C4d within graft capillaries. Adding complement-binding donor-specific anti-HLA antibodies to a traditional risk model improved the stratification of patients at risk for graft failure (continuous net reclassification improvement, 0.75; 95% CI, 0.54 to 0.97).

CONCLUSIONS

Assessment of the complement-binding capacity of donor-specific anti-HLA antibodies appears to be useful in identifying patients at high risk for kidney-allograft loss.

From the Paris Translational Research Center for Organ Transplantation (A.L., C. Lefaucheur, J.-P.D.H., M.D., D.G., C. Legendre, V.F.-B., X.J.), INSERM Unité 970 (A.L., C. Lefaucheur, D.V., C.P., J.-P.D.H., J.-P.E., X.J.), Department of Kidney Transplantation, Hôpital Necker, Université Paris Descartes, and Assistance Publique–Hôpitaux de Paris (AP-HP) (A.L., D.A., C. Legendre), the Departments of Nephrology and Kidney Transplantation (C. Lefaucheur, D.G.), Immunology and Histocompatibility (C.S., D.C.), and Urology (F.D.), Hôpital Saint-Louis, Paris Diderot University, AP-HP, and the Departments of Pathology (J.-P.D.H., D.N., G.S.H.), Immunology (V.F.-B.), and Urology (A.M.), Georges Pompidou European Hospital, INSERM Unité 940 (N.M., D.G.) — all in Paris; the Department of Nephrology and Kidney Transplantation, Foch Hospital, Suresnes, France (M.D.); and the Department of Transplantation Pathology, University of Pittsburgh Medical Center, Pittsburgh (A.Z.). Address reprint requests to Dr. Loupy at alexandreloupy@gmail.com.

Drs. Loupy and Lefaucheur and Drs. Zeevi and Jouven contributed equally to this article.

*Deceased.

N Engl J Med 2013;369:1215-26.
DOI: 10.1056/NEJMoa1302506
Copyright © 2013 Massachusetts Medical Society.

Despite considerable advances in transplantation, the induced alloimmune response remains a major determinant of late kidney-allograft loss.[1-3] In the United States and Europe, thousands of kidney transplants fail each year, and kidney-allograft failure is a major cause of end-stage renal disease, leading to increased morbidity, mortality, and costs.[4,5]

One of the most important advances in transplantation medicine has been the recognition that anti-HLA antibodies are destructive.[6-10] Various studies over the past decade have indicated that the alloimmune response, mediated by anti-HLA antibodies, plays a key role in the failure of kidney allografts; this concept has been extended to heart, lung, and composite tissue transplants.[6] Although anti-HLA antibodies are considered to be harmful, there is a wide spectrum of graft injury related to these antibodies, ranging from no recognizable damage to florid rejection.[11,12] Such a varied effect underscores the need to define distinct graft phenotypes and outcomes according to the presence or absence and characteristics of donor-specific anti-HLA antibodies after transplantation.

Since the pioneering discovery in 1969 that anti-HLA antibodies are lymphocytotoxic,[13] activation of the complement cascade has been considered to be a key component of antibody-mediated allograft rejection, and C4d deposition in renal capillaries has been considered the footprint of antibody-mediated allograft damage.[14-16]

The capacity of anti-HLA antibodies to bind complement fraction C1q, which is the first step in activation of the classic complement cascade, determines the cytotoxic potential of these antibodies, and an assessment of their complement-binding capacity may be useful both for risk stratification and for diagnosis of antibody-mediated rejection. Small studies have suggested that the C1q-binding properties of donor-specific anti-HLA antibodies may be specifically related to antibody-mediated rejection; these findings provide support for the general principle that complement-binding donor-specific anti-HLA antibodies have a role in the pathogenesis of antibody-mediated rejection.[17-24] We hypothesized that the complement-binding properties of donor-specific anti-HLA antibodies detected after transplantation are involved in kidney-allograft failure.

We conducted a study to define the full spectrum of kidney-allograft injury according to the C1q-binding properties of donor-specific anti-HLA antibodies in a large population-based study and to determine whether assessment for the presence of C1q-binding donor-specific anti-HLA antibodies after transplantation might improve risk stratification for kidney-allograft loss.

METHODS

STUDY POPULATION

We enrolled all consecutive patients who underwent kidney transplantation at Necker Hospital and Saint-Louis Hospital (Paris) between January 1, 2005, and January 1, 2011, in this population-based study. Patients were followed until April 15, 2012. We also included an external-validation cohort comprising patients who underwent kidney transplantation at Foch Hospital (Suresnes, France) between January 1, 2004, and January 31, 2010 (see the Methods section in the Supplementary Appendix, available with the full text of this article at NEJM.org). The study was approved by the institutional review boards of Necker Hospital, Saint-Louis Hospital, and Foch Hospital. Written informed consent was obtained from all patients. One Lambda donated reagents but was not otherwise involved in either the conduct of the study or the preparation of the manuscript.

The transplantation allocation system was identical for the three centers and followed the rules of the French national agency for organ procurement (Agence de la Biomédecine). All transplants were compatible with the ABO blood group. A negative result of cross-matching for IgG T-cell and B-cell complement-dependent cytotoxicity was required for all recipients.

CLINICAL DATA

Clinical data on the donors and recipients in the derivation cohort (at Necker and Saint-Louis Hospitals) and the validation cohort (at Foch Hospital) were obtained from two national registries, Données Informatiques Validées en Transplantation (Necker Hospital) and Agence de la Biomédecine (Saint-Louis and Foch Hospitals). Anonymized data from these registries are prospectively entered at specific time points for each patient (on day 0 and 6 months and 1 year after transplantation) and are updated annually thereafter[25,26] (see the Methods section in the Supplementary Appen-

dix). The derivation-cohort data were obtained from the database on April 15, 2012, whereas the validation-cohort data were obtained on December 19, 2012. We documented all cases of acute clinical rejection, defined by deterioration in graft function, proteinuria, or impaired function and histopathological evidence of rejection, according to the consensus rules of the international Banff classification criteria,[27,28]

Immunosuppression protocols and treatment of allograft-rejection episodes after transplantation were similar among the centers.[29,30] The protocols and treatments are described in the Methods section in the Supplementary Appendix.

HISTOLOGIC AND IMMUNOCHEMICAL TESTS

We used specimens from protocol-specified graft biopsies performed 1 year after transplantation in 845 patients without any acute clinical rejection episodes diagnosed in the first year after transplantation, as well as specimens from biopsies performed in 171 patients with acute allograft rejection during the first year after transplantation. All graft-biopsy specimens were scored and graded from 0 to 3 according to the updated Banff criteria[27,28] by three trained pathologists who were unaware of the patient's status with respect to the presence of donor-specific anti-HLA antibodies, C1q-binding status, and clinical course (see the Methods section in the Supplementary Appendix). C4d staining was performed by means of immunochemical analysis on paraffin sections with the use of polyclonal human anti-C4d antibodies (Biomedica Gruppe).

DETECTION AND CHARACTERIZATION OF DONOR-SPECIFIC ANTIBODIES

All patients were tested for the presence of circulating donor-specific anti-HLA antibodies in banked serum samples (at the Jean Dausset Histocompatibility Laboratory, Paris) obtained at the time of transplantation (day 0) and in serum samples obtained at the time of the biopsy (1 year after transplantation or during an episode of acute rejection in the first year after transplantation). The presence of circulating donor-specific anti-HLA-A, -B, -Cw, -DR, -DQ, and -DP antibodies was retrospectively determined with the use of single-antigen flow bead assays (One Lambda) on a Luminex platform.

Serum samples from patients with circulating donor-specific anti-HLA antibodies were analyzed in a blinded fashion at the University of Pittsburgh for the presence of C1q-binding donor-specific anti-HLA antibodies with the use of single-antigen flow bead assays according to the manufacturer's protocol (C1qScreenTM, One Lambda).[17,18,21] For details, see the Methods section in the Supplementary Appendix.

STATISTICAL ANALYSIS

We used means and standard deviations for the description of continuous variables, with the exception of mean fluorescence intensity, for which we used the mean and standard error. We compared means and proportions using Student's t-test and the chi-square test (or Fisher's exact test if appropriate). Survival was analyzed from the time of transplantation to a maximum of 7 years, with kidney-graft loss as the event of interest. Data on graft survival were censored at the time of death.[31] Rates of kidney-allograft survival were plotted on Kaplan–Meier curves and compared according to anti-HLA antibody status with the use of the log-rank test. Cox proportional-hazards models were used to quantify hazard ratios and 95% confidence intervals for kidney-graft loss.

The association of clinical, histologic, functional, and immunologic factors with graft loss was assessed in separate univariate and multivariate Cox regression analyses. The factors identified in these analyses were thereafter included in a final multivariable model with stepwise backward elimination.

The predictive value that C1q-binding status added to a reference risk model (including independent predictors of the final multivariable model plus circulating donor-specific anti-HLA antibodies after transplantation) was evaluated with the use of the C-statistic. This analysis was repeated 1000 times with the use of bootstrap samples to derive 95% confidence intervals for the difference in the C-statistic between models. We calculated the continuous net reclassification improvement and the integrated discrimination improvement associated with the addition of C1q to the reference model.[32,33] Results for complement-binding donor-specific anti-HLA antibodies and kidney-allograft survival were replicated in the independent validation sample. Analyses were conducted with the use of SAS software, version 9.2 (SAS Institute), and R software (version 2.10.1).

All tests were two-sided, and P values less than 0.05 were considered to indicate statistical significance.

RESULTS

BASELINE CHARACTERISTICS OF THE KIDNEY-ALLOGRAFT RECIPIENTS

In total, 1016 patients undergoing renal transplantation (695 at Necker Hospital and 321 at Saint-Louis Hospital) were included in the main analysis. Three distinct populations were identified after transplantation, according to the presence or absence of donor-specific anti-HLA antibodies and complement-binding capacity: 700 patients without circulating donor-specific anti-HLA antibodies, 239 patients with non–complement-binding donor-specific anti-HLA antibodies, and 77 patients with complement-binding donor-specific anti-HLA antibodies. Table 1 shows the characteristics of the donors and recipients at the time of renal transplantation.

KIDNEY-ALLOGRAFT INJURY

In the first year after transplantation, acute clinical rejection developed in 171 patients: 96 patients had T-cell–mediated rejection (56%) and 75 had antibody-mediated rejection (44%). T-cell–mediated rejection occurred in 14 of 77 patients with donor-specific anti-HLA antibodies plus C1q-binding capacity (18%), in 30 of 239 patients with donor-specific anti-HLA antibodies without C1q-binding capacity (13%), and in 52 of 700 patients without donor-specific anti-HLA antibodies (7%) (P<0.001). Antibody-mediated rejection occurred in 37 patients with C1q-binding donor-specific anti-HLA antibodies (48%) and 38 patients with non–C1q-binding donor-specific anti-HLA antibodies (16%) (P<0.001).

Among the patients with C1q-binding donor-specific anti-HLA antibodies, 67 had microvascular inflammation (87%), 28 had tubular and interstitial inflammation scores of 2 or higher (on a scale of 0 to 3, with higher scores indicating more severe abnormality) (36%), 18 had endarteritis (23%), 17 had transplant glomerulopathy (22%), 30 had moderate-to-severe arteriosclerosis (39%), 23 had moderate-to-severe atrophy-scarring lesions (interstitial fibrosis and tubular atrophy) (30%), and 47 had C4d deposition in peritubular capillaries (61%). Patients with C1q-binding donor-specific anti-HLA antibodies had more extensive microvascular inflammation and transplant glomerulopathy and higher scores for graft peritubular capillary C4d deposition than both patients with non–C1q-binding donor-specific anti-HLA antibodies and patients without donor-specific anti-HLA antibodies (Fig. 1). Stratified analyses revealed that these increases applied to samples from protocol-specified biopsies, performed at 1 year, and samples from biopsies performed during an acute-rejection episode in the first year (Fig. S1 in the Supplementary Appendix).

Patients with C1q-binding donor-specific anti-HLA antibodies had a lower estimated glomerular filtration rate (GFR) at 1 year (42±22 ml per minute per 1.73 m^2 of body-surface area) than did patients with non–C1q-binding donor-specific anti-HLA antibodies (51±20 ml per minute per 1.73 m^2) and patients without donor-specific anti-HLA antibodies (54±19 ml per minute per 1.73 m^2) (P<0.001).

KIDNEY-ALLOGRAFT SURVIVAL

The median follow-up after transplantation was 4.8 years (range, 0.2 to 7.0). The median follow-up times were 3.9 years (range, 0.4 to 7.0) in patients with C1q-binding donor-specific anti-HLA antibodies and 4.3 years (range, 0.2 to 7.0) in patients with non–C1q-binding donor-specific anti-HLA antibodies.

Figure 2A shows kidney-allograft survival according to donor-specific anti-HLA antibody status after transplantation. Patients with donor-specific anti-HLA antibodies had significantly worse graft survival than patients without donor-specific anti-HLA antibodies (5-year graft survival after transplantation, 83% vs. 94%; P<0.001 by the log-rank test). When patients with donor-specific anti-HLA antibodies after transplantation were subsequently categorized according to complement-binding capacity, patients with C1q-binding donor-specific anti-HLA antibodies had the poorest 5-year graft survival after transplantation (54%), as compared with patients with non–C1q-binding donor-specific anti-HLA antibodies and patients without donor-specific anti-HLA antibodies (93% and 94%, respectively; P<0.001 for both comparisons) (Fig. 2B). The risk of graft loss according to the donor-specific anti-HLA antibodies–C1q status at day 0 and the status after transplantation revealed that patients with C1q-binding donor-specific anti-HLA antibodies after transplantation had the highest risk of graft loss (Fig. 2C).

Table 1. Baseline Characteristics of the Study Population, According to the Presence or Absence of Donor-Specific Anti-HLA Antibodies and C1q Binding after Transplantation.*

Characteristic	All Patients (N=1016)	Patients without Donor-Specific Antibodies (N=700)	Patients with Donor-Specific Antibodies		P Value†
			Without C1q Binding (N=239)	With C1q Binding (N=77)	
Recipients					
Age — yr	47.6±13	47.8±13	46.5±13	48.3±13	0.29
Male sex — no. (%)‡	599 (59)	436 (62)§	124 (52)	39 (51)	0.85
Retransplantation — no. (%)‡	181 (18)	78 (11)¶	70 (29)	33 (43)	0.03
Time since dialysis — yr ‖	4.9±4.7	4.4±4¶	5.6±5	6.5±6	0.26
Donors					
Age — yr	51.0±16	51.3±16	50.1±17	51.8±16	0.44
Male sex — no. (%)‡	559 (55)	387 (55)	129 (54)	43 (56)	0.77
Deceased — no. (%)‡	835 (82)	550 (79)¶	213 (89)	72 (94)	0.26
Cold-ischemia time — hr	17.0±9.6	16.6±10§	18.1±9	20.3±9	0.06
Immunologic characteristics					
HLA A/B/DR mismatch	3.2±1.5	3.1±1.5¶	3.5±1.4	3.3±1.4	0.18
Recipient blood type — no.‡					0.80
A	461	312	115	34	
B	94	58	27	9	
O	424	305	87	32	
AB	37	25	10	2	
Characteristics of anti-HLA antibodies at time of transplantation					
Anti-HLA antibodies — no. (%)	345 (34)	132 (19)¶	153 (64)	60 (78)	<0.001
Donor-specific anti-HLA antibodies — no. (%)	196 (19)	10 (1)¶	128 (54)	58 (75)	<0.001
HLA class of donor-specific anti-HLA antibodies — no.					0.84
I	125	8	81	36	
II	146	6	98	42	
I and II	82	5	57	20	
C1q-binding donor-specific anti-HLA antibodies — no. (%)	45 (4)	0	22 (9)	23 (30)	<0.001

* Plus–minus values are means ±SD.
† P values are for the comparison between non–C1q-binding donor-specific anti-HLA antibodies and C1q-binding donor-specific anti-HLA antibodies.
‡ Chi-square tests were used for the comparison of categorical variables, and the unpaired t-test was used for the comparison of continuous variables.
§ P<0.01 for the comparison between no donor-specific anti-HLA antibodies and donor-specific anti-HLA antibodies.
¶ P<0.001 for the comparison between no donor-specific anti-HLA antibodies and donor-specific anti-HLA antibodies.
‖ Time since dialysis was determined for 874 patients overall: 593 patients who had no donor-specific anti-HLA antibodies, 208 patients who had non–C1q-binding donor-specific anti-HLA antibodies, and 73 patients who had C1q-binding donor-specific anti-HLA antibodies.

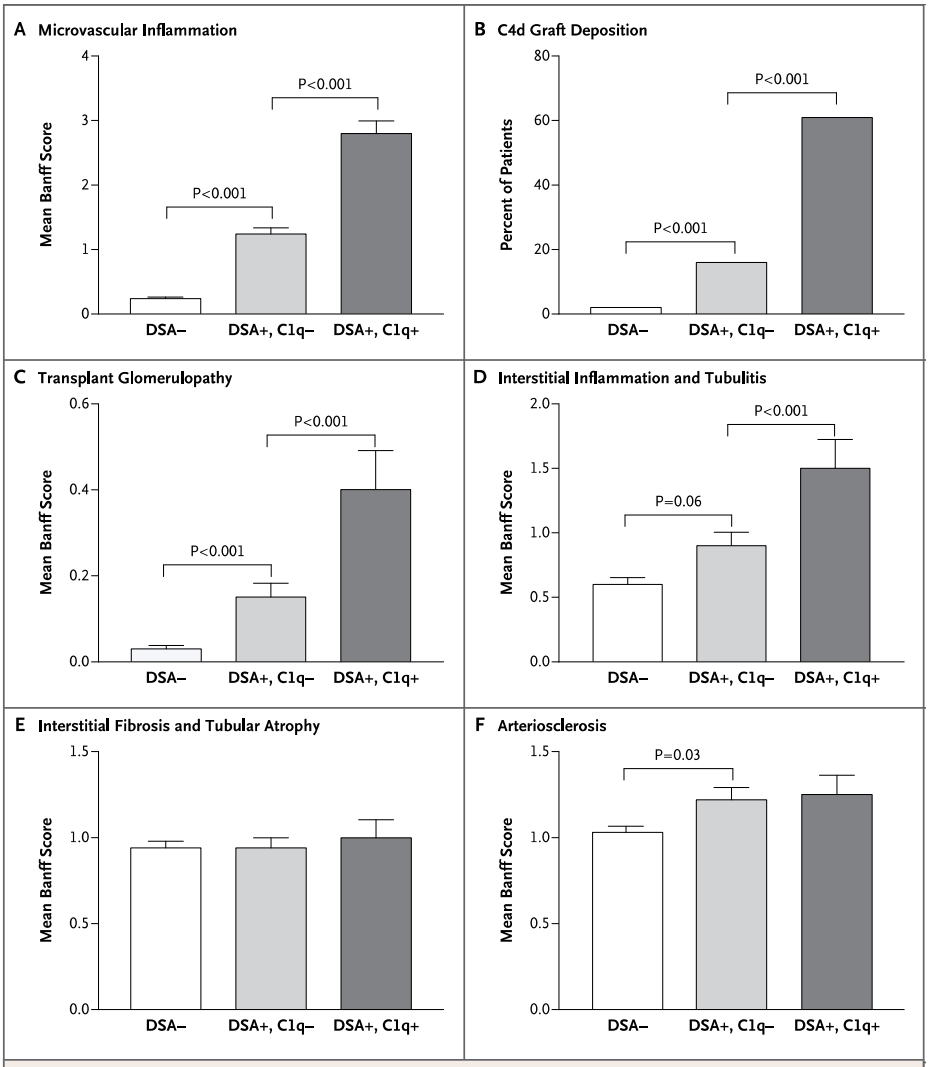


Figure 1. Phenotypic Features of Kidney-Allograft Injury, According to Anti-HLA Antibody Status after Transplantation.

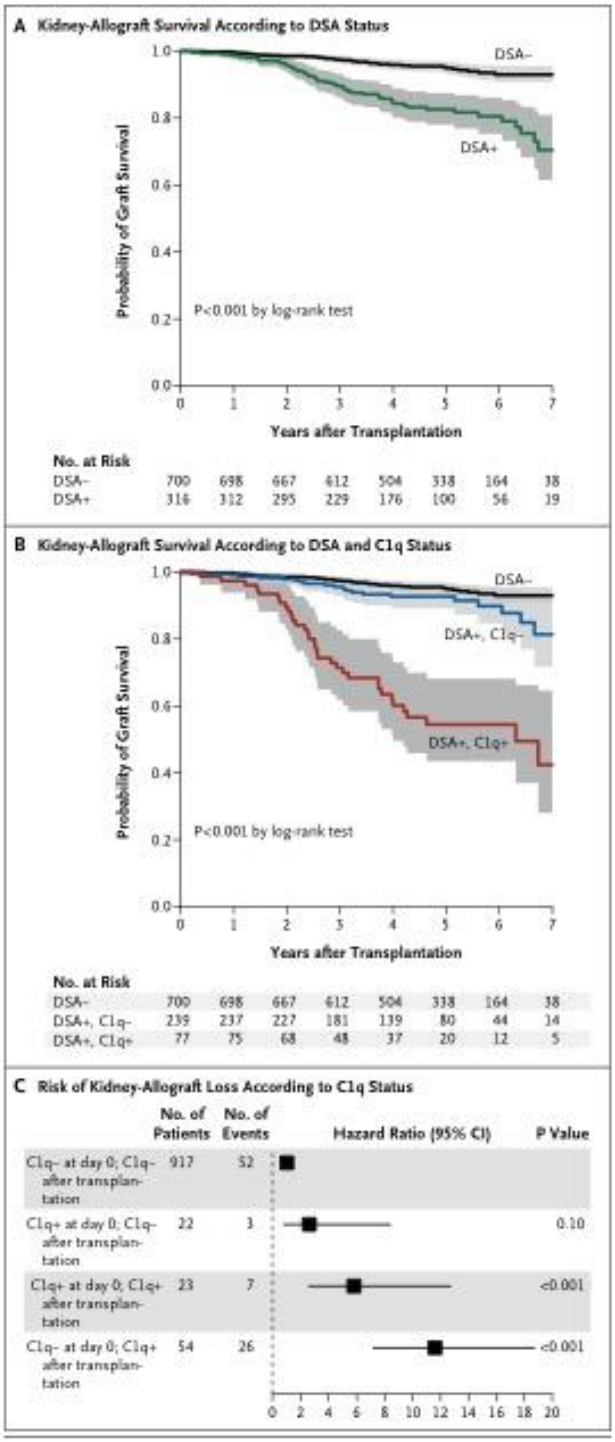
Data are based on 1016 kidney-allograft biopsies performed in the first year after transplantation (845 at 1 year after transplantation and 171 during acute rejection in the first year). The score for microvascular inflammation (Panel A) is the sum of the Banff scores for glomerulitis and peritubular capillaritis. Each of these scores ranges from 0 to 3, with higher scores indicating more severe abnormality. The T bars indicate standard errors. DSA denotes donor-specific anti-HLA antibody.

DETERMINANTS OF KIDNEY-ALLOGRAFT LOSS

The association of clinical, functional, histologic, and immunologic factors with graft loss in univariate and backward-elimination multivariate Cox regression analysis is shown in Tables 2 and 3. The following independent predictors of graft loss were identified: low estimated GFR at 1 year (hazard ratio, 12.49; 95% confidence interval [CI], 5.56 to 28.06; P<0.001), interstitial fibrosis and tubular atrophy (hazard ratio, 2.22; 95% CI, 1.41 to 3.49; P=0.005), glomerular and peritubular inflammation and transplant glomerulopathy (hazard ratio, 2.26; 95% CI, 1.31 to 3.89; P=0.003), and the presence of complement-binding donor-

Figure 2. Kaplan–Meier Curves for Kidney-Graft Survival, According to Donor-Specific Anti-HLA Antibody Status after Transplantation.

Panel A shows the classic approach to determining the probability of graft survival, which is based on the presence or absence of donor-specific anti-HLA antibodies. Panel B shows our approach, which is based on the presence or absence of donor-specific anti-HLA antibodies and their C1q-binding capacity. Panel C shows the risk of graft loss according to C1q status at day 0 and after transplantation.



specific anti-HLA antibodies after transplantation (hazard ratio, 4.78; 95% CI, 2.69 to 8.49; P<0.001). Complement-binding donor-specific anti-HLA antibodies remained independently associated with the risk of kidney-allograft loss after adjustment for the mean fluorescence intensity of donor-specific anti-HLA antibodies (hazard ratio, 4.48; 95% CI, 2.23 to 8.98; P<0.001) (Table S1 in the Supplementary Appendix). The Kaplan–Meier curves for graft survival stratified according to status with respect to donor-specific anti-HLA antibodies, C1q status, and mean fluorescence intensity of donor-specific anti-HLA antibodies showed that patients with C1q-binding donor-specific anti-HLA antibodies had similar graft survival, regardless of whether the mean fluorescence intensity was low (<6000 arbitrary units) or high (≥6000 arbitrary units, P=0.70) (Fig. S2 in the Supplementary Appendix).

SENSITIVITY ANALYSIS

In the sensitivity analysis, we assessed the robustness of our study results by investigating associations separately in each study center and according to kidney function and the timing of biopsies. First, at both centers, patients with complement-binding donor-specific anti-HLA antibodies had the lowest rate of graft survival (Fig. S3 in the Supplementary Appendix). Second, complement-binding donor-specific anti-HLA antibodies after transplantation remained independently associated with graft loss whether they were detected in specimens from protocol-specified biopsies performed at 1 year (hazard ratio, 5.7; 95% CI, 1.7 to 19.4; P=0.005) or in specimens from biopsies performed during acute rejection in the first year (hazard ratio, 4.6; 95% CI, 1.8 to 11.2; P=0.001). Furthermore, the addition of an acute-rejection variable to the final multivariate model did not modify the significant predictors of graft loss. Third, C1q-binding donor-specific anti-HLA

Table 2. Clinical, Functional, Histologic, and Immunologic Factors Associated with Kidney-Graft Loss (Univariate Analysis).*

Variable	No. of Patients	Hazard Ratio (95% CI)	P Value
Clinical factors			
Age per 1-yr increment			
Donor	1016	1.02 (1.01–1.04)	0.001
Recipient	1016	1.01 (1.00–1.03)	0.10
Cold-ischemia time per 1-min increment	1016	1.001 (1.000–1.001)	0.001
Donor type			
Living	181	1.00	
Deceased	835	3.76 (1.52–9.26)	0.004
Donor sex			
Male	559	1.00	
Female	457	1.45 (0.96–2.21)	0.08
Recipient sex			
Male	599	1.00	
Female	417	0.96 (0.63–1.48)	0.87
Graft rank			
First transplant	835	1.00	
Subsequent transplant	181	1.92 (1.22–3.04)	0.005
No. of HLA mismatches	1016	1.04 (0.91–1.20)	0.56
Functional factors			
Estimated GFR at 1 yr†			
≥60 ml/min/1.73 m²	313	1.00	
≥30 and <60 ml/min/1.73 m²	579	2.87 (1.28–6.44)	
<30 ml/min/1.73 m²	111	21.96 (9.86–48.90)	<0.001
Histologic factors‡			
Interstitial fibrosis and tubular atrophy			
Low score: 0 or 1	743	1.00	
High score: ≥2	273	2.55 (1.68–3.87)	<0.001
Arteriosclerosis			
Low score: 0 or 1	653	1.00	
High score: ≥2	363	1.99 (1.31–3.02)	0.001
Interstitial inflammation and tubulitis			
Low score: 0 or 1	832	1.00	
High score: ≥2	184	1.68 (1.06–2.68)	0.028
Glomerular and peritubular inflammation and transplant glomerulopathy			
Low score: 0	820	1.00	
High score: ≥1	196	4.80 (3.16–7.29)	<0.001
Endarteritis			
Low score: 0	949	1.00	
High score: ≥1	64	2.81 (1.56–5.07)	<0.001
C4d graft deposition			
Low score: 0	919	1.00	
High score: ≥1	97	5.91 (3.78–9.24)	<0.001

Table 2. (Continued.)

Variable	No. of Patients	Hazard Ratio (95% CI)	P Value
Immunologic factors			
Donor-specific antibodies at time of transplantation			
No	820	1.00	
Yes	196	3.53 (2.30–5.40)	<0.001
C1q-binding donor-specific antibodies at time of transplantation			
No	971	1.00	
Yes	45	2.95 (1.53–5.70)	0.001
Donor-specific antibodies after transplantation			
No	700	1.00	
Yes	316	3.90 (2.54–5.98)	<0.001
C1q-binding donor-specific antibodies after transplantation			
No	939	1.00	
Yes	77	9.23 (5.99–14.23)	<0.001

* CI denotes confidence interval.
† The estimated glomerular filtration rate (GFR) was calculated with the use of the Modification of Diet in Renal Disease formula.
‡ Banff scores range from 0 to 3, with higher scores indicating more severe abnormality.

antibodies were associated with an increased risk of graft loss in each category of estimated GFR as defined by the National Kidney Foundation[34] (Fig. S4 in the Supplementary Appendix).

PREDICTION OF KIDNEY-ALLOGRAFT LOSS

The inclusion of complement-binding donor-specific anti-HLA antibodies in the reference model significantly improved its discrimination capacity (i.e., its capacity to discriminate between patients with graft loss and those without graft loss) since the C-statistic increased from 0.81 to 0.85 (bootstrap mean difference, 0.013; 95% CI, 0.012 to 0.015), and the integrated discrimination improvement was 0.03 (P = 0.008). Similarly, the addition of complement-binding donor-specific anti-HLA antibodies to the reference model adequately reclassified patients at lower risk for graft loss and those at higher risk, as shown by a continuous net reclassification improvement of 0.75 (95% CI, 0.54 to 0.97).

EXTERNAL VALIDATION

The external-validation cohort was composed of 643 patients; their baseline characteristics are detailed in Table S2 in the Supplementary Appendix. The median follow-up after transplantation was 3.4 years (range, 0.2 to 5.0). The Kaplan–Meier estimate of graft survival confirmed that patients with complement-binding donor-specific anti-HLA antibodies had the highest risk of graft loss as compared with patients with non–complement-binding donor-specific anti-HLA antibodies and patients without donor-specific anti-HLA antibodies (P<0.001) (Fig. S5 in the Supplementary Appendix).

DISCUSSION

In a cohort of 1016 carefully phenotyped kidney-transplant recipients, we observed that the presence of complement-binding donor-specific anti-HLA antibodies detected in the first year after transplantation was an independent predictor of kidney-allograft loss more than 5 years after transplantation and significantly improved individual risk stratification for graft failure. Patients with complement-binding donor-specific anti-HLA antibodies after transplantation had a graft injury phenotype characterized by microvascular inflammation and complement split-product C4d deposition. As compared with a traditional approach

Table 3. Clinical, Functional, Histologic, and Immunologic Factors Associated with Kidney-Graft Loss (Multivariate Analysis).*

Variable	No. of Patients	No. of Events	Hazard Ratio (95% CI)	P Value
Estimated GFR at 1 yr				
≥60 ml/min/1.73 m²	313	7	1.00	
≥30 and <60 ml/min/1.73 m²	579	36	2.45 (1.09–5.53)	
<30 ml/min/1.73 m²	111	42	12.49 (5.56–28.06)	<0.001
Interstitial fibrosis and tubular atrophy†				
Low score: 0 or 1	738	45	1.00	
High score: 2 or 3	265	40	2.22 (1.41–3.49)	0.005
Glomerular and peritubular inflammation and transplant glomerulopathy				
No	809	42	1.00	
Yes	194	43	2.26 (1.31–3.89)	0.003
C1q-binding donor-specific anti-HLA antibodies after transplantation				
No	928	52	1.00	
Yes	75	33	4.78 (2.69–8.49)	<0.001

* Risk factors were identified with the use of backward elimination, with a P value of 0.05 or lower for retention in the model.
† Banff scores range from 0 to 3, with higher scores indicating more severe abnormality.

to predicting graft loss, based only on the presence of donor-specific anti-HLA antibodies after transplantation as a risk factor for graft loss, the approach we used, which integrates the capacity of donor-specific anti-HLA antibodies to bind complement, identified an additional group of patients with an increased risk of graft loss.

It has been known for at least 40 years that cytotoxic anti-HLA antibodies are associated with graft rejection.[13] However, the complement-dependent cytotoxicity assay used to identify these clinically deleterious antibodies lacks sensitivity and specificity and cannot be used on a large scale in transplantation follow-up because of the limited reserves of donor cells. Of the three complement-dependent pathways, the classical pathway involves antibody–C1q fixation. Antibody binding to an antigen and subsequent C1q binding initiate activation of the complement cascade.[35]

A sensitive detection method for C1q-binding anti-HLA antibodies yield findings in addition to those of the single-antigen flow bead test. Although the mean fluorescence intensity of donor-specific anti-HLA antibodies is widely used in clinical practice for risk stratification, our study showed that C1q-binding donor-specific anti-HLA antibodies remained strongly associated with the risk of graft loss regardless of the mean fluorescence intensity of the antibodies. Our results may have implications for other transplanted organs such as the heart, lungs, and small bowel, since accumulating evidence from small studies supports the deleterious effect of C1q-binding donor-specific anti-HLA antibodies on these other transplanted organs.[18-21]

From a prognostic perspective, our results provide support for the finding that C1q and C4d do not provide equivalent predictive information. C1q testing may help identify patients at risk, despite C4d negativity. We and others[36,37] have found that C4d, although specific, lacks sensitivity as an indicator of humoral rejection. Assessment for complement-binding donor-specific anti-HLA antibodies may provide an early indication of the potential for complement-mediated injury, without the functional requirement for progression through the pathway to C4d deposition. Detection of complement-binding donor-specific anti-HLA antibodies may indicate which of the antibodies present have the capacity to activate the complement cascade,

potentially but not inevitably leading to C4d deposition and complement-mediated damage or antibody-mediated injury in vivo. Such findings do not rule out a possible role for complement-independent mechanisms in allograft injury mediated by complement-binding donor-specific anti-HLA antibodies.[12]

There are therapeutic implications of identifying critical pathologic pathways responsible for kidney-allograft loss. Since promising therapeutic agents targeting complement (e.g., a C5 inhibitor [eculizumab] or a C1 inhibitor) are increasingly used in patients undergoing transplantation,[38-40] the present study may provide a basis for future clinical trials. One limitation of our study was that it was not designed to provide kinetics of the capacity of anti-HLA antibodies to bind complement or the effect of treatment on these antibodies. This would require further investigations.

In conclusion, we systematically evaluated immunologic characteristics before and after transplantation in a population-based sample of kidney-allograft recipients, incorporating the full spectrum of graft phenotypes. We found that the presence of complement-binding anti-HLA donor-specific antibodies after transplantation is strongly associated with graft injury and loss and that incorporation of this risk factor improves risk stratification for graft loss.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank the members of the Laboratory of Excellence, Transplantex for helpful discussions.

REFERENCES

1. Nankivell BJ, Kuypers DR. Diagnosis and prevention of chronic kidney allograft loss. Lancet 2011;378:1428-37.
2. Gaston RS, Cecka JM, Kasiske BL, et al. Evidence for antibody-mediated injury as a major determinant of late kidney allograft failure. Transplantation 2010;90:68-74.
3. Sellarés J, de Freitas DG, Mengel M, et al. Understanding the causes of kidney transplant failure: the dominant role of antibody-mediated rejection and nonadherence. Am J Transplant 2012;12:388-99.
4. Organ Procurement and Transplantation Network and Scientific Registry of Transplant Recipients 2010 data report. Am J Transplant 2012;12:Suppl 1:1-156.
5. Rapport 2011 de l'activité de prélèvement et de greffe de l'Agence de la Biomédecine (http://www.agence-biomedecine.fr/Rapport-annuel-2011).
6. Terasaki PI, Ozawa M. Predicting kidney graft failure by HLA antibodies: a prospective trial. Am J Transplant 2004;4:438-43.
7. Hourmant M, Cesbron-Gautier A, Terasaki PI, et al. Frequency and clinical implications of development of donor-specific and non-donor-specific HLA antibodies after kidney transplantation. J Am Soc Nephrol 2005;16:2804-12.
8. Lachmann N, Terasaki PI, Budde K, et al. Anti-human leukocyte antigen and donor-specific antibodies detected by Luminex posttransplant serve as biomarkers for chronic rejection of renal allografts. Transplantation 2009;87:1505-13.
9. Lefaucheur C, Loupy A, Hill GS, et al. Preexisting donor-specific HLA antibodies predict outcome in kidney transplantation. J Am Soc Nephrol 2010;21:1398-406.
10. Nankivell BJ, Alexander SI. Rejection of the kidney allograft. N Engl J Med 2010;363:1451-62.
11. Racusen LC, Haas M. Antibody-mediated rejection in renal allografts: lessons from pathology. Clin J Am Soc Nephrol 2006;1:415-20.
12. Colvin RB. Antibody-mediated renal allograft rejection: diagnosis and pathogenesis. J Am Soc Nephrol 2007;18:1046-56.
13. Patel R, Terasaki PI. Significance of the positive crossmatch test in kidney transplantation. N Engl J Med 1969;280:735-9.
14. Feucht HE, Felber E, Gokel MJ, et al. Vascular deposition of complement-split products in kidney allografts with cell-mediated rejection. Clin Exp Immunol 1991;86:464-70.
15. Collins AB, Schneeberger EE, Pascual MA, et al. Complement activation in acute humoral renal allograft rejection: diagnostic significance of C4d deposits in peritubular capillaries. J Am Soc Nephrol 1999;10:2208-14.
16. Böhmig GA, Exner M, Habicht A, et al. Capillary C4d deposition in kidney allografts: a specific marker of alloantibody-dependent graft injury. J Am Soc Nephrol 2002;13:1091-9.
17. Chen G, Sequeira F, Tyan DB. Novel C1q assay reveals a clinically relevant subset of human leukocyte antigen antibodies independent of immunoglobulin G strength on single antigen beads. Hum Immunol 2011;72:849-58.
18. Chin C, Chen G, Sequeria F, et al. Clinical usefulness of a novel C1q assay to detect immunoglobulin G antibodies capable of fixing complement in sensitized pediatric heart transplant patients. J Heart Lung Transplant 2011;30:158-63.
19. Yabu JM, Higgins JP, Chen G, Sequeira F, Busque S, Tyan DB. C1q-fixing human leukocyte antigen antibodies are specific for predicting transplant glomerulopathy and late graft failure after kidney transplantation. Transplantation 2011;91:342-7.
20. Sutherland SM, Chen G, Sequeira FA, Lou CD, Alexander SR, Tyan DB. Complement-fixing donor-specific antibodies identified by a novel C1q assay are associated with allograft loss. Pediatr Transplant 2012;16:12-7.
21. Zeevi A, Marrari M, Feingold B, Webber S, Duquesnoy RJ. Human leukocyte antigen epitope analysis to assess complement- and non-complement-binding donor-specific antibody repertoire in a pediatric heart transplant recipient. Hum Immunol 2012;73:48-51.
22. Otten HG, Verhaar MC, Borst HP, Hene RJ, van Zuilen AD. Pretransplant donor-specific HLA class-I and -II antibodies are associated with an increased risk for kidney graft failure. Am J Transplant 2012;12:1618-23.
23. Lachmann N, Todorova K, Schulze H, Schonemann C. Systematic comparison of four cell- and Luminex-based methods for assessment of complement-activating HLA antibodies. Transplantation 2013;95:694-700.
24. Freitas MC, Rebellato LM, Ozawa M, et al. The role of immunoglobulin-G subclasses and C1q in de novo HLA-DQ donor-specific antibody kidney transplantation outcomes. Transplantation 2013;15:113-9.
25. Cristal. Agence Biomédecine (http://www.sipg.sante.fr/portail/)
26. Données Informatiques VAlidées en Transplantation (DIVAT) software (http://www.divat.fr/).
27. Racusen LC, Colvin RB, Solez K, et al. Antibody-mediated rejection criteria — an addition to the Banff 97 classification of renal allograft rejection. Am J Transplant 2003;3:708-14.
28. Solez K, Colvin RB, Racusen LC, et al. Banff 07 classification of renal allograft pathology: updates and future directions. Am J Transplant 2008;8:753-60.

29. Lefaucheur C, Nochy D, Andrade J, et al. Comparison of combination Plasmapheresis/IVIg/anti-CD20 versus high-dose IVIg in the treatment of antibody-mediated rejection. Am J Transplant 2009;9:1099-107.
30. Loupy A, Suberbielle-Boissel C, Zuber J, et al. Combined posttransplant prophylactic IVIg/anti-CD 20/plasmapheresis in kidney recipients with preformed donor-specific antibodies: a pilot study. Transplantation 2010;89:1403-10.
31. Lamb KE, Lodhi S, Meier-Kriesche HU. Long-term renal allograft survival in the United States: a critical reappraisal. Am J Transplant 2011;11:450-62.
32. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. Stat Med 2004;23:2109-23.
33. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157-72.
34. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis 2002;39:Suppl 1:S1-S266.
35. Colvin RB, Smith RN. Antibody-mediated organ-allograft rejection. Nat Rev Immunol 2005;5:807-17.
36. Sis B, Jhangri GS, Bunnag S, Allanach K, Kaplan B, Halloran PF. Endothelial gene expression in kidney transplants with alloantibody indicates antibody-mediated damage despite lack of C4d staining. Am J Transplant 2009;9:2312-23.
37. Loupy A, Hill GS, Suberbielle C, et al. Significance of C4d Banff scores in early protocol biopsies of kidney transplant recipients with preformed donor-specific antibodies (DSA). Am J Transplant 2011;11:56-65.
38. Stegall MD, Diwan T, Raghavaiah S, et al. Terminal complement inhibition decreases antibody-mediated rejection in sensitized renal transplant recipients. Am J Transplant 2011;11:2405-13.
39. A pilot study to evaluate the use of C1 esterase inhibitor (human) in patients with acute antibody-mediated rejection. ClinicalTrials.gov (http://clinicaltrials.gov/ct2/show/NCT01147302).
40. Safety & tolerability of Berinert (C1 inhibitor) therapy to prevent rejection. ClinicalTrials.gov (http://clinicaltrials.gov/ct2/show/NCT01134510).

Copyright © 2013 Massachusetts Medical Society.

JOURNAL ARCHIVE AT NEJM.ORG
Every article published by the *Journal* is now available at **NEJM.org**, beginning with the first article published in January 1812. The entire archive is fully searchable, and browsing of titles and tables of contents is easy and available to all. Individual subscribers are entitled to free 24-hour access to 50 archive articles per year. Access to content in the archive is available on a per-article basis and is also being provided through many institutional subscriptions.

VIII.1.3 Démonstration de l'intérêt pronostique du microscope moléculaire dans la stratification du risque de perte du greffon

Loupy A, Lefaucheur C, Vernerey D, Chang J, Hidalgo LG, Beuscart T, Verine J, Aubert O, Dubleumortier S, Duong van Huyen J-P, Jouven X, Glotz D, Legendre C, Halloran PF

Molecular Microscope Strategy to Improve Risk Stratification in Early Antibody-Mediated Kidney Allograft Rejection. **J Am Soc Nephrol JASN** doi:10.1681/ASN.2013111149. (2014)

Impact factor: 9.3

VIII.1.3.1 Positionnement du problème

L'évaluation du rejet humoral par l'histologie, en dépit de son utilité clinique, fait face à de sérieuses limites en termes de précision diagnostique et de prédiction du risque. Ces limites ont clairement été soulignées lors de la dernière conférence internationale de Banff au Brésil, en 2013, dont le consortium d'experts a indiqué le développement d'approches complémentaires dans le domaine du rejet d'organe comme une priorité.

La complexité du rejet humoral tant dans sa présentation clinique que dans son hétérogénéité phénotypique reflète la nécessité d'avoir de meilleurs supports diagnostiques et pronostiques que ceux apportés par l'évaluation histopathologique conventionnelle.

Cependant, plus que l'ajustement des techniques existantes, il semble nécessaire d'ouvrir la voie vers l'utilisation de nouvelles technologies. En effet, les limites des tests conventionnels peuvent être repoussées en considérant l'étude moléculaire des biopsies, c'est-à-dire de l'expression de gènes dans le tissu. Ce diagnostic moléculaire a été développé en utilisant des données de cohortes rigoureusement définies, comme illustré par

l'expérience du cancer. Dans cette spécialité, l 'évaluation moléculaire des biopsies utilise des scores dérivés de populations bien phénotypées qui permettent de définir des profils de risque distincts. Cette approche a ainsi permis de modifier la pratique clinique en identifiant les patients susceptibles de répondre ou non à des thérapeutiques anti-tumorales. En transplantation, les stratégies de profilage moléculaire à l'aide de puces à ADN ont donné des résultats prometteurs avec, comme nouvelle dimension, la possibilité d'évaluer les biopsies de greffons par une cartographie moléculaire qui reflète l'expression de probsets pathogéniques de gènes (PBT's). L'équipe de Phil Halloran à Edmonton a développé depuis plusieurs années un « microscope moléculaire » pour explorer la pathologie du patient transplanté et ainsi apporter un outil complémentaire à l'histologie conventionnelle.

Dans cette étude, nous avons évalué l'intérêt de ce « microscope moléculaire » dans le rejet humoral pour affiner la prédiction du risque de perte du greffon et de progression de la maladie. Pour cela, nous nous sommes appuyés sur des cohortes de malades des Hôpitaux Necker et Saint Louis très bien caractérisées auxquelles les outils conventionnels ont été complétés par l'étude systématique de l'expression de gènes dans le greffon.

VIII.1.3.2 Population de l'étude

A partir de 939 patients qui ont reçu un greffon rénal à l'Hôpital Necker (Paris, France) entre janvier 2004 et octobre 2010, nous avons identifié 101 (10,7%) patients avec un rejet humoral prouvé par une biopsie durant la première année de transplantation. 27 patients ont été exclus en raison de l'absence de matériel adéquat pour réaliser l'analyse moléculaire (14 avec du matériel mais dont la qualité ne permettait pas la réalisation de l'étude moléculaire et 13 patients pour lesquels seul un fragment fixé avait été réalisé). Ainsi, 74 patients avec un rejet humoral aigu ont été inclus dans l'étude.
Nous avons utilisé une cohorte de validation externe composée de 321 patients transplantés à l'Hôpital Saint Louis (Paris, France) entre le 1er janvier 2006 et le 1er janvier 2010. Parmi ces patients, 54 (16,8%) avaient un rejet

humoral précoce prouvé par biopsie et du matériel adéquat pour la réalisation de l'analyse moléculaire.

VIII.1.3.3 Données cliniques

Les données cliniques des donneurs et receveurs de notre cohorte principale (Hôpital Necker) et de la cohorte de validation (Hôpital Saint Louis) ont été obtenues à partir de 2 registres nationaux ; les Données Informatiques Validées en Transplantation (Hôpital Necker) et l'Agence de la Biomédecine (Hôpital Saint Louis) dans lesquels les données sont entrées de manière prospective à des moments spécifiques (Jour 0, 6 mois et 1 an post-transplantation) et complétées annuellement.

Les données qui concernent la détérioration de fonction ou les lésions histologiques des patients ayant un rejet humoral ont été enregistrées en accord avec le consensus de la classification internationale de Banff. Nous avons reconnu comme rejet humoral aigu les rejets C4d-négatifs lorsqu'il y avait une évidence diagnostique de rejet histologique et une présence de DSA circulants.

Le traitement du rejet humoral était standardisé entre les centres. Les évènements mesurés dans cette étude étaient la perte du greffon définie par le retour en dialyse et la progression de la glomérulopathie de l'organe transplanté définie par l'augmentation du score de Banff correspondant, vu en microscopie optique.
Toutes les biopsies ont été relues par 2 anatomopathologistes spécialisés en néphropathologie et en aveugle pour les données cliniques.

La présence de DSA circulants au moment de la transplantation (jour 0) et au moment de la biopsie a été analysée en utilisant la technique single-antigen (One Lambda Canoga Park, CA, USA) sur une plateforme Luminex.

Les caractéristiques moléculaires des biopsies étaient évaluées grâce à des lots de transcrits basés sur la pathogénie (PBTs) qui reflètent les processus biologiques connus comme étant pertinents dans la pathogénie de l'inflammation rénale et des lésions en transplantation.

VIII.1.3.4 Extraction de RNA et procédure d'analyse en microarray

Une biopsie supplémentaire à partir d'un pistolet à biopsie de 16 gauges a été obtenue au moment de la biopsie et conservée à – 80°C dans de l'OCT pour l'analyse de l'expression des gènes.

L'information sur les lots et les algorithmes pour la création du PBT sont disponibles sur la page web : (http://www.atagc.med.ualberta.ca/Research/GeneLists/Pages/default.aspx).

VIII.1.3.5 Approche méthodologique

Un modèle de Cox a été utilisé pour quantifier les HR et les intervalles de confiance à 95% (ICs) pour les facteurs associés à la perte du greffon. Les paramètres avec une p-value ≤ 0,10 ont été utilisés pour construire un modèle multivarié de Cox.
Pour déterminer l'apport du nouveau test (microscope moléculaire) aux tests conventionnels dans la stratification et la reclassification du risque de perte du greffon à 8 ans post-transplantation des patients, nous avons évalué le gain en discrimination par la statistique de concordance ainsi que sa capacité à bien reclasser le risque avec les méthodes du NRI et de l'IDI.

VIII.1.3.6 Résultats

Parmi 939 patients qui ont reçu une transplantation rénale à l'Hôpital Necker entre 2004 et 2010, 472 patients ont eu 583 biopsies réalisées durant la première année suivant la greffe.

Les rejets humoraux prouvés par biopsies ont été diagnostiqués pour 101/939 patients dans la première année, représentant une incidence de 10,7%. Parmi ces 101 rejets humoraux, 74 (73,3%) avaient le matériel nécessaire pour une évaluation moléculaire et ont été inclus dans l'étude. La Table 1 montre les caractéristiques cliniques et les résultats de la biopsie au moment de la transplantation pour les 74 patients qui ont eu un épisode de rejet humoral durant la première année. La médiane de rejet humoral était de 3,1 (étendue interquartile 0,2-12,0) mois post-transplantation.

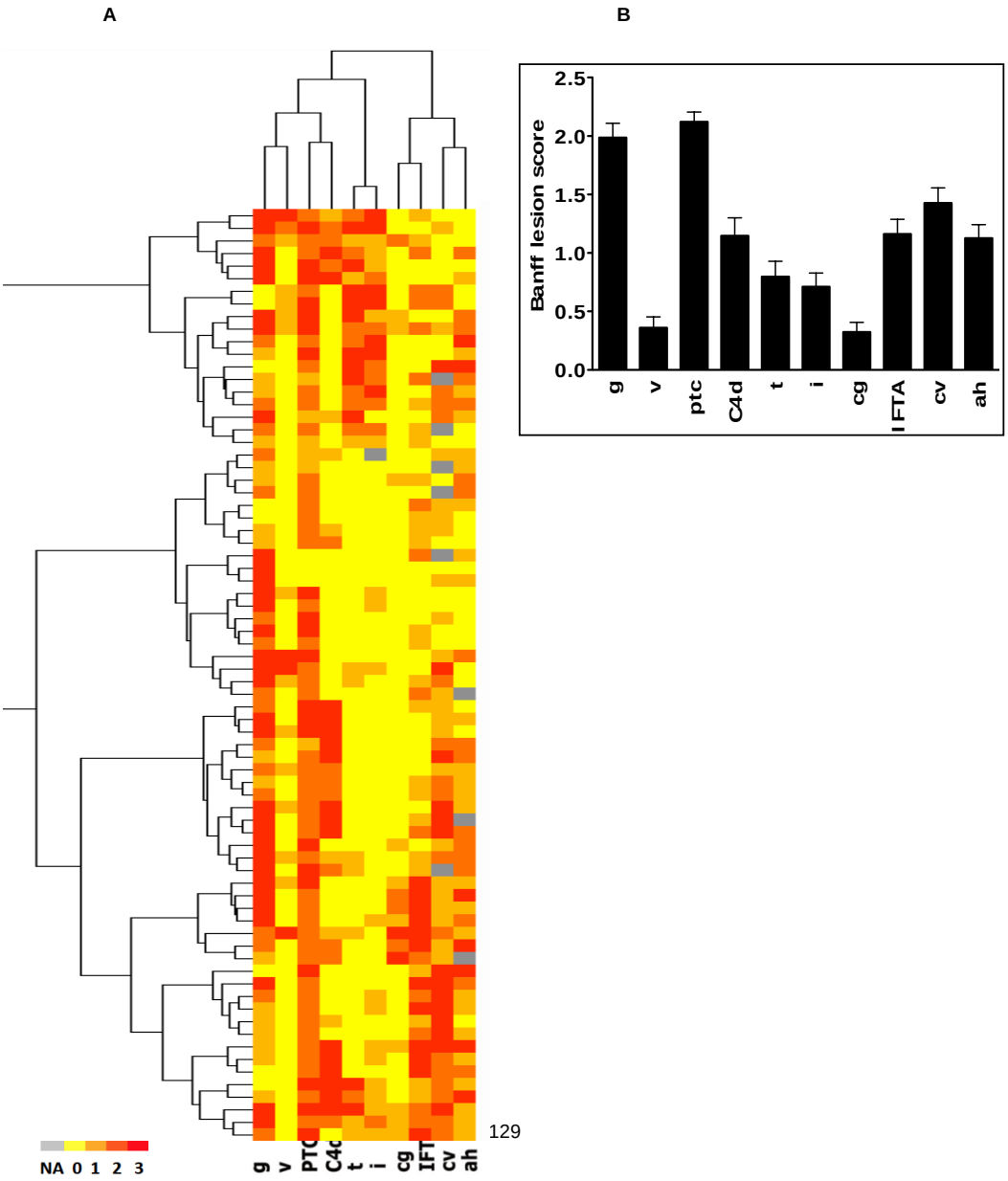
Table 1: Caractéristiques de base des patients avec rejet humoral aigu

Patients	**n=74**
Moyenne d'âge des receveurs *(années)**	48.4 ± 13
Ischémie froide *(heures)**	21.4 ± 12
Créatinine du donneur au moment de la transplantation *(µmol/L)**	83.8 ± 41
Reprise retardée de fonction *(n, %)*	27 (36.4%)
Nombre de dialyses post-transplantation*	1.2 ± 1.9
Nombre de transplantation *(>1)*	34 (48.6%)
Donneurs décédés *(n, %)*	67 (90.5%)
Néphropathie initiale	
Néphropathie diabétique	1 (1.3%)
Hypertension / Néphroangiosclérose	4 (5.4%)
Glomérulite / vascularite	21 (28.4%)
Néphropathie interstitielle / pyélonéphrite	13 (17.6%)
Autres	11 (14.9%)
Indéterminées	24 (32.4%)
Moyenne d'âge des donneurs *(années)**	51.1 ± 14
Sexe du donneur *(male, n, %)*	43 (58.1%)
IMMUNOLOGIE	
HLA-A mismatch*	0.8 ± 0.6
HLA-B mismatch*	1.1 ± 0.7
HLA-DR mismatch*	0.8 ± 0.6
DSA au moment de la transplantation	
PRA Lymphocytotoxique (%, moyenne)*	27 ± 29
cPRA (%, moyenne)*[€]	73
DSA circulants (n,%)	68 (91.9%)
Nombre de DSA (=1 ; ≥2)	44(64.7%) ; 24 (35.3%)
DSA max classe (classe I/ classe II (n,%))	25 (36.8%) / 43 (63.2%)
DSAmax de Classe I MFI**	2641 ± 410
DSAmax de Classe II MFI**	3783 ± 602
DSAmax de classe I ou II MFI**	5548 ± 562
Caractéristiques au moment de la biopsie	**(n=74)**
Creatinine (µmol/L)*	203 ± 142
eGFR (mL/min/1.73 m^2)*	44.8 ± 17
Proteinurie (g/L)* Patients (%) avec protéinurie >0.15 g/L	0.31 ± 0.58 38 (51%)
DSA circulants (n,%)	74 (100%)
Nombre de DSA (=1 ; ≥2)	56 (75.7%) ; 18 (24.3%)
DSA max classe (classe I/ classe II (n,%))	(26, 35.1% / 48, 64.9%)
DSAmax de Classe I MFI**	2315 ± 568
DSAmax de Classe II MFI**	3732 ± 641
DSAmax de Classe I ou II MFI**	3859 ± 471

* variables continues exprimées en moyenne ± déviation standard
** variables continues exprimées en moyenne ± erreur standard
PRA calcul réalisé chez les patients transplantés après le 01/01/2009 n=11

La Figure 1 représente l'histologie du greffon rénal au moment du rejet.

Figure 1: Histologie des greffons au moment du rejet: Carte de chaleur colorée par seuil représentant (A) les caractéristiques histologiques des patients avec un rejet et (B) la moyenne des scores de Banff.

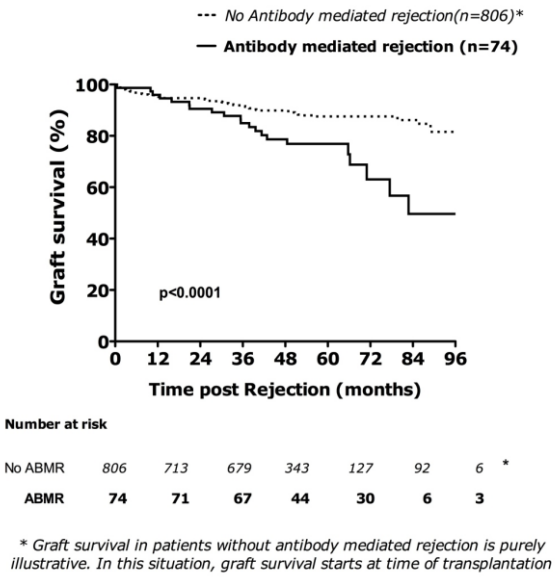


21 dysfonctions du greffon sont survenues après une médiane de suivi de 54 (étendue interquartile 38-66) mois après le diagnostic initial de rejet humoral.

La survie du greffon (après censure pour la mortalité) était de 90,5%, 63,0% et 49,6% à 2, 6 et 8 ans respectivement après le diagnostic de rejet humoral (Figure 2A).

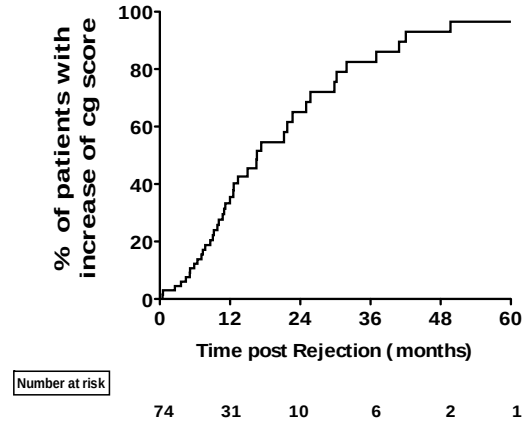
Figure 2: Courbe de survie des greffons selon la présence ou non de rejet humoral (A) incidence cumulée de la progression des lésions de glomérulopathie du transplant après le rejet humoral (B)

A



* Graft survival in patients without antibody mediated rejection is purely illustrative. In this situation, graft survival starts at time of transplantation

B



L'analyse par modèle de Cox univarié des caractéristiques conventionnelles montrait que l'âge du donneur, la fonction rénale (eGFR) au moment du diagnostic du rejet humoral aigu, le score de fibrose interstitielle et d'atrophie tubulaire (FITA), le score de glomérulopathie du greffon (cg) et les dépôts de C4d étaient significativement associés avec la perte du greffon. Le score

histologique de rejet humoral était aussi associé avec l'augmentation du risque de perte du greffon tout comme les paramètres moléculaires qui étaient le score moléculaire de rejet humoral aigu et les transcrits endothéliaux DSA-sélectifs.

Dans le modèle multivarié de Cox, les facteurs indépendamment associés avec la perte du greffon comprenaient les facteurs traditionnels suivants : âge du donneur, le score histologique humoral (g+cpt+v+cg+C4d) et le score moléculaire de rejet humoral aigu (HR = 2,22 ; 95%IC = 1,37 - 3,58 ; p = 0,001) (Table 2).

Table 2: Déterminants de la perte du greffon (modèle multivarié) intégrant le score moléculaire du rejet

		Nombre de patients	Nombre d'évène-ments	HR	95% IC	Valeur de p
Age du donneur (années)	**<60 ans**	54	11	1	-	-
	≥60 ans	20	10	3.84	[1.48 – 9.96]	0.006
eGFR* au moment du rejet	**≥30 mL/min**	52	10	1		
	<30 mL/min	22	11	1.74	[0.70 - 4.35]	0.232
Score humoral histologique (g+ptc+v+cg+c4d)		74	21	1.43	[1.09 – 1.90]	0.011
Score du rejet humoral		74	21	2.22	[1.37 - 3.58]	0.001

L'inclusion du score moléculaire de rejet humoral aigu dans le modèle de référence permettait une augmentation de la capacité discriminative du modèle (c'est-à-dire sa capacité à distinguer les patients qui ont perdu leur greffon et les autres), avec une augmentation significative de la C statistique passant de 0,770 à 0,8102 (Différence moyenne bootstrap = 0,0496505 ; 95%IC = 0,0473 – 0,0519).

L'inclusion du score moléculaire du rejet humoral dans le modèle de référence a permis aussi de reclasser les patients à faible risque (absence de perte du greffon) et les patients à risque élevé (perte du greffon), avec un NRI = 1,0135 (95% IC = 0 ,5696 – 1,4573, p < 0,0001).

L'addition du microscope moléculaire a permis de reclasser dans le bon sens

(diminution du risque) 42 des 53 patients (79,2%) n'ayant pas présenté de perte du greffon et inversement a permis de reclasser dans le bon sens (augmentation du risque) 15 des 21 patients (71,4%) présentant une perte du greffon (Figure 3).

L'IDI était de 0,1579, p = 0,00015 (Figure 4).

Ces analyses démontrent que l'addition du score moléculaire de rejet humoral aigu à l'approche conventionnelle a amélioré la stratification du risque de dysfonction du greffon.

Figure 3: Valeur additionnelle du microscope moléculaire (score du rejet humoral aigu) pour la reclassification du risque de perte du greffon (NRI continu).

Les lignes bleues chez les patients sans perte du greffon indiquent un mouvement de la prédiction du risque par le microscope moléculaire dans la bonne direction (vers le bas) (42/53 = 79.2%). Inversement, les lignes rouges chez les patients avec une perte du greffon indiquent une modification correcte, vers le haut, de l'évaluation du risque lorsqu'on utilise le microscope moléculaire. (15/21 = 71.4%).

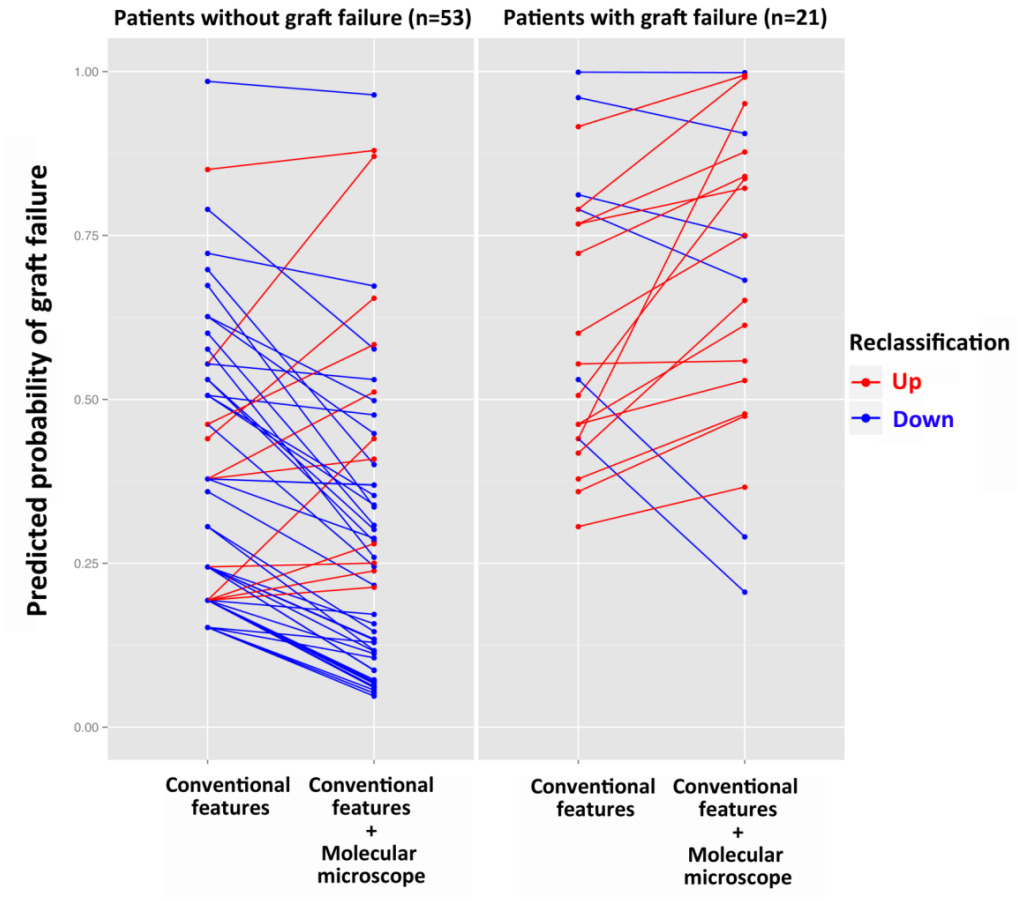
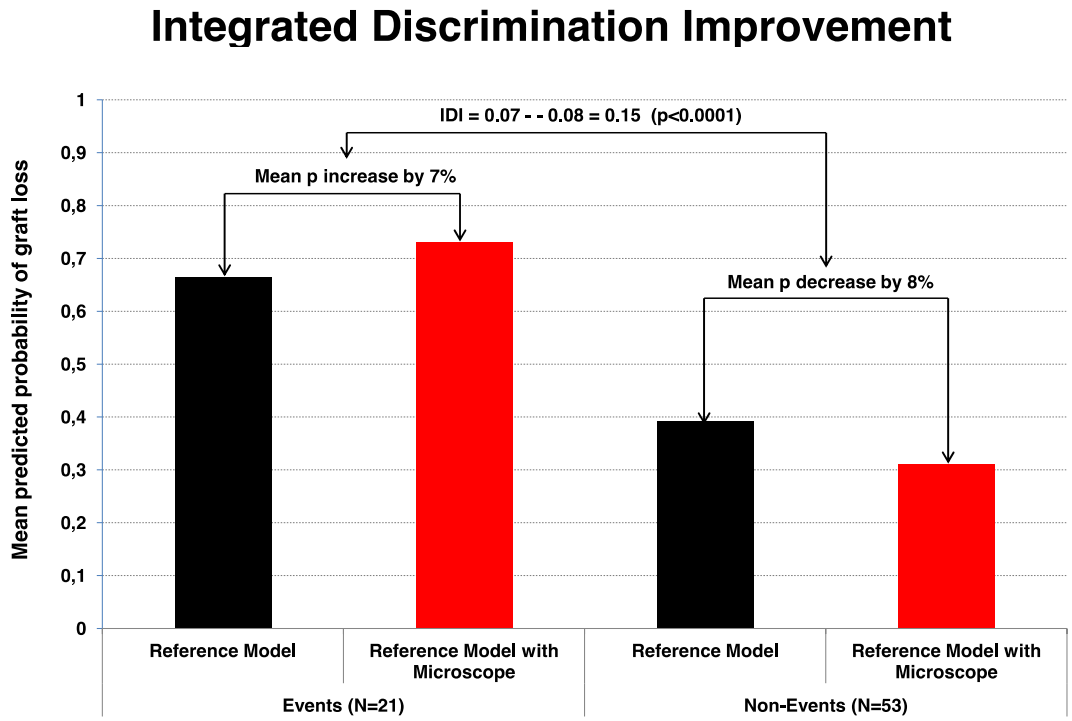


Figure 4: Amélioration de la prédiction du risque de perte du greffon à 8 ans en considérant le microscope moléculaire en complément des paramètres conventionnels.



Cette amélioration de la prédiction du risque est évaluée par l'IDI qui repose sur la quantification par la moyenne des probabilités de perte du greffon dans les deux populations caractérisées par l'occurrence d'un évènement ou non. Le changement de moyenne de la probabilité prédite avec l'ajout du microscope moléculaire est satisfaisant s'il se fait à la hausse chez les patients qui ont réellement perdu le greffon et à l'opposé à la baisse chez ceux qui n'ont pas perdu le greffon.

La cohorte de validation externe était composée de 321 patients transplantés rénaux de l'Hôpital Saint Louis entre le 1er janvier 2006 et le 1er janvier 2010. Elle comportait 54 patients avec un rejet humoral précoce prouvé par une biopsie. Le score moléculaire de rejet humoral aigu était associé de manière indépendante avec la perte du greffon (HR = 4,45 ; 95%IC = 1,01 – 19,95, p = 0,049).

VIII.1.3.7 Conclusion de l'étude

L'évaluation de l'expression des gènes dans le rejet humoral en transplantation rénale, permet d'identifier, indépendamment des paramètres conventionnels, les patients avec un risque élevé de perte du greffon rénal. Le système utilisant le microscope moléculaire, par rapport à l'approche classique histologique, permet une amélioration de la stratification du risque et ainsi d'affiner la prise en charge clinique et aussi améliorer les essais cliniques en transplantation.

VIII.1.3.8 Manuscrit

BASIC RESEARCH www.jasn.org

Molecular Microscope Strategy to Improve Risk Stratification in Early Antibody-Mediated Kidney Allograft Rejection

Alexandre Loupy,*† Carmen Lefaucheur,†‡ Dewi Vernerey,†§ Jessica Chang,‖ Luis G. Hidalgo,‖¶ Thibaut Beuscart,† Jerome Verine,** Olivier Aubert,† Sébastien Dubleumortier,†† Jean-Paul Duong van Huyen,*†‡‡ Xavier Jouven,† Denis Glotz,†‡ Christophe Legendre,*† and Philip F. Halloran‖§§

*Paris Descartes University and Hôpital Necker and ‡Hôpital Saint Louis, Assistance Publique–Hôpitaux de Paris, Paris, France; †Paris Translational Research Center for Organ Transplantation, Institut National de la Santé et de la Recherche Médicale, Unité Mixte de Recherche-S970, Paris, France; §Unit 3181, University Hospital of Besançon, France; ‖Alberta Transplant Applied Genomics Centre, Edmonton, Alberta, Canada; ¶Department of Laboratory Medicine and Pathology and §§Department of Medicine, Division of Nephrology and Transplant Immunology, University of Alberta, Edmonton, Alberta, Canada; **Department of Pathology, Saint Louis Hospital, Paris, France; ††Pateforme Centres de Resources Biologiques, Institut Gustave Roussy, Villejuif, France; and ‡‡Department of Pathology, Necker Hospital, Paris, France

ABSTRACT

Antibody-mediated rejection (ABMR) is the leading cause of kidney allograft loss. We investigated whether the addition of gene expression measurements to conventional methods could serve as a molecular microscope to identify kidneys with ABMR that are at high risk for failure. We studied 939 consecutive kidney recipients at Necker Hospital (2004–2010; principal cohort) and 321 kidney recipients at Saint Louis Hospital (2006–2010; validation cohort) and assessed patients with ABMR in the first 1 year post-transplant. In addition to conventional features, we assessed microarray-based gene expression in transplant biopsy specimens using relevant molecular measurements: the ABMR Molecular Score and endothelial donor-specific antibody-selective transcript set. The main outcomes were kidney transplant loss and progression to chronic transplant injury. We identified 74 patients with ABMR in the principal cohort and 54 patients with ABMR in the validation cohort. Conventional features independently associated with failure were donor age and humoral histologic score (g+ptc+v+cg+C4d). Adjusting for conventional features, ABMR Molecular Score (hazard ratio [HR], 2.22; 95% confidence interval [95% CI], 1.37 to 3.58; P=0.001) and endothelial donor-specific antibody-selective transcripts (HR, 3.02; 95% CI, 1.00 to 9.16; $P<0.05$) independently associated with an increased risk of graft loss. The results were replicated in the independent validation group. Adding a gene expression assessment to a traditional risk model improved the stratification of patients at risk for graft failure (continuous net reclassification improvement, 1.01; 95% CI, 0.57 to 1.46; $P<0.001$; integrated discrimination improvement, 0.16; $P<0.001$). Compared with conventional assessment, the addition of gene expression measurement in kidney transplants with ABMR improves stratification of patients at high risk for graft loss.

J Am Soc Nephrol 25: ●●●–●●●, 2014. doi: 10.1681/ASN.2013111149

Received November 1, 2013. Accepted January 21, 2014.

A.L. and C.L. contributed equally to this work.

Published online ahead of print. Publication date available at www.jasn.org.

Correspondence: Dr. Philip F. Halloran, Department of Medicine, University of Alberta, 250 Heritage Medical Research Center, Edmonton, AB T6G 2S2, Canada. Email: phil.halloran@ualberta.ca

Copyright © 2014 by the American Society of Nephrology

Kidney transplantation is the preferred treatment for patients with ESRD, and it is routinely performed in more than 80 countries.[1] The current standard of care in transplant medicine is largely based on biopsy assessment by histology and immunochemistry supported by laboratory tests, such as HLA antibody measurements.

A key role of the biopsy is to diagnose rejection, particularly antibody-mediated rejection (ABMR), which is the leading cause of kidney transplant failure.[2–4] Kidney allograft loss is now acknowledged to lead to increased morbidity, mortality, and cost.[5,6] In ABMR, the conventional assessment by histopathology, despite its clinical use, faces serious limitations in terms of diagnostic accuracy and risk prediction, which is illustrated by the call for action raised during the International Banff Allograft Rejection Classification meeting held in Brazil in 2013 for the development of complementary approaches in the field of allograft rejection.

ABMR reflects the need for more diagnostic support than is provided by the conventional histologic assessment, but improvements are not likely to be associated with additional adjustments to this existing technique, indicating the need for new technologies. The limitations of conventional tests can be addressed by introducing molecular biopsy measurements developed using data from well defined cohorts, which is illustrated by the experiences in cancer medicine, where molecular biopsy measurements using scores derived from well defined cohorts now provide new levels of information[7–9] and have changed clinical practice.[7–9] In transplant medicine, molecular profiling strategies using microarrays have shown promise as a new dimension for assessing transplant biopsies, which was recently shown by the development of a microarray-based molecular microscope system for testing of biopsies for T cell–mediated rejection,[10] ABMR,[11,12] and AKI.[13] However, to date, studies[11,12] focused on late ABMR (the prevalence of early ABMR in these low-risk cohorts is less than 2% of biopsies with molecular assessment). Moreover, there was heterogeneity in the clinical and therapeutic management of the patients and superimposed diseases in late cases, precluding any conclusions about the clinical relevance of the molecular microscope in early ABMR. As in cancer, homogenous and well phenotyped cohorts are needed to ascertain whether the molecular microscope strategy could be helpful and add to the conventional assessments. To achieve this goal, we used a unique model of early disease (early biopsy-proven ABMR) with intensive follow-up and standardized management with Food and Drug Administration (FDA)–recommended therapeutic strategies.[14] This model led us to focus on the disease progression related to the ongoing humoral process without the interference from superimposed diseases seen in late ABMR cases.

In the present study, we applied the molecular microscope system in addition to conventional clinical histologic and immunologic features to evaluate the potential impact of this approach in terms of prediction of the risk of graft loss and disease progression.

Table 1. Baseline characteristics of the ABMR population

Patient Demographics (*n*=74)	
Mean recipient age, yr[a]	48.4±13
Cold ischemia time, h[a]	21.4±12
Serum creatinine at the time of donation, μmol/L[a]	83.8±41
Delayed graft function, *n*	27 (36.4%)
Number of dialyses post-transplant[a]	1.2±1.9
Rank of transplantation (>1)	34 (48.6%)
Deceased donor, *n*	67 (90.5%)
Primary disease	
Diabetic nephropathy	1 (1.3%)
Hypertension/large vessel disease	4 (5.4%)
GN/v	21 (28.4%)
Interstitial nephritis/pyelonephritis	13 (17.6%)
Others	11 (14.9%)
Unknown etiology	24 (32.4%)
Mean donor age, yr[a]	51.1±14
Donor sex, *n* (men)	43 (58.1%)
Immunology	
HLA-A mismatch[a]	0.8±0.6
HLA-B mismatch[a]	1.1±0.7
HLA-DR mismatch[a]	0.8±0.6
DSA status on the day of transplantation	
Lymphocytotoxic PRA, % (mean)[a]	27±29
cPRA, % (mean)[a]	73
Circulating DSA, *n*	68 (91.9%)
Number of DSA (=1; ≥2)	44 (64.7%); 24 (35.3%)
DSAmax class (class I/class II), *n*	25 (36.8%)/43 (63.2%)
Class I DSAmax MFI[b]	2641±410
Class II DSAmax MFI[b]	3783±602
Class I or II DSAmax MFI[b]	5548±562
Characteristics at the time of biopsy	
Serum creatinine, μmol/L[a]	203±142
eGFR, ml/min per 1.73 m^2[a]	44.8±17
Proteinuria, g/L[a]	0.31±0.58
Patients with urine protein>0.15 g/L	38 (51%)
Circulating DSA, *n*	74 (100%)
Number of DSA (=1; ≥2)	56 (75.7%); 18 (24.3%)
DSAmax class (class I/class II) , *n*	26 (35.1%); 48 (64.9%)
Class I DSAmax MFI[b]	2315±568
Class II DSAmax MFI[b]	3732±641
Class I or II DSAmax MFI[b]	3859±471

Calculated panel-reactive antibody (cPRA) assessed in patients transplanted after January 1, 2009, *n*=11.
[a]Continuous variables expressed as the mean±SD.
[b]Continuous variables expressed as the mean±SEM.

RESULTS

Baseline Characteristics of the Kidney Transplant Recipients

Among 939 patients who received a kidney transplant in the Necker Hospital between 2004 and 2010, 472 patients had 583 biopsies performed in the first 1 year post-transplant. In all, 77 samples (13.2%) showed acute tubular necrosis, 48 samples (8.2%) had isolated interstitial fibrosis tubular atrophy, 51 samples (8.7%) had T cell–mediated rejection, 43 samples (7.4%) had

borderline lesions, 48 samples (8.2%) had calcineurin inhibitor toxic effects, 26 samples (4.4%) had recurrent diseases, 23 samples (3.9%) had BK virus nephropathy, and 56 samples (9.6%) had other diagnoses. Biopsy-proven ABMR was diagnosed in 101 of 939 patients in the first 1 year, representing an incidence of 10.7%. Among these 101 ABMR patients, 74 patients (73.3%) had suitable material for a molecular assessment and were enrolled in this study. Table 1 shows the characteristics at transplantation and the time of biopsy for 74 patients who had an episode of ABMR diagnosed in the first 1 year post-transplant at a median of 3.1 (interquartile range [IQR]=0.2–12.0) months post-transplant. All the patients were sensitized before transplantation with lymphocytotoxic panel-reactive antibodies ranging from 5% to 100%. At the time of transplantation, 68 of 74 patients (91.9%) had detectable donor-specific antibodies (DSAs), with 44 patients (64.7%) having 1 DSA and 24 patients (35.3%) having ≥2 DSAs (range=2–4). The immunodominant mean fluorescence intensity DSA (DSA MFImax) was class I in 25 of 68 patients (36.8%) and class II in 43 of 68 patients (63.2%), and the mean DSA MFImax was 5548±562. The median follow-up time after transplantation was 59.3 (IQR=47–79) months.

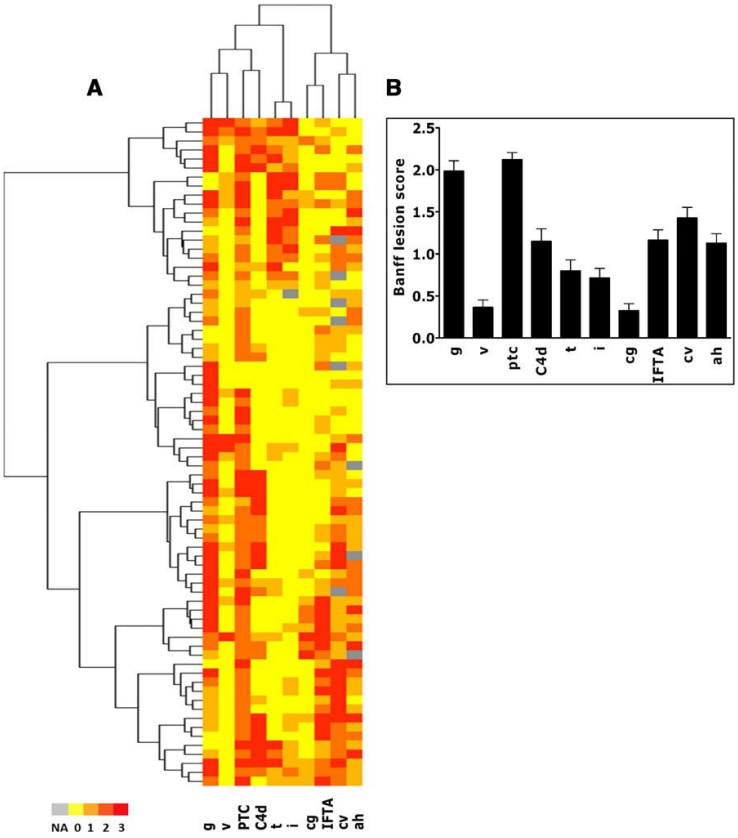


Figure 1. Kidney allograft histopathology at the time of rejection showing the distribution of graft injury lesions in the overall cohort of patients with ABMR. Note that all patients have intense microcirculation injury. Heat map is colored by thresholds representing (A) the histologic characteristics of patients with ABMR and (B) the mean Banff scores. ah, arteriolar hyalinosis; C4d, complement split product; cg, transplant glomerulopathy; cv, arterial fibrous intimal thickening score; g, glomerulitis; i, interstitial inflammation score; IFTA, interstitial fibrosis–tubular atrophy; ptc, peritubular capillary; t, tubulitis score; v, vasculitis.

Histologic and Immunologic Features of Acute ABMR

By histology, all 74 biopsies at the time of diagnosis of ABMR showed intense microvascular inflammation, with a mean glomerulitis (g) score of 1.9±1.1 and mean peritubular capillary inflammation (ptc) score of 2.1±0.7. The complement split product C4d deposition in the peritubular capillaries was present in 34 of 74 patients (45.9%). At the time of diagnosis, 46 of 74 (62.2%) kidneys had some atrophy-fibrotic lesions. Of 68 patients with arteries suitable for analysis, 52 patients (76.4%) had evidence of fibrous intimal thickening in the small arteries (arteriosclerosis), with 21 cases (30.8%) being moderate or severe arteriosclerosis (cv Banff score≥2). Figure 1 depicts the kidney allograft histopathology at the time of rejection. The immunodominant DSA at the time of biopsy-proven ABMR was class I in 26 of 74 patients (35.1%) and class II in 48 of 74 patients (64.9%), with a mean class I or II DSA MFImax of 3859±471.

Gene Expression Assessment in Patients with Acute ABMR

We compared the molecular phenotype with the conventional parameters (Table 2). The ABMR Molecular Score and endothelial DSA-selective transcripts correlated strongly with each other (r=0.74, P<0.001) and with the natural killer (NK) transcripts (r=0.84, P<0.001). The ABMR Molecular Score correlated with the level of circulating DSA at the time of ABMR diagnosis (r=0.44, P=0.01) and with C4d deposition within the peritubular capillaries (r=0.33, P=0.01). The endothelial DSA-selective transcripts showed correlations with the transplant glomerulopathy (cg) score (r=0.24, P=0.04) and mesangial matrix expansion (r=0.29, P=0.04). There was little correlation between the ABMR Molecular Score or the endothelial DSA-selective transcripts and histologic lesions showing microvascular inflammation and interstitial inflammation or tubulitis (Table 2).

Determinants of Kidney Allograft Loss

Twenty-one kidneys failed after a median follow-up post-ABMR diagnosis of 54 (IQR=38–66) months. Death-censored

Table 2. Molecular microscope results at the time of rejection and correlations with concurrent clinical, histologic, and immunologic parameters

Parameters	ABMR Molecular Score	Endothelial DSA-Selective Transcripts
Clinical parameters		
eGFR at time of biopsy, μmol/L	0.002	0.009
Proteinuria at the time of biopsy, g/L	0.030	0.076
Immunologic parameters		
Class II or I DSA MFImax at the time of rejection	0.442[a]	0.273[b]
Histologic parameters		
g score	0.036	0.030
ptc score	0.171	0.207
Interstitial inflammation score	−0.087	−0.031
Tubulitis score	−0.227	−0.094
Interstitial fibrosis/tubular atrophy score	0.247[b]	0.119
Arterial fibrous intimal thickening score	0.112	−0.084
cg score	0.135	0.242[b]
Mesangial matrix expansion score	0.254[b]	0.287[b]
C4d deposition in graft capillaries (C4d Banff score)	0.328[b]	0.181
Humoral histologic score (g+ptc+v+cg+C4d)	0.069	0.163
Molecular parameters		
ABMR Molecular Score	1.000	0.737[a]
Endothelial DSA-selective transcripts	0.737[a]	1.000
NK transcripts	0.841[a]	0.613[a]
T-cell transcripts	0.123	0.172
AKI score	0.171	0.125

[a] $P<0.01$ (Spearman correlations with Bonferroni correction for multiple comparisons).
[b] $P<0.05$ (Spearman correlations with Bonferroni correction for multiple comparisons).

graft survival was 90.5%, 63.0%, and 49.6% at 2, 6, and 8 years, respectively, after the diagnosis of ABMR (Figure 2A). Univariate Cox analysis of the conventional features (Table 3) showed that the donor age, eGFR at the time of ABMR diagnosis, interstitial fibrosis–tubular atrophy scores, cg score, and C4d deposition were significantly associated with kidney transplant loss. The humoral histologic score was also associated with increased risk of graft loss. The molecular parameters associated with failure were the ABMR Molecular Score and endothelial DSA-selective transcripts. The T-cell burden score and the AKI score (injury-repair response associated transcript) were not associated with risk of allograft loss.

In the multivariate Cox model, the factors independently associated with graft loss were certain conventional features (donor age, the humoral histologic score [g+ptc+vasculitis (v) score+cg+C4d], and the ABMR molecular score [hazard ratio (HR), 2.22; 95% confidence interval (95% CI), 1.37 to 3.58; $P=0.001$]) (Table 4) and the endothelial DSA-selective transcripts (HR, 3.02; 95% CI, 1.0 to 9.16; $P<0.05$) (Table 4).

Determinants of Progression to Chronic Allograft Injury

This analysis integrated the follow-up biopsies performed after the initial diagnosis of ABMR. At 1 and 2 years after ABMR, 35% and 68% of patients had an increase of cg lesions defined by an increase of cg Banff score≥1 (Figure 2B).

The median time to cg Banff score increase was 16.6 months after the initial diagnosis of ABMR. Kaplan–Meier analysis shows that the ABMR molecular score, endothelial DSA-selective transcripts, peritubular capillaritis score, and C4d deposition were associated with the time to increase of cg lesions (Figure 3).

Additive Value of the Molecular Microscope for Predicting Kidney Allograft Loss

The inclusion of the ABMR molecular score in the reference model was found to significantly improve the discriminative ability of the model (*i.e.*, its capacity to discriminate between patients who have lost their graft and patients who have not), because the C statistic increased significantly from 0.77 to 0.81 (bootstrap mean difference=0.0496505; 95% CI, 0.0473 to 0.0519). Similarly, the inclusion of the ABMR molecular score in the reference model adequately reclassified patients at lower (no events) or higher (events) risk of graft loss, which was shown by a continuous net reclassification index of 1.0135 (95% CI, 0.57 to 1.46; $P<0.001$). The addition of the molecular microscope reclassified 42 of 53 patients (79.2%) in the right direction in the no event group, whereas it reclassified 15 of 21 patients (71.4%) in the event group (graft loss) (Figure 4). The integrated discrimination improvement was 0.1579 ($P<0.001$).

These analyses show that the addition of the ABMR Molecular Score to kidney biopsies improved the stratification of patients at risk for graft failure.

External Validation

The external validation cohort was composed of 321 kidney recipients transplanted at Saint Louis Hospital between January 1,

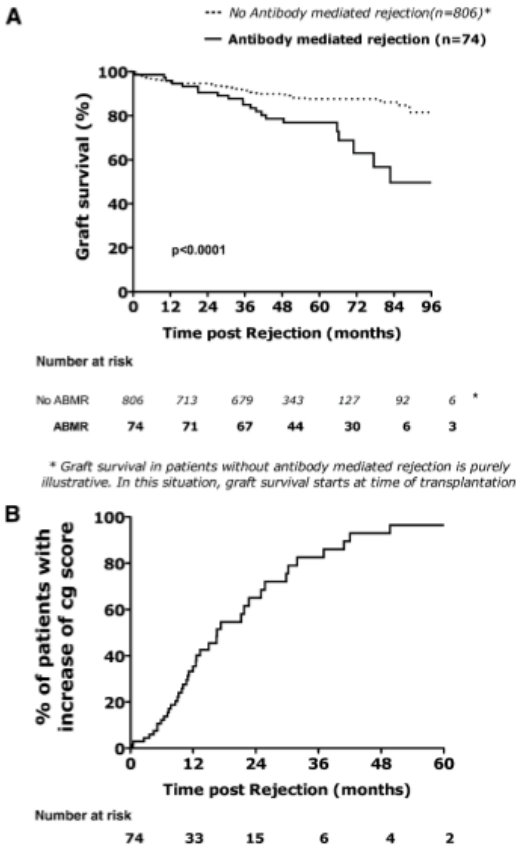


Figure 2. (A) Time to kidney transplant loss and (B) time to the progression of cg lesions after ABMR.

2006, and January 1, 2010, with 54 who had biopsy-proven early ABMR with suitable material for microarray analysis. The baseline characteristics and immunologic characteristics for the validation set are detailed in Supplemental Table 4. The median follow-up time after transplantation was 54.5 (IQR=42–69) months. Both the ABMR Molecular Score and endothelial DSA-selective transcripts were independently associated with graft loss (HR, 4.45; 95% CI, 1.01 to 19.95; P=0.05 for the ABMR Molecular Score [Supplemental Table 5A] and HR, 6.68; 95% CI, 1.36 to 32.93; P=0.02 for the endothelial DSA-selective transcripts [Supplemental Table 5B]).

DISCUSSION

In a carefully phenotyped population of kidney recipients with early biopsy-proven ABMR, we assessed the use of transcript measurement for the molecular interrogation of biopsies. Compared with the traditional approach for predicting graft loss, which is based only on histology and the presence of donor-specific anti-HLA antibodies, the new approach, which combines the traditional approach with gene expression profiling in the kidney allografts, provides a more accurate prediction of graft loss.

Molecular phenotyping of kidney transplant biopsies has provided clues as to the mechanisms of disease independent of conventional assessments. The lack of correlation of the ABMR Molecular Score or endothelial DSA-selective transcripts with the conventional features and the independent correlations of these molecular data with the outcomes establish an independent value for molecular phenotyping in risk assessment, most likely because the ABMR Molecular Score and endothelial DSA-selective transcripts set reflects a change in the microcirculatory endothelium that conventional assessments cannot detect.

Our study is the first to examine the risk of graft loss in a large population of early ABMR patients with extensive follow-up and standardized management using an FDA-approved therapeutic strategy,[14] in which we used contemporary tools for precise allograft phenotyping together with a systematic gene expression assessment in the allograft to represent the full spectrum of ABMR. The present study used a model of early disease using a very high-risk cohort with early ABMR, which led us to focus on the disease progression related to ongoing humoral process without interference with superimposed diseases seen in late ABMR cases.[11,15]

This paper shows the predictive performance of the molecular microscope inside the ABMR disease, where all patients have stereotypical histologic presentation and all have anti-HLA antibodies. We are not using the molecular microscope to confirm the ABMR diagnosis, which has been the subject of previous works,[11,15] but, rather, address a major clinical problem of how to stratify the risk in individuals with early ABMR (*i.e.*, inside the disease).

We used, in the present study, a methodological approach[16] for the reclassification, discrimination, and performance analyses to evaluate the benefit of the molecular microscope system over conventional tools, showing a net individual benefit of the molecular microscope approach when added to the conventional features. Conventional assessment permits to diagnose ABMR but is insufficient for prediction of outcome and risk assessment in patients with ABMR.[17]

Our hypothesis is that the ABMR Molecular Score and the endothelial transcripts reflect active remodeling of the microcirculation because of relatively recent antibody-mediated damage, implying the action of the antibody and its effector mechanism, and thus, can be used as a companion to the conventional features, such as histopathology. The ABMR Molecular Score is the weighted average of a larger set of genes selected using a simple t test comparing two classes of biopsies and then fine tuned using linear discriminant analysis, whereas the endothelial DSA-selective transcripts score is the unweighted average of a gene set selected using domain-specific knowledge of transplant biology. The overlap in these gene sets reflects the convergence of the methods on a similar biology and highlights the important, but not exclusive, participation of antibody-mediated endothelial activation.

Table 3. Determinants of kidney transplant graft outcome after acute ABMR: univariate analysis

Parameters	HR	95% CI	*P* Value
Clinical parameters			
Recipient age, yr	1.02	0.98 to 1.06	0.31
Cold ischemia time, min	1.01	0.97 to 1.05	0.71
Donor age, yr			
<60	1	—	
≥60	3.58	1.45 to 8.87	0.01
Proteinuria at biopsy, g/L			
<0.15	1	—	
≥0.15	1.52	0.82 to 2.81	0.18
Immunologic parameters			
Immunodominant DSA MFI at the time of rejection	1.00	0.99 to 1.00	0.42
Functional parameters			
eGFR[a] at 1 yr			
eGFR≥30	1	—	
eGFR<30	3.04	1.28 to 7.22	0.01
Histologic parameters			
Arterial fibrous intimal thickening score	0.96	0.64 to 1.44	0.86
cg score	1.85	1.18 to 2.90	0.01
Interstitial fibrosis/tubular atrophy score	1.84	1.21 to 2.78	0.01
Interstitial inflammation score	0.75	0.45 to 1.24	0.26
Tubulitis score	0.66	0.39 to 1.11	0.12
v score[b]	1.20	0.45 to 3.08	0.13
g score	1.45	0.89 to 2.35	0.13
ptc score	1.99	0.97 to 4.07	0.06
C4d Banff score	1.47	1.03 to 2.08	0.03
Humoral histologic score[c]	1.36	1.08 to 1.71	0.01
Molecular parameters			
ABMR Molecular Score	8.82	1.82 to 42.73	0.01
Endothelial DSA-selective transcripts	2.94	1.00 to 8.69	0.05
NK transcripts	1.61	0.81 to 3.22	0.18
T-cell transcripts	0.93	0.50 to 1.73	0.83
AKI score	1.17	0.63 to 2.18	0.61

[a]eGFR using the Modification of Diet in Renal Disease formula.
[b]Note that all ABMR episodes with v lesions in the present study were treated by antibody-targeting strategies.[2]
[c]Humoral histologic score defined by g+ptc+v+cg+C4d Banff scores.

This study reinforces the notion that the classic histologic activity criteria including g, capillaritis, C4d, and endarteritis, which are part of the diagnostic criteria of ABMR, may not be as accurate for stratifying the risk inside the disease. This situation clearly differs from the comparison of sick versus well (ABMR versus everything else), where these parameters are consistently and strongly associated with increased risk of graft loss.

In our study, we found weak correlation between the ABMR Molecular Score or the endothelial DSA-selective transcripts and histologic lesions showing microvascular inflammation (Table 2) and found that ABMR patients with similar histopathology may show different levels of molecular signals, reflecting distinct activity and disease state. In addition, we show that, beyond a humoral histologic score defined by g+ptc+v+cg+C4d Banff score, the molecular microscope assessment brings independent risk prediction and additional value for individual risk reclassification (Figure 4, Table 4).

Although there are no doubts about the unmet medical need for improved diagnostics to allow a more personalized approach to therapy after kidney transplantation, our results go far beyond, bringing a new dimension to diagnosis and risk prediction in other solid organ transplants, such as heart, lung, small bowel, pancreas, and composite tissue. The ability of the ABMR Molecular Score and endothelial transcript measurements to reflect the risk of progression may be warranted for the development of therapeutic strategies that respond to the disease activity and stage.

Our study also has some limitations. It was not designed to determine the potential of the molecular microscope system to monitor the response to therapy. This aim would require additional investigations. Although our study was conducted using a reference platform for gene expression studies (Transplant Applied Diagnostic Centre in Alberta) with 10 years of experience with gene expression in the kidneys, established validation procedures, and a powerful biostatistical platform, multigene testing is strictly quantitative and requires normalization against the results of a standard population; it requires the tests to be conducted by a centralized reading service rather than local pathology laboratories, because a specific numerical value must be interpreted with respect to a defined reference set population. Although this model has its drawbacks, central testing and comparison with a reference set have enabled molecular testing in cancer biopsies to offer new insights into prognosis.[7–9] Finally, the costs of microarrays are rapidly dropping; nevertheless, it would be premature to determine whether the molecular microscope strategy is cost-effective until proper pharmacoeconomical studies are done.

The assessment of gene expression in kidney allograft transplant rejection seems to be useful in identifying patients with a high risk for kidney allograft loss. The molecular microscope system provides insight beyond the classic, histology-based approach and allows for risk stratification that may guide clinical management and clinical trials in transplant medicine.

CONCISE METHODS

Participants

From 939 patients who underwent single kidney transplantation at Necker Hospital between January of 2004 and October of 2010, we identified 101 patients (10.7%) with a diagnosis of biopsy-proven

Table 4. Determinants of kidney transplant graft outcome after acute ABMR (multivariate models) using the ABMR Molecular Score and endothelial DSA-selective transcripts

Parameters	Number of Patients	Number of Events	HR	95% CI	*P* Value
Model 1 with ABMR Molecular Score					
Donor age, yr					
<60	54	11	1	—	—
≥60	20	10	3.84	1.48 to 9.96	0.01
eGFR[a] (ml/min) at the time of rejection					
≥30	52	10	1		
<30	22	11	1.74	0.70 to 4.35	0.23
Humoral histologic score (g+ptc+v+cg+C4d)	74	21	1.43	1.09 to 1.90	0.01
ABMR Molecular Score	74	21	2.22	1.37 to 3.58	0.001
Model 2 with endothelial DSA-selective transcripts					
Donor age, yr					
<60	54	11	1	—	—
≥60	20	10	3.03	1.16 to 7.90	0.02
eGFR[a] (ml/min) at the time of rejection					
≥30	52	10	1		
<30	22	11	2.07	0.74 to 5.84	0.17
Humoral histologic score (g+ptc+v+cg+C4d)	74	21	1.22	0.94 to 1.60	0.14
Endothelial DSA-selective transcripts	74	21	3.02	1.00 to 9.16	<0.05

[a]eGFR using the Modification of Diet in Renal Disease formula.

ABMR in the first 1 year post-transplant; 27 patients were excluded because of a lack of suitable material for microarray analysis (14 patients had material that did not provide sufficient quality to run microarray and 13 patients did not have extracore performed at the time of diagnosis) (Supplemental Table 1), leaving 74 patients with ABMR as the study sample.

We used an additional independent validation sample of 321 patients who underwent kidney transplantation at Saint Louis Hospital between January 1, 2006, and January 1, 2010. Of these patients, 54 patients (16.8%) had early biopsy-proven ABMR with suitable material for microarray analysis (Supplemental Material). The patients were followed until April 15, 2013.

All of the transplants were ABO-compatible and had current negative IgG T-cell and B-cell complement-dependent cytotoxicity cross-matching at the time of transplantation. The study was approved by the institutional review board, and written informed consent was obtained from all study patients (registration number 1016618).

Clinical Data

Clinical data for the donors and recipients in the development (Necker Hospital) and validation (Saint Louis Hospital) cohorts were obtained from two national registries, the Données Informatiques Validées en Transplantation (Necker Hospital) and the Agence de la Biomédecine (Saint Louis Hospital), in which data are prospectively entered at specific time points for each patient (0 days, 6 months, and 1 year post-transplantation) and updated annually thereafter[18,19] (Supplemental Material). The development and validation cohort data were retrieved from the database on April 15, 2012.

We recorded the data for all patients with ABMR, which was defined as the association of deterioration in function or impaired function and histopathological lesions according to consensus rules represented by the international Banff classification criteria.[20] We acknowledged C4d-negative ABMR for patients with histologic evidence of ABMR and the presence of circulating DSA.[21]

Patients with ABMR received methylprednisolone pulses (500 mg/d for 3 days) and high-dose intravenous Ig (2 g/kg per week for three rounds). Patients on cyclosporine were switched to tacrolimus. In addition, patients received four plasmapheresis cycles and two weekly doses of rituximab (375 mg/m^2 body surface area; MabThera; Roche, Meylan, France). The treatment of ABMR episodes was standardized between the centers. The outcomes measured in this study were graft loss defined as the return to dialysis and progression of cg defined as the increase of Banff cg score by optical microscopy of kidney allografts.

Patients with ABMR were retrospectively reassessed for the conventional diagnosis criteria with review of histology and immunochemistry for C4d and identification of circulating donor-specific anti-HLA antibodies on sera saved at time of biopsies as well as microarray-based gene expression in allograft biopsies. In addition to the biopsies at the time of ABMR diagnosis, we determined the time to increase of cg score using 178 follow-up biopsies performed after the diagnosis of ABMR in the development cohort; cg score was defined by the duplication of the basement membrane according to Banff classification.[20,22] Electron microscopy was not performed in the present study.

All biopsies were reviewed by two renal pathologists blinded to the clinical information. The biopsies were graded from zero to three according to the Banff[20,22] histologic parameters: g, tubulitis, interstitial inflammation, endarteritis, ptc, cg, interstitial fibrosis, tubular atrophy, arterial fibrous intimal thickening, and arteriolar hyaline thickening. C4d staining was performed by immunohistochemistry on paraffin sections using the human C4d polyclonal antibody (Biomedica Gruppe, Austria).[23] C4d staining was graded from zero to three by the percentage of peritubular capillaries with linear staining, and grades 2 and 3 were considered to be

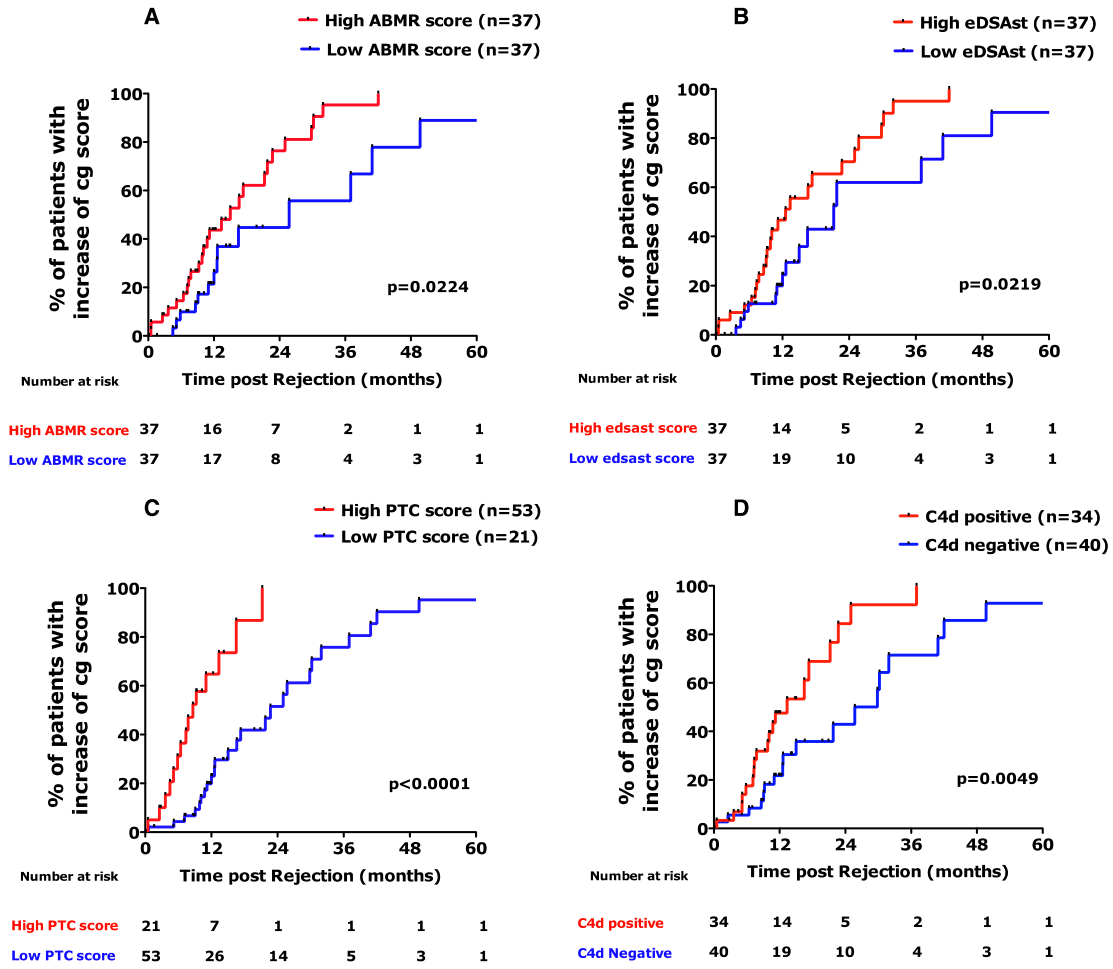


Figure 3. Time of progression of cg lesions after acute ABMR according to (A) ABMR Molecular Score, (B) endothelial DSA-selective transcripts, (C) ptc Banff score, and (D) C4d deposition in the peritubular capillaries. Progression of transplant glomerulopathy lesions is defined by increase of cg Banff score≥1. High and low ABMR Molecular Scores and the endothelial DSA-selective transcripts were defined using the median values for the ABMR and endothelial DSA-selective transcripts as thresholds. A high ptc score was defined by a ptc Banff score≥2 compared with a low ptc score (<2). C4d staining was graded from zero to three by the percentage of ptc having linear staining, with grades 2 or 3 considered to be positive. eDSAst, endothelial DSA-selective transcripts.

positive. We integrated humoral parameters (g, ptc, v, cg, and C4d) in a humoral histologic score defined by the sum of these variables.

Presence of circulating DSA at the time of transplantation (day 0) and the time of biopsy was analyzed using single-antigen flow bead assays (One Lambda, Canoga Park, CA) on the Luminex platform as previously detailed.[24] Beads showing a normalized mean fluorescence intensity (MFI) higher than 500 were considered positive. For each patient, we recorded the number, class, specificities, and MFI of all donor-specific HLA antibodies. The maximum MFI for the DSA (DSA MFImax) was defined as the highest ranked donor-specific bead. HLA typing of donors and recipients was performed using DNA typing (Innolipa HLA Typing Kit; Innogenetics, Belgium).

RNA Extraction and Microarrays Analyses

An additional 16-gauge biopsy core was obtained at the time of the biopsy and stored at −80°C in ornithine carbamyl transferase for gene expression analysis. All available biopsies performed at the time of the ABMR diagnosis were processed for microarray analysis as described previously.[25] RNA extraction, labeling, and hybridization to the HG_U133_Plus_2.0 GeneChip (Affymetrix, Santa Clara, CA) were performed according to the manufacturer's protocols. The

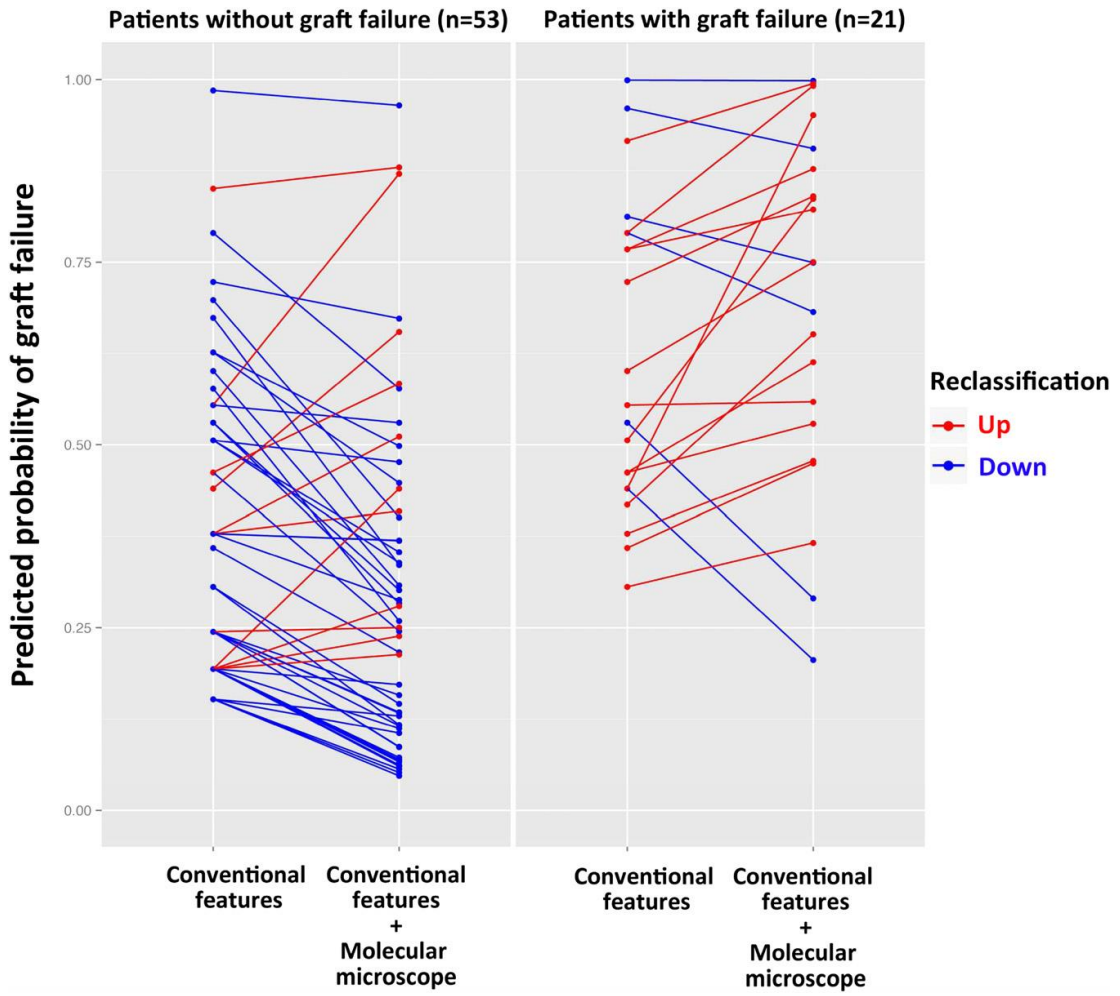


Figure 4. Additive value of the molecular microscope (ABMR Molecular Score) for reclassification of risk of allograft failure (continuous net reclassification improvement). Blue lines in patients without graft failure indicate that the molecular microscope moved risk prediction in the correct (downward) direction (42/53=79.2%). Conversely, red lines in patients with graft failure indicate a correct (upward) change in risk assessment when using the molecular microscope (15/21=71.4%).

microarrays were scanned using the Gene Array Scanner (Affymetrix) and processed with GeneChip Operating Software Version 1.4.0 (Affymetrix). The microarray expression files are posted on the Gene Expression Omnibus website under the accession number GSE36059.

The microarray data files for 126 biopsies were processed using Robust Multiarray Analysis in Bioconductor.[10] The molecular characteristics of the biopsies used pathogenesis-based transcript sets (PBTs) that reflect biologic processes of known relevance for the pathogenesis of renal inflammation and injury in transplants. Here, we applied the following PBTs: the ABMR Molecular Score[12] and endothelial DSA-selective transcripts.[26,27] The top genes for the ABMR Molecular Score and endothelial DSA-selective transcripts include many endothelial transcripts (*e.g.*, *DARC*, *VWF*, and *ROBO4*), NK transcripts, IFN-γ production and inducing transcripts (*e.g.*, *CXCL11*), and cytotoxicity-related transcripts (*e.g.*, *granulysin*[26] and *FGFBP2*[28]), which are summarized in Supplemental Tables 2 and 3. In addition, the NK-cell transcript burden,[29] T-cell transcript burden,[29] and AKI transcripts[13] were evaluated. The information on the probe sets and the algorithms for PBT generation are available at http://www.atagc.med.ualberta.ca/Research/GeneLists/Pages/default.aspx.

Statistical Analyses

We provide the mean (SD) values for the description of continuous variables, with the exception of the MFI, for which we use the mean (SEM) because of its wide distribution. We compared the means and proportions

of kidney transplant phenotypes using *t* tests, ANOVAs, and chi-squared tests (or Fisher's exact tests if appropriate). A Cox proportional hazard model was used to quantify the HRs and 95% CIs for the factors associated with kidney graft loss. Within each group of factors, we performed univariate and multivariate analyses. The selected factors were entered in a single multivariate Cox model (using P value ≤ 0.10 as a threshold) to identify the most predictive independent factors for kidney graft loss.

We further focused on the improvement in model performance because of inclusion of ABMR gene expression comparing two sets of predictions of 8-year graft loss risk probability: one set of predictions based on a Cox proportional hazards model without ABMR gene expression and one set of predictions based on a model with ABMR gene expression.

The discrimination ability and incremental value of ABMR gene expression were evaluated by C statistics. This analysis was repeated 1000 times using bootstrap samples to derive 95% CIs for the difference in the C statistic between models.

We also used net reclassification improvement and integrated discrimination improvement to quantify the performance and the net benefit of the addition of ABMR Molecular Score to the reference model.[30,31]

The proportionality assumption for the Cox model was verified using the log graphic method. The association between the molecular features and kidney transplant survival was subsequently replicated and confirmed in the independent validation sample.

The analyses were conducted using SAS 9.2 (SAS, Cary, NC). All of the tests were two-sided, and P values <0.05 were regarded as significant.

Ethical Approval

This study was approved by the Comité de Protection des Personnes Ile de France II. Each patient from the present study has given written informed consent to be included in the French national registry agency (Agence de la Biomédecine) databases CRISTAL (https://www.sipg.sante.fr/portail/) and DIVAT (https://www.divat.fr). The Données informatisées Validées en Transplantation (DIVAT) and CRISTAL database networks have been approved by the National French Commission for bioinformatics data and patients liberty: DIVAT: CNIL, registration number 1016618, validated June 8, 2004; CRISTAL: CNIL, registration number 363505, validated April 3, 1996.

ACKNOWLEDGMENTS

We are indebted to Vido Ramassar and Anna Hutton for excellent technical support. We thank Juliane Posson for data retrieval and Jerôme Verine from Saint Louis Hospital "Tumorothèque" for providing Hôpital Saint Louis samples. We thank the members of the Laboratory of Excellence, Transplantex for helpful discussions.

This research is supported by funding and/or resources from Novartis Pharma France and Canada and Astellas France. In the past, research was supported by Genome Canada, the University of Alberta Hospital Foundation, Roche Molecular Systems, Hoffmann-La Roche Canada Ltd., the Alberta Ministry of Advanced Education and Technology, the Roche Organ Transplant Research Foundation, and Astellas. P.F.H. held a Canada Research Chair in Transplant Immunology until 2008 and currently holds the Muttart Chair in Clinical Immunology.

The authors declare no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years and no other relationships or activities that could seem to have influenced the submitted work. All authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

DISCLOSURES

P.F.H. has shares in TSI, a company with an interest in molecular diagnostics; the other authors of this manuscript have no conflicts of interest to disclose.

REFERENCES

1. Garcia GG, Harden P, Chapman J; World Kidney Day Steering Committee 2012: The global role of kidney transplantation. *Lancet* 379: e36–e38, 2012
2. Lefaucheur C, Loupy A, Vernerey D, Duong-Van-Huyen JP, Suberbielle C, Anglicheau D, Vérine J, Beuscart T, Nochy D, Bruneval P, Charron D, Delahousse M, Empana JP, Hill GS, Glotz D, Legendre C, Jouven X: Antibody-mediated vascular rejection of kidney allografts: A population-based study. *Lancet* 381: 313–319, 2013
3. Einecke G, Sis B, Reeve J, Mengel M, Campbell PM, Hidalgo LG, Kaplan B, Halloran PF: Antibody-mediated microcirculation injury is the major cause of late kidney transplant failure. *Am J Transplant* 9: 2520–2531, 2009
4. Gaston RS, Cecka JM, Kasiske BL, Fieberg AM, Leduc R, Cosio FC, Gourishankar S, Grande J, Halloran P, Hunsicker L, Mannon R, Rush D, Matas AJ: Evidence for antibody-mediated injury as a major determinant of late kidney allograft failure. *Transplantation* 90: 68–74, 2010
5. Nankivell BJ, Alexander SI: Rejection of the kidney allograft. *N Engl J Med* 363: 1451–1462, 2010
6. Montgomery RA, Lonze BE, King KE, Kraus ES, Kucirka LM, Locke JE, Warren DS, Simpkins CE, Dagher NN, Singer AL, Zachary AA, Segev DL: Desensitization in HLA-incompatible kidney recipients and survival. *N Engl J Med* 365: 318–326, 2011
7. Hartmann S, Gerber B, Elling D, Heintze K, Reimer T: The 70-gene signature as prognostic factor for elderly women with hormone receptor-positive, HER2-negative breast cancer. *Breast Care (Basel)* 7: 19–24, 2012
8. Reed SD, Lyman GH: Cost effectiveness of gene expression profiling for early stage breast cancer: A decision-analytic model. *Cancer* 118: 6298–6299, 2012
9. Prat A, Parker JS, Fan C, Cheang MC, Miller LD, Bergh J, Chia SK, Bernard PS, Nielsen TO, Ellis MJ, Carey LA, Perou CM: Concordance among gene expression-based predictors for ER-positive breast cancer treated with adjuvant tamoxifen. *Ann Oncol* 23: 2866–2873, 2012
10. Reeve J, Sellarés J, Mengel M, Sis B, Skene A, Hidalgo L, de Freitas DG, Famulski KS, Halloran PF: Molecular diagnosis of T cell-mediated rejection in human kidney transplant biopsies. *Am J Transplant* 13: 645–655, 2013
11. Halloran PF, Pereira AB, Chang J, Matas A, Picton M, De Freitas D, Bromberg J, Serón D, Sellarés J, Einecke G, Reeve J: Microarray diagnosis of antibody-mediated rejection in kidney transplant biopsies: An international prospective study (INTERCOM). *Am J Transplant* 13: 2865–2874, 2013
12. Sellarés J, Reeve J, Loupy A, Mengel M, Sis B, Skene A, de Freitas DG, Kreepala C, Hidalgo LG, Famulski KS, Halloran PF: Molecular diagnosis of antibody-mediated rejection in human kidney transplants. *Am J Transplant* 13: 971–983, 2013
13. Famulski KS, de Freitas DG, Kreepala C, Chang J, Sellares J, Sis B, Einecke G, Mengel M, Reeve J, Halloran PF: Molecular phenotypes of acute kidney injury in kidney transplants. *J Am Soc Nephrol* 23: 948–958, 2012

14. Archdeacon P, Chan M, Neuland C, Velidedeoglu E, Meyer J, Tracy L, Cavaille-Coll M, Bala S, Hernandez A, Albrecht R: Summary of FDA antibody-mediated rejection workshop. *Am J Transplant* 11: 896–906, 2011
15. Sellarés J, de Freitas DG, Mengel M, Reeve J, Einecke G, Sis B, Hidalgo LG, Famulski K, Matas A, Halloran PF: Understanding the causes of kidney transplant failure: The dominant role of antibody-mediated rejection and nonadherence. *Am J Transplant* 12: 388–399, 2012
16. Pencina MJ, D'Agostino RB Sr., Steyerberg EW: Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 30: 11–21, 2011
17. Lefaucheur C, Nochy D, Hill GS, Suberbielle-Boissel C, Antoine C, Charron D, Glotz D: Determinants of poor graft outcome in patients with antibody-mediated acute rejection. *Am J Transplant* 7: 832–841, 2007
18. Données Informatiques Validées en Transplantation (DIVAT): Available at: http://www.divat.fr/. Accessed April 15, 2012
19. Agence Biomédecine: Cristal. Available at: http://www.sipg.sante.fr/portail/. Accessed April 15, 2012
20. Sis B, Mengel M, Haas M, Colvin RB, Halloran PF, Racusen LC, Solez K, Baldwin WM 3rd, Bracamonte ER, Broecker V, Cosio F, Demetris AJ, Drachenberg C, Einecke G, Gloor J, Glotz D, Kraus E, Legendre C, Liapis H, Mannon RB, Nankivell BJ, Nickeleit V, Papadimitriou JC, Randhawa P, Regele H, Renaudin K, Rodriguez ER, Seron D, Seshan S, Suthanthiran M, Wasowska BA, Zachary AA, Zeevi A: Banff '09 meeting report: Antibody mediated graft deterioration and implementation of Banff working groups. *Am J Transplant* 10: 464–471, 2010
21. Haas M, Sis B, Racusen LC, Solez K, Glotz D, Colvin RB, Castro MC, David DS, David-Neto E, Bagnasco SM, Cendales LC, Cornell LD, Demetris AJ, Drachenberg CB, Farver CF, Farris AB 3rd, Gibson IW, Kraus E, Liapis H, Loupy A, Nickeleit V, Randhawa P, Rodriguez ER, Rush D, Smith RN, Tan CD, Wallace WD, Mengel M; the Banff Meeting Report Writing Committee: Banff 2013 meeting report: Inclusion of c4d-negative antibody-mediated rejection and antibody-associated arterial lesions. *Am J Transplant* 14: 272–283, 2014
22. Mengel M, Sis B, Haas M, Colvin RB, Halloran PF, Racusen LC, Solez K, Cendales L, Demetris AJ, Drachenberg CB, Farver CF, Rodriguez ER, Wallace WD, Glotz D; Banff Meeting Report Writing Committee: Banff 2011 Meeting report: New concepts in antibody-mediated rejection. *Am J Transplant* 12: 563–570, 2012
23. Loupy A, Hill GS, Suberbielle C, Charron D, Anglicheau D, Zuber J, Timsit MO, Duong JP, Bruneval P, Vernerey D, Empana JP, Jouven X, Nochy D, Legendre CH: Significance of C4d Banff scores in early protocol biopsies of kidney transplant recipients with preformed donor-specific antibodies (DSA). *Am J Transplant* 11: 56–65, 2011
24. Lefaucheur C, Loupy A, Hill GS, Andrade J, Nochy D, Antoine C, Gautreau C, Charron D, Glotz D, Suberbielle-Boissel C: Preexisting donor-specific HLA antibodies predict outcome in kidney transplantation. *J Am Soc Nephrol* 21: 1398–1406, 2010
25. Mueller TF, Einecke G, Reeve J, Sis B, Mengel M, Jhangri GS, Bunnag S, Cruz J, Wishart D, Meng C, Broderick G, Kaplan B, Halloran PF: Microarray analysis of rejection in human kidney transplants using pathogenesis-based transcript sets. *Am J Transplant* 7: 2712–2722, 2007
26. Hidalgo LG, Sis B, Sellares J, Campbell PM, Mengel M, Einecke G, Chang J, Halloran PF: NK cell transcripts and NK cells in kidney biopsies from patients with donor-specific antibodies: Evidence for NK cell involvement in antibody-mediated rejection. *Am J Transplant* 10: 1812–1822, 2010
27. Sis B, Jhangri GS, Bunnag S, Allanach K, Kaplan B, Halloran PF: Endothelial gene expression in kidney transplants with alloantibody indicates antibody-mediated damage despite lack of C4d staining. *Am J Transplant* 9: 2312–2323, 2009
28. Ogawa K, Tanaka K, Ishii A, Nakamura Y, Kondo S, Sugamura K, Takano S, Nakamura M, Nagata K: A novel serum protein that is selectively produced by cytotoxic lymphocytes. *J Immunol* 166: 6404–6412, 2001
29. Hidalgo LG, Sellares J, Sis B, Mengel M, Chang J, Halloran PF: Interpreting NK cell transcripts versus T cell transcripts in renal transplant biopsies. *Am J Transplant* 12: 1180–1191, 2012
30. Steyerberg EW: *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*, 1st Ed., New York, Springer Science + Business Media, LLC, 2009
31. Pencina MJ, D'Agostino RB: Overall C as a measure of discrimination in survival analysis: Model specific population value and confidence interval estimation. *Stat Med* 23: 2109–2123, 2004

This article contains supplemental material online at http://jasn.asnjournals.org/lookup/suppl/doi:10.1681/ASN.2013111149/-/DCSupplemental.

VIII.2 Annexe 2:

Fiteni F*, **Vernerey D***, Bonnetain F, Vaylet F, Sennélart H, Trédaniel J, Moro-Sibilot D, Herman D, Laizé H, Masson P, Derollez M, Clément-Duchêne C, Milleron B, Morin F, Zalcman G, Quoix E, Westeel V (2016) Prognostic value of health-related quality of life for overall survival in elderly non-small-cell lung cancer patients. Eur J Cancer Oxf Engl 1990 52: 120–128, doi:10.1016/j.ejca.2015.10.004.

*Co first authorship

Valeur pronostique de la qualité de vie relative à la santé pour la survie globale au diagnostic chez les patients âgés atteints d'un cancer du poumon avancé non à petites cellules

Contexte

Nous avons étudié si les informations véhiculées par la qualité de vie relative à la santé (QdV) peuvent être pronostiques de la survie globale (SG) chez des patients âgés atteints d'un cancer du poumon avancé non à petites cellules (CBNPC) au diagnostic.

Méthodes

Au total, 451 patients âgés de 70 à 89 ans atteints de CBNPC avancés ont été inclus dans l'essai IFCT-0501. Le questionnaire QLQ-C30 administré au diagnostic a été utilisé pour étudier la valeur pronostique additionnelle aux paramètres conventionnels de la QdV sur la SG. Une modélisation de Cox a été réalisée pour l'analyse univariée et multivariée. L'évaluation de l'ajout en termes de discrimination et de reclassification du risque de décès de la QdV a ensuite été réalisée.

Résultats

La dimension de santé globale (GH) au diagnostic est associée à une SG favorable après ajustement sur les facteurs cliniques, fonctionnels et histologique (hazard ratio [HR]: 0.986; IC 95% : 0.980-0.992).

Nous avons distingué 3 groupes selon le niveau de score GH : élevé (GH < 46), intermédiaire (46 ≤ GH ≤ 67), et faible (GH > 67) risque de décès. La médiane de SG était respectivement de 14.5, 8.2 et 5.3 mois dans les groupes à risque faible, intermédiaire et élevé de décès (log-rank P < 0.0001). Dans le groupe à risque élevé, la bi-chimiothérapie n'était pas associé à une survie favorable (HR = 0.70; IC 95% : 0.49-1.003; P = 0.052), alors que dans le groupe intermédiaire et le groupe à faible risque, la bi-chimiothérapie était associé avec une survie globale favorable (HR = 0.72; IC 95%: 0.54-0.96; P = 0.023 et HR = 0.50; IC 95%: 0.30-0.84; P = 0.0089, respectivement).

Conclusion

Cette étude démontre la valeur pronostique additionnelle de la QdV pour prédire la survie globale au diagnostic. La QdV permet d'identifier des sous-groupes de populations à risque de décès différents. De plus, la QdV pourrait permettre de sélectionner les patients qui peuvent tirer d'un bénéfice de l'administration d'une bi-chimiothérapie.

European Journal of Cancer 52 (2016) 120–128

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Prognostic value of health-related quality of life for overall survival in elderly non-small-cell lung cancer patients

Frédéric Fiteni [a,b,c,*,1], Dewi Vernerey [a,1], Franck Bonnetain [a,c,d], Fabien Vaylet [e], Hélène Sennélart [f], Jean Trédaniel [g], Denis Moro-Sibilot [h], Dominique Herman [i], Hélène Laizé [j], Philippe Masson [k], Marc Derollez [l], Christelle Clément-Duchêne [m], Bernard Milleron [n], Franck Morin [n], Gérard Zalcman [n,o], Elisabeth Quoix [n,p], Virginie Westeel [n,q]

[a] *University Hospital of Besançon, Methodology and Quality of Life in Oncology Unit, Besançon, France*
[b] *University Hospital of Besançon, Department of Medical Oncology, Besançon, France*
[c] *EA 3181 University of Franche-Comté, Besançon, France*
[d] *Platform Quality of Life and Cancer, France*
[e] *Instruction Hospital of the Armies, Department of Pneumology, Percy-Clamart, France*
[f] *CLCC René Gauducheau, Department of Medical Oncology, Nantes, France*
[g] *St Joseph Hospital, Department of Medical Oncology, Paris, France*
[h] *University Hospital of Grenoble, Pneumology Department, Grenoble, France*
[i] *Hospital of Nevers, Pneumology Department, Nevers, France*
[j] *Hospital of Sceaux, Pneumology Department, Sceaux, France*
[k] *Hospital of Cholet, Pneumology Department, Cholet, France*
[l] *Private Hospital, Pneumology Department, Maubeuge, France*
[m] *University Hospital of Nancy, Pneumology Department, Nancy, France*
[n] *Intergroupe Francophone de Cancérologie Thoracique (IFCT), Paris, France*
[o] *University Hospital of Caen, Pneumology Department, Caen, France*
[p] *University Hospital of Strasbourg, Pneumology Department, France*
[q] *University Hospital of Besançon, Pneumology Department, France*

Received 13 August 2015; received in revised form 20 September 2015; accepted 4 October 2015
Available online xxx

* *Corresponding author*: University Hospital of Besançon, Department of Medical Oncology, 3 Boulevard Fleming, 25000, Besançon, France. Tel.: +333 81668796.
E-mail address: fredericfiteni@gmail.com (F. Fiteni).
[1] Shared authorship.

http://dx.doi.org/10.1016/j.ejca.2015.10.004
0959-8049/© 2015 Elsevier Ltd. All rights reserved.

KEYWORDS
Quality of life;
Prognostic factor;
Methodology;
Lung cancer

Abstract ***Background:*** We investigated whether the health-related quality of life (HRQoL) score is a prognostic factor for overall survival (OS) in elderly patients with advanced non-small-cell lung cancer (NSCLC).

Methods: We included 451 NSCLC patients aged 70–89 years enrolled in the Intergroupe Francophone de Cancérologie Thoracique 0501 trial, using scores of the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 at baseline to investigate the prognostic value of HRQoL for OS, in addition to conventional factors. Cox regression model was used for both univariate and multivariate analyses of OS.

Results: Global health status (GH) dimension score at baseline was associated with favourable OS when adjusted for clinical, functional, and histological factors (hazard ratio [HR]: 0.986; 95% confidence interval [CI]: 0.980–0.992).

We distinguished three groups according to GH score: high (GH <46), intermediate (46 ≤GH ≤67), and low (GH >67) mortality risk. The median OS values were 14.5, 8.2, and 5.3 months in the low-, intermediate-, and high-risk categories, respectively (log-rank P <0.0001).

In the high-risk group, doublet chemotherapy was not associated with favourable OS (HR: 0.70; 95% CI: 0.49–1.003; P = 0.052), whereas in the intermediate- and low-risk groups, doublet chemotherapy was associated with favourable OS (HR: 0.72; 95% CI: 0.54–0.96; P = 0.023 and HR: 0.50; 95% CI: 0.30–0.84; P = 0.0089, respectively).

Conclusion: This study supports the additional prognostic value of HRQoL data at diagnosis to identify vulnerable subpopulations in elderly NSCLC patients. HRQoL could thus be valuable in selecting patients who will benefit from doublet chemotherapy.
© 2015 Elsevier Ltd. All rights reserved.

1. Background

The number of studies using health-related quality of life (HRQoL) assessment has been growing over the last decade. The Food and Drug Administration considers HRQoL to be an end-point for assessing direct clinical benefits for the patient [1–3]. Moreover, there has been evidence to suggest that assessing baseline HRQoL dimension scores in cancer patients improves the prediction of overall survival (OS) [4–9]. Quinten et al. carried out a meta-analysis involving over 10,000 cancer patients (16% lung cancer), revealing that baseline HRQoL was a prognosticator of longer survival [10]. In non-small-cell lung cancer (NSCLC) patients, several studies have demonstrated that HRQoL represents a significant prognosticator of favourable OS [6,9,11–13]. Sloan et al. prospectively observed 2,442 patients with stage I–IV NSCLC, 47% ≤65 years old and 53% >65 years old, all completing a single-item measure of overall HRQoL from the Lung Cancer Symptom Scale questionnaire within the first 6 months post-diagnosis. They demonstrated that QoL deficits at diagnosis were significantly associated with shorter OS (hazard ratio [HR]: 1.55; P < 0.001). Yet no study has specifically focused on elderly advanced NSCLC patients.

We sought to investigate the additional prognostic value of baseline HRQoL assessed by European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire Core 30 (QLQ-C30) in elderly advanced NSCLC patients treated with chemotherapy in the randomised Intergroupe Francophone de Cancérologie Thoracique (IFCT) 0501 trial.

2. Methods

2.1. Sample

The IFCT 0501 study design has been previously described [14]. Patients aged 70–89 years with stage IV NSCLC or stage III unsuitable for radical radiation therapy and performance status (PS) ≤2 were eligible for this phase III trial. They were randomly assigned 1:1 to four 28-day cycles of monthly carboplatin plus weekly paclitaxel or five 21-day cycles of single agent vinorelbine or gemcitabine, on days 1 and 8 of each cycle. Patients were stratified by centre, World Health Organization (WHO) PS score (0–1 versus 2), stage (III versus IV), and age (≤80 versus >80 years).

The protocol was approved by the *Comité de Protection des Personnes* of Ile-de-France X, Aulnay-sous-Bois, France, the trial being authorised by the French National Authority for Health. All patients provided written informed consent.

2.2. Health-related quality of life

HRQoL was assessed using EORTC QLQ-C30 questionnaire [15] at randomisation, then at 6 and 18 weeks.

The QLQ-C30 is a cancer-specific tool composed of 30 items [16–18]. Five functional scores (physical, role, cognitive, social, and emotional) have been developed, rated on a global health score ranging from 0 (worst) to 100 (best), as well as nine symptom scores (nausea, pain [PA], fatigue, dyspnoea, difficulty sleeping, anorexia, constipation, diarrhoea, and perceived financial difficulties), ranging from 0 (best) to 100 (worse).

2.3. Statistical analysis

We used means and standard deviations for continuous variables, and proportions for qualitative variables. We compared means and proportions using Student's t-test and chi-squared test, or Fisher's exact test if appropriate.

Patient characteristics were described based on the completion of baseline QoL questionnaire, ensuring that any non-random missing patient profiles were detected.

OS was defined as the time from randomisation to death from any cause. Survival was censored at last follow-up or time of analysis. OS was estimated using the Kaplan-Meier method and presented as median with 95% confidence interval (CI). Follow-up was calculated using a reverse Kaplan–Meier estimation.

The association of non-HRQoL characteristics at diagnosis and baseline HRQoL dimensions in terms of OS was assessed using univariate Cox regression analysis, followed by multivariate analysis for those exhibiting $P < 0.1$. The factors identified with a $P < 0.1$ in multivariate analysis were thereafter included in a final multivariate model with stepwise backward elimination ($P < 0.05$).

The proportionality assumption for the Cox model was verified using the log graphic method. Collinearity of baseline HRQoL scores with other covariates was examined using a multiple linear regression model.

The Hosmer and Lemeshow goodness-of-fit statistics test, adapted for survival analysis, was used to evaluate the final model's calibration.

Internal validation using a bootstrap procedure was performed to assess the final model's robustness, analysing hazard regression uncertainty for parameters involved in the final model [19].

The prognostic value of HRQoL scores added to a reference risk model, including the non-HRQoL characteristics enrolled in the final multivariate model, was evaluated using C-statistics. Harrell's C-index estimates discriminate capability, i.e. the ability to distinguish between high-risk and low-risk patients, the C-index varying from 0.5 (no discrimination) to 1 (perfect discrimination). This analysis was repeated 1000 times using bootstrap samples to derive 95% CIs for the between-model difference in C-statistics.

We used continuous net reclassification improvement (cNRI) and integrated discrimination improvement (IDI) to quantify the performance and net benefit of adding HRQoL scores to the reference model at 24 months post-randomisation [20,21]. The cNRI quantifies the direction of change, and the IDI the magnitude of change. When significantly greater than 0, IDI and cNRI suggest the existence of a net benefit through adding the marker of interest to the reference model.

To implement HRQoL scoring into clinical practice, we determined a cut-off value via an unsupervised method using the Q1 and Q3 interquartiles.

We performed sensitivity analyses. We first included HRQoL and clinical factors with a $P < 0.1$ in the univariate analysis using a stepwise backward elimination procedure. We then conducted a stepwise multivariate model with the treatment variable as stratification factor for the final model construction. As some data from the activities of daily living (ADL), mini-mental state (MMS), and HRQol questionnaires could have been redundant by the time of analysis, we eventually conducted a stepwise multivariable model excluding the ADL score, then the MMS score, and eventually both.

Tests were two-sided, with P-values <0.05 considered significant. The analyses were conducted using SAS 9.2 (SAS, Cary, NC) and R software (Version 2.10.1).

3. Results

3.1. Study population

Between April 2006 and December 2009, 451 patients were enrolled. The number of patients who completed the entire questionnaire at baseline was 361 (80.04%), the number of available questionnaires (i.e. with at least one QoL score that could be calculated) being 421 (93.3%). At baseline, the patients who completed the entire QoL questionnaire and those who did not were found to display similar clinical characteristics (Table 1). The baseline HRQoL scores for each dimension have been presented in Table 2. Median follow-up was 30.3 months (range: 8.6–45.2). There were 199 (88%) deaths under monotherapy versus 177 (78.6%) under doublet chemotherapy.

3.2. Association between baseline HRQoL scores and overall survival

Data pertaining to the association of clinical and HRQoL dimension scores in terms of OS is shown in Tables 3A and B. In the final model, the following characteristics were independent favourable prognosticators of OS: increasing global health status (GH) score (HR: 0.986; 95% CI: 0.980–0.992; $P < 0.0001$), PS 0–1 (HR: 0.63; 95% CI: 0.49–0.81; $P < 0.0001$), doublet chemotherapy (HR: 0.65; 95% CI: 0.52–0.80; P

Table 1
Patient characteristics according to Quality of Life Questionnaire Core 30 completion.

	Patients who completed the questionnaire at baseline (n=361)		Patients who did not complete the questionnaire at baseline (n=90)		Fisher's exact test P-value
	n	(%)	n	(%)	
Age					
<77	180	49.9	48	53	0.56
≥77	181	50.1	42	47	
Gender					
Male	270	74.8	63	70	0.35
Female	91	25.2	27	30	
Performance status					
0–1	267	74	63	70	0.15
2	93	25.7	27	30	
Unknown	1	0.3	0		
Smoking status					
Never smoked	73	20	21	23	0.51
Ever smoked	288	80	69	77	
MMS					
≤20	29	8	6	7	0.82
>20	330	91	76	84	
Unknown	2	1	8	9	
ADL					
<6	66	18	22	24	0.12
6	288	80	62	69	
Unknown	7	2	6	7	
CCI					
≤2	268	74	73	81	0.17
>2	93	26	17	19	
BMI					
≤20	43	12	9	10	0.77
20<BMI≤30	276	76	72	80	
>30	42	12	9	10	
Stage					
IIIA–IIIB	70	19	17	19	0.9
IV	291	81	73	81	
Histology					
Adenocarcinoma	184	51	45	50	0.63
Squamous	123	34	28	21	
Other	54	15	17	19	

MMS = mini-mental state score; ADL = activities of daily living score; CCI = Charlson comorbidity index; BMI = body mass index.

<0.0001), never smoked status (HR: 0.58; 95% CI: 0.43–0.78; P=0.0003), adenocarcinoma (HR: 0.68; 95% CI: 0.50–0.93; P=0.047), increasing ADL score (HR: 0.73; 95% CI: 0.61–0.88; P=0.0011), and increasing MMS score (HR: 0.97; 95% CI: 0.94–0.99, P=0.044) (Table 4). The model's calibration was acceptable (Hosmer–Lemeshow with deciles p=0.1). The internal validation HR uncertainties reflected its robustness, especially the association between GH score and OS (Table 4).

Entering the GH score into the reference model was found to significantly improve its discriminative ability, notably its capacity to discriminate between patients who died and those who did not, as C-index

Table 2
Health-related quality of life scores at baseline by treatment arm.

QLQ-C30 scores	All patients (n=451)		
	N*	Mean (SD)	Median (minimum–maximum)
Global health status	420	56.8 (18.7)	58.3 (8.3–100)
Physical functioning	420	69.2 (22.5)	73.3 (0–100)
Role functioning	420	66 (34.2)	66.7 (0–100)
Emotional functioning	420	72.2 (22.6)	75 (0–100)
Cognitive functioning	421	82.4 (21.1)	82.3 (0–100)
Social functioning	411	78 (30.5)	100 (0–100)
Fatigue	420	43.5 (27.8)	33.3 (0–100)
Nausea/vomiting	421	5.5 (15.2)	0 (0–100)
Pain	420	27.9 (29.8)	16.7 (0–100)
Dyspnoea	419	44.2 (34.7)	33.3 (0–100)
Insomnia	420	28.9 (32.3)	33.3 (0–100)
Appetite loss	419	35.8 (37.2)	33.3 (0–100)
Constipation	420	25.6 (32)	0 (0–100)
Diarrhoea	417	7.5 (18.4)	0 (0–100)
Financial problems	416	4.4 (15.2)	0 (0–100)

SD = standard deviation; QLQ-C30 = Quality of Life Questionnaire Core 30.

* Number of patients with HRQoL score at baseline that can be calculated high score for a functional scale represents a *high/healthy level of functioning*, high score for the global health status/HRQoL represents a *high HRQoL*, but high score for a symptom scale/item represents a *high level of symptomatology/problems*.

statistics significantly increased from 0.66 to 0.69 (bootstrap mean difference: 0.0253; 95% CI: 0.0248–0.0259). Similarly, including the GH score into the reference model adequately reclassified patients into lower (no events) or higher (events) mortality risk, as demonstrated by a continuous net reclassification index of 0.38 (95% CI: 0.13–0.64) at 24 months post-randomisation. When adding the GH score adequately reclassified 42/68 patients (61.8%) into the 'no event' group, whereas it reclassified 196/341 patients (57.4%) into the 'event' group (Fig. 1). The IDI was 0.03 (P=0.0006). Adding the GH score to the classical risk factors improved the stratification of patients at risk of death.

The factors affecting baseline GH scores were explored using multiple linear regression. Increasing baseline GH score was associated with a PS=0–1 (P <0.0001), increasing ADL score (P=0.0061), and increasing body mass index (BMI) (P=0.023). Nevertheless, this model exhibited R^2 statistics of 12%, indicating that the GH score was not completely accounted for by PS, ADL score, and BMI.

3.3. Sensitivity analyses

In the multivariate Cox model including all variables, GH dimension score remained associated with OS (HR: 0.98; 95% CI: 0.97–0.99; P=0.0002) (Supplementary Table 1). All variables significantly influencing OS in the final model were also found to be statistically

Table 3A
Univariate and multivariate Cox regression with clinical parameters associated with overall survival.

	Number of patients	Number of events	Univariate analysis (n=451)		Multivariate analysis (n=393)	
			HR (95% CI)	P-value	HR (95% CI)	P-value
Gender						
Male	333	282	1.00		1.00	
Female	118	95	0.77 (0.61–0.97)	0.026	1.053 (0.76–1.47)	0.76
Age (years)	451	377	0.99 (0.97–1.02)	0.60		
Treatment						
Monochemotherapy	225	199	1.00		1.00	
Doublet chemotherapy	226	178	0.64 (0.52–0.78)	<0.0001	0.61 (0.50–0.77)	<0.0001
Performance status score						
0–1	327	262	1.00		0.59 (0.46–0.75)	
2	123	114	2.10 (1.67–2.60)	<0.0001	1.00	<0.0001
Smoking status						
Never smoked	87	68	0.65 (0.50–0.84)		0.58 (0.39–0.85)	
Ever smoked	364	309	1.00	0.001	1.00	0.0034
Disease stage						
III	82	71	1.00			
IV	349	306	1.05 (0.81–1.36)	0.71		
Histology						
Adenocarcinoma	219	178	0.55 (0.41–0.73)		0.65 (0.48–0.87)	
Squamous	146	131	0.75 (0.56–1.01)		0.76 (0.56–1.03)	
Other	66	68	1.00	<0.0001	1.00	0.0015
MMS	441	377	0.96 (0.93–0.98)	0.0013	0.97 (0.94–1.00)	0.057
ADL	438	377	0.64 (0.55–0.74)	<0.0001	0.70 (0.59–0.83)	<0.0001
CCI	451	377	1.06 (0.99–1.13)	0.086	0.98 (0.96–1.01)	0.77
BMI	451	377	0.98 (0.95–1.00)	0.067	1.01 (0.94–1.09)	0.17

HR = hazard ratio; CI = confidence interval; MMS = mini-mental state; ADL = activities of daily living; CCI = Charlson comorbidity index; BMI = body mass index.

significant in this model. Only one variable, not included into the final model, significantly and negatively correlated with OS: the PA dimension score (HR: 1.004; 95% CI: 1.00–1.01; P=0.046).

In the model including treatment as stratification variable, GH score, PS 0–1, never smoked status, ADL score, and MMS score were favourably associated with OS (Supplementary Table 2).

GH score remained statistically significant in all sensitivity analyses, conducted without ADL, MMS and both scores. All other covariates significantly associated with OS in our final model were found to be statistically significant prognosticators of OS (Supplementary Tables 3–5). Physical functioning (PF) score was favourably associated with OS when the ADL score was not included.

Table 3B
Univariate and multivariate Cox regression only with health-related quality of life dimensions scores associated with overall survival.

	Number of patients	Number of events	Univariate analysis (n=451)		Multivariate analysis (n=451)	
			HR (95% CI)	P-value	HR (95% CI)	P-value
Global health status	420	377	0.98 (0.98–0.99)	<0.0001	0.986 (0.98–0.992)	<0.0001
Physical functioning	420	377	0.98 (0.98–0.99)	<0.0001	0.989 (0.984–0.995)	0.0003
Role functioning	420	377	0.99 (0.987–0.993)	<0.0001	0.994 (0.991–0.998)	0.001
Emotional functioning	420	377	0.989 (0.989–0.998)	0.0055	0.996 (0.991–1.000)	0.068
Cognitive functioning	421	377	0.992 (0.987–0.997)	0.0011	0.996 (0.991–1.001)	0.12
Social functioning	411	377	0.993 (0.989–0.996)	<0.0001	0.997 (0.993–1.001)	0.12
Fatigue	420	377	1.011 (1.007–1.015)	<0.0001	1.007 (1.003–1.011)	0.0012
Nausea and vomiting	421	377	1.003 (0.996–1.010)	0.40		
Pain	420	377	1.010 (1.006–1.013)	<0.0001	1.007 (1.003–1.011)	0.0003
Dyspnoea	419	377	1.007 (1.004–1.010)	<0.0001	1.004 (1.000–1.007)	0.031
Insomnia	420	377	1.003 (1.000–1.006)	0.0983	1.002 (0.998–1.005)	0.33
Appetite loss	419	377	1.007 (1.004–1.010)	<0.0001	1.005 (1.001–1.008)	0.006
Constipation	420	377	1.004 (1.000–1.007)	0.025	1.002 (0.999–1.006)	0.22
Diarrhoea	417	377	1.004 (0.998–1.009)	0.23		
Financial difficulties	416	377	1.005 (0.998–1.012)	0.14		

HR = hazard ratio; CI = confidence interval.

Table 4
Clinical and health-related quality of life dimension scores associated with overall survival.

	Number of patients	Number of events	HR (95% CI)	P-value
GH	420	377	0.986 (0.980–0.992)	<0.0001
Treatment				
Monochemotherapy	225	199	1.00	
Doublet chemotherapy	226	198	0.65 (0.52–0.80)	<0.0001
Performance status score				
0–1	327	262	0.63 (0.49–0.81)	
2	123	114	1.00	0.0003
Smoking status				
Never smoked	87	68	0.58 (0.43–0.78)	
Ever smoked	264	309	1.00	0.0003
Histology				
Adenocarcinoma	219	178	0.68 (0.50–0.93)	
Squamous	146	131	0.80 (0.58–1.01)	
Other	66	68	1.00	0.047
MMS	441	377	0.97 (0.94–1.00)	0.043
ADL	438	377	0.81 (0.66–0.99)	0.0011

HR = hazard ratio; CI = confidence interval; GH = global health score; MMS = mini-mental state; ADL = activities of daily living.

3.4. Proposal for implementing HRQoL

When using the interquartile ranges (GH $\leq$46, 46<GH<67, and GH $\geq$67), GH score remained associated with OS in stepwise multivariable Cox regression (P <0.0001) (Table 5). We thus distinguished three groups, categorised as high (GH <46), intermediate (46$\leq$GH$\leq$67), or low (GH >67) mortality risk. Median OS values were 5.3, 8.2, and 14.5 months in the low-, intermediate-, and high-risk groups, respectively (log-rank P <0.0001) (Fig. 2).

In the high-risk subgroup, doublet chemotherapy was not associated with favourable OS (HR: 0.70; 95% CI: 0.49–1.003; P=0.052). In the intermediate- and low-risk groups, doublet chemotherapy was associated with favourable OS (HR: 0.72; 95% CI: 0.54–0.96; P=0.023, and HR: 0.50; 95% CI: 0.30–0.84; P=0.0089, respectively).

4. Discussion

To our knowledge, this is the first analysis of HRQoL data derived from the EORTC QLQ-C30 questionnaire as prognostic markers of OS in elderly advanced NSCLC patients. Based on our data, the GH dimension score provided significant value in addition to PS, treatment type, smoking status, histology, and both ADL and MMS scores.

This is in line with other studies investigating HRQoL in NSCLC patients [6,14–16]. Sloan et al. [12] and Jacot et al. [13] demonstrated that overall HRQoL deficits, at lung cancer diagnosis, were significantly associated with poor OS (HR: 1.55; P <0.001 and 2.20; P <0.001, respectively). Yet these studies did not take account of disease stage or age.

Among the Charlson comorbidity index (CCI), MMS score, and ADL score, no geriatric index has been found to be able to guide thoracic oncologists in decision-making for elderly NSCLC patients. In our final model including GH, the MMS and ADL scores were both associated with OS, whereas CCI was not, as previously-published [14]. Our results indicate that HRQoL by the EORTC QLQ-C30 questionnaire could provide a useful tool, the GH score being statistically significant in all sensitivity analyses. Moreover, with PF being statistically associated with OS in the model without ADL, these results suggest that the EORTC QLQ-C30 questionnaire could even surpass the ADL score, reflecting the same characteristics while adding the global GH evaluation. The collinearity of baseline HRQoL scores with other covariates was examined using a multiple linear regression model. This model exhibited R^2 statistics of 12%, indicating that the GH score was much more than a simple amalgamation of PS, ADL score and BMI. Moreover, the PS and ADL scores are evaluated by the physician rather than being self-reported. Therefore, the use of the HRQoL questionnaire could limit the interpretation by the physicians. Finally, the idea would be to use only the HRQoL questionnaire rather than two or more questionnaires (ADL, MMS etc) for the geriatric evaluation to help physicians in their decision-making, which is often difficult in elderly patients with lots of comorbidities.

Subgroup analyses suggested the baseline GH score to be a predictor of treatment effect, with 46 being the cut-off value. The ESOGIA-Groupe Français de Pneumo-Cancérologie 0802 trial assessed the integration of the comprehensive geriatric assessment (CGA) in treatment-decision-making in stage IV NSCLC patients over 70 years old [22]. The study failed to prove the superiority of a CGA-based strategy compared to PS-guided strategy of treatment allocation in terms of time to treatment failure. CGA has never

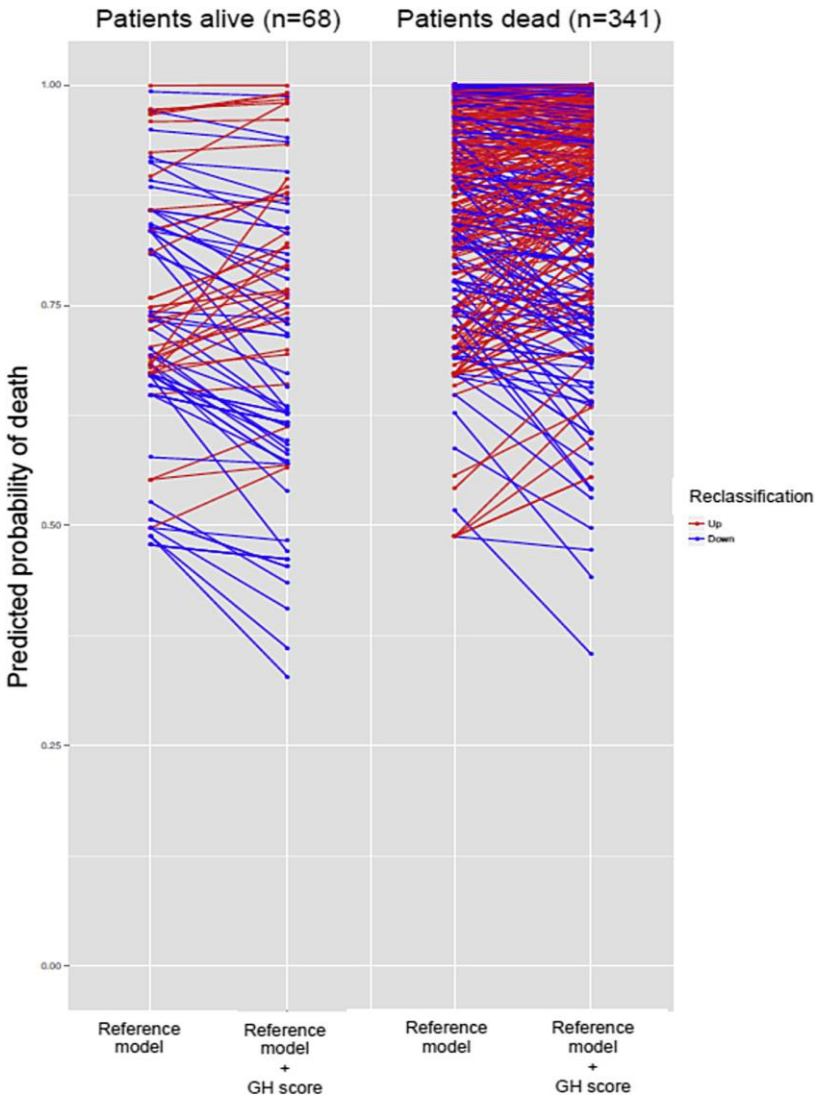


Fig. 1. Additive value of the GH dimension score for reclassifying mortality risk at 24 months post-randomisation (continuous net reclassification improvement). Blue lines in patients who did not die indicate that the GH score had the correct (downward) influence on risk prediction (42/68 = 61.8%). Conversely, red lines in patients who died indicate a correct (upward) change in risk assessment when using GH score (196/341 = 57.4%). GH = global health.

proven able to predict treatment efficacy in elderly lung cancer patients, meaning that HRQoL could represent a better tool to identify patients likely to benefit from doublet chemotherapy. However, further research is warranted to validate the questionnaire's predictive value and to define cut-off values.

Our study displayed several limitations. First, the specific lung cancer module QLQ-LC13 questionnaire, which could have improved the HRQoL's prognostic value, was not employed. Secondly, data must be replicated using external validation study and confirmed in a prospectively recruited cohort. Furthermore, the Instrumental Activities Daily Living index, which explores patients' ability to use public transportation, telephone, drive, etc., was not administered. Finally, our study was not designed to predict the treatment type to be given to the patients, based on HRQoL score.

Our study provides evidence of the additional prognostic value of HRQoL data to identify vulnerable elderly NSCLC subpopulations. The EORTC QLQ-C30

Table 5
Clinical and health-related quality of life dimension scores associated with overall survival using the Q1 and Q3 interquartiles of global health score.

	Number of patients	Number of events	HR (95% CI)	P-value
GH <46	136	123	1.00	<0.0001
46 ≤GH ≤67	228	193	0.42 (0.31–0.59)	
GH >67	87	61	0.67 (0.53–0.85)	
Treatment				
Monochemotherapy	225	199	1.00	
Doublet chemotherapy	226	198	0.64 (0.52–0.79)	<0.0001
Performance status score				
0–1	327	262	0.66 (0.51–0.84)	
2	123	114	1.00	0.0003
Smoking status				
Never smoked	87	68	0.58 (0.44–0.77)	
Ever smoked	264	309	1.00	0.0003
Histology				
Adenocarcinoma	219	178	0.67 (0.50–0.89)	
Squamous	146	131	0.77 (0.57–1.05)	
Other	66	68	1.00	0.025
MMS	441	377	0.97 (0.95–1.01)	0.059
ADL	438	377	0.73 (0.61–0.87)	0.0003

HR = hazard ration; CI = confidence interval; GH = global health score; MMS = mini-mental state; ADL = activities of daily living.

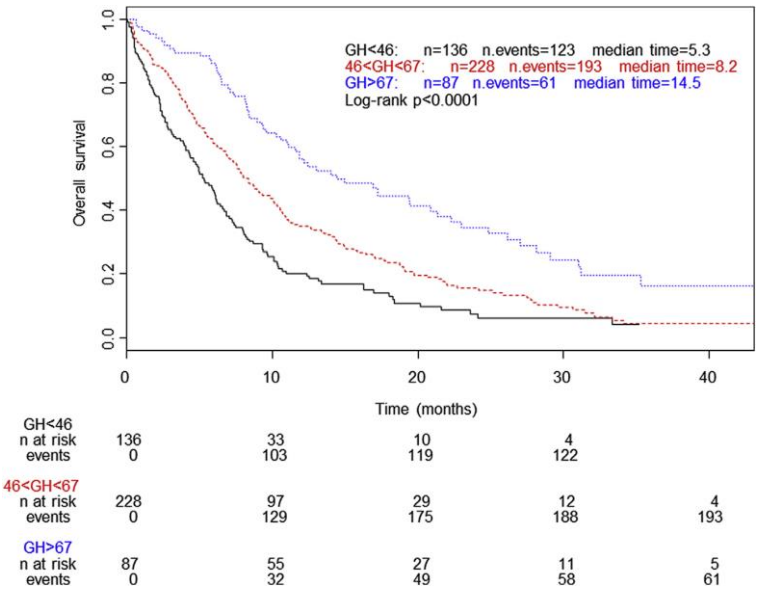


Fig. 2. Kaplan–Meier survival curves according to GH score. GH = global health.

questionnaire could constitute a valuable tool for selecting patients likely to benefit from doublet chemotherapy.

Conflict of interest statement

Frédéric Fiteni declares no conflict of interest. Dewi Vernerey declares no conflict of interest. Franck Bonnetain reports grants and personal fees from ROCHE, grants, personal fees and non-financial support from NOVARTIS, personal fees from MERCK SERONO, and personal fees from NESTLE. Fabien Vaylet declares no conflict of interest. Hélène Sennelart declares no conflict of interest. Jean Tredaniel declares no conflict of interest. Denis Moro-Sibilot declares participation to Roche, Eli Lilly, AstraZeneca, Novartis, Pfizer, Boehringer Ingelheim, Amgen, and BMS boards. Dominique Herman declares no conflict of interest. Hélène Laizé declares no conflict of interest. Philippe Masson declares no conflict of interest. Marc Derollez declares no conflict of interest. Christelle Clément-Duchêne declares no conflict of interest. Bernard

Milleron received personal fees outside the submitted work from AstraZeneca, BMS, Chugai, Lilly and Roche. Franck Morin declares no conflict of interest. Gérard Zalcman declares no conflict of interest. Elisabeth Quoix declares no conflict of interest. Virginie Westeel declares no conflict of interest.

Funding

Funding was received from Ligue Contre Le Cancer and Intergroupe Francophone de Cancérologie Thoracique.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ejca.2015.10.004.

References

[1] Beitz J, Gnecco C, Justice R. Quality-of-life end points in cancer clinical trials: the U.S. food and drug administration perspective. J Natl Cancer Inst Monogr 1996;20:7–9.
[2] Johnson JR, Temple R. Food and drug administration requirements for approval of new anticancer drugs. Cancer Treat Rep 1985;69(10):1155–9.
[3] Fiteni F, Westeel V, Pivot X, Borg C, Vernerey D, Bonnetain F. Endpoints in cancer clinical trials. J Visc Surg 2014;151(1):17–22.
[4] Montazeri A. Quality of life data as prognostic indicators of survival in cancer patients: an overview of the literature from 1982 to 2008. Health Qual Life Outcomes 2009;7:102.
[5] Eton DT, Fairclough DL, Cella D, Yount SE, Bonomi P, Johnson DH. Early change in patient-reported health during lung cancer chemotherapy predicts clinical outcomes beyond those predicted by baseline report: results from Eastern Cooperative Oncology Group Study 5592. J Clin Oncol 2003;21(8):1536–43.
[6] Djärv T, Metcalfe C, Avery KN, Lagergren P, Blazeby JM. Prognostic value of changes in health-related quality of life scores during curative treatment for esophagogastric cancer. J Clin Oncol 2010;28(10):1666–70.
[7] Luoma ML, Hakamies-Blomqvist L, Sjöström J, Pluzanska A, Ottoson S, Mouridsen H, et al. Prognostic value of quality of life scores for time to progression (TTP) and overall survival time (OS) in advanced breast cancer. Eur J Cancer 2003;39(10):1370–6.
[8] Movsas B, Moughan J, Sarna L, Langer C, Werner-Wasik M, Nicolaou N, et al. Quality of life supersedes the classic prognosticators for long-term survival in locally advanced non-small-cell lung cancer: an analysis of RTOG 9801. J Clin Oncol 2009;27(34):5816–22.
[9] Fielding R, Wong WS. Quality of life as a predictor of cancer survival among Chinese liver and lung cancer patients. Eur J Cancer 2007;43(11):1723–30.
[10] Quinten C, Coens C, Mauer M, Comte S, Sprangers MA, Cleeland C, et al. Baseline quality of life as a prognostic indicator of survival: a meta-analysis of individual patient data from EORTC clinical trials. Lancet Oncol 2009;10(9):865–71.
[11] Efficace F, Bottomley A, Smit EF, Lianes P, Legrand C, Debruyne C, Schramel F, et al. Is a patient's self-reported health-related quality of life a prognostic factor for survival in non-small-cell lung cancer patients? A multivariate analysis of prognostic factors of EORTC study 08975. Ann Oncol 2006;17(11):1698–704.
[12] Sloan JA, Zhao X, Novotny PJ, Wampfler J, Garces Y, Clark MM, et al. Relationship between deficits in overall quality of life and non-small-cell lung cancer survival. J Clin Oncol 2012;30(13):1498–504.
[13] Jacot W, Colinet B, Bertrand D, Lacombe S, Bozonnat MC, Daurès JP, et al. Quality of life and comorbidity score as prognostic determinants in non-small-cell lung cancer patients. Ann Oncol 2008;19(8):1458–64.
[14] Quoix E, Zalcman G, Oster JP, Westeel V, Pichon E, Lavolé A, et al. Carboplatin and weekly paclitaxel doublet chemotherapy compared with monotherapy in elderly patients with advanced non-small-cell lung cancer: IFCT-0501 randomised, phase 3 trial. Lancet 2011;378(9796):1079–88.
[15] Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85(5):365–76.
[16] Ringdal GI, Ringdal K. Testing the EORTC Quality of Life Questionnaire on cancer patients with heterogeneous diagnoses. Qual Life Res 1993;2(2):129–40.
[17] Anderson RT, Aaronson NK, Wilkin D. Critical review of the international assessments of health-related quality of life. Qual Life Res 1993;2(6):369–95.
[18] Hjermstad MJ, Fossa SD, Bjordal K, Kaasa S. Test/retest study of the European Organization for Research and Treatment of Cancer Core Quality-of-Life Questionnaire. J Clin Oncol 1995;13(5):1249–54.
[19] Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15(4):361–87.
[20] Pencina MJ, D'Agostino RB, Vasan RS. Statistical methods for assessment of added usefulness of new biomarkers. Clin Chem Lab Med 2010;48(12):1703–11.
[21] Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. Stat Med 2004;23(13):2109–23.
[22] Corre R, Chouaid C, Barlesi F, Le Caer H, Dansin E, Vergnenegre A, et al. Study ESOGIA-GFPC 08–02: phase III, randomized, multicenter trial involving subjects over age 70 with stage IV non-small cell lung cancer and comparing a "classical" strategy of treatment allocation (dual-agent therapy based on carboplatin or monotherapy with docetaxel alone), based on performance status and age, with an "optimized" strategy allocating the same treatments according to a simplified geriatric screening scale, plus a more thorough geriatric evaluation if necessary. Paper presented at the 2011 ASCO Annual Meeting. Available at: http://meetinglibrary.asco.org/content/83335-102.

VIII.3 Annexe 3:

Loupy A*, **Vernerey D***, Viglietti D*, Aubert O, Duong Van Huyen J-P, Empana J-P, Bruneval P, Glotz D, Legendre C, Jouven X, Lefaucheur C (2015) Determinants and Outcomes of Accelerated Arteriosclerosis: Major Impact of Circulating Antibodies. Circ Res 117: 470–482, doi:10.1161/CIRCRESAHA.117.306340.

*Co first authorship

Déterminants et conséquences cliniques de la formation accélérée de l'artériosclérose : impact majeur des anticorps circulants

Contexte:

Le rôle des anticorps circulants en complément des facteurs de risque cardiovasculaires traditionnels dans le développement accéléré de l'artériosclérose et leurs conséquences cliniques à long terme n'ont jamais été investigués et démontrés en population.

Objectif:

Cette étude explore le rôle des anticorps circulants dans la formation accélérée de l'artériosclérose ainsi que l'influence de l'artériosclérose associé à l'immunologie sur la survie des greffons de rein, des patients ainsi que sur la survenue d'événements cardiovasculaires majeurs.

Méthodes et résultats:

Nous proposons dans ce travail une étude observationnelle prospective incluant une cohorte de 1 065 patient ayant réalisés une greffe de rein entre 2004 et 2010 (cohorte principale, n = 744; cohorte de validation, n = 321).Tous les patients ont été évalués pour les facteurs de risque cardio-vasculaire traditionnels et la présence d'anticorps anti HLA circulants. Des biopsies ont été réalisées chez tous les patients afin d'observer la présence de lésions d'artériosclérose dans le greffon.

Dans la cohorte principale, 250 patients (33,6%) présentaient de l'artériosclérose qualifiée de sévère (rétrécissement luminal > 25%). La présence d'anticorps anti-HLA circulants spécifiques du donneur étaient significativement associés à la présence d'artériosclérose sévère dans le greffon (odds ratio, 2.9; P < 0,0001), indépendamment des facteurs de risque traditionnels. Les patients présentant de l'artériosclérose sévère ainsi que des anticorps anti-HLA (n = 91, 12,2%) ont montré une activation endothéliale dans le greffon, de l'endartérite, et du dépôt du complément.

Les niveaux élevés d'anticorps anti-HLA et leur capacité à fixer le complément ont été trouvés comme étant associés à une augmentation significative de la sévérité de l'artériosclérose.

Les patients qui présentaient une artériosclérose sévère associés à la présence d'anticorps avaient une survie de leur greffon diminuée ainsi qu'une augmentation de risque de décès (P < 0,0001); Pour ces patients, une augmentation du risque d'événements cardiovasculaires majeurs de l'ordre de 2.5 par rapport aux patients présentant de l'artériosclérose sévère sans anticorps et de 4,1 comparé aux patients avec une artériosclérose minimale, a été mise en évidence (P < 0,0005).
Les anticorps anti-HLA circulants spécifiques du donneur étaient aussi significativement associés à la survenue d'évènements cardiovasculaires majeurs (hazard ratio, 2,4; P = 0,0004), ceci toujours indépendamment des facteurs de risque traditionnels.

Conclusions:

La présence d'anticorps circulants est donc un facteur pronostique majeur dans le développement de l'artériosclérose sous sa forme sévère et aussi pour la survenue d'évènements cardiovasculaires majeurs, indépendamment des facteurs de risque cardio-vasculaire traditionnels.

Determinants and Outcomes of Accelerated Arteriosclerosis
Major Impact of Circulating Antibodies

Alexandre Loupy,* Dewi Vernerey,* Denis Viglietti,* Olivier Aubert, Jean-Paul Duong Van Huyen, Jean-Philippe Empana, Patrick Bruneval, Denis Glotz, Christophe Legendre, Xavier Jouven,* Carmen Lefaucheur*

Rationale: **The role of circulating antibodies in addition to traditional cardiovascular risk factors in the development of accelerated arteriosclerosis and their long-term clinical consequences have not been demonstrated.**

Objective: **We investigated the role of circulating antibodies in accelerated arteriosclerosis and the role of immune-associated arteriosclerosis in graft and patient survival and the occurrence of major adverse cardiovascular events.**

Methods and Results: **This was an observational prospective cohort study that included 1065 kidney transplant patients (principal cohort, n=744; validation cohort, n=321) between 2004 and 2010. Participants were assessed for traditional cardiovascular risk factors and circulating anti–human leukocyte antigen (HLA) antibodies. All patients underwent allograft biopsies to assess arteriosclerotic lesions and endothelial activation, endarteritis, and complement deposition. In the principal cohort, 250 (33.6%) patients had severe arteriosclerosis (luminal narrowing >25% via fibrointimal arterial thickening). Circulating donor-specific anti-HLA antibodies were significantly associated with severe allograft arteriosclerosis (hazard ratio, 2.9; *P*<0.0001), independently of traditional risk factors. Patients with severe arteriosclerosis and anti-HLA antibodies (n=91, 12.2%) demonstrated allograft endothelial activation, endarteritis, and complement deposition. High levels of anti-HLA antibodies and their complement binding capacity were associated with increased severity of arteriosclerosis. Patients with antibody-associated severe arteriosclerosis had decreased allograft survival and increased mortality (*P*<0.0001); they exhibited a 2.5- and 4.1-fold increased risk of major adverse cardiovascular events compared with patients who had severe arteriosclerosis without antibodies and patients with minimal arteriosclerosis, respectively (*P*<0.0005). Circulating donor-specific anti-HLA antibodies were significantly associated with occurrence of major adverse cardiovascular events (hazard ratio, 2.4; *P*=0.0004), independently of traditional risk factors.**

Conclusions: **Circulating antibodies are major determinants of severe arteriosclerosis and major adverse cardiovascular events, independent of traditional cardiovascular risk factors. (*Circ Res*. 2015;117:470-482. DOI: 10.1161/CIRCRESAHA.117.306340.)**

Key Words: antibodies ■ arteriosclerosis ■ cardiovascular diseases ■ transplantation

Arteriosclerosis is a pathological condition characterized by fibrointimal thickening of the arteries; it leads to a progressive form of vascular obliteration and dysfunction of various organs. Arteriosclerosis is also a component of a more general arterial disease called atherosclerosis, which describes a heterogeneous disease with a continuum between pure arteriosclerosis and fibrolipidic plaques (atheroma).[1–3]

Editorial, see p 398

Arteriosclerosis is recognized as a primary cause of end-stage renal disease[4] and loss of kidney[5,6] and heart transplants.[7] Arteriosclerosis is also part of a more general burden of cardiovascular disease that is responsible for heart failure and cerebral strokes, leading to increased morbidity and mortality.[8–10]

The prevailing hypothesis holds that the main contributors to arteriosclerosis include conventional cardiovascular risk factors, such as hypertension, tobacco consumption, and aging.[8] However, it has been suggested that immune and inflammatory

Original received February 26, 2015; revision received June 4, 2015; accepted June 8, 2015. In May 2015, the average time from submission to first decision for all original research papers submitted to Circulation Research was 15.49 days.

From the Paris Translational Research Center for Organ Transplantation and Cardiovascular Disease, INSERM, UMR-S970, Paris, France (A.L., D. Viglietti, O.A., J.-P.D.V.H., J.-P.E., P.B., D.G., C. Legendre, X.J., C. Lefaucheur); Paris Descartes University, Sorbonne Paris Cité and Necker Hospital (A.L., J.-P.D.V.H., C. Legendre) and Department of Transplantation, Saint-Louis Hospital (D. Viglietti, D.G., C. Lefaucheur), Assistance Publique - Hôpitaux de Paris, Paris, France; Methodology Unit (EA 3181) CHRU de Besançon, Besançon, France (D. Vernerey); Department of Pathology, Necker Hospital, Paris, France (J.-P.D.V.H.); and Department of Pathology (P.B.) and Department of Cardiology (X.J.), Georges Pompidou European Hospital, Paris, France.

*These authors contributed equally to this article.

The online-only Data Supplement is available with this article at http://circres.ahajournals.org/lookup/suppl/doi:10.1161/CIRCRESAHA.117.306340/-/DC1.

Correspondence to Alexandre Loupy, MD, PhD, Paris Descartes University, Sorbonne Paris Cité & Necker Hospital, Assistance Publique - Hôpitaux de Paris, Paris, France. E-mail alexandreloupy@gmail.com

© 2015 American Heart Association, Inc.

Nonstandard Abbreviations and Acronyms	
DSA	donor-specific antibodies
HLA	human leukocyte antigen
MACE	major adverse cardiovascular events
MFI	mean fluorescence intensity

processes play a role in arteriosclerosis and more generally in atherosclerosis, as supported by animal experimental data and clinical observations in autoimmune diseases.[2,11–16] Animal models are limited by the fact that they do not recapitulate the human pathophysiology of immune-mediated vascular injury,[17–19] whereas human studies lack access to patient arteries for histopathologic analysis and exhibit slow disease progression. As a consequence, we have an incomplete understanding of the respective contributions of nonimmune and immune factors in arteriosclerosis and their local (organ failure) and more general consequences (patient death and cardiovascular events). Solid-organ transplant offers the unique opportunity to address the general condition of arteriosclerosis disease by using kidney transplantation as a human prototype for the antibody-dependent immune response that coexists with traditional risk factors.

In view of the profound interactions between the immune system and the vessels that characterize solid organ transplantation,[2,20–25] we used a population-based approach to assess the determinants of premature arteriosclerosis in a cohort of extensively phenotyped kidney transplant patients. Although observations have suggested an association between antibodies and local allograft arteriosclerosis,[5,26] the extent to which antibodies are independent risk factors in premature arteriosclerosis and occurrence of major cardiovascular events (MACE), beyond traditional risk factors, is still unknown.

The aim of the present population-based study was to use a model of premature vascular injury to evaluate the respective contributions of immune and nonimmune factors and to integrate these factors into the more general process of arteriosclerosis disease. We used a multidisciplinary approach in a large prospective cohort of kidney recipients. Patients were assessed for traditional cardiovascular risk factors as well as circulating anti–human leukocyte antigen (HLA) antibodies and their properties (specificity, HLA class, strength, and complement-binding capacity) and underwent prospective kidney allograft biopsies to assess arteriosclerotic lesions and markers of endothelial activation, endarteritis, and local complement deposition. We assessed the determinants of local severe arteriosclerosis (in allografts) and the occurrence of MACE.

Methods

Multidisciplinary Network

This study involved the Paris Translational Research Center for Organ Transplantation, which is a multidisciplinary collaborative research project with the aim of advancing basic knowledge about the determinants of vascular diseases. This network includes experts in the fields of cardiology, nephrology, transplant immunology, basic science, histopathology, and epidemiology.

Study Population

This population-based study included consecutive patients who underwent kidney transplantation at Necker Hospital (Paris, France) between January 1, 2004 and January 1, 2010 (n=1012). We included all patients who underwent a systematic evaluation of arteriosclerosis on kidney biopsy at 1 year post transplantation (n=744). We excluded 144 patients who did not reach the 1-year follow-up (graft failure, n=53; patient death, n=26) or who were lost to follow-up in the first year post transplant (n=65). Patients who did not undergo a screening biopsy (contraindication, refusal, or not conducted) or for whom the screening biopsy was performed but was not adequate according to the international classification for kidney allograft histopathology were also excluded from the analysis (n=124). We also used an additional independent validation sample of 321 kidney recipients from Saint-Louis Hospital (Paris, France; see Methods, Independent validation cohort). The study was approved by the Institutional Review Boards of Necker Hospital and Saint-Louis Hospital.

Data Collection

Clinical data were obtained from a prospective database, Données Informatiques Validées en Transplantation (DIVAT clinical prospective cohort, official website: www.divat.fr), in which data were entered at specific points for each patient (at day 0, 6 months, and 1 year post transplantation); follow-up was updated annually thereafter.[27] For the DIVAT database, all data were prospectively entered by research assistants under the supervision of transplant clinicians, and the database underwent an annual external audit for data accuracy. For each patient, follow-up data after 1 year were entered based on a complete review of the medical charts and the collection of clinical events, including cardiovascular complications. Data were retrieved from the database on April 15, 2012.

We performed an extensive evaluation of potential nonimmune (classical) and immune (antibody-related) risk factors that might be associated with arteriosclerosis. We analyzed the following data: (1) donor cardiovascular profile, including age at donation; sex; cerebrovascular cause of death (including ischemia and hemorrhagic cerebral stroke); the presence of hypertension, diabetes mellitus, or hypercholesterolemia; tobacco consumption; and renal function at donation; (2) recipient cardiovascular profile at the time of transplantation, including age, sex, body mass index, hypertension, diabetes mellitus, hypercholesterolemia, tobacco consumption, past history of MACE (myocardial infarction, unstable angina, revascularization, cerebral stroke, and severe peripheral arterial disease that required revascularization), and the use of calcineurin inhibitors and corticosteroid dose; (3) recipient cardiovascular profile assessed at 6 month post transplant, including blood pressure, diabetes mellitus, low-density lipoprotein levels, antihypertensive therapy, and the use of antiaggregant or statin therapy; (4) immunologic parameters (HLA mismatch, rejection and the presence of circulating anti-HLA donor-specific antibodies [DSA] post transplant); and (v) transplant characteristics (cold ischemia time, graft rank, and delayed graft function).

Definition of Arteriosclerosis and Allograft Injury Phenotype

The severity of allograft arteriosclerosis and injury phenotypes was defined on the basis of allograft biopsies taken at 1 year post transplant in all 744 included patients. Allograft arteriosclerosis was defined and graded based on the criteria established by the international Banff classification on allograft pathology[28,29] and was assessed by trained pathologists (J.-P.D.V.H., P.B) who were blinded to the clinical information.

The Banff arteriosclerosis score reflects the degree of narrowing of the luminal area in the worst involved artery in the biopsy as follows: cv0, no chronic vascular changes; cv1, vascular luminal narrowing of ≤25% via fibrointimal arterial thickening; cv2, vascular luminal narrowing of 26% to 50%; and cv3, vascular narrowing of ≤50%.[30] We used the Banff classification to separate minimal arteriosclerosis (Banff scores of 0 and 1) from severe arteriosclerosis (Banff scores of 2 and 3).[28–30]

In addition to the degree of arteriosclerosis, we assessed the presence of microcirculatory inflammation, arteritis, arterial hyalinosis, and atrophic scarring in all biopsies.[28,29] We also evaluated the deposition of complement in the graft capillaries and arteries using complement split product C4d immunohistochemical staining on paraffin

sections with a human polyclonal antibody (Biomedica). We recorded biopsy-proven acute rejection, as defined according to the international classification,[31] during the first year post transplant.

Detection and Characterization of Antibodies Against Donor-Specific HLA Molecules

All patients included in this study were screened for the presence of circulating anti-HLA antibodies in serum collected at 6 month post transplant. In addition, patients with evidence of acute allograft rejection were tested for circulating anti-HLA antibodies at the time of acute rejection. Antibodies against the HLA-A, HLA-B, HLA-Cw, HLA-DR, HLA-DQ, and HLA-DP epitopes were tested using single-antigen flow bead assays (One Lambda Inc, Canoga Park, CA) on a Luminex platform, as previously described.[32,33] All beads with a normalized mean fluorescence intensity (MFI) of >500 were considered positive. In patients with multiple anti-HLA DSA, the immunodominant anti-HLA DSA (the one with the highest MFI) was selected for analysis. HLA typing of the recipients was performed via molecular biology (Innolipa HLA typing kit, Innogenetics, Belgium). For all kidney transplant donors, tissue typing was initially performed using the microlymphocytotoxicity technique with One Lambda Inc tissue typing trays at transplantation and was confirmed by molecular biology.

Serum samples from patients with circulating donor-specific anti-HLA antibodies were analyzed in a blinded fashion for the presence of C1q-binding donor-specific anti-HLA antibodies using single-antigen flow bead assays according to the manufacturer's protocol (C1qScreenTM, One Lambda).[33]

Outcome Parameters

The outcomes measured in this study were time to graft loss, patient death, and time to the development of MACE. Kidney allograft loss was defined by a return to dialysis, whereas MACE included the following complications: acute coronary syndrome (myocardial infarction, unstable angina), stroke, severe peripheral arterial disease, and death from cardiovascular causes, including sudden death.

Independent Validation Cohort

The external validation population included 321 consecutive kidney transplant recipients who had been transplanted at Saint-Louis Hospital. Kidney transplants were performed between January 1, 2006 and January 1, 2010. Data from Saint-Louis Hospital were collected from the French National Registry Agency's (Agence de la Biomédecine) Cristal database (official website: https://www.sipg.sante.fr/portail/).[34] Data were retrieved from the database on April 15, 2012. The same inclusion criteria as were used for the principal cohort (Necker hospital) were applied to the validation cohort; these included the performance of a protocol allograft biopsy at 1 year post transplant and screening for circulating antibodies after transplantation.

Statistical Analysis

Continuous variables are described in terms of the mean and SD values. We compared the means and proportions using Student *t* test and the χ^2 test (or Fisher exact test if appropriate).

First, within each group of potential risk factors, including the recipient cardiovascular profile, the donor cardiovascular profile, the transplant characteristics, and the post-transplant immunologic profile, we assessed the determinants of severe allograft arteriosclerosis at 1 year post transplant. In the main analyses, severe allograft arteriosclerosis was constructed according to the cv score as a binary variable (2–3, severe versus 0–1, minimal) to allow the use of a simple logistic regression analysis to identify the determinants of arteriosclerosis. Two sensitivity analyses were further conducted[1]: using the arteriosclerosis score in 4 categories (grade 0, 1, 2, and 3) and performing a multinomial logistic regression analysis using the grade cv0 score as the reference category; and[2] using the arteriosclerosis grade as a continuous (ordinal) variable and constructing a generalized linear model.

Second, we determined whether severe antibody-associated arteriosclerosis was associated with long-term allograft loss and the future development of cardiovascular events.

The graft survival and patient survival analyses were performed for a maximum follow-up period of 7 years from the time of transplantation. Graft survival, patient survival, and the occurrence of MACE were plotted on Kaplan–Meier curves according to the severity of arteriosclerosis and the presence of anti-HLA antibodies and were compared using the log-rank test. Finally, we assessed the determinants of MACE in a Cox proportional hazards model.

For the logistic, linear, and Cox regression models, we first performed univariate analysis. The *P* value threshold of 0.10 for entering variable in the multivariate model was corrected using the Bonferroni correction method (0.1/31=0.003). Then, significant risk factors were entered in a single multivariate model using backward stepwise elimination.

Analyses were conducted using SAS 9.2 (Statistical Analysis System, Cary, NC) and R (version 2.10.1). All tests were 2-sided, and, except for the univariate analyses, all *P* values <0.05 were regarded as statistically significant.

Results

Baseline Characteristics of the Study Population

Table 1 summarizes the recipient and donor baseline profiles and the transplant characteristics in the overall population. The mean age of the 744 recipients was 47.8±12.9 years. A total of 429 (58%) recipients were men, 603 (81%) recipients had hypertension, 83 (11%) recipients had diabetes mellitus, 161 (22%) recipients had hypercholesterolemia, and 108 (15%) recipients presented with MACE before transplantation. A total of 276 (37%) recipients were active or past smokers at the time of transplantation. The mean donor age was 50.6±15.9 years. A total of 407 (55%) donors were men. The prevalence rates of hypertension, diabetes mellitus, and hypercholesterolemia in the donors were 23%, 4%, and 7%, respectively. A total of 203 (38%) donors were active or past smokers at the time of donation.

Two distinct populations were identified according to the arteriosclerotic allograft lesions identified at 1 year post transplantation: (1) kidney transplant recipients with minimal arteriosclerosis (0%–25% arterial luminal narrowing; n=494, 66.4%) and (2) kidney transplant recipients with severe arteriosclerosis (>25% arterial luminal narrowing; n=250, 33.6%). Compared with the patients with minimal arteriosclerosis, the patients with severe arteriosclerosis were significantly older (45.4±12.5 versus 52.4±12.5 years, *P*<0.0001), received allografts from older donors (46.9±15.8 versus 57.8±13.5 years, *P*<0.0001), and were more likely to have received allografts from donors who died from cerebrovascular disease (50% versus 64%, *P*<0.001; Table 1).

Determinants of Severe Arteriosclerosis

Among the immune and nonimmune parameters tested, the following factors were significantly associated with severe arteriosclerosis at 1 year post transplantation in the univariate analysis: (1) recipient cardiovascular parameters at the time of transplantation (day 0)—age at transplantation (*P*<0.0001); (2) donor cardiovascular parameters—age (*P*<0.0001), hypertension (*P*<0.0001), diabetes mellitus (*P*<0.0001), hypercholesterolemia (*P*=0.0014), and donor death from cerebrovascular disease (*P*<0.0001); (3) transplant parameters—cold ischemia time (*P*<0.0001); (4) post-transplant recipient

Table 1. Patient Characteristics at the Time of Transplantation

	All Patients (n=744)		Patients With Minimal Arteriosclerosis (n=494)		Patients With Severe Arteriosclerosis (n=250)		*P* Value
	n		n		n		
Recipient cardiovascular profile at time of transplantation							
Age, y	744	47.8±12.9	494	45.4±12.5	250	52.4±12.5	<0.0001
Sex (men), n (%)	744	429 (58%)	494	285 (58%)	250	144 (58%)	0.9999
Hypertension, n (%)	744	603 (81%)	494	393 (80%)	250	210 (84%)	0.1439
Diabetes mellitus, n (%)	744	83 (11%)	494	49 (10%)	250	34 (14%)	0.1320
Hypercholesterolemia, n (%)	744	161 (22%)	494	98 (20%)	250	63 (25%)	0.0934
Tobacco, n (%)	737	276 (37%)	488	186 (38%)	249	90 (36%)	0.6012
Body mass index, kg/m^2	741	23.4±4.0	491	23.3±3.9	250	23.4±4.3	0.8143
Past history of MACE, n (%)	744	108 (15%)	494	69 (14%)	250	39 (16%)	0.5505
Cause of ESRD	744	…	494	…	250	…	…
Diabetes mellitus, n (%)	…	59 (8%)		41 (8%)		18 (7%)	…
Vascular, n (%)	…	47 (6%)		28 (6%)		19 (8%)	…
Other, n (%)	…	638 (86%)		425 (86%)		213 (85%)	0.5352
Time since onset of dialysis, y	667*	5.1±4.5	417	4.8±4.3	222	5.4±4.8	0.1636
Donor cardiovascular profile							
Age, y	744	50.6±15.9	494	46.9±15.8	250	57.8±13.5	<0.0001
Sex (men), n (%)	744	407 (55%)	494	253 (51%)	250	154 (62%)	0.0113
Hypertension, n (%)	744	169 (23%)	494	80 (16%)	250	89 (36%)	<0.0001
Diabetes mellitus, n (%)	744	31 (4%)	494	8 (2%)	250	23 (9%)	<0.0001
Hypercholesterolemia, n (%)	744	49 (7%)	494	22 (4%)	250	27 (11%)	0.0053
Tobacco, n (%)	530	203 (38%)	327	107 (33%)	203	77 (38%)	0.2207
Donor type							
Living/dead, n (%)	744	144 (19%) 600 (81%)	494	117 (24%) 377 (76%)	250	27 (11%) 223 (89%)	<0.0001†
Cerebrovascular cause of death, n (%)	600	329 (55%)	377	187 (50%)	223	142 (64%)	
Other cause of death, n (%)	600	271 (45%)	377	190 (50%)	223	81 (36%)	
Transplant characteristics							
Graft rank (previous transplantation), n (%)	744	143 (19%)	494	84 (17%)	250	59 (24%)	0.0383
Cold ischemia time, h	720	18.3±10.0	476	17.1±10.1	244	20.6±9.5	<0.0001
HLA A/B/DR mismatch	733	2.8±1.4	487	2.7±1.4	246	2.9±1.4	0.1606
Anti-HLA DSA at day 0	733	145 (20%)	489	80 (16%)	244	65 (27%)	0.001568
Immunodominant anti-HLA DSA class at day 0 (class I/class II)	…	n=48/n=97	…	n=23/n=57	…	n=25/n=40	0.287
Recipient blood group type A/B/O/AB	744	338/57/326/23	494	252/34/195/13	250	118/18/107/7	0.7954

DSA indicates donor-specific antibodies; ESRD, end stage renal disease; HLA, human leukocyte antigen; and MACE, major adverse cardiovascular events.
*Seventy-seven patients had preemptive transplantation and did not undergo dialysis before kidney transplantation.
†Global χ^2 test comparing the distribution of donor type in the 2 groups of patients.

cardiovascular parameters (assessed at 6 month post transplant in all patients)—systolic blood pressure (P=0.0007) and hypertension requiring antihypertensive therapy (P<0.0001); and (5) immunologic and functional parameters—presence of circulating DSA (P<0.0001) and kidney function (P<0.0001). Associations with male donor sex (P=0.0102), graft rank (P=0.0317), recipient diabetes mellitus post transplant (P=0.0254), and recipient blood low-density lipoprotein concentration post transplant (P=0.0063) were above the threshold limits for P values of <0.003 (Table 2A).

In the multivariate logistic regression model (Table 2B), factors independently associated with severe arteriosclerosis at 1 year post transplantation were donor age (per year increment: odds ratio [OR], 1.05; 95% confidence interval [CI], 1.03–1.06; P<0.0001), donor diabetes mellitus (OR, 3.56; 95% CI, 1.5–8.6; P=0.0046), allograft cold ischemia time per 1-hour increment (OR, 1.0004; 95% CI, 1.000–1.001; P=0.0121), post transplant recipient hypertension requiring antihypertensive therapy (OR, 2.17; 95% CI, 1.06–4.42; P=0.033), and the presence of circulating DSA (OR, 2.9; 95% CI, 1.96–4.3; P<0.0001).

Finally, we defined the characteristics of the immunodominant circulating DSA and allograft injury phenotypes according to arteriosclerosis scores (Table 3). We found that patients with increasing arteriosclerosis scores had higher circulating

Table 2. Determinants of Severe Arteriosclerosis

	Number of Patients	Number of Patients With Severe Arteriosclerosis	OR	95% CI	*P* Value
Univariate analysis					
Recipient cardiovascular profile at the time of transplantation					
Age (per 1-y increment)	744	250	1.046	1.032–1.059	<0.0001
Sex					
Female	312	105	1	…	…
Male	432	144	0.977	0.717–1.331	0.8819
Hypertension					
No	141	40	1	…	…
Yes	603	210	1.349	0.902–2.018	0.1449
Diabetes mellitus					
No	656	216	1	…	…
Yes	88	34	1.430	0.896–2.280	0.1334
Hypercholesterolemia					
No	583	187	1	…	…
Yes	161	63	1.361	0.949–1.953	0.0941
Tobacco					
No	461	159	1	…	…
Yes	276	90	0.919	0.670–1.261	0.6012
Body mass index (per 1 kg/m^2 increment)	741	250	1.005	0.967–1.043	0.8135
Past history of MACE*					
No	636	211	1	…	…
Yes	108	39	1.138	0.744–1.743	0.5507
Donor cardiovascular profile					
Age (per 1-y increment)	735	248	1.050	1.038–1.063	<0.0001
Sex					
Female	327	91	1	…	…
Male	417	154	1.512	1.103–2.072	0.0102
Hypertension					
No	575	161	1	…	…
Yes	169	89	2.860	2.010–4.069	<0.0001
Diabetes mellitus					
No	713	227	1	…	…
Yes	31	23	6.155	2.712–13.972	<0.0001
Hypercholesterolemia					
No	695	223	1	…	…
Yes	49	27	2.597	1.447–4.661	0.0014
Tobacco					
No	346	126	1	…	…
Yes	184	77	1.256	0.872–1.811	0.2211
Donor type					
Living	144	27	1	…	…
Cerebrovascular cause of death	329	142	3.291	2.053–5.275	…
Other cause of death	271	77	1.896	1.154–3.116	<0.0001
Transplant characteristics					
Graft rank					
1	601	191	1	…	…
>1	143	59	1.508	1.037–2.193	0.0317
Cold ischemia time (per 1-h increment)	720	244	1.036	1.019–1.053	<0.0001
HLA A/B/DR mismatch	733	246	1.080	0.970–1.203	0.1606

(*Continued*)

Table 2. Continued

	Number of Patients	Number of Patients With Severe Arteriosclerosis	OR	95% CI	*P* Value
Calcineurin inhibitor treatment					
No	25	4	1	…	…
Yes	719	246	2.730	0.927–8.042	0.0684
Corticosteroid dose (per mg increment)	704	243	1.003	0.984–1.023	0.7511
Recipient cardiovascular profile post transplant					
Systolic blood pressure, mm Hg	700	245	1.016	1.007–1.025	0.0007
Diastolic blood pressure, mm Hg	700	245	0.998	0.984–1.011	0.7303
Diabetes mellitus					
No	594	188	1	…	…
Yes	150	62	1.522	1.053–2.199	0.0254
Blood LDL, mmol/L	707	244	1.373	1.094–1.723	0.0063
Antihypertensive treatment					
No	117	14	1	…	…
Yes	622	236	4.496	2.514–8.039	<0.0001
Antiaggregant treatment					
No	565	190	1	…	…
Yes	179	60	0.995	0.697–1.421	0.9786
Statin treatment					
No	422	131	1	…	…
Yes	322	119	1.302	0.959–1.769	0.0910
Immunologic and functional profiles post transplant					
Delay in graft function					
No	578	187	1	…	…
Yes	161	63	1.344	0.937–1.929	0.1086
Allograft GFR†					
≥60	239	64	1	…	…
30–60	443	154	1.457	1.030–2.061	…
≤30	48	29	4.174	2.189–7.958	<0.0001
Acute rejection episode					
No	400	118	1	…	…
Yes	344	132	1.488	1.096–2.020	0.0108
Circulating donor-specific antibodies					
No	567	159	1	…	…
Yes	177	91	2.715	1.919–3.841	<0.0001
Multivariate‡ analysis (n=664)					
Donor age (per 1-y increment)	664	233	1.047	1.033–1.060	<0.0001
Donor diabetes mellitus					
No	635	212	1	…	…
Yes	29	21	3.557	1.480–8.552	0.0046
Cold ischemia time (per 1-h increment)	664	233	1.0004	1.000–1.001	0.0121
Recipient hypertension post transplant requiring antihypertensive therapy					
No	72	11	1	…	…
Yes	592	222	2.168	1.064–4.417	0.0332
Circulating donor-specific antibodies					
No	496	147	1	…	…
Yes	168	86	2.902	1.960–4.296	<0.0001

CI indicates confidence interval; GFR, glomerular filtration rate; HLA, human leukocyte antigen; LDL, low-density lipoprotein; and OR, odds ratio.

*Major adverse cardiovascular events.

†GFR (MDRD formula) assessed in all patients at 6 month post transplant.

‡The final multivariate logistic regression model was obtained by entering risk factors that reached $P \leq 0.003$ in the univariate analysis. The final multivariate model was adjusted for the following parameters using a backwards procedure: recipient age, donor hypertension, donor hypercholesterolemia, donor type, recipient systolic blood pressure, and allograft GFR post transplant.

DSA MFI levels (P=0.0243), a higher prevalence of complement-binding DSA (P<0.0001), increased allograft endothelial activation (microcirculation inflammation, P<0.0001), and increased complement fraction C4d deposition in capillaries (P=0.011).

Phenotype of Antibody-Associated Severe Arteriosclerosis

The histopathology of the 744 allograft biopsies taken at 1 year post transplant revealed that patients with severe antibody-associated arteriosclerosis exhibited a distinct phenotype compared with patients with severe arteriosclerosis without antibodies and patients with minimal arteriosclerosis. The pathology in patients with severe antibody-associated atherosclerosis was characterized by (1) increased microcirculatory inflammation (75% versus 18% and 21%, respectively; P<0.0001), (2) a higher rate of complement fraction C4d deposition in allograft capillaries (35% versus 7% and 8%, respectively; P<0.0001) and (3) a higher rate of arterial intimal inflammation (5% versus 2% and 1%, respectively; P=0.023; Figure 1A). Figure 1B depicts the phenotype of antibody-associated arteriosclerosis, demonstrating severe fibrointimal arterial narrowing (A), complement fraction C4d deposition in the arterial intima (B), endothelial activation (microcirculation inflammation, C), and C4d deposition in capillaries (D). The distribution of DSA HLA class in patients with antibody-associated arteriosclerosis was class I in 21/91 patients (23%), class II in 42/91 patients (46%), and class I+II in 28/91 patients (31%). The mean MFI level of the immunodominant DSA was 5091±465, and these patients were more likely to have preformed persisting DSA (64/91, 70%) than de novo DSA (27/91, 30%). Of the 91 patients with antibody-associated severe arteriosclerosis, 34 (37%) exhibited C1q-binding capacity.

Outcome of Antibody-Associated Severe Arteriosclerosis

We investigated the long-term consequences of severe antibody-associated arteriosclerosis.

Patients and Allograft Survival

Figure 2 shows kidney allograft (A) and patient (B) survival. The 3- and 7-year graft survival rates were, respectively, 93.2% and 76.9% in the group of patients with severe antibody-associated arteriosclerosis, 97.5% and 92.7% in the group with severe arteriosclerosis without antibodies, and 97.9% and 93.1% in patients with minimal arteriosclerosis (log rank, P=0.001).

The 3- and 7-year patient survival rates were, respectively, 95.6% and 80.8% in the group of patients with antibody-associated severe arteriosclerosis, 96.9% and 93.2% in the group with severe arteriosclerosis without antibodies, and 98.4% and 93.4% in the group of patients with minimal arteriosclerosis (log-rank P=0.001).

Occurrence of MACE

A total of 72 (10%) patients presented with episodes of MACE after 1 year post transplant, including 29 (40%) patients with acute coronary syndrome, 7 (10%) patients with stroke, 29 (40%) patients with severe peripheral arterial disease, and 7 (10%) deaths from fatal cardiovascular causes. The median time to the occurrence of MACE was 2.5 years (IQR, 0.99–3.45). The risk of MACE significantly differed between patients with minimal arteriosclerosis, patients with severe arteriosclerosis without antibodies, and patients with severe antibody-associated arteriosclerosis (Figure 3). Furthermore, two-by-two comparisons indicated that patients with severe antibody-associated arteriosclerosis exhibited a 2.5-fold increased risk of MACE (95% CI, 1.3–4.9; P=0.005) compared with patients with severe arteriosclerosis without antibodies and a 4.1-fold increased risk of MACE (95% CI, 2.4–7.2; P<0.0001) compared with patients with minimal arteriosclerosis.

Determinants of MACE

In the univariate Cox model, the following traditional factors were significantly associated with an increased risk of MACE: recipient age (P<0.0001), past history of MACE (P=0.0002), recipient diabetes mellitus post transplant (P=0.0009), and kidney function (P=0.0018). In addition, the cold ischemia time (P<0.0001) and presence of circulating DSA (P=0.0002) were associated with the risk of MACE (Online Table I). In the multivariate Cox model (Table 4), factors independently associated with the occurrence of MACE were recipient age (per year increment: hazard ratio [HR], 1.029; 95% CI, 1.007–1.052; P=0.0092), recipient diabetes mellitus post transplant (HR, 1.767; 95% CI, 1.067–2.926; P=0.0269), allograft cold ischemia time (per 1-hour increment: HR, 1.037; 95% CI, 1.011–1.063; P=0.0050), past history of MACE (HR, 1.924; 95% CI, 1.117–3.313; P=0.0184), and the presence of circulating DSA (HR, 2.424; 95% CI, 1.490–3.945; P=0.0004).

Among the patients with circulating DSA, we defined the characteristics of the immunodominant DSA according to the occurrence of MACE (Online Table II). We found that compared with patients without MACE, the patients with MACE had higher circulating DSA MFI levels (3903±231 versus 6307±895, respectively; P=0.0035) and a higher prevalence of complement-binding DSA (20% versus 62%, respectively; P<0.0001).

Acute coronary syndrome (P=0.007) and death from cardiovascular causes (P=0.001) were overrepresented in the patients with DSA compared with the patients without, whereas the occurrence of peripheral arterial disease and stroke remained similar between DSA-positive and negative patients.

Sensitivity Analysis

First, when using other thresholds to define arteriosclerosis at 1 year, including contrasting each grade from 1 to 3 against grade 0 in a multinomial logistic regression (Online Table III) or using the cv score as a continuous variable in the generalized linear model (Online Table IV), we found consistent associations between circulating DSA and the severity of arteriosclerosis. We also showed that when the preimplantation biopsy evaluation of arteriosclerosis was included, circulating donor-specific anti-HLA antibodies remained strongly and independently associated with severe allograft arteriosclerosis at 1 year post transplant (OR, 3.058; 95% CI, 1.69–5.52; P=0.0002), as shown in Online Table V.

Second, when the level of renal function (assessed by glomerular filtration rate) was included in the final multivariate

Table 3. Characteristics of Circulating Anti-HLA DSA and Allograft Injury Phenotypes Based on Arteriosclerosis Scores

	cv0 (n=281)		cv1 (n=213)		cv2 (n=189)		cv3 (n=61)		
	n		n		n		n		*P* Value
DSA characteristics									
DSA, n (%)	281	47 (17%)	213	39 (18%)	189	63 (33%)	61	28 (46%)	<0.0001
HLA class of DSA (n, %)									
I	47	21 (45%)	39	14 (36%)	63	15 (24%)	28	11 (39%)	0.129
II		26 (55%)		25 (64%)		48 (76%)		17 (61%)	
DSA MFI, mean±SEM	47	3204±543	39	3760±576	63	4892±589	28	5540±735	0.0243
Complement-binding DSA, n (%)	281	8 (3%)	213	6 (3%)	189	25 (13%)	61	9 (15%)	<0.0001
Allograft injury									
C4d deposition, n (%)	194	16 (8%)	155	13 (8%)	158	28 (18%)	57	10 (18%)	0.011
Microcirculatory inflammation, n (%)	281	58 (21%)	211	46 (22%)	189	68 (36%)	61	28 (46%)	<0.0001
Arteritis, n (%)	276	2 (1%)	210	2 (1%)	186	7 (4%)	60	0 (0%)	0.063
Arterial hyalinosis, n (%)	273	100 (37%)	208	140 (67%)	177	137 (77%)	58	47 (81%)	<0.0001
Fibrosis-atrophy, n (%)	280	156 (56%)	210	141 (67%)	184	144 (78%)	58	47 (81%)	<0.0001

cv0 indicates no chronic vascular changes; cv1, vascular luminal narrowing of ≤25% via fibrointimal arterial thickening; cv2, vascular luminal narrowing of 26% to 50%; cv3, vascular narrowing of ≤50%; DSA, donor-specific antibodies; HLA, human leukocyte antigen; and MFI, mean fluorescence intensity.

model, the independent determinants of arteriosclerosis and MACE remained unchanged.

Third, when peripheral arterial diseases were excluded from the MACE definition, the set of risk factors identified in the multivariate Cox model of the determinants of MACE remained unchanged, with circulating DSA being strongly associated with increased risk of MACE, independent of traditional risk factors (HR, 3.175; 95% CI, 1.726–5.843; *P*=0.0002), as shown in Online Table VI.

Fourth, when we included the degree of allograft arteriosclerosis in the Cox model of the determinants of MACE, the presence of circulating anti-HLA DSA remained strongly associated with the occurrence of MACE (HR, 2.3; *P*=0.0014). The degree of allograft arteriosclerosis was not associated with the occurrence of MACE (*P*=0.16).

External Validation

The external validation cohort was composed of 321 patients whose baseline characteristics (compared with those of the principal cohort) are presented in Online Table VII.

The multivariate logistic regression model confirmed the independent association of circulating anti-HLA DSA with severe arteriosclerosis at 1 year post transplant (Online Table VIII).

The analysis of the allograft biopsies performed at 1 year confirmed that patients with antibody-associated severe arteriosclerosis had significantly increased microcirculatory inflammation (*P*<0.0001), complement fraction C4d deposition in the allograft capillaries (*P*=0.032), and arteritis lesions (*P*=0.001; Online Table IX).

The Kaplan–Meier survival analysis confirmed that patients with severe antibody-associated arteriosclerosis had decreased patient and allograft survival compared with the group of patients with severe arteriosclerosis without antibodies and minimal arteriosclerosis, respectively (*P*=0.009 and *P*<0.001, respectively; see Online Figure I).

Discussion

Using an integrated and population-based approach that combined contemporary immunologic and histopathologic techniques, we demonstrated the major role of circulating anti-HLA antibodies in the development of severe allograft arteriosclerosis, independent of traditional cardiovascular risk factors, and renal function. Antibody-associated severe arteriosclerosis has a distinct phenotype that is characterized by endothelial activation, endarteritis, and local complement deposition. We showed that the amount of circulating anti-HLA antibodies together with their complement-binding capacity is strongly correlated with the severity of arteriosclerosis. Finally, we demonstrated that circulating anti-HLA antibodies is a determinant of MACE independently of traditional cardiovascular risk factors. The recognition of an immune component of arterial disease (antibody-associated severe arteriosclerosis) identifies a local disease with general cardiovascular consequences and increased mortality.

Kidney transplantation provides a unique opportunity to address the unmet need of defining the respective roles of nonimmune and immune factors in the development of premature arteriosclerosis. First, kidney transplantation is a prototypic antibody-dependent alloimmune response, with primary microcirculatory and macrocirculatory endothelial injuries that can be characterized and precisely phenotyped using allograft biopsies.[25,26,35] Second, traditional cardiovascular risk factors are over-represented in kidney transplant recipients, in whom cardiovascular disease represents the main cause of death.[36,37] Kidney transplantation thus provides the possibility of using this population as a prototype for premature local arteriosclerosis and atherosclerotic diseases in general.

In this study, we showed the importance of traditional donor-related cardiovascular risk factors (age, diabetes mellitus, and hypertension) in the development of severe allograft arteriosclerosis, which suggests that post-transplant accelerated arteriosclerosis may be partially donor-transmitted. We also demonstrated the relevance of traditional recipient risk factors,

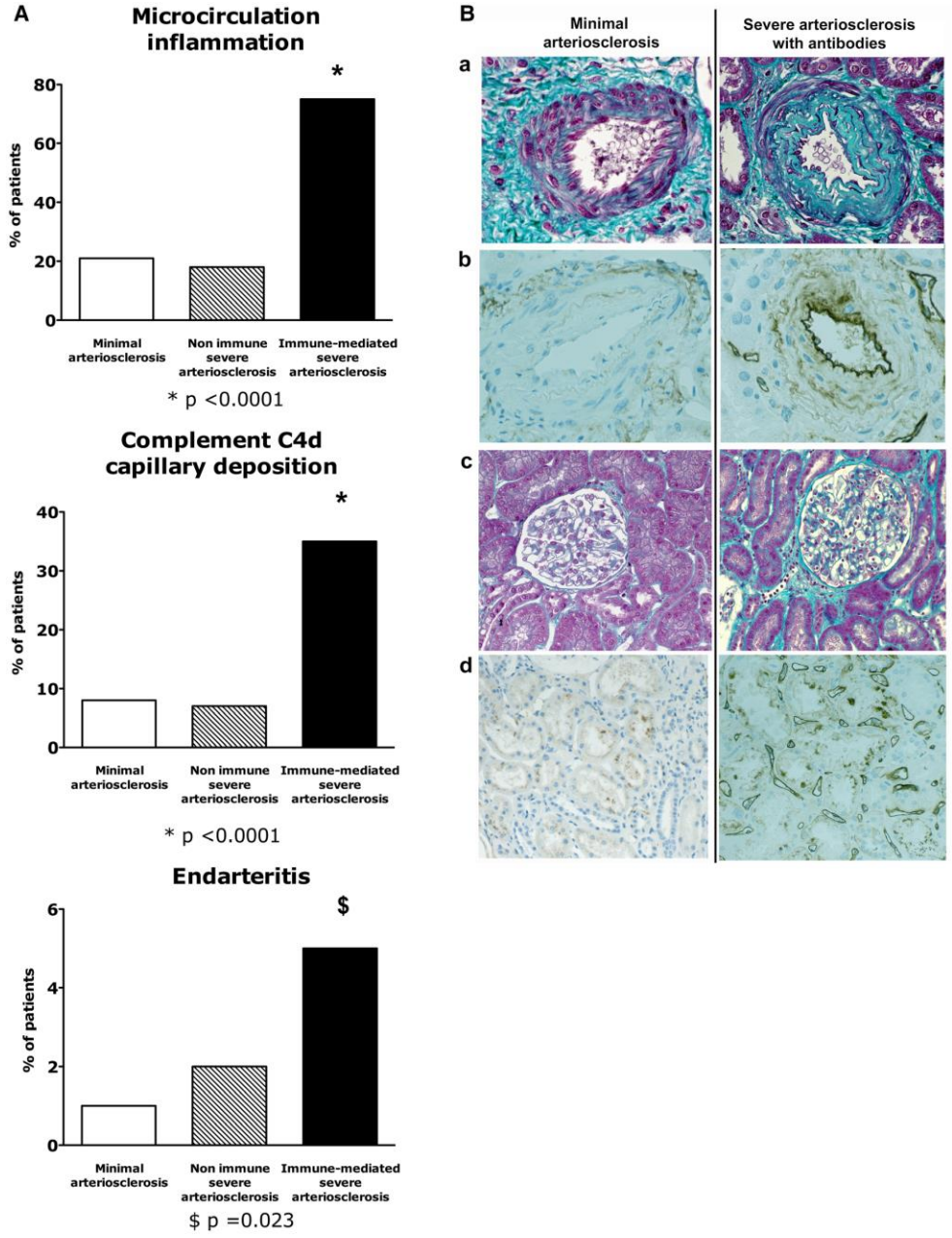


Figure 1. A, Allograft injury phenotypes according to arteriosclerotic lesions at 1 year: prevalence of microcirculatory inflammation, complement C4d deposition and endarteritis**. B**, Allograft injury phenotypes according to arteriosclerotic lesions at 1 year: fibrointimal arterial narrowing (a), complement fraction C4d deposition in the arterial intima (b), endothelial activation (microcirculatory inflammation, c), and C4d capillary deposition (d)**.**

including age, past history of MACE, diabetes mellitus, hypercholesterolemia, tobacco use, and renal allograft function in the occurrence of post-transplant cardiovascular events.

Among the 174 studies addressing an association between antibodies and arteriosclerosis in the current literature, only 3 were prospective population-based studies evaluating antiphosphorylcholine,[38] antioxidized low-density lipoprotein,[39] and antiphospholipid[40] antibodies. None of these studies integrated the simultaneous assessment of circulating antibodies, arteries, traditional cardiovascular risk factors,

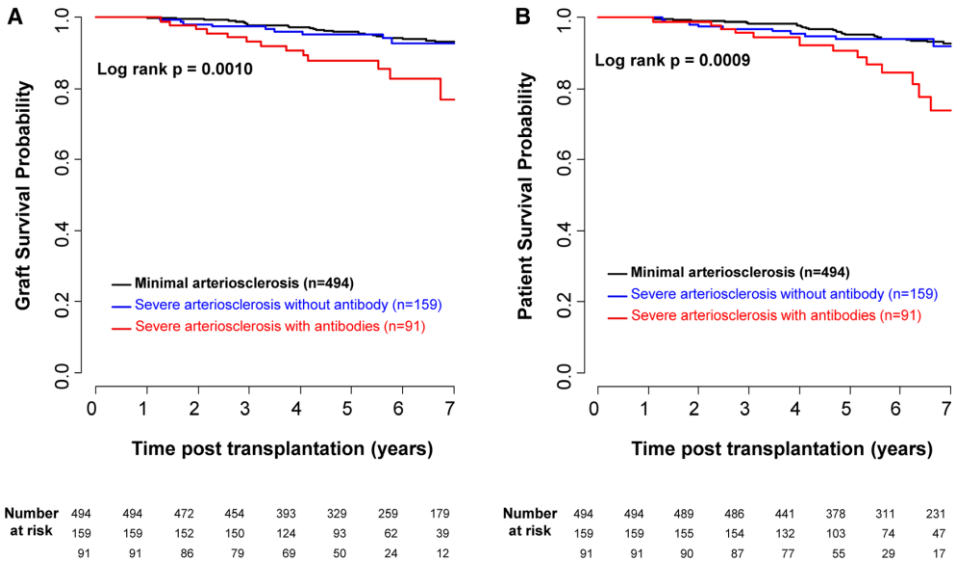


Figure 2. A, Kidney allograft survival**. B**, Patient survival**.**

and clinical events (Research in context in Online Data Supplement).

We defined a distinct disease entity, antibody-associated arteriosclerosis, which is characterized by circulating donor-specific anti-HLA antibodies, prominent arterial intimal thickening and hypercellularity associated with allograft endothelial activation (microcirculation inflammation), and complement fraction C4d deposition. There is abundant experimental evidence for the pathophysiologic underpinnings of antibody-associated arteriosclerosis in animal models. Ligation of anticlass I and anticlass II antibodies on the vascular endothelium and smooth muscle cells induces the release of a variety of cytokines and stimulates the proliferation of myofibroblasts, inducing neointimal hyperplasia and transplant vasculopathy.[6,22,41–43] In the present human study, we found that both class I and II anti-HLA antibodies were associated with allograft arteriosclerosis. The severity of arteriosclerosis was correlated with increasing circulating DSA levels and their ability to bind complement. In addition, the severity of arteriosclerosis was correlated with a gradient of endothelial

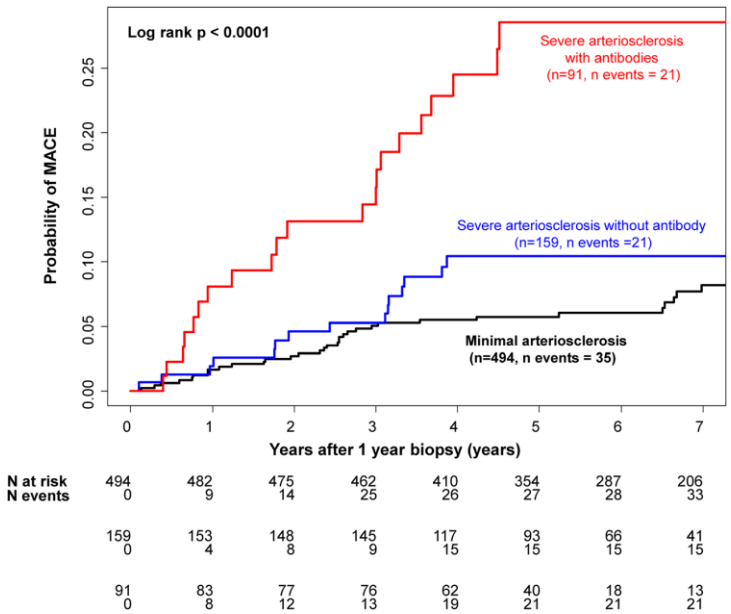


Figure 3. Risk of major adverse cardiovascular events (MACE) after 1 year post transplant.

Table 4. Determinants of MACE After 1 Year Post Transplant (Multivariate Cox Model, n=709)

	Number of Patients	Number of MACE	HR	95% CI	*P* Value
Recipient age (per 1-y increment)	709	70	1.029	1.007–1.052	0.0092
Recipient history of MACE					
No	602	49	1	…	…
Yes	107	21	1.924	1.117–3.313	0.0184
Recipient diabetes mellitus post transplant					
No	561	45	1	…	…
Yes	148	25	1.767	1.067–2.926	0.0269
Cold ischemia time (per 1-h increment)	709	70	1.037	…	0.0050
Circulating donor-specific antibodies					
No	540	41	1	1.011–1.063	…
Yes	169	29	2.424	1.490–3.945	0.0004

The final multivariate Cox model was obtained by entering risk factors from the univariate model that reached $P \leq 0.003$, the threshold used in the single multivariate proportional hazards model. CI indicates confidence interval; HR, hazard ratio; and MACE, major adverse cardiovascular events.

activation and complement allograft deposition. Our results suggested analogies with the heart transplant field in terms of an immune-mediated component in arteriosclerosis. Heart transplant recipients with late-onset allograft rejection develop accelerated chronic allograft vasculopathy (a fibrointimal thickening of epicardial coronary arteries), supporting a link between alloantibodies and damage to large vessels.[44]

We demonstrate a robust and consistent association between circulating DSA and arteriosclerosis, which permits us to move further than simple associations, approaching causation by providing the following: (1) the strength of the association between circulating DSA and the severity of arteriosclerosis; (2) a biological gradient: greater exposure to higher levels of circulating DSA (MFI level) and complement-binding capacity lead to higher incidence of severe arteriosclerosis; (3) temporality: the detection of circulating DSA in patient sera occurs before the assessment of arteriosclerosis; (iv) independency with potential confounders: mathematical integration of traditional cardiovascular risk factors and adjustments to these parameters in robust epidemiological models showing strong and independent associations between circulating DSA and the severity of arteriosclerosis; (v) biological plausibility of the hypothesis and the mechanism previously shown in small series[5,6] and further supported by animal models showing that circulating anti-HLA antibodies induce an immune response involving endothelial activation and accompanying arteriosclerosis[21,45]; and (vi) analogies with models of autoimmunity (lupus) with circulating autoantibodies promoting the development of atherosclerosis and MACE.[14] Finally, the validation of the results in an independent cohort reinforces the robustness of the study and the exportability of our findings.

Notably, we show that circulating DSA is a strong determinant of MACE, independent of traditional cardiovascular risk factors. Our data support the notion that the severe arteriosclerosis observed in the allograft may be only a part of a more general vascular process operating in patients with circulating DSA. In the absence of direct access to patient arteries outside the allograft, this study cannot make conclusions about the mechanisms involved outside the allograft. A plausible hypothesis is that circulating DSA trigger allograft endothelial activation, endothelial injury, complement cascade activation, and the release of cytokines and inflammatory mediators that extend to a more general vascular injury process.[1,11,17,46–48] Local antibody-mediated injury may participate in a more systemic vascular injury process, demonstrating an interplay between immune activation and vascular remodeling. This process can lead to plaque progression and destabilization, as suggested by animal models and human studies of autoimmune disease.[1,17,18,46,47,49,50] This is also consistent with recent findings that establish a link between genetic susceptibility to coronary artery disease (eg, 9p21 DNA variants)[51] and the altered responses of vascular cells to both inflammatory cytokines (eg, Ifng) and proliferation/survival factors.[47,52,53]

Strengths and Limitations

The simultaneous evaluation of pretransplant and post-transplant immunologic, histopathologic, and cardiovascular risk factors allowed us to address the multiple dimensions of arteriosclerosis development in renal allograft recipients. In particular, this unique approach permitted us to investigate the role of circulating anti-HLA antibodies, taking into account a large set of potentially confounding and mediating factors. However, we acknowledge that despite the robustness of our study findings, the study design did not permit us to provide a complete demonstration of causality between the presence of circulating anti-HLA antibodies, the development of arteriosclerosis, and the occurrence of MACE, and additional potential residual confounders may not have been taken into account in the assessment of these associations.

Conclusions

Using a multidisciplinary population-based approach, we demonstrate that circulating anti-HLA antibodies are major determinants of severe arteriosclerosis, independent of traditional cardiovascular risk factors. We also show that circulating anti-HLA antibodies exert long-term detrimental effects that are characterized by an increased risk of MACE. Thus, it is prudent to identify patients with circulating anti-HLA

antibodies to screen them for cardiovascular diseases and aggressively treat traditional risk factors. Given the recent development of new immune-modulating agents, our study also opens avenues for the potential use of antibody targeting strategies in various human arterial diseases.

Disclosures

None.

References

1. Libby P. Vascular biology of atherosclerosis: overview and state of the art. *Am J Cardiol*. 2003;91:3A–6A.
2. Libby P, Zhao DX. Allograft arteriosclerosis and immune-driven angiogenesis. *Circulation*. 2003;107:1237–1239.
3. Virmani R, Kolodgie FD, Burke AP, Farb A, Schwartz SM. Lessons from sudden coronary death: a comprehensive morphological classification scheme for atherosclerotic lesions. *Arterioscler Thromb Vasc Biol*. 2000;20:1262–1275.
4. Ortiz A, Covic A, Fliser D, Fouque D, Goldsmith D, Kanbay M, Mallamaci F, Massy ZA, Rossignol P, Vanholder R, Wiecek A, Zoccali C, London GM; Board of the EURECA-m Working Group of ERA-EDTA. Epidemiology, contributors to, and clinical trials of mortality risk in chronic kidney failure. *Lancet*. 2014;383:1831–1843. doi: 10.1016/S0140-6736(14)60384-6.
5. Hill GS, Nochy D, Bruneval P, Duong van Huyen JP, Glotz D, Suberbielle C, Zuber J, Anglicheau D, Empana JP, Legendre C, Loupy A. Donor-specific antibodies accelerate arteriosclerosis after kidney transplantation. *J Am Soc Nephrol*. 2011;22:975–983. doi: 10.1681/ASN.2010070777.
6. Hill GS, Nochy D, Loupy A. Accelerated arteriosclerosis: a form of transplant arteriopathy. *Curr Opin Organ Transplant*. 2010;15:11–15. doi: 10.1097/MOT.0b013e3283342684.
7. Loupy A, Cazes A, Guillemain R, Amrein C, Hedjoudje A, Tible M, Pezzella V, Fabiani JN, Suberbielle C, Nochy D, Hill GS, Empana JP, Jouven X, Bruneval P, Duong Van Huyen JP. Very late heart transplant rejection is associated with microvascular injury, complement deposition and progression to cardiac allograft vasculopathy. *Am J Transplant*. 2011;11:1478–1487. doi: 10.1111/j.1600-6143.2011.03563.x.
8. Go AS, Mozaffarian D, Roger VL, et al; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics–2014 update: a report from the American Heart Association. *Circulation*. 2014;129:e28–e292. doi: 10.1161/01.cir.0000441139.02102.80.
9. Lorenz MW, Polak JF, Kavousi M, et al; PROG-IMT Study Group. Carotid intima-media thickness progression to predict cardiovascular events in the general population (the PROG-IMT collaborative project): a meta-analysis of individual participant data. *Lancet*. 2012;379:2053–2062. doi: 10.1016/S0140-6736(12)60441-3.
10. Polak JF, Pencina MJ, O'Leary DH, D'Agostino RB. Common carotid artery intima-media thickness progression as a predictor of stroke in multi-ethnic study of atherosclerosis. *Stroke*. 2011;42:3017–3021. doi: 10.1161/STROKEAHA.111.625186.
11. Hansson GK, Hermansson A. The immune system in atherosclerosis. *Nat Immunol*. 2011;12:204–212. doi: 10.1038/ni.2001.
12. Hansson GK, Robertson AK, Söderberg-Nauclér C. Inflammation and atherosclerosis. *Annu Rev Pathol*. 2006;1:297–329. doi: 10.1146/annurev.pathol.1.110304.100100.
13. Packard RR, Lichtman AH, Libby P. Innate and adaptive immunity in atherosclerosis. *Semin Immunopathol*. 2009;31:5–22. doi: 10.1007/s00281-009-0153-8.
14. Roman MJ, Shanker BA, Davis A, Lockshin MD, Sammaritano L, Simantov R, Crow MK, Schwartz JE, Paget SA, Devereux RB, Salmon JE. Prevalence and correlates of accelerated atherosclerosis in systemic lupus erythematosus. *N Engl J Med*. 2003;349:2399–2406. doi: 10.1056/NEJMoa035471.
15. Descombes E, Droz D, Drouet L, Grünfeld JP, Lesavre P. Renal vascular lesions in lupus nephritis. *Medicine (Baltimore)*. 1997;76:355–368.
16. Nochy D, Daugas E, Droz D, Beaufils H, Grünfeld JP, Piette JC, Bariety J, Hill G. The intrarenal vascular lesions associated with primary antiphospholipid syndrome. *J Am Soc Nephrol*. 1999;10:507–518.
17. Ludewig B, Freigang S, Jäggi M, Kurrer MO, Pei YC, Vlk L, Odermatt B, Zinkernagel RM, Hengartner H. Linking immune-mediated arterial inflammation and cholesterol-induced atherosclerosis in a transgenic mouse model. *Proc Natl Acad Sci U S A*. 2000;97:12752–12757. doi: 10.1073/pnas.220427097.
18. Getz GS, Reardon CA. Animal models of atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2012;32:1104–1115. doi: 10.1161/ATVBAHA.111.237693.
19. Bentzon JF, Falk E. Atherosclerotic lesions in mouse and man: is it the same disease? *Curr Opin Lipidol*. 2010;21:434–440. doi: 10.1097/MOL.0b013e32833ded6a.
20. Avery RK. Cardiac-allograft vasculopathy. *N Engl J Med*. 2003;349:829–830. doi: 10.1056/NEJMp038124.
21. Wehner J, Morrell CN, Reynolds T, Rodriguez ER, Baldwin WM 3rd. Antibody and complement in transplant vasculopathy. *Circ Res*. 2007;100:191–203. doi: 10.1161/01.RES.0000255032.33661.88.
22. Schmauss D, Weis M. Cardiac allograft vasculopathy: recent developments. *Circulation*. 2008;117:2131–2141. doi: 10.1161/CIRCULATIONAHA.107.711911.
23. Mehra MR, Crespo-Leiro MG, Dipchand A, Ensminger SM, Hiemann NE, Kobashigawa JA, Madsen J, Parameshwar J, Starling RC, Uber PA. International Society for Heart and Lung Transplantation working formulation of a standardized nomenclature for cardiac allograft vasculopathy-2010. *J Heart Lung Transplant*. 2010;29:717–727. doi: 10.1016/j.healun.2010.05.017.
24. Nankivell BJ, Alexander SI. Rejection of the kidney allograft. *N Engl J Med*. 2010;363:1451–1462. doi: 10.1056/NEJMra0902927.
25. Colvin RB. Antibody-mediated renal allograft rejection: diagnosis and pathogenesis. *J Am Soc Nephrol*. 2007;18:1046–1056. doi: 10.1681/ASN.2007010073.
26. Lefaucheur C, Loupy A, Vernerey D, et al. Antibody-mediated vascular rejection of kidney allografts: a population-based study. *Lancet*. 2013;381:313–319. doi: 10.1016/S0140-6736(12)61265-3.
27. Données Informatiques VAlidées en Transplantation: http://di vat.idbc.fr/.
28. Solez K, Colvin RB, Racusen LC, et al. Banff 07 classification of renal allograft pathology: updates and future directions. *Am J Transplant*. 2008;8:753–760. doi: 10.1111/j.1600-6143.2008.02159.x.
29. Sis B, Mengel M, Haas M, et al. Banff '09 meeting report: antibody mediated graft deterioration and implementation of Banff working groups. *Am J Transplant*. 2010;10:464–471. doi: 10.1111/j.1600-6143.2009.02987.x.
30. Racusen LC, Solez K, Colvin RB, et al. The Banff 97 working classification of renal allograft pathology. *Kidney Int*. 1999;55:713–723. doi: 10.1046/j.1523-1755.1999.00299.x.
31. Haas M, Sis B, Racusen LC, et al; Banff meeting report writing committee. Banff 2013 meeting report: inclusion of c4d-negative antibody-mediated rejection and antibody-associated arterial lesions. *Am J Transplant*. 2014;14:272–283. doi: 10.1111/ajt.12590.
32. Lefaucheur C, Loupy A, Hill GS, Andrade J, Nochy D, Antoine C, Gautreau C, Charron D, Glotz D, Suberbielle-Boissel C. Preexisting donor-specific HLA antibodies predict outcome in kidney transplantation. *J Am Soc Nephrol*. 2010;21:1398–1406. doi: 10.1681/ASN.2009101065.
33. Loupy A, Lefaucheur C, Vernerey D, et al. Complement-binding anti-HLA antibodies and kidney-allograft survival. *N Engl J Med*. 2013;369:1215–1226. doi: 10.1056/NEJMoa1302506.
34. CRISTAL, Agence Biomédecine: https://http://www .sipg.sante.fr/portail/.
35. Loupy A, Hill GS, Jordan SC. The impact of donor-specific anti-HLA antibodies on late kidney allograft failure. *Nat Rev Nephrol*. 2012;8:348–357. doi: 10.1038/nrneph.2012.81.
36. Young JB, Neumayer HH, Gordon RD. Pretransplant cardiovascular evaluation and posttransplant cardiovascular risk. *Kidney Int Suppl*. 2010:S1–S7.
37. Lentine KL, Costa SP, Weir MR, Robb JF, Fleisher LA, Kasiske BL, Carithers RL, Ragosta M, Bolton K, Auerbach AD, Eagle KA; American Heart Association Council on the Kidney in Cardiovascular Disease and Council on Peripheral Vascular Disease; American Heart Association; American College of Cardiology Foundation. Cardiac disease evaluation and management among kidney and liver transplantation candidates: a scientific statement from the American Heart Association and the American College of Cardiology Foundation: endorsed by the American Society of Transplant Surgeons, American Society of Transplantation, and National Kidney Foundation. *Circulation*. 2012;126:617–663. doi: 10.1161/CIR.0b013e31823eb07a.
38. Gigante B, Leander K, Vikström M, et al; IMPROVE Study Group. Low levels of IgM antibodies against phosphorylcholine are associated with fast carotid intima media thickness progression and cardiovascular risk in men. *Atherosclerosis*. 2014;236:394–399. doi: 10.1016/j.atherosclerosis.2014.07.030.
39. Gómez M, Vila J, Elosua R, Molina L, Bruguera J, Sala J, Masià R, Covas MI, Marrugat J, Fitó M. Relationship of lipid oxidation with subclinical atherosclerosis and 10-year coronary events in

general population. *Atherosclerosis*. 2014;232:134–140. doi: 10.1016/j.atherosclerosis.2013.10.026.
40. Arvanitakis Z, Brey RL, Rand JH, Schneider JA, Capuano AW, Yu L, Leurgans SE, Bennett DA, Levine SR. Relation of antiphospholipid antibodies to postmortem brain infarcts in older people. *Circulation*. 2015;131:182–189. doi: 10.1161/CIRCULATIONAHA.114.012479.
41. Cai J, Terasaki PI. Humoral theory of transplantation: mechanism, prevention, and treatment. *Hum Immunol*. 2005;66:334–342. doi: 10.1016/j.humimm.2005.01.021.
42. Bian H, Reed EF. Anti-HLA class I antibodies transduce signals in endothelial cells resulting in FGF receptor translocation, down-regulation of ICAM-1 and cell proliferation. *Transplant Proc*. 2001;33:311.
43. Bieri M, Oroszlan M, Farkas A, Ligeti N, Bieri J, Mohacsi P. Anti-HLA I antibodies induce VEGF production by endothelial cells, which increases proliferation and paracellular permeability. *Int J Biochem Cell Biol*. 2009;41:2422–2430. doi: 10.1016/j.biocel.2009.06.009.
44. Frank R, Molina MR, Wald JW, Goldberg LR, Kamoun M, Lal P. Correlation of circulating donor-specific anti-HLA antibodies and presence of C4d in endomyocardial biopsy with heart allograft outcomes: a single-center, retrospective study. *J Heart Lung Transplant*. 2013;32:410–417. doi: 10.1016/j.healun.2012.12.005.
45. Uehara S, Chase CM, Cornell LD, Madsen JC, Russell PS, Colvin RB. Chronic cardiac transplant arteriopathy in mice: relationship of alloantibody, C4d deposition and neointimal fibrosis. *Am J Transplant*. 2007;7:57–65. doi: 10.1111/j.1600-6143.2006.01599.x.
46. Mach F, Schönbeck U, Sukhova GK, Atkinson E, Libby P. Reduction of atherosclerosis in mice by inhibition of CD40 signalling. *Nature*. 1998;394:200–203. doi: 10.1038/28204.
47. Lundberg AM, Hansson GK. Innate immune signals in atherosclerosis. *Clin Immunol*. 2010;134:5–24. doi: 10.1016/j.clim.2009.07.016.
48. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med*. 2005;352:1685–1695. doi: 10.1056/NEJMra043430.
49. Hermansson A, Ketelhuth DF, Strodthoff D, Wurm M, Hansson EM, Nicoletti A, Paulsson-Berne G, Hansson GK. Inhibition of T cell response to native low-density lipoprotein reduces atherosclerosis. *J Exp Med*. 2010;207:1081–1093. doi: 10.1084/jem.20092243.
50. Blake GJ, Ostfeld RJ, Yucel EK, Varo N, Schönbeck U, Blake MA, Gerhard M, Ridker PM, Libby P, Lee RT. Soluble CD40 ligand levels indicate lipid accumulation in carotid atheroma: an in vivo study with high-resolution MRI. *Arterioscler Thromb Vasc Biol*. 2003;23:e11–e14.
51. Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, Heintzman N, Ren B, Fu XD, Topol EJ, Rosenfeld MG, Frazer KA. 9p21 DNA variants associated with coronary artery disease impair interferon-γ signalling response. *Nature*. 2011;470:264–268. doi: 10.1038/nature09753.
52. Middleton J, Patterson AM, Gardner L, Schmutz C, Ashton BA. Leukocyte extravasation: chemokine transport and presentation by the endothelium. *Blood*. 2002;100:3853–3860. doi: 10.1182/blood.V100.12.3853.
53. Duewell P, Kono H, Rayner KJ, et al. NLRP3 inflammasomes are required for atherogenesis and activated by cholesterol crystals. *Nature*. 2010;464:1357–1361. doi: 10.1038/nature08938.

Novelty and Significance

What Is Known?

- Traditional cardiovascular risk factors do not fully explain arteriosclerosis pathogenesis and its long-term clinical consequences.
- Experimental data suggest that immune processes play a role in arteriosclerosis.
- Past studies in humans did not aim to understand the contribution of immune determinants of arteriosclerosis beyond traditional cardiovascular risk factors.
- Kidney transplantation represents a unique human model where traditional cardiovascular risk factors coexist with immune responses.

What New Information Does This Article Contribute?

- Circulating antibodies (anti–human leukocyte antigen) are major determinants of severe allograft arteriosclerosis, independent of traditional cardiovascular risk factors.
- Circulating antibodies are associated with an increased risk of cardiovascular events, independent of traditional cardiovascular risk factors.

Using a population-based approach in a large prospective cohort of kidney transplant recipients, we demonstrated that circulating antibodies (anti–human leukocyte antigen antibodies) are major risk factors for the development of arteriosclerosis and cardiovascular events, independent of traditional risk factors. We identified antibody-associated arteriosclerosis as an accelerated form of the disease, characterized by a distinct immunologic and histological phenotype. High levels of circulating anti–human leukocyte antigen antibodies and their complement-binding capacity were associated with increased severity of arteriosclerosis and increased incidence of cardiovascular events. Circulating anti–human leukocyte antigen antibodies were also associated with major adverse cardiovascular events independent of allograft arteriosclerosis. Overall, our results support the notion of immunologic development of arteriosclerosis in humans. We identified a population of kidney transplant recipients at high risk of graft loss, major cardiovascular events and death, for whom specific screening and therapeutic strategies need to be developed. This idea extends beyond renal transplantation to reports of immunologic damage to arteries in heart transplantations and other solid organs. The recognition of an immune component in the pathogenesis of severe arteriosclerosis could lead to the development of immune-modulating therapeutic strategies in solid organ transplant recipients and potentially in the general population.