POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Estimation du pronostic de survie en fonction des caractéristiques du greffon et du receveur d'une transplantation rénale

PIERRE-LUC BOIVIN
Département de mathématiques et de génie industriel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Mathématiques appliquées

Juin 2024

© Pierre-Luc Boivin, 2024.

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Ce mémoire intitulé :

Estimation du pronostic de survie en fonction des caractéristiques du greffon et du receveur d'une transplantation rénale

présenté par **Pierre-Luc BOIVIN**
en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

Julie CARREAU, présidente
Jonathan JALBERT, membre et directeur de recherche
Héloïse CARDINAL, membre et codirectrice de recherche
Lévis THÉRIAULT, membre et codirecteur de recherche
Quentin CAPPART, membre

REMERCIEMENTS

Je tiens à accentuer le support et la motivation qui m'ont été alloués par mes superviseurs Jonathan Jalbert, Héloïse Cardinal et Lévis Thériault. Leur patience et leur confiance en moi ne sont pas passées inaperçues. Merci à ma famille, à mes amis et à ma copine pour croire en moi et de m'encourager à viser haut.

RÉSUMÉ

Bien que la greffe rénale soit le meilleur traitement de la maladie rénale chronique terminale, il persiste un écart entre le besoin d'organes pour la greffe et la disponibilité de ceux-ci. L'un des moyens d'expandre le bassin d'organe est d'accepter de greffer des reins ayant une longévité attendue réduite. La majorité des études montrent qu'il y a un bénéfice de survie associé à la transplantation d'un rein à longévité réduite par rapport à demeurer en dialyse. Cependant, les patients doivent également être informés des risques potentiels de cette option. Il existe peu de modèles personnalisés à la combinaison caractéristiques du patient et du greffon permettant d'estimer les survies du patient et du greffon après la greffe lorsque des reins à longévité réduite sont utilisés.

Le but de notre étude était de créer un modèle d'estimation de la survie du greffon et du patient ayant une excellente calibration et permettant de comparer ces survies si un rein à longévité réduite vs un rein «standard» était transplanté. Nous avons effectué une étude de cohorte rétrospective incluant des patients ayant reçu une première transplantation rénale entre le 1er Janvier 2000 et le 1er Juin 2021 dans le registre américain United Network for Organ Sharing (UNOS). Les patients ont été suivis de la date de greffe jusqu'au décès ou la perte de greffon, définie comme le retour en dialyse ou la regreffe. Le critère de jugement primaire était la perte de greffon (décès, retour en dialyse ou regreffe). Le critère de jugement secondaire était la mortalité. L'approche proposées repose sur les arbres de régression qui séparent de façon automatique des sous-groupes de patients ayant des survies similaires sur la base des diverses variables indépendantes, jusqu'à atteindre un critère d'arrêt. Une métrique originale basée sur le test de Kolmogorov-Smirnov a été utilisée comme critère de partitionnement.

Au total, 159 140 patients ont été inclus, avec un suivi médian de 6.9 ans. Les variables incluses comptaient entre autres le Kidney Donor Risk Index du donneur, l'âge du candidat et du donneur, le temps passé par le candidat en dialyse, la cause de la maladie rénale, et bien d'autres. En limitant le nombre minimal d'événements dans les feuilles à 500, on obtient des arbres dont la profondeur varie de 29 et 40 niveaux et une largeur d'arbre entre 129 et 171 feuilles. Nous avons démontré que les arbres décisionnels et les forêts aléatoires généraient des prédictions de survie très bien calibrées. Ces estimations permettront aux néphrologues transplanteurs d'informer les patients en ce qui a trait aux offres de rein à longévité réduite.

ABSTRACT

Although transplantation is the best treatment for end-stage kidney disease (EKD), there remains an urgent need to increase the number of transplants and organ availability. One solution is to increase the organ pool and accept extended criteria donors (ECD). Most studies show a benefit in survival for candidates who are transplanted with organs from ECD. However, this requires patients to be informed of the potential risks of choosing that option. There are very few models to predict individual survival curves of kidney allografts and patients.

Our study's goal is to create a well-calibrated graft and patient survival estimation model to inform the candidates about the implications of accepting an ECD offer in a personalized fashion. We have performed a retrospective cohort study including patients who received a first renal transplant between January 1st, 2000, and June 1st, 2021 in the United Network for Organ Sharing (UNOS) American transplant registry. Patients had follow-up from the date of transplant to death or graft loss, which was defined as a return to dialysis or a re-transplant.

Our approach includes a novel algorithm that finds at each iteration the ideal splits based on a variety of independent variables to group patients that share similar survival tendencies until a stopping criterion is met. In total, 159,140 patients were included, with a median follow-up time of 6.9 years. The included variables are, among others, the donor's KDRI, the donor's and recipient's age, the diagnosed cause of EKD, and more. By limiting the number of events in the leaves to 500, we obtain trees with a depth varying from 29 to 40 levels and a size between 129 and 171. We have demonstrated that this novel model's survival predictions are well-calibrated. Thus, decision-making can be based on its results with confidence.

This method allows for patient and graft survival estimates that are calibrated and individualized based on the patient's and donor's characteristics. After validating our approach with data from the province of Quebec, transplant nephrologists and transplant candidates will have a tool to inform them more precisely of the risks and value of accepting offers from ECD.

TABLE DES MATIÈRES

REMERCIEMENTS iii

RÉSUMÉ iv

ABSTRACT v

TABLE DES MATIÈRES vi

LISTE DES TABLEAUX viii

LISTE DES FIGURES ix

LISTE DES SIGLES ET ABRÉVIATIONS xi

LISTE DES ANNEXES xii

CHAPITRE 1 INTRODUCTION 1

CHAPITRE 2 REVUE DE LITTÉRATURE 5
2.1 Allocation rénale au Québec 5
2.1.1 Score d'attribution 5
2.1.2 Prise de décision 6
2.1.3 Donneurs à longévité réduite 7
2.2 Informations supportant la prise de décision 7
2.3 Modèles de prédiction 9
2.4 Calibration d'un modèle de prédiction 10

CHAPITRE 3 CADRE THÉORIQUE 13
3.1 Analyse de survie 13
3.1.1 Données de survie 13
3.1.2 Kaplan-Meier 14
3.2 Test d'adéquation de Kolmodrov-Smirnov 14
3.3 Arbres de décision 15
3.3.1 L'index de Gini 15
3.3.2 Gain d'information 16
3.3.3 Somme des erreurs au carré 17

3.3.4 Critère d'arrêt 17

CHAPITRE 4 DONNÉES 18
4.1 Pré-traitement 18
4.2 Variables indépendantes 18
4.3 Variables dépendantes 22
4.4 Gestion des valeurs manquantes 23

CHAPITRE 5 MÉTHODE 24
5.1 Arbres de décision 24
5.2 Critère de partitionnement 24
5.3 Prédictions 26
5.4 Forêt aléatoire 26
5.5 Hyper-paramètres 28
5.6 Calibration 29

CHAPITRE 6 RÉSULTATS 31
6.1 Échec du greffon 31
6.1.1 Arbre obtenu avec la cohorte d'entraînement 31
6.1.2 Forêt obtenue avec la cohorte d'entraînement 32
6.2 Décès du receveur 35
6.2.1 Forêt obtenue avec la cohorte d'entraînement 35
6.3 Analyse des arbres et des forêts obtenus 36
6.4 Calibration D 39
6.5 Évaluation de la calibration 39
6.6 Effet de l'âge et du KDRI 40

CHAPITRE 7 DISCUSSION 44

CHAPITRE 8 CONCLUSION 48

RÉFÉRENCES 49

ANNEXES 59

LISTE DES TABLEAUX

Tableau 4.1 Description des étapes de pré-traitement à partir des données brutes. 19
Tableau 4.2 Description des variables indépendantes catégorielles. 20
Tableau 4.3 Description des variables indépendantes continues. 22
Tableau 4.4 Description des variables dépendantes. 23
Tableau 5.1 Valeurs sélectionnées pour les différents hyper-paramètres. 29

LISTE DES FIGURES

Figure 2.1 Exemple de graphiques de calibration modérée pour trois scénarios différent, lorsque les prédictions sont calibrées (A), lorsque les estimés de risque sont trop extrêmes (B) et lorsque les risques sont sous-estimés dans l'ensemble (C). 11

Figure 5.1 Représentation simplifiée du processus de prédiction avec un arbre de survie unique. 27

Figure 5.2 Représentation simplifiée du processus de prédiction une forêt à cinq arbres. 27

Figure 5.3 Courbe de Kaplan-Meier décrivant l'utilité du greffon censuré à la dernière date de mise à jour si le premier événement est le décès (cible 1). 30

Figure 5.4 Visualisation d'un exemple fictif de calibration D tiré de Haider *et al.* (2020). 30

Figure 6.1 Structure de l'arbre unique entraîné pour prédire la fonction de la greffe. 33

Figure 6.2 Courbes de survie du greffon estimées par l'estimateur de Kaplan-Meier pour les deux sous-groupes obtenus, à savoir les patients âgés de 72 ans et plus et les plus jeunes. Le partitionnement à l'âge de 72 ans produit la plus grande différence de survie selon la distance de Kolmogorov-Smirnov entre les courbes, indiquée par la flèche. 34

Figure 6.3 Graphique de calibration D d'un arbre unique en fonction de l'échec de la greffe. 34

Figure 6.4 Graphique de calibration d'une forêt à cinq arbres en fonction de l'échec de la greffe. 34

Figure 6.5 Graphique de calibration D d'un arbre unique en fonction du décès du receveur. 36

Figure 6.6 Graphique de calibration d'une forêt à cinq arbres en fonction du décès du receveur. 36

Figure 6.7 Structure de l'arbre unique entraîné pour prédire la survie du receveur. 37

Figure 6.8 Courbes de survie du patient estimées par l'estimateur de Kaplan-Meier pour les deux sous-groupes obtenus, à savoir les patients âgés de 72 ans et plus et les plus jeunes. Le partitionnement à l'âge de 72 ans produit la plus grande différence de survie selon la distance de Kolmogorov-Smirnov entre les courbes, indiquée par la flèche. 38

Figure 6.9 Exemple de prédiction de survie du greffon faite par le modèle de forêt à 5 arbres pour un candidat âgé de 68 ans ayant accepté une offre de KDRI inférieur à 1.2. 42
Figure 6.10 Exemple de prédiction de survie du greffon faite par le modèle de forêt à 5 arbres pour un candidat âgé de 30 ans ayant accepté une offre de KDRI inférieur à 1.2. 42
Figure 6.11 Exemple de prédiction de survie du greffon faite par le modèle de forêt à 5 arbres pour un candidat âgé de 68 ans ayant accepté une offre de KDRI supérieur à 2.4. 42
Figure 6.12 Exemple de prédiction de survie du greffon faite par le modèle de forêt à 5 arbres pour un candidat âgé de 30 ans ayant accepté une offre de KDRI supérieur à 2.4. 42
Figure 6.13 Courbe Kaplan-Meier des candidats de la cohorte ayant eu des offres de rein provenant de donneurs ayant des KDRI de 2.3 à 2.5, donc de moindre qualité. 43
Figure 6.14 Courbe Kaplan-Meier des candidats de la cohorte ayant eu des offres de reins provenant de donneurs ayant des KDRI de 1.1 à 1.3, donc de qualité supérieure. 43
Figure A.1 Représentation de l'arbre #1 de la forêt pour prédire l'échec du greffon. 54
Figure A.2 Représentation de l'arbre #2 de la forêt pour prédire l'échec du greffon. 55
Figure A.3 Représentation de l'arbre #3 de la forêt pour prédire l'échec du greffon. 56
Figure A.4 Représentation de l'arbre #4 de la forêt pour prédire l'échec du greffon. 57
Figure A.5 Représentation de l'arbre #5 de la forêt pour prédire l'échec du greffon. 58
Figure A.6 Représentation de l'arbre #1 de la forêt pour prédire le décès du candidat. 60
Figure A.7 Représentation de l'arbre #2 de la forêt pour prédire le décès du candidat. 61
Figure A.8 Représentation de l'arbre #3 de la forêt pour prédire le décès du candidat. 62
Figure A.9 Représentation de l'arbre #4 de la forêt pour prédire le décès du candidat. 63
Figure A.10 Représentation de l'arbre #5 de la forêt pour prédire le décès du candidat. 64

LISTE DES SIGLES ET ABRÉVIATIONS

cPRA	*Calculated Panel Reactive Antibody*
DCE	Donneur à Critères Étendus
EPTS	*Estimated Post-Transplant Survival*
ESP	*Eurotransplant Senior Program*
HLA	*Human Leukocyte Antigens*
KDPI	*Kidney Donor Profile Index*
KDRI	*Kidney Donor Risk Index*
KM	Kaplan-Meier
KS	Kolmodrov-Smirnov
MRC	Maladie Rénale Chronique
OPTN	Organ Procurement and Transplantation Network

LISTE DES ANNEXES

Annexe A Visualisation des modèles d'arbres et de forêts 53

CHAPITRE 1 INTRODUCTION

La maladie rénale chronique (MRC) est une condition sévère affectant près de 50 000 Canadiens et dont le taux d'incidence ne cesse d'augmenter (Canadian Organ Replacement Register, 2023). Le traitement de ces patients représente 1.2% des dépenses en soins de santé du pays, constituant un fardeau pour le système de santé canadien (Manns *et al.*, 2007; Wolfe *et al.*, 1999; Laupacis *et al.*, 1996). Une personne souffrant de MRC terminale (de stade 5) a généralement trois options en ce qui concerne les traitements possibles : l'hémodialyse, la dialyse péritonéale et la greffe rénale.

L'hémodialyse consiste à filtrer le sang du patient par diffusion le long d'une membrane semi-perméable dans un dialyseur. Pouvant durer de 2.5 à 5 heures à chaque séance, ce traitement est généralement effectué 3 fois par semaine et peut être fait dans un établissement médical (le plus souvent) ou à la maison. L'hémodialyse est une solution qui, quoique plus répandue, nécessite un investissement de temps considérable de la part des patients. De plus, l'ultrafiltration (extraction de volume effectuée pour éviter la surcharge volémique en présence de diminution du volume urinaire en MRC) peut être moins bien tolérée par les patients plus âgés ou qui souffrent de maladie cardio-vasculaire (Danovitch, 2017).

La dialyse péritonéale est un traitement moins répandu que l'hémodialyse. Il consiste à installer un cathéter au travers de la paroi abdominale du patient par chirurgie plusieurs semaines avant le début de la dialyse. Les patients apprennent à faire leur propre traitement chaque jour en injectant par le tube de 1.5 à 3 L de fluide de dialyse péritonéale. Ils ferment ensuite le cathéter et laissent le soluté en place pour 4 à 5 heures. Les échanges entre le sang et le soluté se font par diffusion via les capillaires de la membrane péritonéale pendant ce temps. Ensuite, le patient peut ouvrir le cathéter et drainer le soluté, ce qui entraînera la clairance des molécules urémiques ainsi qu'une extraction volémique puisque des solutés osmotiques sont présents dans le dialysat. La procédure sera répétée manuellement 4 à 5 fois par jour ou via un appareil appelé cycler. Celui-ci permet une dialyse automatisée la nuit. Les échanges durent alors de 20 à 45 minutes pour un total de 8 à 10 heures (Danovitch, 2017).

La dernière option est la greffe rénale. Elle est reconnue pour être le traitement idéal de la MRC puisqu'elle offre une meilleure survie à long terme par rapport aux 2 autres modalités de remplacement de la fonction rénale. Par exemple, les receveurs d'une greffe bénéficient d'une diminution du risque relatif de décès de 65 % après 1 an par rapport à la dialyse. La greffe rénale est également moins coûteuse que la dialyse, ce qui est un avantage pour le

système de santé canadien. Aux États-Unis, l'hémodialyse coûte près de 75 000 $ annuellement par patient, tandis que la transplantation coûte 100 000 $ la première année, puis approximativement 10 000 $ par année. Donc, la transplantation est bien moins coûteuse à long terme (Danovitch, 2017).

Les patients ayant reçu une greffe rapportent généralement avoir une meilleure qualité de vie que ceux et celles qui sont en hémodialyse ou en dialyse péritonéale. La satisfaction à l'égard de la vie, le bien-être physique et émotionnel et la capacité de retourner au travail sont significativement meilleurs pour les patients transplantés. Souvent, la transplantation corrige ou améliore des complications liées à l'urémie qui ne sont généralement pas renversées par la dialyse telles que l'anémie, la neuropathie périphérique, la neuropathie autonome et la dysfonction sexuelle (Danovitch, 2017).

Même si la transplantation présente bien des avantages, il y a un besoin pressant d'augmenter le nombre d'organes transplantés puisque plus de 3000 Canadiens sont présentement en attente de greffe. Il existe effectivement un écart important entre le besoin d'organe pour la greffe et la disponibilité de ceux-ci (Canadian Organ Replacement Register, 2023). Une façon d'augmenter le nombre de greffe serait d'accepter des organes de donneurs qui présentent des caractéristiques pouvant limiter la durée de fonction du greffon, tels que l'âge et le degré d'hypertension du donneur (Pascual *et al.*, 2008).

Plusieurs étiquettes ont été apposées aux offres de rein ayant une longévité potentielle réduite. Par exemple, ces reins ont été appelés *reins provenant de donneurs à critères étendus*. Les donneurs à critères étendus sont soit 1) âgés de 60 ans et plus ou 2) âgés de 50 à 59 ans en présence d'au moins deux caractéristiques parmi les suivantes :

— un historique d'hypertension ;
— un niveau de créatinine sérique supérieur ou égal à 1.5 mg/dL (132 umol/L) ;
— le décès causé par un accident vasculaire cérébral.

Il arrive plus souvent que les organes provenant de donneurs à critères étendus ne soient pas utilisés pour la greffe en comparaison avec ceux des donneurs standards, en raison d'une crainte associée à une durée de vie du greffon sous-optimale (Argani, 2022). Cependant, malgré qu'ils soient associés à une longévité moindre, 50 % des reins de donneurs à critères étendus sont encore fonctionnels 5 ans après la greffe. De plus, la survie des receveurs de ces reins est meilleure que s'ils demeurent en dialyse sur la liste d'attente (Argani, 2022; Ojo *et al.*, 2001). Ce gain se produit 1.5 ans après la transplantation. Un plus grand pourcentage de ces donneurs à critères étendus pourraient être utilisés, ce qui aiderait à contrer le manque d'organe (Ojo *et al.*, 2001). En Espagne, l'inclusion des donneurs plus âgés a permis d'augmenter le nombre de transplantations rénales effectuées, diminuant ainsi le nombre de

patients en attente de greffe et le temps d'attente jusqu'à la greffe pour ceux-ci (Port *et al.*, 2002).

Plus récemment, ces reins à longévité réduite ont été appelés *reins provenant de donneurs à Kidney Donor Risk Index (KDRI) élevé*. Le KDRI est un score dérivé d'un modèle proportionnel de Cox entraîné sur les données américaines de transplantation rénale. Le KDRI utilise des caractéristiques du donneur afin de prédire la durée de survie des greffons rénaux. Tout comme pour la définition précédente, la greffe de reins provenant de donneurs à KDRI élevé est associée à une longévité réduite de fonction du greffon (Rao *et al.*, 2009). Cependant, un avantage de survie demeure présent par rapport à l'option de demeurer sur la liste d'attente en dialyse dans la grande majorité des cas (Massie *et al.*, 2014).

Puisque d'accepter les offres de reins à longévité réduite peut résulter en un gain de survie, il est important que ces offres soient considérées avec soin et acceptées dans la mesure du possible. Cependant, les travaux antérieurs de notre équipe montrent que les patients désirent obtenir de l'information sur le risque de longévité réduite associé à l'utilisation de ces reins (Khalili *et al.*, 2022). Une autre équipe canadienne a démontré le même besoin d'information ressenti par les patients (Rosaasen *et al.*, 2022). De plus, les participants à cette étude, qui étaient des candidats à la greffe âgés, ont nommé qu'une espérance de fonction de 3-5 ans était suffisante pour accepter une offre de rein à longévité réduite.

Il apparaît donc qu'il est préférable d'accepter plus de reins à longévité réduite afin d'augmenter l'accès à la greffe. Pour ce faire, il faut informer les candidats à la greffe de façon adéquate. Cependant, il n'existe présentement pas d'outil personnalisé permettant aux médecins transplanteurs de bien informer les patients de leur pronostic de survie (patient et greffon) s'ils acceptent une offre d'organe à longévité réduite plutôt qu'à longévité standard.

Le but des travaux présentés dans ce mémoire consiste à offrir au patient en attente d'une greffe de rein et au néphrologue transplanteur soignant des estimations des pronostics de survie du greffon et du patient. Ces estimations permettront à l'équipe patient-médecin de prendre une décision éclairée sur l'acceptation ou le refus de telles offres. Afin d'obtenir des pronostics individualisés concernant les caractéristiques du rein et du patient, les arbres de régression permettant de stratifier la cohorte en sous-groupes homogènes par rapport à la survie du patient et du greffon sont utilisés. Les objectifs de ce travail sont les suivants :

1. Préparer les données ;
2. Définir une cible, soit la survie du greffon ou la survie du patient ;
3. Identifier une métrique appropriée pour classifier les patients de la cohorte selon la cible choisie ;
4. Implémenter les arbres de régression avec la métrique identifiée pour estimer la survie

du patient en fonction de l'offre à longévité réduite ;

5. Vérifier la bonne calibration des estimations obtenues par le modèle.

Les estimations obtenues permettront potentiellement d'augmenter le nombre de transplantations d'offres à longévité réduite.

Le reste du mémoire est organisé de la façon suivante. Le chapitre 2 présente une revue de littérature sur l'aide à la décision pour les dons d'organes. Le chapitre 3 présente le cadre théorique des concepts de survie et d'arbres de décision pertinent pour la lecture de ce mémoire. Le chapitre 4 décrit l'ensemble de données utilisé et les étapes de pré-traitement faites pour obtenir la cohorte pour notre analyse. Le chapitre 5 explique le processus d'entraînement du modèle et des algorithmes implémentés pour obtenir nos résultats. Le chapitre 6 présente les résultats obtenu à partir de notre modèle. Le chapitre 7 contient la discussion et les réflexions qui sont tirés de ces résultats. Finalement, le chapitre 7 conclut l'entièreté de ce mémoire.

CHAPITRE 2 REVUE DE LITTÉRATURE

2.1 Allocation rénale au Québec

L'allocation de rein du Québec est un processus servant à offrir les organes de façon équitable aux candidats en attente d'une transplantation. Les critères utilisés pour déterminer l'ordre de priorité ont été établis par des professionnels (néphrologues et chirurgiens transplanteurs) des centres de transplantations du Québec qui siègent sur le comité "Rein-Pancréas" de Transplant Québec. Dans son entièreté, Transplant Québec est responsable de l'allocation des organes de donneurs décédés – dont les reins – pour la province de Québec.

2.1.1 Score d'attribution

Après que certaines priorités (urgence médicale avec impossibilité d'offrir une thérapie de remplacement rénal, candidat ayant un taux d'anticorps anti-HLA très élevé, greffe organe combinée (ie, rein-pancréas, foie-rein), candidat pédiatrique) aient été prises en compte, les donneurs décédés identifiés par Transplant Québec sont attribués aux candidats inscrits sur la liste d'attente par pointage. Cette liste centralisée comprend des patients souffrant de maladie rénale chronique terminale qui sont suivis (en hémodialyse, dialyse péritonéale ou pré-dialyse) dans tous les centres hospitaliers du Québec et dont la candidature a été acceptée dans l'un des 5 centres hospitaliers effectuant des greffes rénales au Québec (Centre hospitalier de l'Université de Montréal (CHUM), Centre universitaire de santé McGill (CUSM), Centre intégré en santé et services sociaux de l'Est de l'Ile-Hôpital Maisonneuve-Rosemont (HMR), Centre hospitalier universitaire de Sherbrooke (CHUS) et Centre hospitalier universitaire de Québec (CHUQ)). Sur la liste par pointage, l'ordre de priorité des candidats est principalement déterminé par un score d'attribution calculé à l'aide de plusieurs caractéristiques reliées au candidat et au donneur. Ces caractéristiques sont : le temps passé par le candidat en dialyse, le degré de compatibilité HLA entre le donneur et le candidat, le taux élevé d'anticorps anti-HLA du candidat, appelé "Calculated Panel Reactive Antibody" (cPRA), l'écart d'âge entre le donneur et le candidat et l'âge du candidat (Transplant Québec, 2020). Ces critères ont été choisis afin de créer un certain équilibre entre « l'utilité » de la greffe (i.e. s'assurer que les reins attribués aient la durée de fonction la plus longue possible) et la justice, en particulier pour les patients qui attendent une offre depuis longtemps ou qui possèdent des anticorps contre une grande proportion de donneurs (cPRA élevé), réduisant le nombre d'offres potentielles. Les candidats ayant un groupe sanguin compatible avec celui du donneur sont classés

par ordre décroissant du score d'attribution, en considérant que celui ou celle ayant le score le plus élevé est prioritaire. Les coordonnateurs de Transplant Québec appellent ensuite les néphrologues transplanteurs de garde aux centres où les candidats sont inscrits en ordre de priorité (ordre décroissant du score). Les coordonnateurs de Transplant Québec transfèrent donc l'histoire médicale, les données cliniques et de laboratoire du donneur au néphrologue transplanteur. Celui-ci évalue la qualité de l'offre sur la base de facteurs de risque qui ont précédemment été associés avec la durée de fonction de la greffe. Si l'offre est jugée tout à fait inacceptable par le néphrologue transplanteur, il refusera l'offre de façon unilatérale. Dans la très grande majorité des cas, cependant, le candidat à la greffe pour lequel l'offre est faite sera contacté par le néphrologue transplanteur et l'information pertinente à l'offre sera transmise au candidat. Une décision quant à l'acceptation de l'offre sera prise de façon conjointe par le néphrologue transplanteur et le candidat à la greffe. La décision sera communiquée aux coordonnateurs de Transplant Québec. Ceux-ci continueront le processus d'appel jusqu'à ce que les 2 reins d'un donneur soient acceptés ou refusés pour tous les candidats. À noter qu'un seul rein est greffé pour chaque candidat – un donneur crée donc une greffe rénale chez 2 receveurs différents dans la très grande majorité des cas.

2.1.2 Prise de décision

Accepter ou refuser une offre est une décision qui doit être prise rapidement par les candidats et les néphrologues transplanteurs. C'est un choix qui implique des enjeux importants pour les patients. La prise de décision partagée est reconnue pour améliorer la satisfaction et la qualité décisionnelle pour les receveurs de rein (Ladin, 2022), quoiqu'elle n'est pas appliquée systématiquement au Québec. Elle consiste à impliquer les patients en les informant des risques et des bénéfices du traitement pour choisir l'option qui satisfait leurs attentes et leurs désirs. La greffe rénale est le meilleur traitement de la maladie rénale chronique car elle prolonge l'espérance de vie et améliore la qualité de vie en comparaison avec la dialyse. Cependant, toutes les offres de rein ne sont pas de qualité égale et il est démontré que les caractéristiques du donneur sont associées avec la durée de fonction de la greffe rénale. Ainsi, lorsqu'un offre de rein provenant d'un donneur jeune et en parfait santé est faite, la décision d'accepter est simple. Au contraire, une offre provenant d'un donneur avec des caractéristiques extrêmement défavorables en terme d'âge et de co-morbidités sera facile à refuser. Entre ces 2 cas de figure se situent la majorité des offres de rein. Dans ces cas, les médecins transplanteurs doivent informent les patients des risques possibles d'accepter ou de refuser l'offre en fonction de leur expérience professionnelle et des données scientifiques disponibles dans la littérature médicale en transplantation rénale.

2.1.3 Donneurs à longévité réduite

Il a été bien démontré que certaines caractéristiques des donneurs sont associées à une longévité moindre de la greffe rénale (Rao *et al.*, 2009; Port *et al.*, 2002). Au cours des 20 dernières années, la nomenclature a évolué à ce sujet. Alors que le terme « donneur marginal » a été utilisé sans définition opérationnelle précise initialement, deux autres termes, « donneur à critères étendu » ou plus récemment, « donneur à KDPI élevé », sont associés à des définitions précises. A cause de changements macroscopiques et microscopiques causés par l'âge, comme la néphroangiosclérose et l'accumulation de matrice extra-cellulaire, la fonction rénale se détériore avec le temps. Cela s'exprime par une diminution du débit de filtration glomérulaire (la façon dont la fonction rénale est mesurée) de 1 mL par année à partir de l'âge de 40 ans (Argani, 2022). Un donneur à critères étendus (DCE) est soit âgé de plus de 60 ans ou âgé de 50 à 59 ans avec deux des conditions suivantes : une histoire d'hypertension, un niveau de créatinine sérique supérieur ou égal à 1.5 mg/dL (132 umol/L), ou un décès par accident vasculaire cérébral. Les reins de DCE sont associés à une survie du greffon (longévité) moindre par rapport à un organe de donneur standard. Cependant, les greffes effectuées avec des reins de DCE offrent généralement une meilleure survie et qualité de vie aux receveurs par rapport à demeurer en dialyse (Argani, 2022). En effet, des études démontrent qu'en moyenne, accepter une offre de DCE est associé à une meilleure survie du receveur que de refuser ces offres en espérant un donneur de meilleure qualité et pas Massie (Merion *et al.*, 2005). Quoique ces offres ont tendance à représenter un plus haut risque de longévité réduite du greffon rénal, elles offrent une diminution significative du risque de mortalité à long terme vs la dialyse, particulièrement chez les receveurs plus âgés (50 ans et plus) (Merion *et al.*, 2005). Comparativement à la dialyse, les receveurs de DCE ont une augmentation de l'espérance de vie attendue de 5 ans par rapport à ceux qui restent en dialyse sans transplantation (Ojo *et al.*, 2001). Plusieurs solutions ont été adoptées pour tenter d'optimiser l'allocation d'organes de DCE. Des programmes comme le « Eurotransplant Senior Program » (ESP) tentent d'apparier des donneurs âgés avec des receveurs âgés, peu importe leur compatibilité d'antigènes HLA. Dans ce cas, le taux de rejet augmente de 5% à 20%, mais la disponibilité d'organes de donneurs âgés double et le temps d'attente moyen diminue pour les candidats âgés (Argani, 2022). Il a été démontré que 10% plus d'organes pourraient être transplantés en favorisant l'utilisation des reins provenant de DCE (Argani, 2022).

2.2 Informations supportant la prise de décision

Le « Kidney Donor Profile Index » (KDPI) est un score dérivant du « Kidney Donor Risk Index » (KDRI), calculé à partir des caractéristiques du donneur et qui sert à estimer la

qualité de l'offre. Plus le KDPI s'approche de 100%, plus la longévité anticipée du greffon est courte, donc moins bonne est la qualité de l'offre. Cette mesure est utilisée aux États-Unis dans l'allocation d'organe, mais pas au Québec. Par exemple aux Etats-Unis, les reins de meilleure qualité (KDPI 20% et moins) se voient attribués préférentiellement aux receveurs qui ont les meilleures chances de survie à long terme. Pour identifier ces candidats, un score appelé « Expected Post-Transplant Survival » (EPTS) a été développé (Organ Procurement and Transplantation Network, 2020a). L'EPTS est une mesure numérique servant à estimer le risque de décès d'un receveur de rein en fonction de son âge, son statut diabétique, son temps passé en dialyse et s'il a été transplanté auparavant. Comme le KDPI, l'EPTS est normalisé par rapport à l'EPTS médian d'une année de référence précédente, puis représenté sous forme de proportion (entre 0 et 100%). L'EPTS et le KDPI sont donc utilisés en tandem afin d'apparier les candidats et les donneurs dans le but d'offrir aux candidats ayant une grande longévité potentielle des organes qui ont aussi des caractéristiques associées à une grande longévité, maximisant ainsi l'utilité de la greffe (Organ Procurement and Transplantation Network, 2020b). Le « Organ Procurement and Transplantation Network » (OPTN), une association regroupant les professionnels impliqués dans le système de dons et de transplantation d'organes aux État-Unis, gère le réseau de transplantation d'organes et s'occupe de définir les critères d'allocation comme l'appariement du KDPI et de l'EPTS. Le KDRI est un score comprenant 10 variables indépendantes reliées au donneur : âge, ethnicité, taille, poids, hypertension, diabète, hépatite C, type de décès (décès neurologique, après arrêt cardio-circulatoire), cause du décès et niveau de créatinine circulant. Le KDRI a une discrimination adéquate (C-index 0.62) pour son association avec la survie du greffon dans une cohorte rétrospective de patients transplantés du rein pour une première greffe entre 1995 et 2005 dans le registre United Network for Organ Sharing (UNOS). Le KDPI est ensuite calculé en exprimant le rang du KDRI d'un donneur x par rapport à la distribution du KDRI de tous les donneurs d'une année de référence précédente. Le KDRI du donneur x est alors converti en son percentile par rapport à la distribution de référence. Au final, le KDPI représente le pourcentage des donneurs de l'année de référence qui ont un meilleur (plus petit) KDRI que celui du donneur x. Par exemple, un KDPI de 80% signifie que 80% des donneurs de l'année de référence précédente avaient une meilleure qualité (Organ Procurement and Transplantation Network, 2020b). Une étude récente montre que l'équation initiale du KDRI est robuste sur une cohorte de validation plus récente, soit les patients greffés du rein aux Etats-Unis entre 2000 et 2016 (Zhong *et al.*, 2019). Le KDRI a été validé dans la population ontarienne de patients transplantés du rein et ajoute de l'information pour prédire la survie du greffon par rapport à un modèle contenant seulement l'âge du donneur (Young *et al.*, 2018). Dans la population de patients greffés du rein de Colombie-Britannique, le KDRI n'ajoutait pas

d'information au-delà de celle fournie par l'âge du donneur (Rose *et al.*, 2018). Tout comme pour les reins provenant de DCE (Merion *et al.*, 2005), des études ont été menées en utilisant le KDPI comme mesure de la qualité de l'offre afin de vérifier s'il y a un gain de survie potentiel à accepter une offre de donneur à KDPI élevé par rapport à demeurer en dialyse pour attendre une meilleure offre. Ces études montrent que de façon générale, il est préférable d'accepter de telles offres que de demeurer en dialyse (Massie *et al.*, 2014; Wey *et al.*, 2018; Bae *et al.*, 2019), surtout pour les candidats âgés qui n'ont pas encore débuté la dialyse (Kadatz *et al.*, 2023; Jay *et al.*, 2017). En résumé, malgré que certaines caractéristiques du donneur soient associées avec une longévité réduite de la greffe, il est généralement bénéfique d'accepter des offres de rein de donneurs ayant ces caractéristiques plutôt que de demeurer en dialyse. Cependant, les travaux de notre équipe ainsi que ceux de collègues de l'Ouest canadien montrent que les candidats et les receveurs de greffe rénale veulent être informés des enjeux qui pourraient affecter la longévité du greffon (Khalili *et al.*, 2022; Rosaasen *et al.*, 2022). En particulier, une majorité de candidats âgés indiquent qu'une longévité de greffe de 3-5 ans était acceptable pour prendre le risque de recevoir une greffe de DCE ou de donneurs à KDPI élevé (Rosaasen *et al.*, 2022). Il apparait donc crucial, dans le contexte de la pénurie d'organe pour la greffe et de l'augmentation du nombre et de l'âge des patients qui attendent une greffe rénale (Schold *et al.*, 2019; Canadian Organ Replacement Register, 2023), de bien informer les candidats à la greffe de la longévité attendue de l'organe et de leur survie globale selon la qualité de l'offre qu'ils recevront. Dans les sections suivantes, je présenterai différents modèles permettant d'estimer la survie afin de bien informer les patients.

2.3 Modèles de prédiction

En analyse de survie, il existe plusieurs façons de modéliser des données de survie et plusieurs formes de prédictions. Le modèle de risques proportionnels de Cox (Cox, 1972) est un exemple de modèle qui calcule une valeur ponctuelle (un risque) pour chaque candidat d'une cohorte. Il s'agit donc d'un modèle discriminant et dont la performance peut être évaluée grâce à l'index de concordance. Un modèle de Cox ajusté permet de déterminer quelles variables ont le plus d'impact sur le risque de subir un événement, et ce, en tenant en compte des données censurées. Une donnée est dite censurée si un candidat est perdu de vue durant l'étude. On sait donc qu'il n'a pas subi d'événement jusqu'à son dernier temps de suivi. Cependant, le modèle de Cox est contraignant puisqu'il fonctionne en posant l'hypothèse que les risques de deux populations sont toujours proportionnels entre eux. Il suppose aussi qu'il existe un risque de base de subir un événement qui est inconnu, mais qui peux être estimé grâce à méthodes comme l'estimateur de Nelson-Aalen. Une valeur de risque n'est

malheureusement pas très informative pour un candidat qui souhaite décider d'accepter ou de refuser une offre de rein. D'autres modèles, tels que celui de Kaplan-Meier (KM) (Kaplan et Meier, 1958), sont utilisés pour modéliser statistiquement la survie d'une cohorte entière. Une courbe de KM présente la probabilité d'événement pour tous les points dans le temps $t \geq 0$ en considérant le nombre de candidats à risque pour chaque moment où l'un d'entre eux subit un événement. Ce type de représentation peut être très informatif dans un contexte où l'on veut comparer la survie de plusieurs populations, mais n'est pas adapté pour offrir des courbes de survie individualisées. Les types de modèles qui nous intéressent dans ce mémoire sont ceux qui produisent des courbes de survie individualisées. Ce type de prédiction peut être obtenu grâce à des modèles tels que « Accelerated Failure Time Model » (Kalbfleisch et Prentice, 2002), Cox avec une extension estimant le risque de base, la forêt de survie aléatoire avec l'extension KM (Ishwaran *et al.*, 2008) et la régression logistique multi-tâche (Yu *et al.*, 2011). Cependant, la majorité de ces méthodes ne présentent pas un niveau de calibration satisfaisant pour leur utilisation clinique.

2.4 Calibration d'un modèle de prédiction

Un modèle de prédiction de la survie ayant une bonne discrimination n'implique pas nécessairement des prédictions fiables. Par exemple, le C-index de prédictions entièrement sous-estimées peut sembler satisfaisant à première vue. En effet, les algorithmes standards de prédiction de la survie peuvent être sur-entraînés, produisant ainsi une forte discrimination, mais une généralisation médiocre, d'où l'importance de la validation externe. Généralement, les modèles plus complexes présentent un plus haut risque de sur-entraînement (Van Calster *et al.*, 2019). Pour démontrer la validité des prédictions, en plus de la discrimination, un modèle doit être calibré. Un modèle est dit calibré si les prédictions s'accordent avec le nombre observé d'événements (Haider *et al.*, 2020; Van Calster *et al.*, 2019, 2016). Plus concrètement, si pour des patients ayant un risque prédit de R%, en moyenne R sur 100 d'entre eux devraient en effet vivre l'événement d'intérêt (Van Calster *et al.*, 2016). Des revues systématiques ont prouvé que la calibration est grandement négligée, malgré qu'elle soit particulièrement importante lorsque les prédictions supportent une décision médicale, même si le modèle présente une moins bonne discrimination (Van Calster *et al.*, 2019).

Il existe plusieurs niveaux de calibration, chacun ayant différentes utilités. Notons la calibration d'ensemble, la calibration faible, modérée et forte (Van Calster *et al.*, 2019, 2016). La calibration d'ensemble demande simplement que le taux observé d'événement corresponde avec le risque prédit moyen. Pour ce faire, une régression logistique est appliquée en fixant la pente à 1, puis un test de vraisemblance à 1 degré de liberté permet vérifier si l'intersection

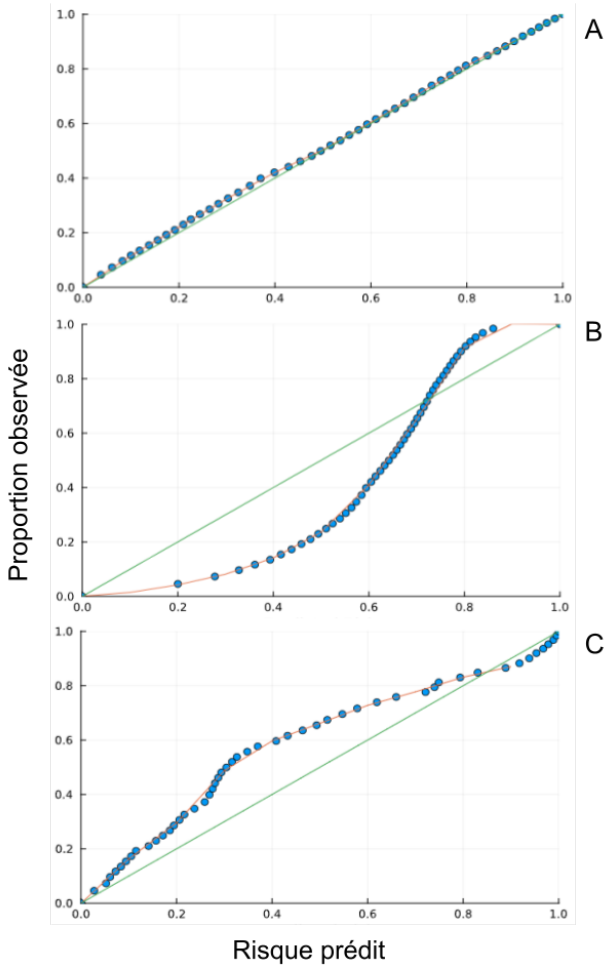


FIGURE 2.1 Exemple de graphiques de calibration modérée pour trois scénarios différent, lorsque les prédictions sont calibrées (A), lorsque les estimés de risque sont trop extrêmes (B) et lorsque les risques sont sous-estimés dans l'ensemble (C).

de la régression est égale à 0. Dans le cas échéant, les prédictions seraient donc calibrées dans leur ensemble. Cependant, ce niveau de calibration est souvent insuffisant pour démontrer la validité des prédictions (Van Calster *et al.*, 2016, 2019).

La calibration faible veut dire qu'en moyenne, le modèle ne sur- ou sous-estime pas les risques et qu'il ne donne pas de prédictions trop extrêmes ou trop modestes (Van Calster *et al.*, 2016, 2019). On le fait en appliquant une régression aux proportions de risque observés par rapport aux risques prédits, la courbe résultante est une courbe de calibration dont la pente devrait être 1 et l'intersection de 0 si les prédictions sont belles et bien calibrées. On peut tester la déviation par un test de vraisemblance de degré 2 (Van Calster *et al.*, 2016). Cette méthode manque de flexibilité puisque la courbe de survie est résumée avec seulement deux paramètres, mais elle est tout de même pratique pour la validation externe et avec des petits ensembles de données (Van Calster *et al.*, 2016).

La calibration modérée est généralement celle à laquelle on se réfère le plus souvent quand on parle de calibration. Si un modèle a une bonne calibration modérée, on s'attend à ce que parmi les patients avec un même risque prédit, le taux d'événement observé soit égal au risque prédit. On évalue la calibration modérée grâce à une courbe de calibration ou en catégorisant les risques prédits, préférablement en rapportant l'intervalle de confiance (Van Calster *et al.*, 2016). Il existe plusieurs façons d'évaluer la calibration modérée d'un modèle de prédiction de la survie en fonction du type de prédiction (probabilité, risque, courbe de survie, etc.). La figure 2.1 présente trois exemples de graphiques de calibrations pour des modèles fictifs qui prédisent des risques d'événement. L'exemple A présente un modèle calibré modérément puisque le risque prédit s'accorde avec la proportion d'événement observé sur l'ensemble des prédictions. Pour l'exemple B, le modèle surestime le risque d'évènement lorsque celui-ci est faible mais sous-estime le risque lorsqu'il est élevé. L'exemple C montre une sous-estimation des risques par rapport aux observations à tous les niveaux de risque d'évènement. Dans le cas des modèles B et C, un test statistique de Hosmer-Lemeshow permettrait de démontrer statistiquement qu'ils ne sont pas modérément calibrés. Des études ont démontré que la calibration modérée augmente l'utilisation clinique(Haider *et al.*, 2020).

Finalement, la calibration forte est présente lorsque les risques prédits correspondent au taux d'événement observé pour chaque combinaison de valeurs de variables prédictrices. On utilise alors la même méthode que pour la calibration modérée, mais les "bins" se limitent à chaque valeur possible de nos variables prédictrices (Van Calster *et al.*, 2016).

CHAPITRE 3 CADRE THÉORIQUE

Ce chapitre introduit les concepts qui seront utilisés dans les chapitres suivants. D'abord, l'analyse de survie, une branche de la statistique qui s'intéresse au temps avant un événement, sera introduite. Ensuite, le test d'adéquation de Kolmogorov-Smirnov sera décrit car la statistique du test sera utilisée pour mesurer la différence de survie entre deux cohortes. Finalement, les arbres de régression sont présentés. Ceux-ci seront employés pour stratifier la cohorte en sous cohortes homogènes.

3.1 Analyse de survie

L'analyse de survie est une branche de la statistique utilisée en épidémiologie dont la variable d'intérêt est le temps avant l'occurrence d'un événement. L'événement peut être le décès du patient, la défaillance d'une machinerie ou même une catastrophe naturelle. Dans tous ces cas, le temps avant un événement est nommé la survie et elle peut être visualisée par une courbe de survie ou fonction de survie qui présente la probabilité d'événement en fonction du temps. Il est possible de modéliser la survie de façon à évaluer l'impact de variables explicatives sur la survie d'un groupe cible. En santé, l'analyse de survie sert généralement à évaluer l'incidence une maladie dans une population ou la survie de personnes malades. Les variables explicatives peuvent compter des caractéristiques démographiques d'une population cible comme sa taille, sa composition son taux de mortalité par exemple. Elles peuvent aussi compter des facteurs de santé des individus qui compose la population comme la prévalence d'une maladie, l'âge, l'exposition à des produits chimiques, etc.

3.1.1 Données de survie

En santé, une population étudiée se nomme une cohorte et elle est composée d'un échantillon d'individus sélectionnées pour représenter une plus grande population. Par exemple, si l'on souhaite étudier le taux d'incidence d'intoxication au plomb dans une région métropolitaine durant une période de 10 ans, il est presque impossible de faire une étude exhaustive de la population cible. C'est particulièrement le cas si l'étude se prolonge sur une longue période de temps. Un échantillon d'individus est donc sélectionnée pour représenter la population de la ville cible sous forme de cohorte. Ces individus sont étudiés à intervalles réguliers afin de dresser un portrait temporel de la population. Durant la période de l'étude, si l'événement cible se produit chez des membres de la cohorte, comme des cas d'intoxication au plomb, le

moment où s'est produit l'événement est noté et servira à produire la variable d'intérêt de l'analyse. Le temps avant l'événement peut être influencé par plusieurs variables, certaines sont catégorielles (présence ou non d'un virus), et d'autres sont continues (âge) et même ordinales (niveau de douleur de 1 à 10). Cependant, il est fréquent que certains participants de cohortes de recherche soient perdus de vue pour diverses raisons. On considère ces données comme étant censurées à droite, c'est-à-dire qu'on sait jusqu'en date du dernier suivi que le participant n'a pas subi l'évènement d'intérêt, mais qu'après ce moment, l'état du participant est inconnu. C'est pourquoi notre variable d'intérêt est représentée avec un temps d'événement (ou de censure) t ainsi qu'une valeur binaire (0 ou 1) s indiquant si le participant est censuré (0) au moment t ou si ce temps correspond à un événement. Sous ce contexte temporel, il est important de spécifier le temps de référence à partir duquel commence le suivi. Dans notre travail, le début du suivi est la date de transplantation rénale.

3.1.2 Kaplan-Meier

La méthode de Kaplan-Meier (KM, Kaplan et Meier, 1958) permet d'obtenir une estimation non paramétrique de la fonction de survie des patients de la cohorte. C'est une estimation marginale qui ne prends pas en compte les variables explicatives. Pour évaluer la survie d'une population, l'estimateur de KM permet de visualiser sous forme de courbe de probabilité d'événement en fonction du temps la survie de la cohorte cible, tout en tenant compte des données censurées. La survie estimée au temps t, soit $\hat{S}(t)$ est donnée par l'estimateur de KM :

$$\hat{S}(t) = \prod_{\{i:t_i \leq t\}} \left(1 - \frac{d_i}{n_i}\right); \tag{3.1}$$

où t_i est un temps où au moins un événement s'est produit, d_i est le nombre d'événements qui se sont produits au temps t_i et n_i est le nombre de sujets à risque tout juste avant t_i.

3.2 Test d'adéquation de Kolmodrov-Smirnov

Le test d'hypothèse de Kolmodrov-Smirnov, initialement proposé par Kolmogorov (1933), puis repris par Smirnov (1939), est un test statistique permettant d'évaluer l'ajustement de deux échantillons en comparant les fonctions de répartitions empiriques des deux échantillons et évalue la probabilité que les échantillons proviennent de la même distribution. Soit $\boldsymbol{x} = (x_1, \ldots, x_n)$ et $\boldsymbol{y} = (y_1, \ldots, y_m)$ deux échantillons aléatoires de tailles $n > 0$ et $m > 0$ indépendants et identiquement distribués respectivement selon les fonctions de répartition F

et G. Les hypothèses du test sont les suivantes :

$\mathcal{H}_0$: $F = G$, *i.e.* les fonctions de répartition F et G sont les mêmes ;

$\mathcal{H}_1$: $F \neq G$, *i.e.* les fonctions de répartition F et G sont différentes.

Soit $\hat{F}_n$ et $\hat{G}_m$ les fonctions de répartition empiriques estimées à partir des deux échantillons et soit $D_{n,m}$ la distance maximale entre ces deux fonctions de répartition empiriques

$$D_{n,m} = \sup_x \left| \hat{F}_n(x) - \hat{G}_m(x) \right|. \tag{3.2}$$

Sous l'hypothèse nulle, la statistique du test

$$K_{n,m} = \sqrt{\frac{mn}{m+n}}\, D_{n,m}$$

est distribuée selon la distribution de Kolmogrov (Smirnov, 1948) dont les valeurs sont tabulées. Alors si la statistique observée du test dépasse la valeur critique de la distribution de Kolmogorov correspondant à un seuil fixé, l'hypothèse nulle d'égalité des distributions doit être rejetée.

3.3 Arbres de décision

En apprentissage automatique, l'arbre de décision est un modèle de régression ou de classification cherchant à subdiviser un ensemble de données en sous-groupes homogènes grâce à un algorithme de partitionnement récursif. Un tel modèle est composé d'arêtes reliant des noeuds par relation de parenté. Chaque noeud est parent à deux enfants (dans un arbre binaire), à l'exception de ceux qui sont au bout de l'arbre, nommées les feuilles. Commençant par la racine, ou l'ensemble de données de départ, la boucle récursive consiste à évaluer la meilleure séparation des données parmi toutes les variables disponibles à chaque niveau de l'arbre de façon à optimiser un critère de partitionnement. Les critères les plus souvent utilisés pour des arbres de classification sont l'index de Gini et le gain d'information. C'est la somme des erreurs au carré (SSE) qui est le critère préférable pour des arbres de régression.

3.3.1 L'index de Gini

L'index de Gini mesure la probabilité de faussement classifier un élément sélectionné au hasard dans une feuille. Il mesure la pureté des feuilles, ou l'homogénéité. La formule de

l'index de Gini :

$$Gini = 1 - \sum_{i=1}^{k} P(i)^2;$$

où k représente le nombre de classes de la variable d'intérêt et $P(i)$ représente la proportion de la classe i parmi l'ensemble des classes. L'index de Gini varie entre 0 (homogène) et 0.5 (hétérogène). Donc, si une feuille ne contient des éléments que d'une seule classe parmi deux classes (disons vrai et faux), l'index de Gini est égal à 0. Au contraire, si pour deux classes, l'échantillon compte autant de *Vrai* que de *Faux*, l'index est de 0.5.

Avec l'index de Gini comme critère de partitionnement de branches dans un arbre de décision, on effectue à chaque niveau de l'arbre une séparation qui minimise l'index de Gini. Le processus est répété pour chaque sous-groupes (enfants) produits par la séparation.

3.3.2 Gain d'information

Le critère de gain d'information vise à minimiser l'entropie dans les feuilles. L'entropie est calculée par

$$Entropie = -\sum_{i=1}^{k} p_i \log_2(p_i)$$

où p_i est la probabilité de sélectionner aléatoirement la classe i pour une feuille dont la variable d'intérêt contient k classes. L'entropie mesure la pureté (homogénéité) d'un ensemble et varie de 0 (pureté) à 1 (impureté). Tout comme l'index de Gini, dans un contexte où la variable d'intérêt contient deux classes, l'entropie d'un échantillon qui ne contient que des éléments d'une même classe est de 0, tandis que si on y compte autant d'éléments de chaque classe, l'entropie est de 1.

Le gain d'information est ensuite calculé avec la différence d'entropie du parent et l'entropie moyen des enfants, soit

$$\text{Gain d'information} = Entropie_{Parent} - \frac{1}{n}\sum_{i=1}^{n} Entropie_{Enfants}(i);$$

où n est le nombre d'enfants. Pour un arbre de décision, les séparations sont choisies de façon à maximiser le gain d'information à chaque itération.

3.3.3 Somme des erreurs au carré

Dans le cas d'un arbre de décision pour la régression, donc pour une variable dépendante continue, on cherche à minimiser la somme des erreurs au carré suivante :

$$SSE = \sum_{S_k \in S} \sum_{i \in S_k} (y_i - \hat{y}_{S_k})^2;$$

où S_k correspond à l'ensemble des données dans la feuille, S correspond à l'ensemble des feuilles, y_i est la valeur cible de i dans S_k et $\hat{y}_{S_k}$ est la moyenne des valeurs des éléments de la feuille S_k. La SSE est calculé à chaque récursion et l'arbre est approfondie si une stratification de plus implique une diminution de la SSE.

3.3.4 Critère d'arrêt

Un des défis reliés à l'utilisation d'arbres de décision est de définir un critère d'arrêt pour éviter un sur-apprentissage des données. En effet, le sur-apprentissage est inévitable si l'arbre est trop pur. C'est-à-dire que lorsqu'on tend vers un arbre de pureté parfaite, ou que chaque feuille ne contient qu'un seul élément, on perd tout pouvoir prédictif. Souvent, le premier critère d'arrêt est l'absence d'amélioration du gain d'information et de l'index de Gini. Parmi les solutions, on peut aussi définir des hyperparamètres, comme une profondeur maximale et un nombre minimum de données dans les feuilles, qui peuvent être sélectionnés par validation.

CHAPITRE 4 DONNÉES

4.1 Pré-traitement

Les données utilisées pour notre analyse proviennent du registre national américain de transplantation du «United Network for Organ Sharing» (UNOS). Un fichier de données appelé STAR File (Standard Transplant Analysis and Research) nous a été transmis sur demande par UNOS. Puisqu'aux États-Unis, le système de partage d'organes est centralisé, on compte dans ce registre toutes les transplantations d'organe effectuées au pays depuis le 1er octobre 1987 ainsi que toutes les inscriptions en liste d'attente pour une greffe d'organe solide. On compte dans le fichier qui nous a été transmis 1 051 757 transplantations de rein et/ou pancréas et 486 colonnes avant toute opération de pré-traitement. Après examen, 131 colonnes de variables continues, catégorielles ou sous forme de dates ont été jugées pertinentes pour ce projet. Le projet dans son entièreté est rédigé avec le langage de programmation Julia. Certaines étapes du pré-traitement concernent donc la conversion de certaines variables pour faciliter leur manipulation, telles que les dates. Le tableau 4.1 présente ces manipulations en ordre ainsi que le nombre de données restantes à chaque étape.

4.2 Variables indépendantes

Parmi les 131 variables indépendantes décrivant les transplantations du registre, 19 d'entre elles ont été sélectionnées en fonction de l'avis de d'experts comme étant les plus pertinentes dans le contexte de notre étude. Le tableau 4.2 et le tableau 4.3 présentent les variables indépendantes sélectionnées, leurs descriptions et leurs valeurs possibles. Certaines valeurs sont encodées, comme l'historique de fumeur du donneur (ACT, EX, etc.), et d'autres sont représentées telles qu'elles le sont dans l'ensemble de données initiales, comme le diagnostic de la maladie rénale du candidat et l'éthnicité du candidat.

TABLEAU 4.1 Description des étapes de pré-traitement à partir des données brutes.

N°	**Description de l'opération**	**Nombre de données restant**
1.	Nombre initial de données (d'observations).	1051757
2.	Inclure seulement les transplantations de reins.	507116
3.	Convertir les dates (texte) en objet *DateTime*.	507116
4.	Exclure les lignes manquant une ou plusieurs dates utilisées pour calculer les variables dépendantes d'intérêt.	504577
5.	Inclure seulement la première transplantation rénale pour chaque receveur.	459208
6.	Inclure seulement les transplantations effectuées entre 1er janvier 2000 et 1er janvier 2020.	307881
7.	Exclure les transplantations de rein doubles ou séquentielles.	302393
8.	Exclure les transplantations faites à des candidats ayant déjà eu une ou plusieurs autres transplantations d'organe solide dans le passé.	292070
9.	Exclure les transplantations provenant de donneurs après arrêt circulatoire non-contrôlé.	291485
10.	Exclure les transplantations provenant de donneurs vivants.	183928
11.	Exclure les transplantations de candidats âgés de moins de 18 ans ou dont l'âge est manquant.	176037
12.	Exclure les transplantations ABO incompatibles donneur-receveur.	174747
13.	Exclure les transplantations faites conjointement avec d'autres organes solides.	166200
14.	Exclure les transplantations de donneurs âgés de moins de 5 ans ou dont l'âge est manquant.	164423
15.	Exclure les transplantations de donneurs dont le niveau de créatinine, la taille ou le poids est manquant.	164228

TABLEAU 4.1 Description des étapes de pré-traitement à partir des données brutes (suite et fin).

16.	Inclure seulement les transplantations de donneurs dont le taux de créatinine est entre 0.2 et 23 mg/dl.	164034
17.	Inclure seulement les transplantations de donneurs dont la taille est entre 30 et 250 cm.	164034
18.	Inclure seulement les transplantations de donneurs pesant entre 30 et 250 cm.	164034
19.	Convertir les pourcentages en valeurs décimales.	164034
20.	Exclure les transplantations dont le mismatch A, B, DR est manquant, dont le mismatch HLA est manquant ou dont statut décès neurologique versus après arrêt cardio-circulatoire est manquant.	163996
21.	Convertir les variables catégoriques binaires en "P", "N" ou "IND" (si manquant).	163996
22.	Exclure les transplantations manquant une des variables indépendantes d'intérêt.	159140

TABLEAU 4.2 Description des variables indépendantes catégorielles.

Nom	**Description**	**Valeur**
DIAB	Statut diabétique du candidat.	Y : diabétique N : non diabétique IND : indéterminé
DIAG_KI	Diagnostic du candidat. Cause de la maladie rénale chronique.	Maladie autoimmune Diabète Glomérulonéphrite Hypertension vasculaire Maladie rénale polykystique Autre ou inconnu
ETHCAT	Ethnicité du candidat.	Blanc Noir Hispanique Asiatique Autre

TABLEAU 4.2 Description des variables indépendantes catégorielles (suite et fin).

PERIP_VASC	Maladie rénale périphérique.	Y : présente N : absente IND : indéterminée
GENDER	Sexe du candidat.	M : homme F : femme
GENDER_DON	Sexe du donneur.	M : homme F : femme
HCV_REC	Historique de sérologie du candidat contre l'hépatite C.	P : positif N : négatif
TABAC_DON	Historique de fumeur du donneur.	ACT : actif EX : ex-fumeur NO : non fumeur IND : indéterminé
MALIG	Historique de malignité du candidat.	Y : oui N : non IND : indéterminé
MM	Nombre de "mismatchs" ABDR entre le receveur et le donneur.	[1,2,3,4,5,6]
CMV_LOW	Sérologies du cytomégalovirus du receveur et du donneur toutes deux négatives.	1 : oui 0 : non
CMV_MED	Sérologies du cytomégalovirus du receveur et du donneur toutes deux positives ou sérologie du receveur positive.	1 : oui 0 : non
CMV_HIGH	Sérologies du cytomégalovirus positive chez le donneur et négative chez le receveur.	1 : oui 0 : non
EBV_MM	Sérologie du virus d'Epstein-Barr positive chez le donneur et négative chez le receveur.	1 : positif 0 : négatif

TABLEAU 4.3 Description des variables indépendantes continues.

AGE	Âge du candidat en années.	Min. : 18 ans Médiane : 54 ans Max. : 96 ans
TIME_ON- _DIALYSIS	Temps qu'a passé le candidat en dialyse avant la transplantation en années.	Min. : 0 an Médiane : 3.61 ans Max. : 72.36 ans
IMC	Indice de masse corporelle du candidat en kg/m^2.	Min. : 11.24 kg/m^2 Médiane : 27.83 kg/m^2 Max. : 74.22 kg/m^2
PRA_PRE	PRA calculé du candidat avant la transplantation en pourcentage.	Min. : 0.0 % Médiane : 0.0 % Max. : 100.0 %
KDRI_RAO	"Kidney Donor Risk Index" ou KDRI. Un score estimant le risque de perte du greffon en fonction d'informations sur le donneur (sans unités).	Min. : 0.60 Médiane : 1.17 Max. : 3.95

4.3 Variables dépendantes

Pour notre étude, on s'intéresse au temps avant un événement pour des candidats à la greffe. L'événement d'intérêt peut être de nature diverse. Dans notre cas, on s'intéresse au décès du receveur et à la perte de greffon suite à la transplantation. Le tableau 4.4 présente une description de ces variables d'intérêt. Tous les temps sont calculés à partir de la date de la transplantation rénale. D'abord, on s'intéresse à la perte du greffon qui se définit comme le décès du candidat, le retour en dialyse ou la regreffe (la date d'évènement retenue étant celle du premier des 3 évènements qui se produit). On juge que pour informer un candidat à la greffe rénale, le plus pertinent est de prédire le temps pendant lequel il/elle bénéficiera d'une greffe fonctionnelle. C'est pourquoi on choisit comme cible un tel événement composite. À noter que l'on fait la différence entre deux variantes de cette cible. Les informations de l'ensemble de données proviennent de diverses sources et institutions. Certaines informations sont mieux documentées que d'autres. Par exemple, il est possible (quoique peu probable) qu'un candidat retourne en dialyse ou qu'il soit re-transplanté sans que cela soit enregistré dans le fichier STAR, tandis que le décès des receveurs est mis à jour dans la base de données même

plusieurs années après le dernier suivi et revérifié avec les registres nationaux de mortalité. Il se peut qu'entre le dernier suivi d'un candidat et son décès, il subisse une regreffe ou une re-transplantation non-répertorié. Pour tenter d'éviter ce problème, la première cible du tableau 4.4 est censurée à la date du dernier suivi si le premier événement est le décès.

TABLEAU 4.4 Description des variables dépendantes.

N°	**Événement**	**Description**
1.	Echec du greffon.	Décès du candidat, retour en dialyse ou regreffe. Le premier qui se produit. Censuré à la fin de l'étude si aucun événement ne se produit. Si le premier événement qui se produit est le décès, mais que la dernière mise à jour c'est faite avant, on censure à la date de dernière mise à jour et le patient n'est pas considéré comme ayant eu un évènement.
2.	Décès.	Décès du candidat. Censuré à la fin de l'étude si aucun événement ne se produit.

4.4 Gestion des valeurs manquantes

Pour plusieurs variables catégorielles, comme on le voit dans le tableau 4.2, les données manquantes sont classées sous la valeur "IND" pour indéterminé. Les KDRI_RAO manquants sont calculés dans les cas où aucunes des variables nécessaires pour le faire ne manquaient.

CHAPITRE 5 MÉTHODE

5.1 Arbres de décision

La base du modèle proposé est un arbre de décision qui partitionne de façon optimale la cohorte en sous-groupes dont la survie des candidats (ou des greffons) est homogène. L'algorithme 1 décrit le processus récursif de partitionnement du modèle. L'ensemble d'entraînement est un échantillon aléatoire de 70% de la cohorte et 30% de la cohorte est conservé pour la validation et les tests. Le processus d'entraînement débute par la création de la racine qui contient l'ensemble des données d'entraînement. Puis, on itère sur toutes les variables indépendantes et sur toutes les partitions binaires possibles pour identifier la stratification qui maximise la différence de survie entre les deux sous-groupes. Pour des variables catégorielles ou ordinales, la fonction *PartitionsBinaires* retourne une liste de toutes les façons possibles de partitionner les valeurs en deux ensembles. Par exemple, si une variable *Type de Diabète* pouvait avoir les valeurs $\{0, 1, 2\}$, les partitions possibles seraient $\{(0), (1, 2)\}$, $\{(0, 1), (2)\}$ et $\{(0, 2), (1)\}$. Pour des variables continues, la fonction *PartitionsBinaires* retourne tous les partitionnements binaires possibles selon les quartiles de la valeur ou selon la résolution choisie. La résolution fait partie des hyper-paramètres du modèle et elle est choisie par validation. Par exemple, si la résolution est de 0.01, l'algorithme teste les partitionnements à chaque intervalle de taille 0.01 de l'étendue de la variable en question.

5.2 Critère de partitionnement

Le critère de partitionnement consiste en la différence entre les courbes de survie des deux sous-groupes. La différence est calculée par la distance utilisée dans le test de Kolmodrov-Smirnov à deux échantillons exprimée à l'équation (3.2) (Kolmogorov, 1933; Smirnov, 1939, 1948) et l'estimateur de Kaplan-Meier (Kaplan et Meier, 1958). Pour chaque partition possible, les courbes de Kaplan-Meier des deux sous-groupes sont calculées, puis la statistique du test d'adéquation de Kolmodrov-Smirnov permet de quantifier la différence entre la survie des deux sous-groupes. Cela est possible puisque les courbes de survies correspondent au complément de la fonction de répartition représentant la proportion des candidats ayant eu un événement. La partition ayant la statistique du test de Kolmodrov-Smirnov maximale est choisie pour cette variable indépendante. La variable dont le partitionnement rapporte le plus haut score est utilisée pour définir les deux sous-groupes qui seront sujets à la prochaine récursion de l'arbre. Le processus est répété avec chacun de ces sous-groupes jusqu'à

Algorithme 1 : Entraînement d'un arbre de décision.

Données : X sont les variables indépendantes de dimensions (N, n) et $Y = \{T, E\}$ est la variable dépendante de dimensions $(N, 2)$ avec T le temps d'événement $(E = 1)$ ou de censure $(E = 0)$.

Soit $parent = \{X = X, Y = Y, parent = null, enfants = null\}$ la racine de l'arbre contenant l'ensemble de données initial.

fonction EntrainerArbre ($parent$) **début**

- $meilleur\_s \leftarrow 0$; `// Meilleur score`
- $meilleur\_x_G \leftarrow []$; `// Meilleur partition à inférieure de` x.
- $meilleur\_y_G \leftarrow []$; `// Meilleur partition à inférieure de` y.
- $meilleur\_x_D \leftarrow []$; `// Meilleur partition à supérieure de` x.
- $meilleur\_y_D \leftarrow []$; `// Meilleur partition à supérieure de` y.
- $X \leftarrow parent.X$; `// Les variables indépendantes du noeud` $parent$.
- $Y \leftarrow parent.Y$; `// Les variables dépendantes du noeud` $parent$.
- **pour** x_i ***dans*** X **faire**
  - `/* Pour toutes les partitions binaires possibles de` x_i. `*/`
  - **pour** (x_G, x_D, y_G, y_D) ***in*** $PartitionsBinaires(x_i, Y)$ **faire**
    - `/* Évaluer le score du partitionnement */`
    - $score \leftarrow ScoreStrat(x_G, x_D, y_G, y_D)$
    - `/* Le meilleur partitionnement est retenu. */`
    - **si** $score > meilleur\_s$ **alors**
      - $meilleur\_s \leftarrow score$
      - $meilleur\_x_G \leftarrow x_G$
      - $meilleur\_x_D \leftarrow x_D$
    - **fin**
  - **fin**
- **fin**
- `/* Créer les noeuds associés à la meilleure partition inférieure et supérieure */`
- $noeud_G \leftarrow \{X = meilleur\_x_G, Y = meilleur\_y_G, parent = parent, children = null\}$
- $noeud_D \leftarrow \{X = meilleur\_x_D, Y = meilleur\_y_D, parent = parent, children = null\}$ `/* Arrêter si le critère d'arrêt est satisfait */`
- **si** $evaluation(parent, noeud_G, noeud_D)$ **alors**
  - **retourner** $parent$
- **fin**
- `/* Continuer l'arbre récursivement */`
- $parent.enfants \leftarrow [EntrainerArbre(noeud_G), EntrainerArbre(noeud_D)]$
- **retourner** $parent$

fin

ce qu'un des critères d'arrêt soit satisfait. L'algorithme 1 illustre la procédure récursive de partitionnement utilisée afin de diviser la cohorte en sous-groupes homogènes concernant leur survie.

5.3 Prédictions

Dans le but d'offrir aux candidats et aux néphrologues des prédictions de la survie du greffon ou du candidat qui sont faciles à interpréter, notre modèle offre des prédictions sous forme de courbe de survie comme à la figure 5.3. À partir d'un ensemble de données, l'algorithme de Kaplan-Meier produit une courbe de survie qui présente la probabilité d'avoir un événement en fonction du temps après un moment de référence, dans notre cas le moment de la greffe rénale. Cette représentation instinctive de la survie permet d'en tirer toutes sortes d'information pertinentes comme le temps de survie moyen ou médian, la probabilité de survie à un temps t et même le risque instantané à un temps t à partir de la pente en ce point. Pour effectuer des prédictions à partir d'un arbre de survie entraîné, un nouveau candidat de l'ensemble de test est passe à travers l'arbre à partir de la racine. À chaque branchement, on observe à quel enfant du noeud il est assigné, et ce, jusqu'à ce qu'il atteigne une feuille. La courbe de Kaplan-Meier du groupe qui correspond à cette feuille est utilisée comme prédiction individualisée de sa survie.

5.4 Forêt aléatoire

Un modèle en ensemble consiste à entraîner plus d'un modèle de prédiction et d'effectuer un consensus de leurs prédictions pour augmenter la généralisation et la précision(Ali *et al.*, 2012). Un ensemble est particulièrement utile lorsqu'on a accès à un grand nombre de données. Une forêt aléatoire est un modèle d'ensemble que l'on obtient en combinant des modèles d'arbre de décision. Cela permet de compenser pour la haute variabilité d'un modèle d'arbre unique. Une forêt est obtenue en entraînant plusieurs arbres en excluant aléatoirement un échantillon différent à chaque arbre (communément 1/5). Les résultats des arbres sont mis en commun par une moyenne ou par vote dans le cas de classification. La mise en commun pour notre modèle d'arbre de survie consiste à concaténer les sous-groupes des feuilles de chaque arbre de l'ensemble en ne conservant que les lignes uniques. Puis, encore une fois, la courbe de Kaplan-Meier de ce nouveau groupe sert de prédiction individualisée de la survie. Normalement, avec un seul arbre, le nombre de courbes uniques possibles se limite au nombre de feuille de l'arbre. La méthode d'ensemble permet d'éviter cette contrainte. Une représentation du processus de prédiction avec un modèle de forêt est présenté à la figure 5.2.

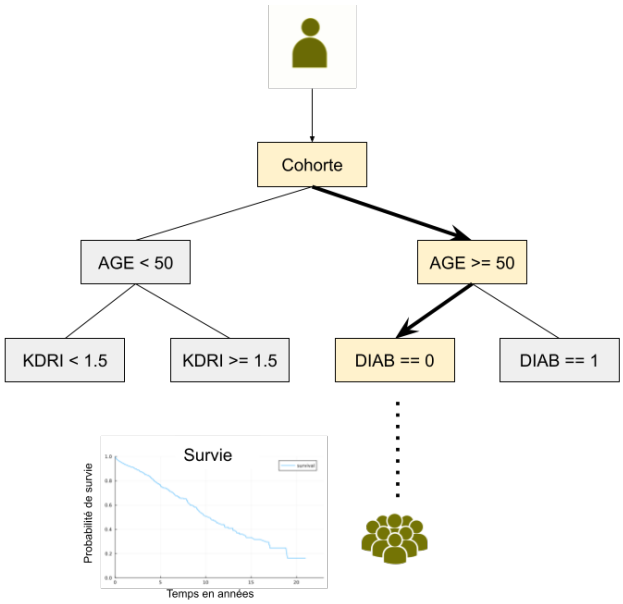


FIGURE 5.1 Représentation simplifiée du processus de prédiction avec un arbre de survie unique.

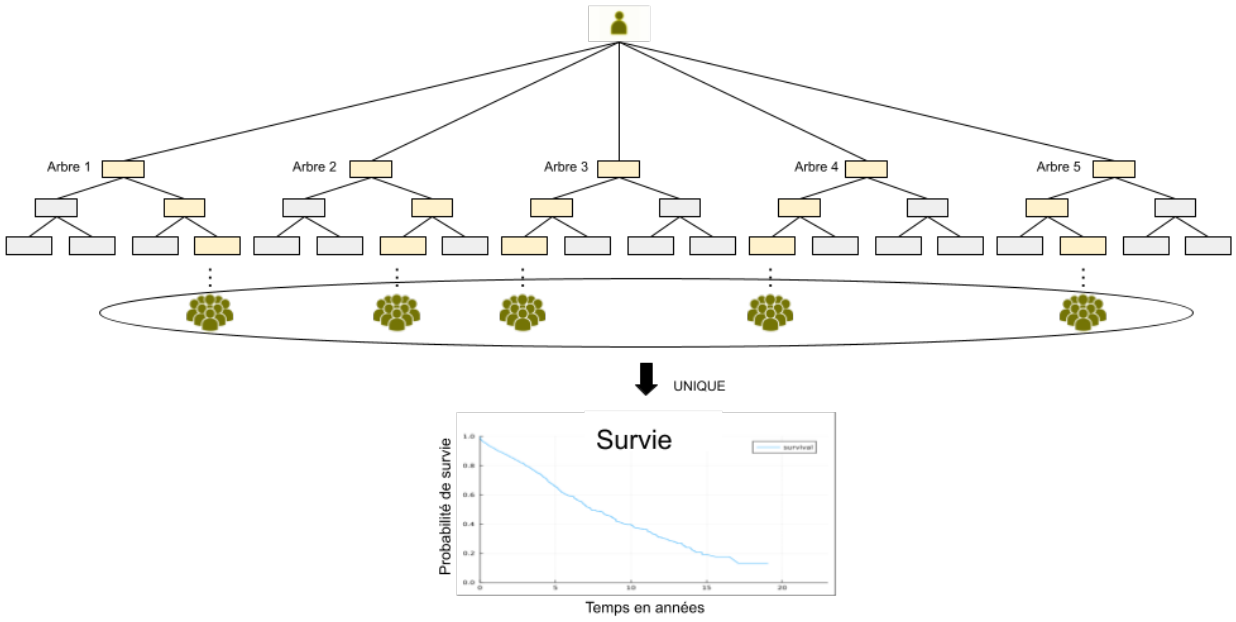


FIGURE 5.2 Représentation simplifiée du processus de prédiction une forêt à cinq arbres.

5.5 Hyper-paramètres

Une série de paramètres en lien avec l'architecture et le format des données (des hyper-paramètres) doivent être ajustés à l'aide d'un ensemble de validation. Certains d'entre eux peuvent être définis arbitrairement si le contexte le permet. Parmi ceux du modèle de forêt, on compte le nombre d'arbres dans la forêt, la profondeur maximale des arbres, le nombre minimal d'événements dans une feuille pour poursuivre le partitionnement et la résolution des stratification des variables continues. La méthode de recherche en grille permet de tester toutes les combinaisons d'hyper-paramètres possibles parmi un choix de valeurs possibles pour chacun d'eux. Cependant, cela demande beaucoup de temps d'exécution. En raison du nombre important d'hyperparamètres et de la puissance de calcul à notre disposition, nous avons choisi des valeurs raisonnables pour les hyperparamètres (voir tableau 5.1) spécifiées avec l'expertise des praticiens.

TABLEAU 5.1 Valeurs sélectionnées pour les différents hyper-paramètres.

Hyper-paramètres	**Valeurs**
Nombre d'arbres.	1 (arbre) et 5 (forêt)
Profondeur maximale.	50
Nombre minimal d'événement dans une feuille.	500
Résolution des stratifications.	0.05

5.6 Calibration

Le modèle est validé par la calibration D (Haider *et al.*, 2020), un algorithme qui assure de pouvoir faire confiance aux courbes prédites. Il consiste à mesurer la probabilité de survie prédite au temps d'événement de chaque donnée de l'ensemble de test et de les regrouper en paniers en fonction de leur probabilité prédite. Si la variable d'intérêt était le décès du candidat par exemple, s'il est décédé après 12 mois et que sa probabilité de survie prédite était de 35%, il serait assigné au panier des 30% à 40%. Le même processus est répété pour tous les candidats de l'ensemble de test jusqu'à obtenir dix paniers contenant une certaine proportion de l'ensemble de test. Si le suivi du candidat est censuré car il ne présente pas l'évènement d'intérêt pendant la période d'observation de l'étude, sa probabilité d'événement est distribuée entre les paniers de probabilité inférieure (40% et moins). La figure 5.4 résume ce processus et présente un exemple fictif. Ensuite, si on observe 10% de l'ensemble de test dans chaque panier, on peut confirmer que le modèle est calibré D dans cette cohorte et que l'on peut faire confiance à ses prédictions. Cette méthode est celle que nous avons utilisée pour valider le modèle d'arbre et de forêt de survie dans ce mémoire.

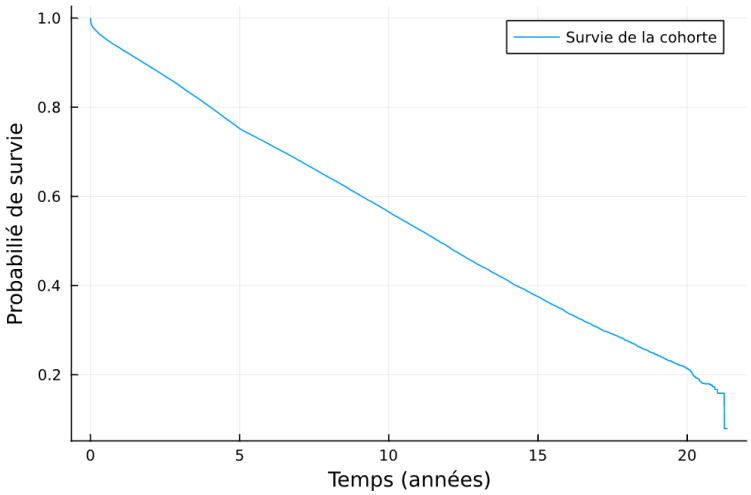


FIGURE 5.3 Courbe de Kaplan-Meier décrivant l'utilité du greffon censuré à la dernière date de mise à jour si le premier événement est le décès (cible 1).

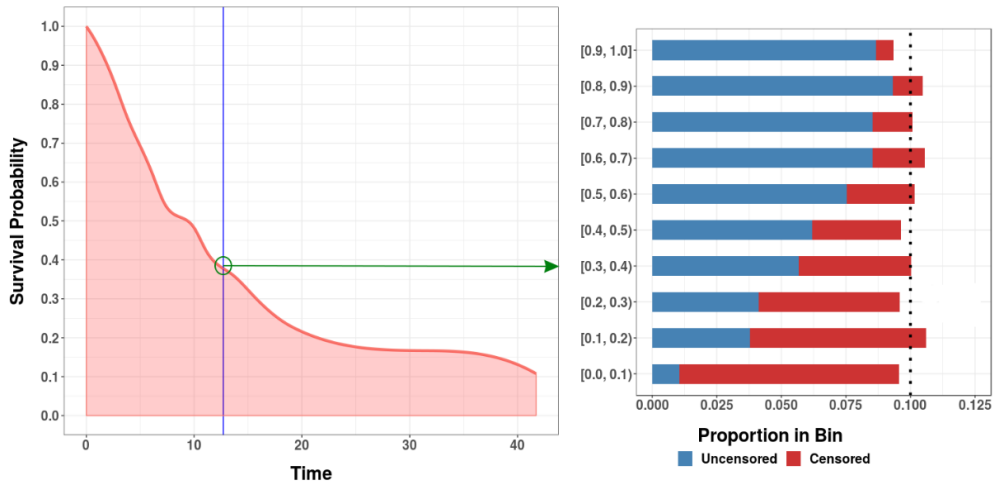


FIGURE 5.4 Visualisation d'un exemple fictif de calibration D tiré de Haider *et al.* (2020).

CHAPITRE 6 RÉSULTATS

Cette section présente les arbres et les forêts produits par l'algorithme récursif présenté au chapitre précédent, utilisant un partitionnement utilisant la distance de Kolmodrov-Smirnof entre les courbes de survie des deux sous groupe comme critère à maximiser pour le partitionnement. L'approche proposée a été utilisée pour les deux cibles, soit l'échec du greffon et la survie du receveur. La validité des prédictions de ces modèles est démontrée notamment par la calibration D.

On s'intéresse particulièrement aux premiers partitionnements sélectionnées par l'algorithme puisque ces variables et ces valeurs sont celles qui maximisent la statistique de Kolmodrov-Smirnof équivalent à la distance entre les courbes de survie des sous-groupes. On peut les interpréter comme étant celles qui sont le plus déterminantes de la survie.

6.1 Échec du greffon

6.1.1 Arbre obtenu avec la cohorte d'entraînement

La figure 6.1 présente l'arbre obtenu pour prédire la survie du greffon [1]. Les variables sélectionnées par l'algorithme pour effectuer les premières partitions à partir de la cohorte sont :

Cohorte

- ↳ Âge du candidat $\geq$ 72 ans
  - ↳ **KDRI RAO $\geq$ 2.0883**
  - ↳ KDRI RAO $<$ 2.0883
    - ↳ Temps en dialyse $\geq$ 0.4219 ans ...
    - ↳ **Temps en dialyse $<$ 0.4219 ans**
- ↳ Âge du candidat $<$ 72 ans
  - ↳ Âge du candidat $\geq$ 69 ans
    - ↳ Temps en dialyse $\geq$ 0.5293 ans ...
    - ↳ **Temps en dialyse $<$ 0.5293 ans**
  - ↳ Âge du candidat $<$ 69 ans
    - ↳ Âge du candidat $\geq$ 67 ans ...

1. Dans le mémoire en format numérique, il est possible d'agrandir l'arbre autant que nécessaire.

↳ Âge du candidat < 67 ans ...

Les points de suspension ("...") indiquent la continuité de l'arbre. Les prochaines variables les plus significatives sont encore une fois le temps passé par le candidat en dialyse, l'âge du candidat, le statut diabétique du candidat, le diagnostic du candidat et l'éthnicité du candidat. Plus loin dans l'arbre, on retrouve la maladie vasculaire périphérique et le match des sérologies du cytomégalovirus donneur et candidat. À noter que le noeud contenant les candidats âgés de 72 ans et plus avec une offre de KDRI supérieure à 2.0883 est terminal et ne progresse pas plus loin. Ce dernier et les autres noeuds terminaux sont identifiées en gras. L'arbre généré compte 167 feuilles et 40 niveaux de profondeur. À titre d'exemple, la figure 6.2 illustre les courbes de survie de chacun des deux sous-groupes obtenus après le premier partitionnement. On constate que la variable «âge» influence grandement la survie du greffon dans cette cohorte.

6.1.2 Forêt obtenue avec la cohorte d'entraînement

La forêt générée pour prédire l'échec du greffon compte cinq arbres dont le nombre de feuilles varie entre 129 et 138 feuilles et dont le niveau de profondeur varie de 29 à 37. Les figures des arbres qui composent la forêt sont présentés en annexe A. Le premier partitionnement de ces arbres se fait à l'âge du candidat supérieur et inférieur à 71 ou 72 ans, similairement à l'arbre seul. On observe une tendance de croissance similaire entre les arbres. Ainsi, les branches ne progressent pas aussi loin pour les candidats plus âgés que pour ceux qui sont plus jeunes, ce qui fait que les partie inférieures des arbres s'étendent davantage. On remarque donc que les facteurs déterminants de la survie du greffon se diversifient davantage pour les candidats plus jeunes. Le temps passé en dialyse demeure encore une fois un facteur important pour prédire la survie du greffon.

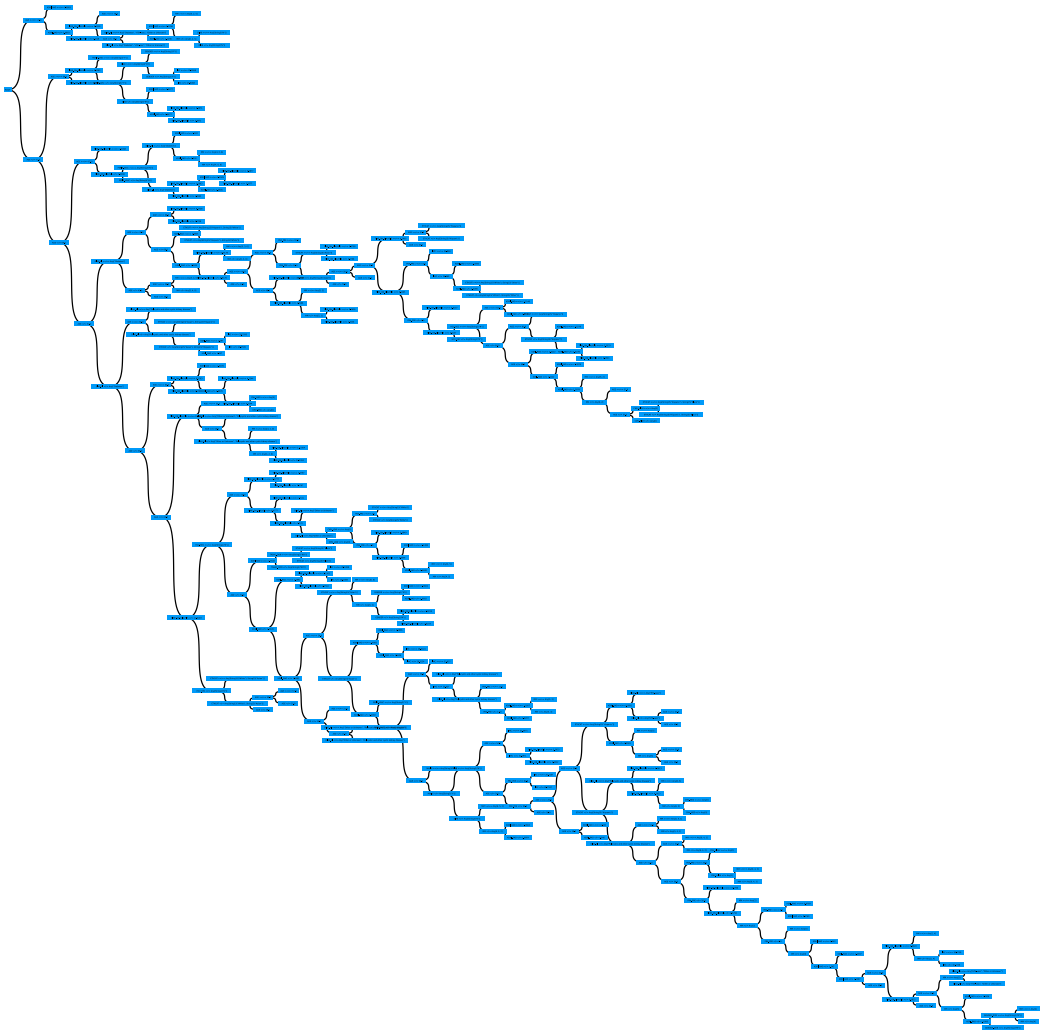


FIGURE 6.1 Structure de l'arbre unique entraîné pour prédire la fonction de la greffe.

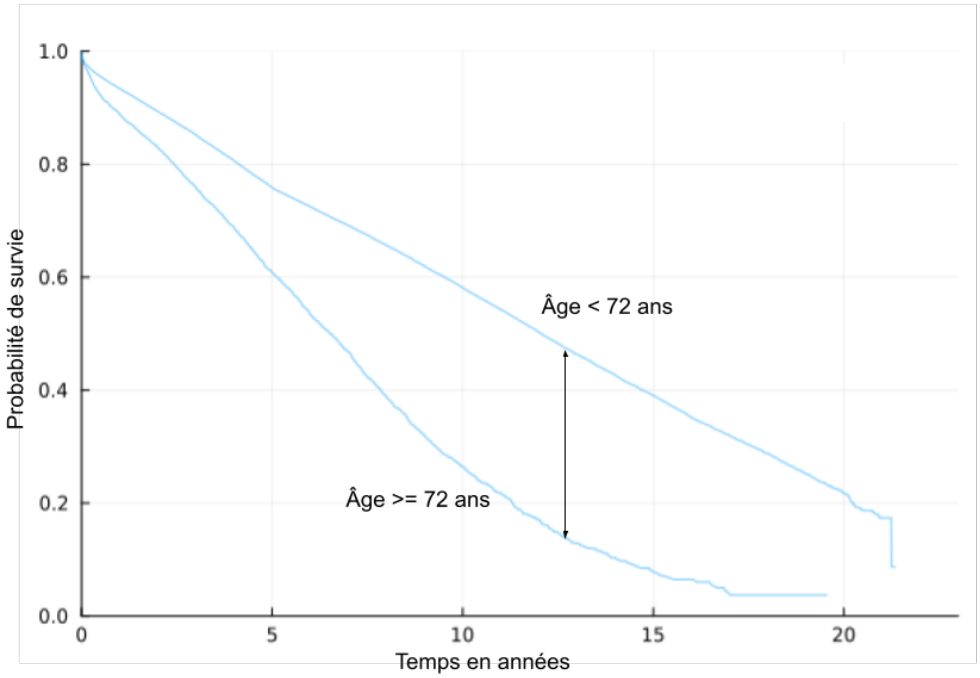


FIGURE 6.2 Courbes de survie du greffon estimées par l'estimateur de Kaplan-Meier pour les deux sous-groupes obtenus, à savoir les patients âgés de 72 ans et plus et les plus jeunes. Le partitionnement à l'âge de 72 ans produit la plus grande différence de survie selon la distance de Kolmogorov-Smirnov entre les courbes, indiquée par la flèche.

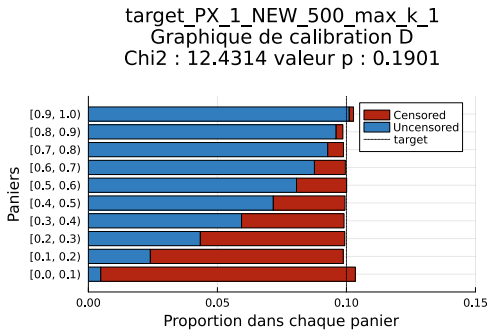


FIGURE 6.3 Graphique de calibration D d'un arbre unique en fonction de l'échec de la greffe.

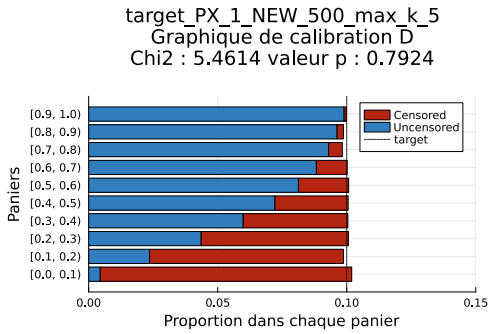


FIGURE 6.4 Graphique de calibration d'une forêt à cinq arbres en fonction de l'échec de la greffe.

6.2 Décès du receveur

La figure 6.7 présente l'arbre obtenu pour le décès du receveur (qu'on appelle aussi candidat). Les premières variables sélectionnées par l'algorithme pour effectuer les premières partitions à partir de la cohorte sont :

Cohorte

- ↳ Âge du candidat $\geq$ 72 ans
  - ↳ **Âge $\geq$ 78 ans**
  - ↳ Âge du candidat $<$ 78 ans
    - ↳ KDRI RAO $\geq$ 2.0754 ...
    - ↳ KDRI RAO $<$ 2.0754 ...
- ↳ Âge du candidat $<$ 72 ans
  - ↳ Âge du candidat $\geq$ 29 ans
    - ↳ Âge du candidat $\geq$ 34 ans ...
    - ↳ Âge du candidat $<$ 34 ans ...
  - ↳ Âge du candidat $<$ 29 ans
    - ↳ Diagnostic "Maladie autoimmune" ou "Diabète" ...
    - ↳ Diagnostic "Glomérulonéphrite", "Hypertension", "Maladie rénale polykystique" ou "Autre ou inconnu" ...

Les prochaines variables les plus présentes sont encore l'âge du candidat, le statut diabétique et l'éthnicité du candidat. Plus loin dans l'arbre, on retrouve fréquemment le temps passé en dialyse et quelques fois la sérologie d'hépatite C, la maladie vasculaire périphérique et le match de sérologies du cytomégalovirus donneur-receveur. À l'instar des arbres obtenus pour la survie du greffon, le noeud contenant les candidats âgés de 72 ans et plus est terminal et ne progresse pas plus loin. L'arbre généré compte 171 feuilles et 32 niveaux de profondeur. La figure 6.8 présente les courbes de survie des sous-groupes obtenus après le premier partitionnement.

6.2.1 Forêt obtenue avec la cohorte d'entraînement

La forêt générée pour prédire le décès du receveur compte cinq arbres dont le nombre de feuilles varie entre 130 et 141 feuilles et dont le niveau de profondeur varie de 23 à 32. Les figures des arbres sont présentées en annexe. Contrairement à la survie du greffon, on observe dans les arbres de la forêt pour prédire la survie du candidat que les premiers partitionnements se font autour de l'âge de 29 ans dans quatre des cinq arbres. Dans ceux-ci, le second

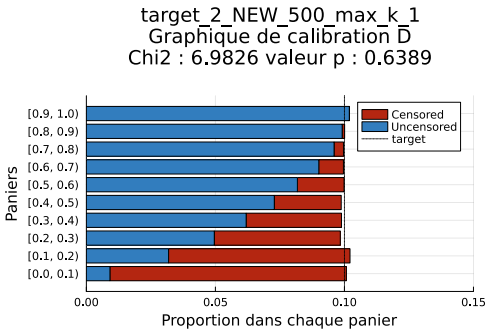


FIGURE 6.5 Graphique de calibration D d'un arbre unique en fonction du décès du receveur.

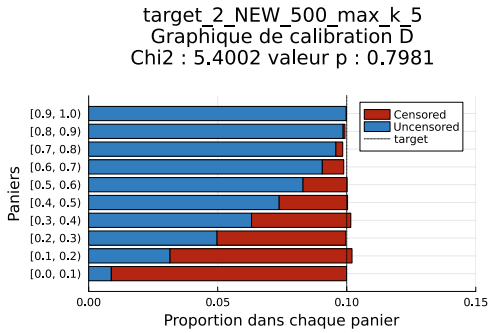


FIGURE 6.6 Graphique de calibration d'une forêt à cinq arbres en fonction du décès du receveur.

partitionnement se fait à 72 ans et la tendance de partitionnement à des âges graduellement plus petits reprend. Pour ces candidats âgés entre 29 et 72 ans, on voit que l'âge est toujours un facteur particulièrement important. On remarque que les branches produites à partir des candidats âgés de moins de 29 ans ne considèrent plus l'âge passé ce noeud. Ce sont plutôt le diagnostic (cause de la maladie rénale chronique), le KDRI et l'ethnicité qui sont les facteurs déterminants de la survie de ces jeunes candidats. Pour ce qui est de cet arbre unique aux quatre autres, immédiatement après le premier partitionnement à 72 ans, on retrouve une branche des candidats âgés de plus et de moins de 33 ans. Encore une fois, la survie des candidats plus jeunes est majoritairement déterminée par des facteurs comme le diagnostic, le KDRI de l'offre, l'éthnicité, mais également l'indice de masse corporelle du candidat.

6.3 Analyse des arbres et des forêts obtenus

On observe aux figures 6.2 et 6.8 que l'algorithme performe bel et bien tel qu'attendu. On voit que le premier partitionnement se fait par rapport à l'âge de 72 ans pour nos deux cibles. La branche de candidats âgés de plus de 72 ans ne progresse que de quelques niveaux après la première itération dans les deux arbres, contrairement aux candidats plus jeunes. Cela est attendu puisque non seulement peu de transplantations sont faites chez des patients aussi âgés, mais également parce que la cause première de perte de greffon pour ceux-ci est le décès. C'est donc cohérent de voir ce partitionnement dans les deux arbres.

Il est intéressant d'observer que l'âge est une variable abondamment présente dans l'arbre prédisant le décès du candidat. Cela suggère que les caractéristiques de la greffe et du donneur ont un impact moins significatif sur la survie du receveur que sur la survie du greffon. On

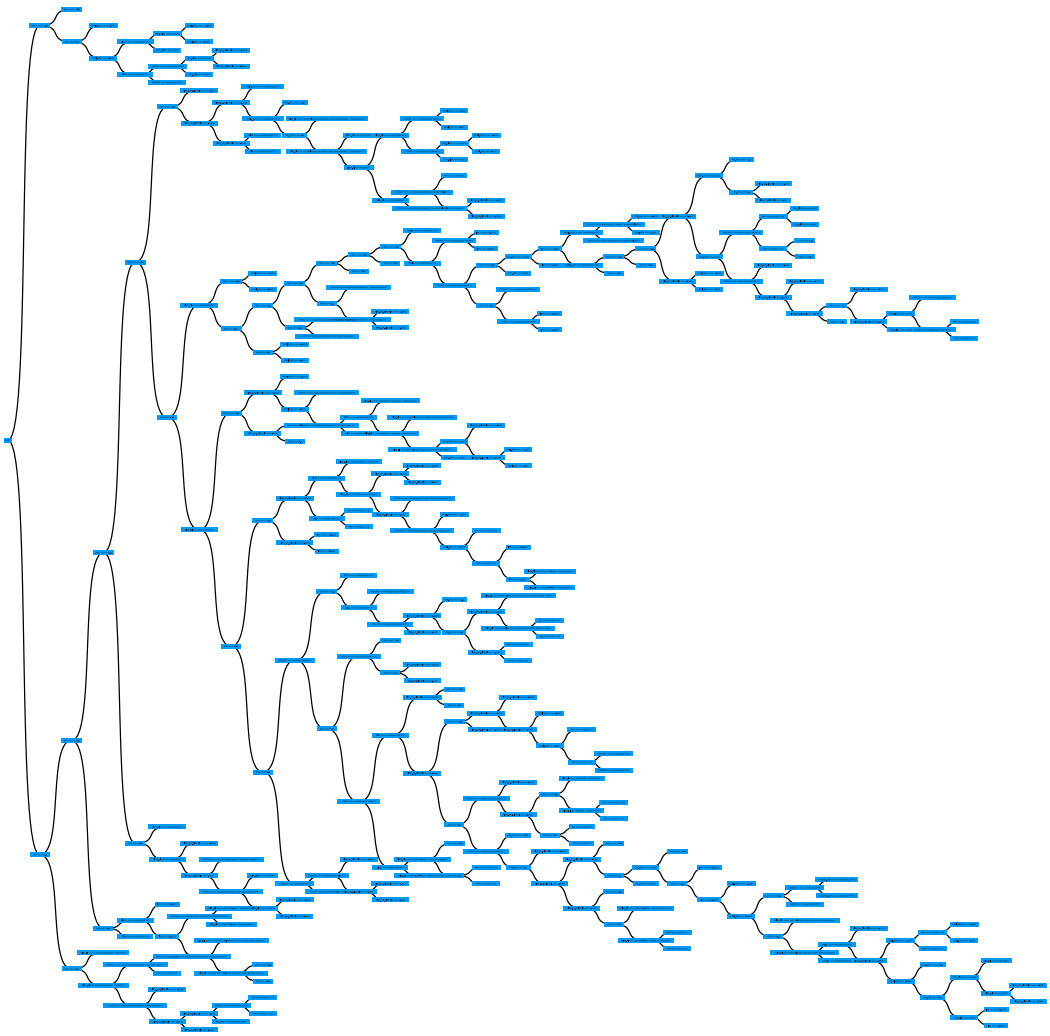


FIGURE 6.7 Structure de l'arbre unique entraîné pour prédire la survie du receveur.

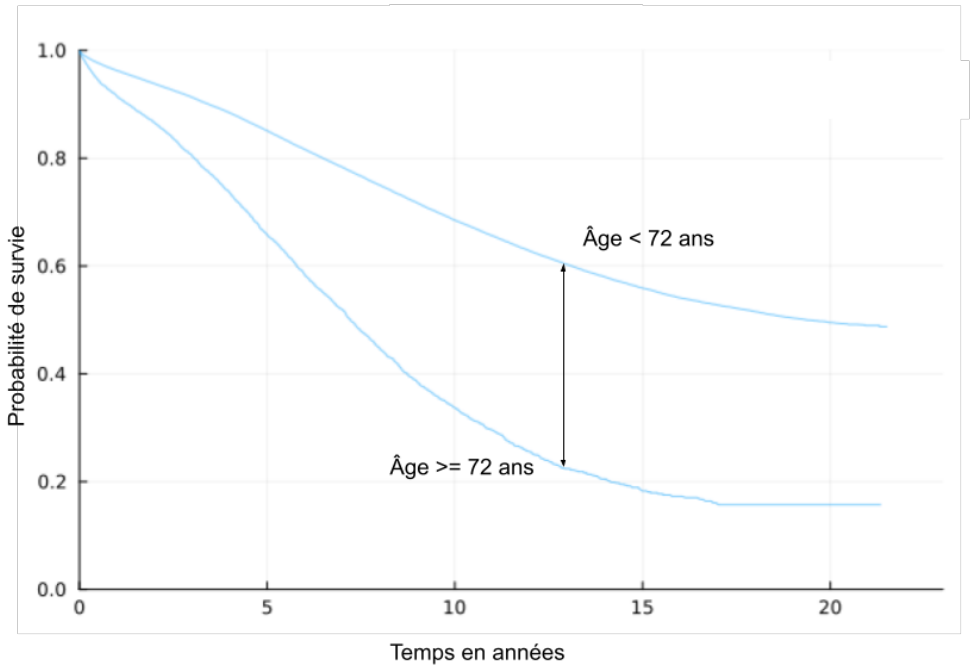


FIGURE 6.8 Courbes de survie du patient estimées par l'estimateur de Kaplan-Meier pour les deux sous-groupes obtenus, à savoir les patients âgés de 72 ans et plus et les plus jeunes. Le partitionnement à l'âge de 72 ans produit la plus grande différence de survie selon la distance de Kolmogorov-Smirnov entre les courbes, indiquée par la flèche.

s'y attend puisque, particulièrement chez les candidats plus âgés, des décès après seulement quelques années est assez fréquent pour dépasser les autres facteurs déterminants. Dans l'arbre à la figure 6.2, puisque le critère d'arrêt est d'avoir un minimum de 500 événements dans les sous-groupes pour continuer le partitionnement, on observe que les chemins les plus profonds contiennent les types de candidats et de donneurs les plus abondants, comme les candidats dont l'âge est près de la médiane (54 ans) ou ayant un faible cPRA, par exemple.

Une autre observation est que dans l'arbre de la figure 6.7, puisque le décès est censuré à la fin de l'étude si aucun événement ne se produit, les branches des candidats plus jeunes (< 34 ans) dans le bas de l'arbre ne progressent pas aussi loin que les autres puisque le critère d'arrêt est satisfait plus rapidement dû au bas taux de décès, entraînant un taux élevé de censure. À noter que le décès chez les patients plus jeunes est souvent causé par des comorbidités ou des complications. C'est pourquoi on observe la cause de la maladie rénale chronique comme étant un facteur déterminant de la survie pour ces patients, particulièrement le diabète. Il a également été démontré que le temps qu'un candidat passe en dialyse affecte grandement sa survie post-greffe. Il est donc attendu que le temps passé en dialyse soit une variable souvent sélectionnée par l'algorithme pour prédire la survie du receveur.

6.4 Calibration D

Les graphiques de calibration D servent à s'assurer que notre modèle produise bel et bien des courbes de survie valides et qu'on peut leur faire confiance pour estimer la survie. Les graphiques présentent la distribution des paniers des modèles d'arbre et de forêt pour les deux variables cible, soient l'échec du greffon et le décès du receveur. Les figures 6.5, 6.6, 6.3 et 6.4 présentent ces graphiques de calibration D. L'objectif est que chacune des 10 classes contienne 10% de la cohorte, menant à un alignement avec la distribution uniforme représentée par la ligne verticale sur ces figures. La statistique du test de khi carré de degré 9 permet d'évaluer l'adéquation des classes obtenues à la loi uniforme attendue si le modèle est bien calibré.

6.5 Évaluation de la calibration

Le test de khi carré permet dans le contexte de calibration D d'évaluer si les paniers d'un graphique de calibration suivent une distribution uniforme afin d'affirmer que les prédictions de notre modèle sont bien calibrées. Le test va comme suit :

$\mathcal{H}_0$: L'échantillon suit une distribution uniforme ;

$\mathcal{H}_1$: L'échantillon ne suit pas une distribution uniforme.

On rejette l'hypothèse nulle si la statistique du test dépasse le seuil d'une distribution khi carré de degré 9 avec $\alpha = 0.05$, soit 16.919.

Les figures 6.5, 6.6, 6.3 et 6.4 nous mènent à affirmer que nos modèles d'arbres et de forêts sont calibrés D et que l'on peut faire confiance à leurs prédictions pour estimer l'échec du greffon et le décès du candidat dans un contexte médical. En effet, dans tous nos graphiques de calibration D, on observe qu'aucune des statistiques de test ne dépasse le seuil. On ne peut donc pas rejeter l'hypothèse nulle et on conclut que les échantillons suivent la loi uniforme. La nature du test servant également de mesure de distance entre les échantillons et la distribution, une plus basse statistique du test signifie un meilleur ajustement de l'échantillon à la distribution uniforme. Les forêts présentent donc une amélioration de la calibration par rapport aux arbres seuls parce qu'elles impliquent une diminution de la statistique de test.

6.6 Effet de l'âge et du KDRI

Un constat intéressant se fait lorsqu'on compare l'impact de l'âge du candidat et de la qualité de l'organe sur l'échec de la greffe. Notre modèle de forêt suggère que le KDRI de l'offre a peu d'influence sur la survie chez les candidats plus âgés, tandis que les candidats plus jeunes bénéficient davantage d'une offre de bonne qualité. Cela peut être expliqué par le fait que la cause première de perte de greffon chez les patients âgés est le décès. Cette tendance est observée sur l'ensemble des candidats du jeu de validation qui satisfont ces critères. Les figures 6.10 et 6.12 présentent les survies prédites pour un même candidat âgé de 30 ans pour deux offres de qualités différentes, soit de KDRI de 1.2 et de 2.4 respectivement. Parallèlement, les figures 6.9 et 6.11 montrent les survies prédites pour une même candidat âgé de 68 ans pour deux offres de qualités différentes, encore une fois ayant des KDRI 1.2 et 2.4 respectivement. On observe une amélioration de la survie du candidat plus jeune d'environ 20% sur l'ensemble de sa survie dès 5 ans après la transplantation. Contrairement à ce dernier, le candidat plus âgé ne bénéficie presque pas d'une offre de meilleure qualité puisque sa courbe de survie prédite varie très peu en fonction du KDRI de l'offre. Il y a donc place à considération des offres de donneurs ayant un KDRI élevé pour les candidats agés, afin d'augmenter le nombre de transplantations effectuées en élargissant le bassin de donneurs décédés.

Pour renforcer davantage les bénéfices de notre approche, la figure 6.14 et 6.13 présentent les courbes Kaplan-Meier de l'ensemble des candidats de la cohorte ayant reçu des offres de KDRI entre 1.1 et 1.3 puis entre 2.3 et 2.5, respectivement. À l'heure actuelle, ce serait à l'aide de ces graphiques qu'un néphrologue informerait le candidat de la survie du greffon attendue si de telles offres étaient reçues. Cependant, les résultats de notre modèle suggèrent que les caractéristiques du candidat et certaines informations qui ne sont pas utilisées dans le calcul

du KDRI sont également nécessaires pour prédire adéquatement la survie du greffon. En effet, cela s'observe par le fait que les survies prédites de deux candidats différents ayant eu des offres similaires (figures 6.10, 6.9, 6.12 et 6.11) diffèrent grandement des courbes présentées aux figures 6.14 et 6.13. Le KDRI n'est donc pas suffisant pour informer les candidats de leur survie attendue ou de celle du greffon.

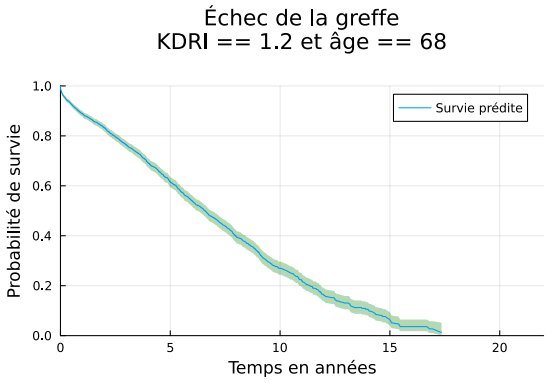


FIGURE 6.9 Exemple de prédiction de survie du greffon faite par le modèle de forêt à 5 arbres pour un candidat âgé de 68 ans ayant accepté une offre de KDRI inférieur à 1.2.

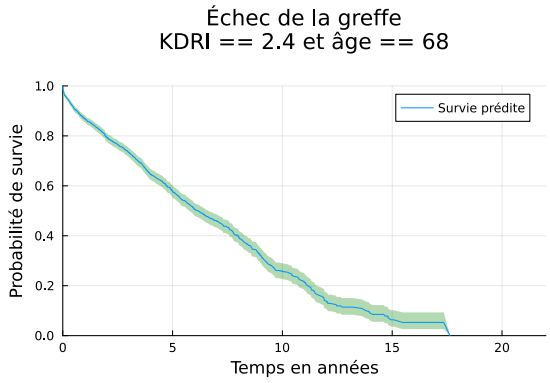


FIGURE 6.11 Exemple de prédiction de survie du greffon faite par le modèle de forêt à 5 arbres pour un candidat âgé de 68 ans ayant accepté une offre de KDRI supérieur à 2.4.

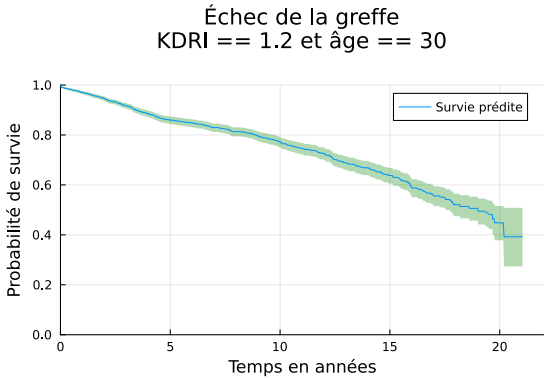


FIGURE 6.10 Exemple de prédiction de survie du greffon faite par le modèle de forêt à 5 arbres pour un candidat âgé de 30 ans ayant accepté une offre de KDRI inférieur à 1.2.

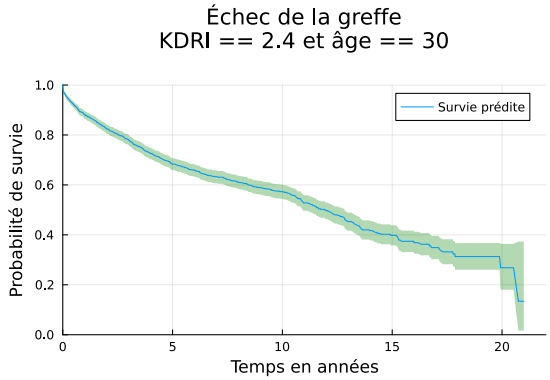


FIGURE 6.12 Exemple de prédiction de survie du greffon faite par le modèle de forêt à 5 arbres pour un candidat âgé de 30 ans ayant accepté une offre de KDRI supérieur à 2.4.

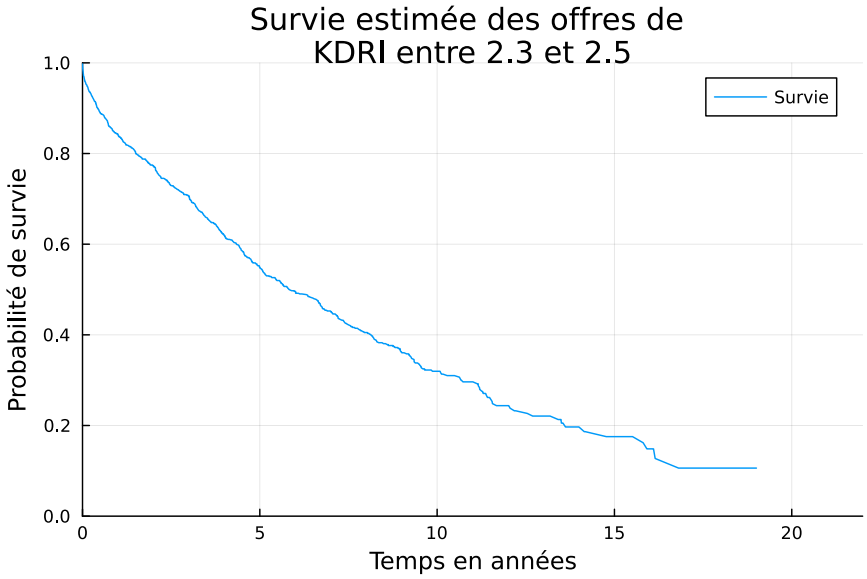


FIGURE 6.13 Courbe Kaplan-Meier des candidats de la cohorte ayant eu des offres de rein provenant de donneurs ayant des KDRI de 2.3 à 2.5, donc de moindre qualité.

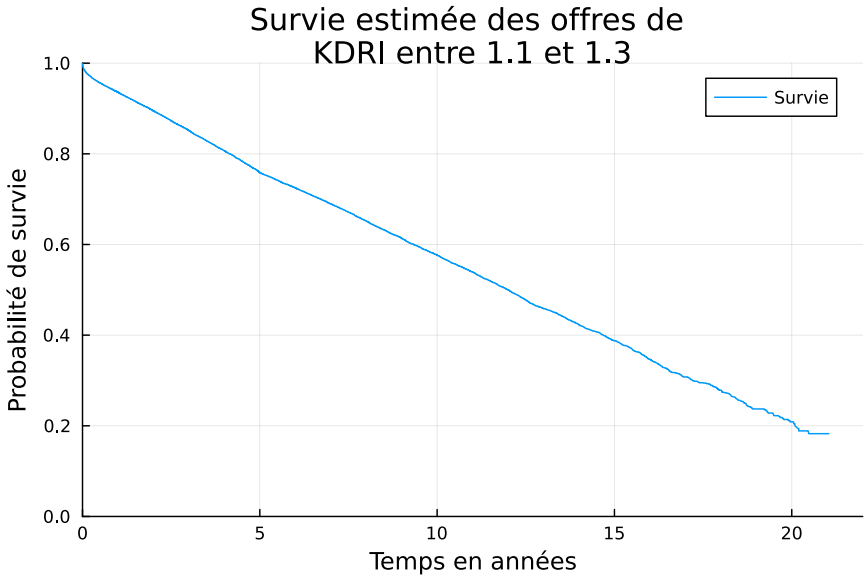


FIGURE 6.14 Courbe Kaplan-Meier des candidats de la cohorte ayant eu des offres de reins provenant de donneurs ayant des KDRI de 1.1 à 1.3, donc de qualité supérieure.

CHAPITRE 7 DISCUSSION

La calibration D a permis d'évaluer la validité des prédictions de nos modèles. On peut donc déduire que tous les arbres et les forêts produites par notre algorithme sont en effet calibrés. L'objectif initial étant d'obtenir un modèle de prédiction pouvant aider la prise de décision clinique, soit d'accepter ou de refuser une offre rénale de donneur décédé, on en conclut que les modèles proposés ont le potentiel de satisfaire cet objectif. C'est par leur capacité à produire des prédictions individualisées tout en atteignant une calibration adéquate que nous sommes en mesure de tirer de telles conclusions.

Cependant, il ne faut pas négliger que cet exercice a été effectué dans un contexte précis et dans un environnement fixe. Entre autres, les hyperparamètres utilisés ont été sélectionnés selon le savoir des experts. Les hyperparamètres n'ont donc pas été optimisés en fonction des performances sur un ensemble de validation. En raison de ressources limitées, une recherche exhaustive de la combinaison d'hyperparamètres idéale n'a pas été possible. Cette étape aurait potentiellement permis d'améliorer le niveau de calibration de nos modèles en retrouvant l'accord idéal entre le nombre de feuilles des arbres et leur capacité de généralisation, par exemple. Autrement dit, nous ne savons pas présentement si d'autres valeurs d'hyperparamètres auraient pu produire une meilleure calibration. L'algorithme de recherche en grille (*grid-search*), avec assez de temps d'exécution et de puissance de calcul, pourrait éventuellement permettre d'identifier la combinaison idéale des hyperparamètres qui maximisera le niveau de calibration dans des travaux futurs. Ces derniers serait ensuite le sujet de l'avis d'experts pour assurer la validité du modèle final. Néanmoins, le fait d'utiliser le savoir d'experts pour la spécification des hyperparamètres n'est pas sans intérêt. Cette approche permet d'impliquer directement les experts dans le développement de l'approche et d'induire confiance envers la méthode.

Le nombre minimal d'événements dans les feuilles des arbres est l'exemple d'un hyperparamètre initialement fixé qui pourrait être sujet à changement après avoir testé différentes valeurs et identifié la meilleure. Le nombre minimal d'événements dans les feuilles influence non seulement la résolution des courbes de survie prédites, mais également la profondeurs des arbres puisque c'est un critère d'arrêt. Si ce nombre est trop élevé, les arbres seraient trop petits et produiraient des prédictions qui ne sont pas assez ajustées aux caractéristiques du candidat et du donneur. D'une autre part, si ce nombre est trop petit, les arbres seraient plus profonds, les courbes de survies résultants seraient trop saccadées et on ferait face à un modèle sur-entraîné et plus complexe. Pour compenser avec la complexité d'un tel arbre, on

peut réduire la profondeur maximale. Cela risque toute fois de nuire à la calibration puisqu'on obtiendrait des feuilles avec beaucoup de candidats, offrant une estimation très généralisé de ces patients, et d'autres feuilles avec très peu de candidats qui produiraient des prédictions sur-ajustées. Aussi, une haute résolution des partitionnement (≤ 0.05) permet d'identifier les effets non-linéaires des variables explicatives, mais pourrait mener à un sur-entraînement. La spécification des hyperparamètres est donc un processus exigeant en temps de calcul. L'optimisation de la sélection serait nécessaire pour assurer d'avoir un modèle bien ajuster pour notre cohorte lorsqu'il sera entraîné avec des données locales.

Comme dans bien des solutions en apprentissage automatique, le temps d'entraînement est toujours un enjeu. Dû à la très grande cohorte d'entraînement et au nombre de variables explicatives, l'entraînement d'un arbre peut s'étendre jusqu'à une heure et l'entraînement d'une forêt composée de cinq arbres peut durer jusqu'à trois heures. L'intégration de l'algorithme pourrait être optimisée afin d'accélérer certains processus comme le calcul de la statistique de Kolmogorov-Smirnof à chaque partitionnement possible. Puisque cette partie du code est exécutée un très grand nombre de fois, la moindre amélioration de son exécution influencerait grandement le temps d'entraînement.

Heureusement, un projet d'intégration du modèle dans un système utilisé en clinique n'est pas hors de portée. En effet, il serait possible d'implémenter un processus de ré-entraînement automatisé du modèle lorsqu'un certain nombre de nouvelles données sont disponibles. Par exemple, lorsque l'information de 50 ou 100 nouvelles transplantations devient disponible, le modèle pourrait être ré-entraîné après les avoir ajoutées à l'ensemble d'entraînement initial. Des nouvelles transplantations et des mises à jour de l'ensemble de données sont des événements assez rares pour que le ré-entraînement ne soit nécessaire que quelques fois par années. Un temps d'entraînement de trois heures devient donc relativement négligeable si on le compare au taux de disponibilité de nouvelles données et à la fréquence de ré-entraînement nécessaire.

Une autre méthode de généralisation et de simplification d'un modèle d'arbre est l'élagage. Elle consiste à jumeler ou supprimer des branches de l'arbre entraîner qui n'améliore que peu les prédictions. On cherche de cette façon à réduire la complexité du modèle, ce qui accélère le temps d'inférence, et de compenser pour le possible sur-entraînement. Un algorithme d'élagage pourrait donc contribuer rendre notre approche plus robuste.

Un autre élément à considérer est que les données utilisées pour ce mémoire viennent d'un registre américain et qu'il existe des différences entre le devenir des patients greffés du rein entre le Canada et les États-Unis. La survie du greffon rénal est généralement meilleure au Canada. Le système de santé américain, étant géré par des organisations privées, est teinté

d'une discrimination systématique affectant les personnes moins aisées financièrement souffrant de maladie rénale chronique qui ne peuvent pas se permettre les coût des médicaments anti-rejets dont la couverture cesse souvent 3 ans post greffe (Kulkarni *et al.*, 2019; Harding *et al.*, 2018; Laster *et al.*, 2019). Du même coup, les personnes de minorités visibles sont affectées davantage par ce problème étant donné l'association entre l'ethnicité et le statut socio-économique. Le registre des transplantations rénales américain reflète ces inégalités. Si on souhaite que notre modèle soit en mesure de refléter la démographie et les tendances présentes dans le système de santé du Québec, il serait nécessaire de faire le processus d'entraînement et de validation entier en utilisant des données locales. Aussi, le niveau de calibration obtenu provient du grand nombre d'événements présents dans l'immense jeu de données américain, tandis que les données québécoises sont moins nombreuses et plus clairsemées. On doit donc s'attendre à un moins bon niveau de calibration avec les données locales, malgré que ce soit une nécessité pour pouvoir implanter ce modèle au Québec.

Malgré ces points à améliorer, on peut tout de même conclure que l'approche proposée représente un progrès dans la bonne direction par rapport à des méthodes précédentes. Par exemple, la méthodologie proposée par Luck *et al.* (2017) introduisait une approche utilisant un réseau de neurones pour estimer la fonction de survie des patients. Une fonction d'erreur permettait de prendre en compte des données censurées durant l'entraînement du modèle. Pour l'estimation des paramètres, la méthode impose que la probabilité de survie des candidats après 20 ans, soit la date de fin de l'étude, soit fixée à 0. Évidemment, des candidats vivent avec un greffon fonctionnel bien au delà de 20 ans. Si on performait la calibration D de ce modèle, on verrait que les classes associées à une basse probabilité prédite, $[0.0, 0.1]$ par exemple, contiendrait bien trop de candidats. Cela serait dû au fait que tous les candidats ayant subit un événement à un temps s'approchant de la date de fin d'étude se verrait prédit une survie près de zéro.

L'approche proposée n'a pas ce problème puisque les prédictions sont basées sur les estimations de Kaplan-Meier et la probabilité de survie à la fin de l'étude n'est pas forcément zéro. La méthode de Luck *et al.* (2017) ont validé leur approche avec l'index de concordance, ou l'AUROC. C'est en effet la métrique communément utilisée pour évaluer la discrimination d'un modèle de prédiction, mais elle n'assure pas la calibration des prédictions. En effet, un modèle peut données des prédictions bien ordonnées, puisque c'est ce que l'index de concordance mesure, mais en même temps sous-estimer ou sur-estimer la probabilité réelle d'événements pour l'ensemble de la cohorte. En pratique dans un contexte d'aide à la décision pour un patient, le modèle proposé qui est calibré est donc plus adéquat pour supporter la décision clinique d'accepter ou de refuser une offre rénale.

L'implantation d'un tel modèle dans le processus de prise de décision ajoute un aspect éthique qui doit être considéré et analysé. Devrait-on prioriser les prédictions du modèle ? Ou devrions-nous plutôt les utiliser seulement pour supporter la décision ? Si le modèle offre des prédictions calibrées, pourquoi conserver l'avis d'experts et des candidats qui pourraient biaiser la prise de décision ? Pour y répondre, il faut se rappeler de l'objectif du support à la prise de décision partagée. On cherche d'abord à offrir aux candidats une façon de leur remettre en main la décision qui ultimement ne concerne qu'eux. C'est-à-dire qu'on ne cherche pas à les inciter à prendre la décision que l'algorithme leur suggère, mais plutôt de prendre en considération ses résultats. Encore une fois, on a pour but d'informer les candidats et les professionnels. Certains candidats pourraient même vouloir prendre un risque et refuser une offre satisfaisante si c'est ce qu'il ou elle désire profondément. Il ne faut donc pas se fier strictement aux prédictions pour choisir d'accepter ou de refuser une offre.

CHAPITRE 8 CONCLUSION

Dans ce mémoire, un nouvel algorithme de prédiction de la survie a été développé, entraîné sur des données américaines et validé par la calibration dans le but d'offrir aux candidats et aux néphrologues transplanteurs une référence pour la prise de décision d'accepter ou de refuser une offre rénale. En effet, en débutant par la préparation ses données américaines du UNOS, des cibles modélisant la survie du greffon et la survie du patient ont été établies. Grâce à une métrique de partitionnement basée sur la statistique de test de Kolmodrov-Smirnof, des modèles d'arbres décisionnels et de forêts aléatoires ont été entraînés en évaluant les partitionnements optimaux pour produire des sous-groupes homogènes. À partir de ceux-ci, on a produit des courbes de survie individualisées pour une combinaison candidat-donneur. Ces modèles se sont montrés calibrés par l'algorithme de calibration D. Cela implique donc que leurs prédictions sont cliniquement viables et que les candidats et les professionnels peuvent s'y fier pour s'informer des risques associés à l'acceptation d'une offre. On peut donc conclure que les objectifs initiaux ont été satisfaits et qu'un tel outil pourrait mener à améliorer l'allocation des reins et d'offrir aux candidats la chance de s'approprier davantage la prise de décision. Dans le futur, les modèles devront être ré-entraînés à partir de données locales qui représente la démographie québécoise et les pratiques en transplantations utilisées au Québec pour évaluer et ultimement renforcer la validité des modèles. On observerait une grande valeur ajoutée à exécuter une recherche en grille des meilleures combinaisons d'hyperparamètres en fonction de ces nouveaux jeux de données. Après tout cela, la vision d'implémenter un tel modèle dans le processus d'allocation pourrait devenir réalité. Si utilisé à grande échelle, cet outil aurait non seulement le potentiel de normaliser et d'améliorer l'allocation, d'offrir aux candidats un portrait réaliste de leur condition, mais également de les éclairer afin de choisir l'option qui les satisfait le mieux pour s'assurer une meilleure qualité de vie. Passons le flambeau à ceux qui sauront pousser cette recherche plus loin. Ceux-ci pourraient produire un modèle de prédiction de survie généraliste en l'entraînant sur une plus grande population et en incluant plus de variables explicatives. On pourrait inclure le contexte socio-économique des candidats et des donneurs ou même la localisation des transplantations. Cela permettrait d'étendre la portée des bénéfices associés à l'utilisation d'un tel outil.

RÉFÉRENCES

Ali, J., Khan, R., Ahmad, N. et Maqsood, I. (2012). Random forests and decision trees. *International Journal of Computer Science Issues*, 9.

Argani, H. (2022). Expanded criteria donors. *Experimental and Clinical Transplantation*, 20(Suppl 4):13–19.

Bae, S., Massie, A. B., Thomas, A. G., Bahn, G., Luo, X., Jackson, K. R., Ottmann, S. E., Brennan, D. C., Desai, N. M., Coresh, J., Segev, D. L. et Garonzik Wang, J. M. (2019). Who can tolerate a marginal kidney ? predicting survival after deceased donor kidney transplant by donor–recipient combination. *American Journal of Transplantation*, 19(2):425–433.

Canadian Organ Replacement Register (2023). Annual statistics on organ replacement in canada, 2013 to 2022. Rapport technique, Canadian Institute for Health Information.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society*.

Danovitch, G. M. (2017). *Handbook of kidney transplantation*. Wolters kluwar, sixième édition.

Haider, H., Hoehn, B., Davis, S. et Greiner, R. (2020). Effective ways to build and evaluate individual survival distributions. *Journal of Machine Learning Research*, 21:63.

Harding, K., Mersha, T. B., Pham, P.-T., Waterman, A. D., Webb, F. J., Vassalotti, J. A. et Nicholas, S. B. (2018). Health disparities in kidney transplantation for african americans. *American Journal of Nephrology*, 46(2):165–175.

Ishwaran, H., Kogalur, U. B., Blackstone, E. H. et Lauer, M. S. (2008). Random survival forests. *Annals of Applied Statistics*, 2(3).

Jay, C. L., Washburn, K., Dean, P. G., Helmick, R. A., Pugh, J. A. et Stegall, M. D. (2017). Survival benefit in older patients associated with earlier transplant with high kdpi kidneys. *Transplantation*, 101(4):867–872.

Kadatz, M. J., Gill, J., Gill, J., Lan, J. H., McMichael, L. C., Chang, D. T. et Gill, J. S. (2023). The benefits of preemptive transplantation using high–kidney donor profile index kidneys. *Clinical Journal of the American Society of Nephrology*, 18(5):634–643.

Kalbfleisch, J. D. et Prentice, R. L. (2002). The statistical analysis of failure time data. *Wiley New York*.

Kaplan, E. L. et Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.

Khalili, M., Cardinal, H., Ballesteros, F. et Fortin, M.-C. (2022). Kidney transplant candidates' and recipients' perspectives on the decision-making process to accept or refuse a deceased donor kidney offer : Trust and graft survival matter. *Clinical Transplantation*, 36(5):e14604.

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Instituto Italiano degli Attuari*, 4:83–91.

Kulkarni, S., Ladin, K., Haakinson, D., Greene, E., Li, L. et Deng, Y. (2019). Association of racial disparities with access to kidney transplant after the implementation of the new kidney allocation system. *Journal of the American Medical Association Surgery*, 154(7):618–625.

Ladin, K. (2022). The elusive promise of shared decision making : A step forward. *American Journal of Kidney Diseases*, 79.

Laster, M., Shen, J. I. et Norris, K. C. (2019). Kidney disease among african americans : A population perspective. *American Journal of Kidney Diseases*, 72(5).

Laupacis, A., Keown, P., Pus, N., Krueger, H., Ferguson, B., Wong, C. et Muirhead, N. (1996). A study of the quality of life and cost-utility of renal transplantation. *Kidney International*, 50:235–242.

Luck, M., Sylvain, T., Cardinal, H., Lodi, A. et Bengio, Y. (2017). Deep learning for patient-specific kidney graft survival analysis. arXiv :1705.10245 [cs, stat].

Manns, B. J., Mendelssohn, D. C. et Taub, K. J. (2007). The economics of end-stage renal disease care in canada : incentives and impact on delivery of care. *International Journal of Health Care, Finance and Economics*, pages 149–169.

Massie, A. B., Luo, X., Chow, E. K. H., Alejo, J. L., Desai, N. M. et Segev, D. L. (2014). Survival benefit of primary deceased donor transplantation with high-kdpi kidneys. *American Journal of Transplantation*, 14(10):2310–2316.

Merion, R. M., Ashby, V. B., Wolfe, R. A., Distant, D. A., Hulbert-Shearon, T. E., Metzger, R. A., Ojo, A. O. et Port, F. K. (2005). Deceased-donor characteristics and the survival benefit of kidney transplantation. *Journal of the American Medical Association*.

Ojo, A. O., Hanson, J. A., Meier-Kriesche, H. U., Okechukwu, C. N., Wolfe, R. A., Leichtman, A. B., Agodoa, L. Y., Kaplan, B. et Port, F. K. (2001). Survival in recipients of marginal cadaveric donor kidneys compared with other recipients and wait-listed transplant candidates. *Journal of the American Society of Nephrology*, 12(3).

Organ Procurement and Transplantation Network (2020a). A guide to calculating and interpreting estimated post-transplant survival (epts) score used in the kidney allocation system (kas).

Organ Procurement and Transplantation Network (2020b). A guide to calculating and interpreting the kidney donor profile index (kdpi).

Pascual, J., Zamora, J. et Pirsch, J. D. (2008). A systematic review of kidney transplantation from expanded criteria donors. *American Journal of Kidney Diseases*.

Port, F. K., Bragg-Gresham, J. L., A, M. R., Dykstra, D. M., Gillespie, B. W., Young, E. W., Delmonico, F. L., Wynn, J. J., Merion, R. M., Wolfe, R. A. et Held, P. J. (2002). Donor characteristics associated with reduced graft survival : an approach to expanding the pool of kidney donors. *Transplantation*, 74(9).

Rao, P. S., Schaubel, D. E., Guidinger, M. K., Andreoni, K. A., Wolfe, R. A., Merion, R. M., Port, F. K. et Sung, R. S. (2009). A comprehensive risk quantification score for deceased donor kidneys : The kidney donor risk index. *Transplantation*, 88(2): 231–236.

Rosaasen, C., Rosaasen, N., Mainra, R., Trachtenberg, A., Ho, J., Parsons, C., Delaney, S. et Mansell, H. (2022). Waitlisted and transplant patient perspectives on expanding access to deceased-donor kidney transplant : A qualitative study. *Canadian Journal of Kidney Health and Disease*, 9:205435812211002.

Rose, C., Sun, Y., Ferre, E., Gill, J., Landsberg, D. et Gill, J. (2018). An examination of the application of the kidney donor risk index in british columbia. *Canadian Journal of Kidney Health and Disease*, 5:205435811876105.

Schold, J. D., Arrigain, S., Flechner, S. M., Augustine, J. J., Sedor, J. R., Wee, A., Goldfarb, D. A. et Poggio, E. D. (2019). Dramatic secular changes in prognosis for

kidney transplant candidates in the united states. *American Journal of Transplantation*, 19(2):414–424.

Smirnov, N. V. (1939). Estimate of deviation between empirical distribution functions in two independent samples. *Bulletin Moscow University*, 2(2):3–16.

Smirnov, N. V. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, 19:279–281.

Transplant Québec (2020). Procédure d'attribution normalisée : Attribution rénale.

Van Calster, B., Nieboer, D., Vergouwe, Y., De Cock, B., Pencina, M. J. et Steyerberg, E. W. (2016). A calibration hierarchy for risk models was defined : from utopia to empirical data. *Journal of Clinical Epidemiology*, 74:167–176.

Van Calster, B., van Smeden, M., Wynants, L. et Steyerberg, E. W. (2019). Calibration : the achilles heel of predictive analytics. *BioMed Central Medicine*, 17(1):230.

Wey, A., Salkowski, N., Kremers, W. K., Schaffhausen, C. R., Kasiske, B. L., Israni, A. K. et Snyder, J. J. (2018). A kidney offer acceptance decision tool to inform the decision to accept an offer or wait for a better kidney. *American Journal of Transplantation*, 18(4):897–906.

Wolfe, R. A., Ashby, V. B., Milford, E. L., Ojo, A. O., Ettenger, R. E., Agodoa, L. Y., Held, P. J. et Port, F. K. (1999). Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *New England Journal of Medicine*, 341:1725–1730.

Young, A., Knoll, G. A., McArthur, E., Dixon, S. N., Garg, A. X., Lok, C. E., Lam, N. N. et Kim, S. J. (2018). Is the kidney donor risk index a useful tool in non-us patients ? *Canadian Journal of Kidney Health and Disease*, 5:205435811879114.

Yu, C.-N., Greiner, R., Lin, H.-C. et Baracos, V. (2011). Learning patient-specific cancer survival distributions as a sequence of dependent regressors. *In NIPS'11 : Proceedings of the 24th International Conference on Neural Information Processing Systems*. Curran Associates Inc.

Zhong, Y., Schaubel, D. E., Kalbfleisch, J. D., Ashby, V. B., Rao, P. S. et Sung, R. S. (2019). Reevaluation of the kidney donor risk index. *Transplantation*, 103(8):1714–1721.

ANNEXE A VISUALISATION DES MODÈLES D'ARBRES ET DE FORÊTS

Visualisation des forêts pour l'échec du greffon

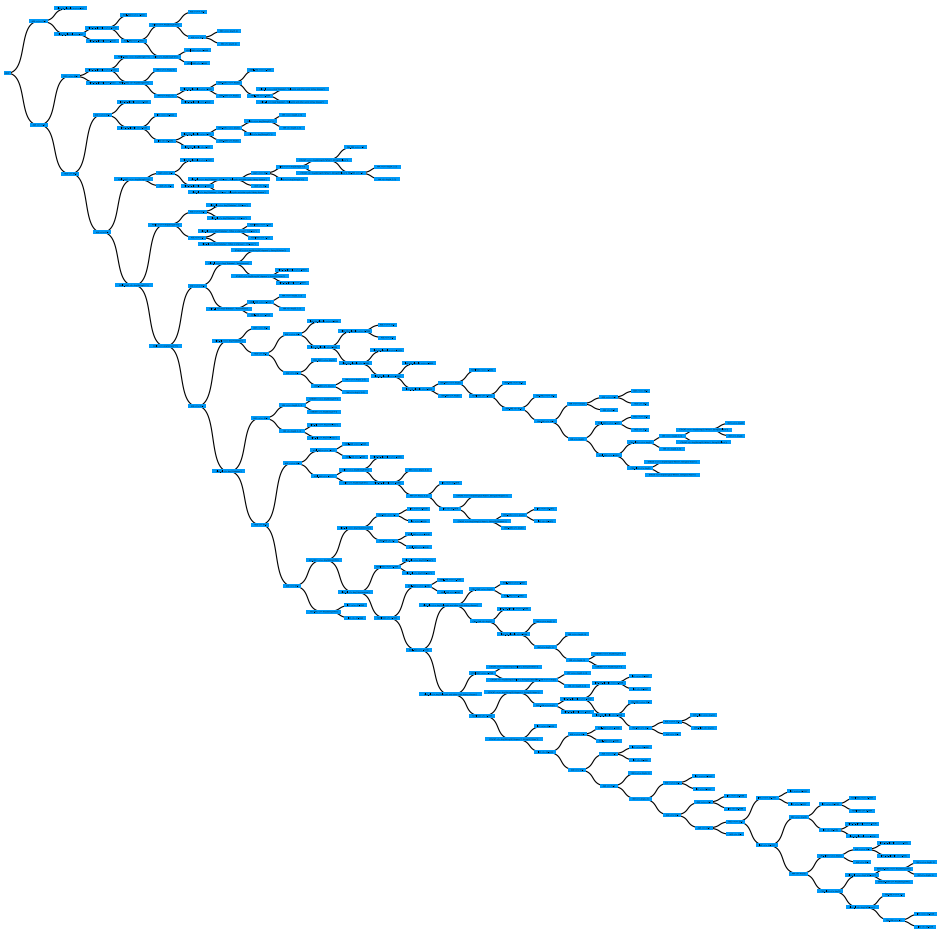


FIGURE A.1 Représentation de l'arbre #1 de la forêt pour prédire l'échec du greffon.

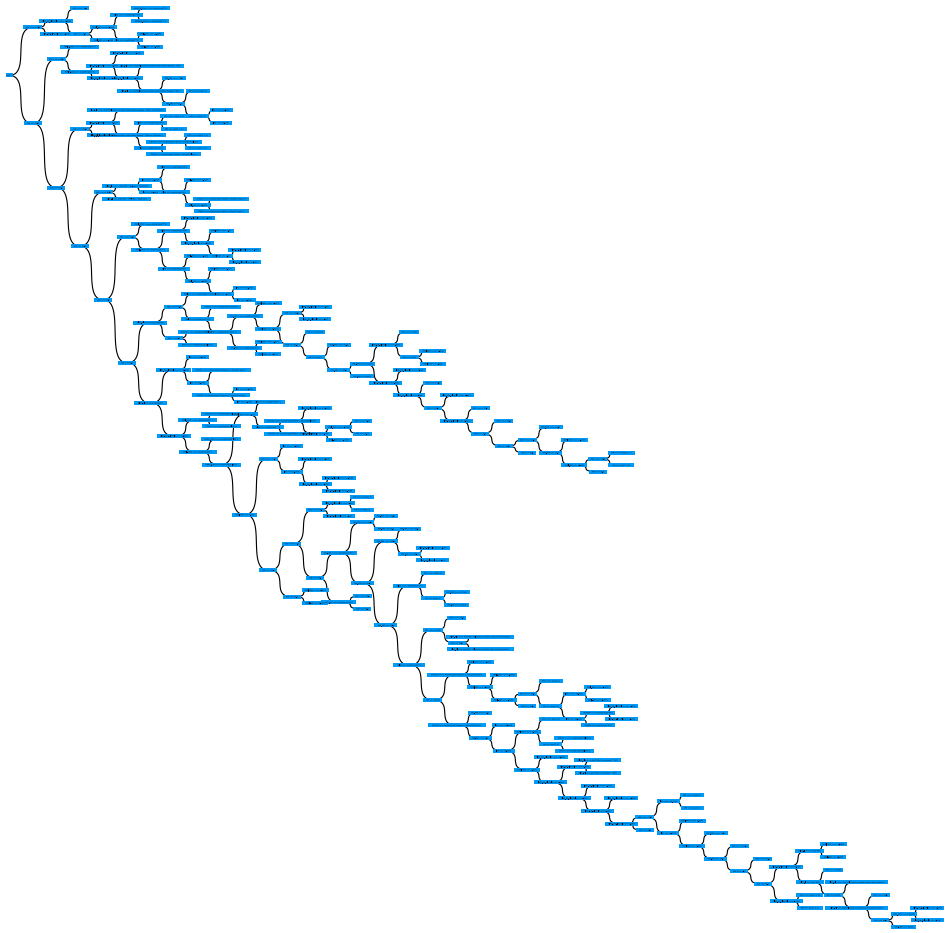


FIGURE A.2 Représentation de l'arbre #2 de la forêt pour prédire l'échec du greffon.

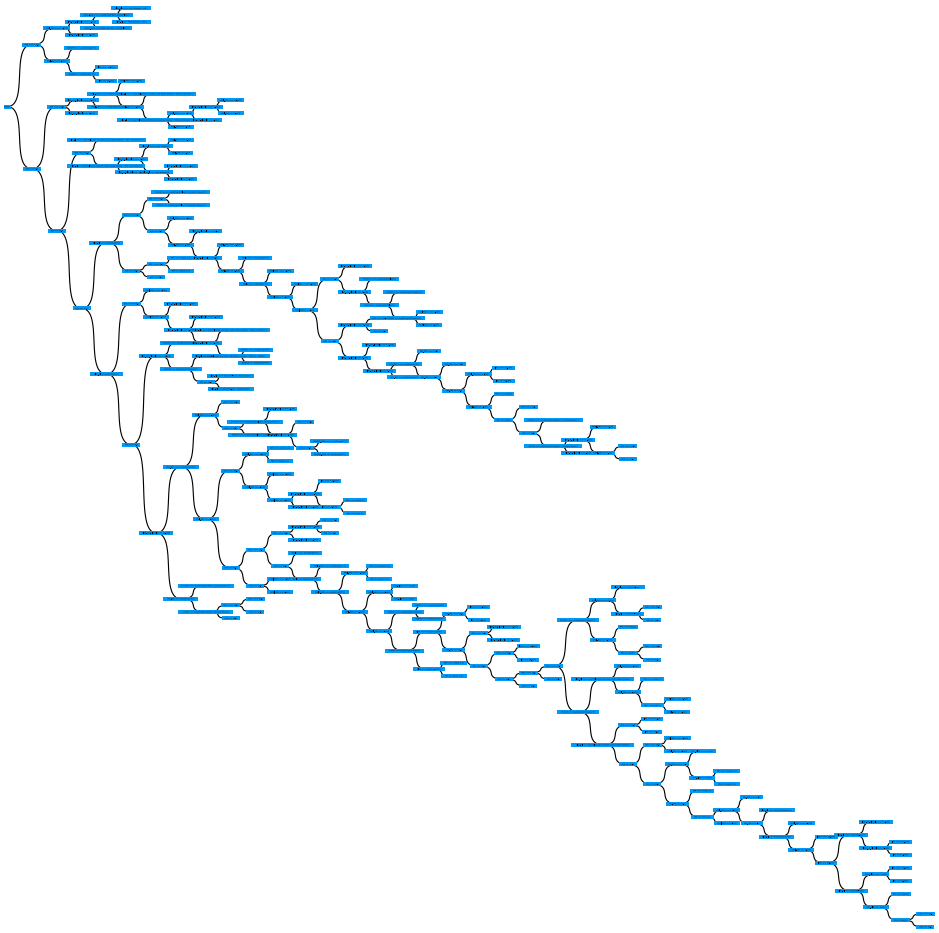


FIGURE A.3 Représentation de l'arbre #3 de la forêt pour prédire l'échec du greffon.

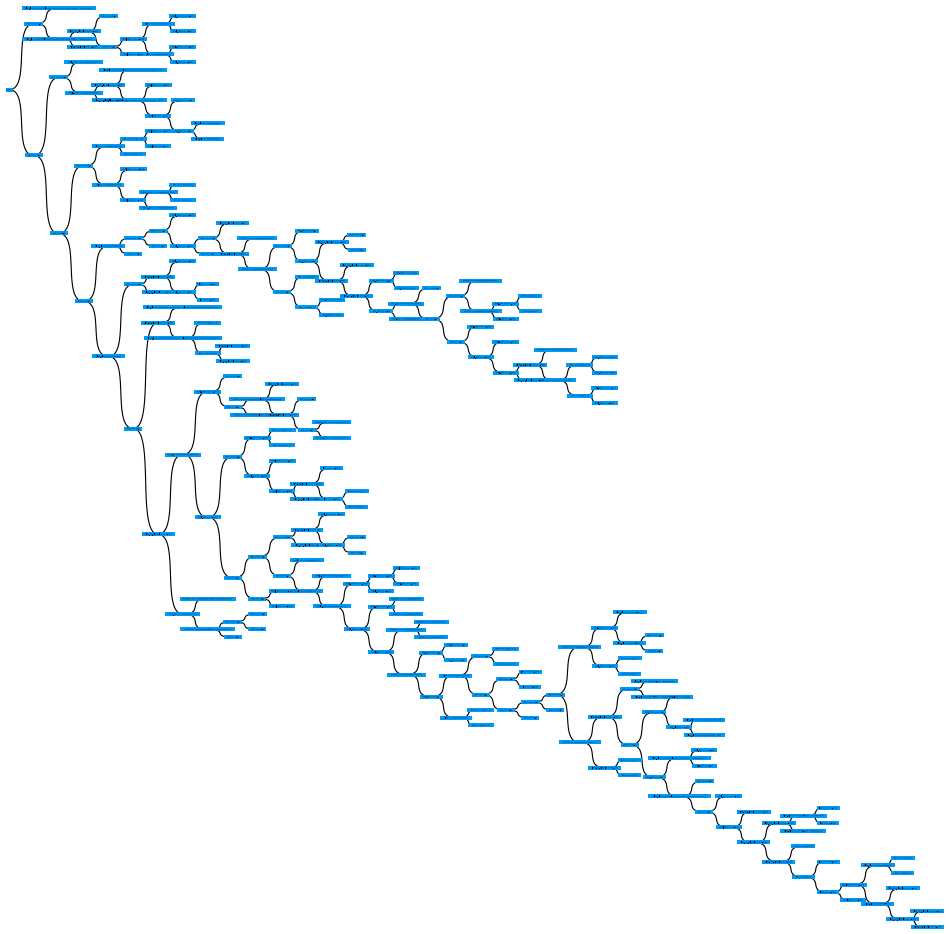


FIGURE A.4 Représentation de l'arbre #4 de la forêt pour prédire l'échec du greffon.

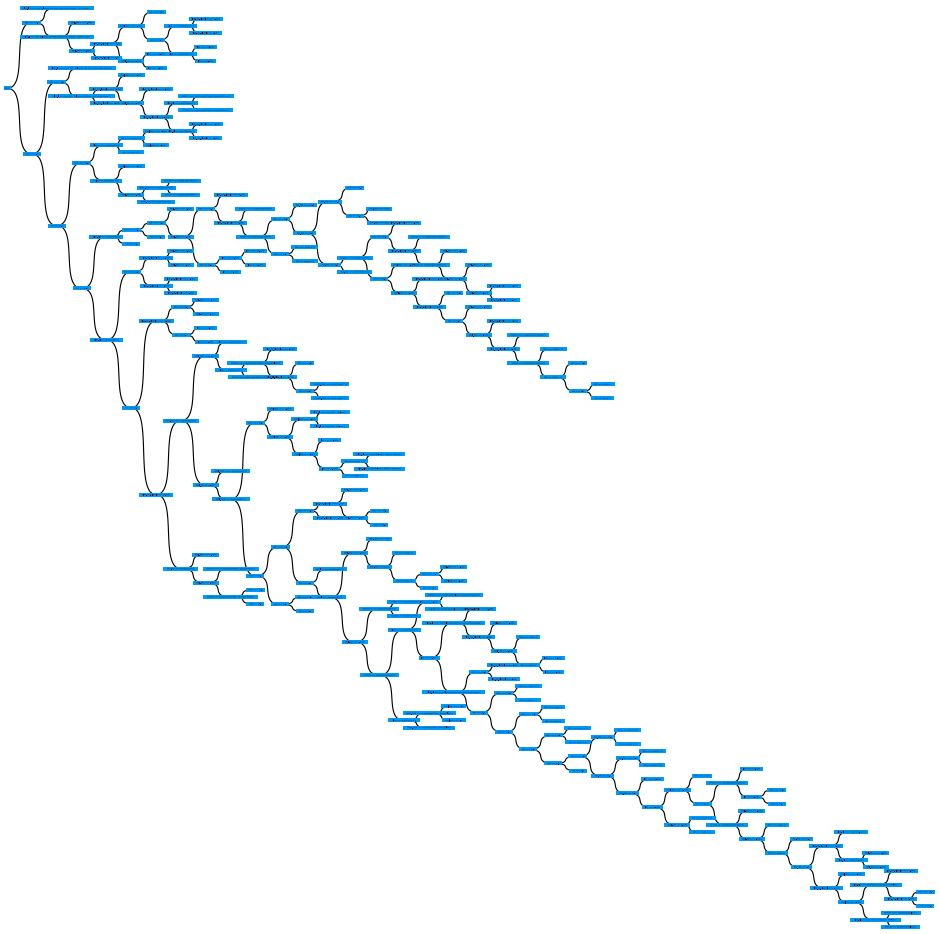


FIGURE A.5 Représentation de l'arbre #5 de la forêt pour prédire l'échec du greffon.

Visualisation de la forêt pour le décès du candidat

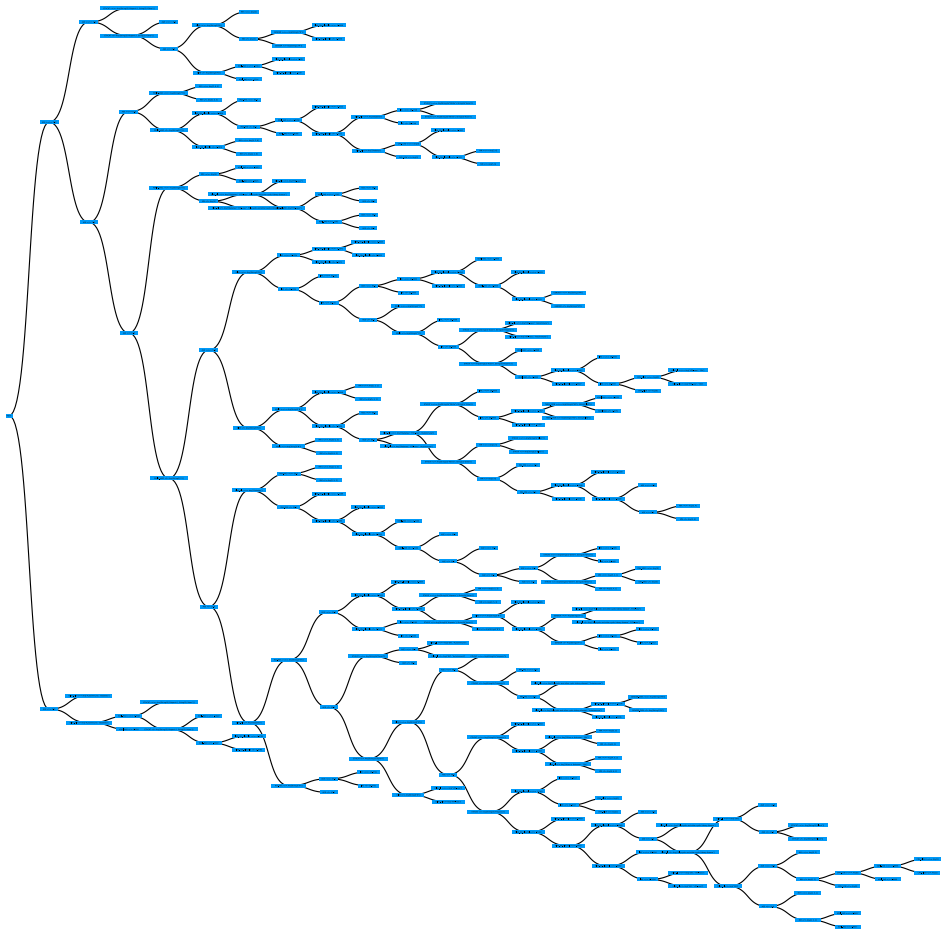


FIGURE A.6 Représentation de l'arbre #1 de la forêt pour prédire le décès du candidat.

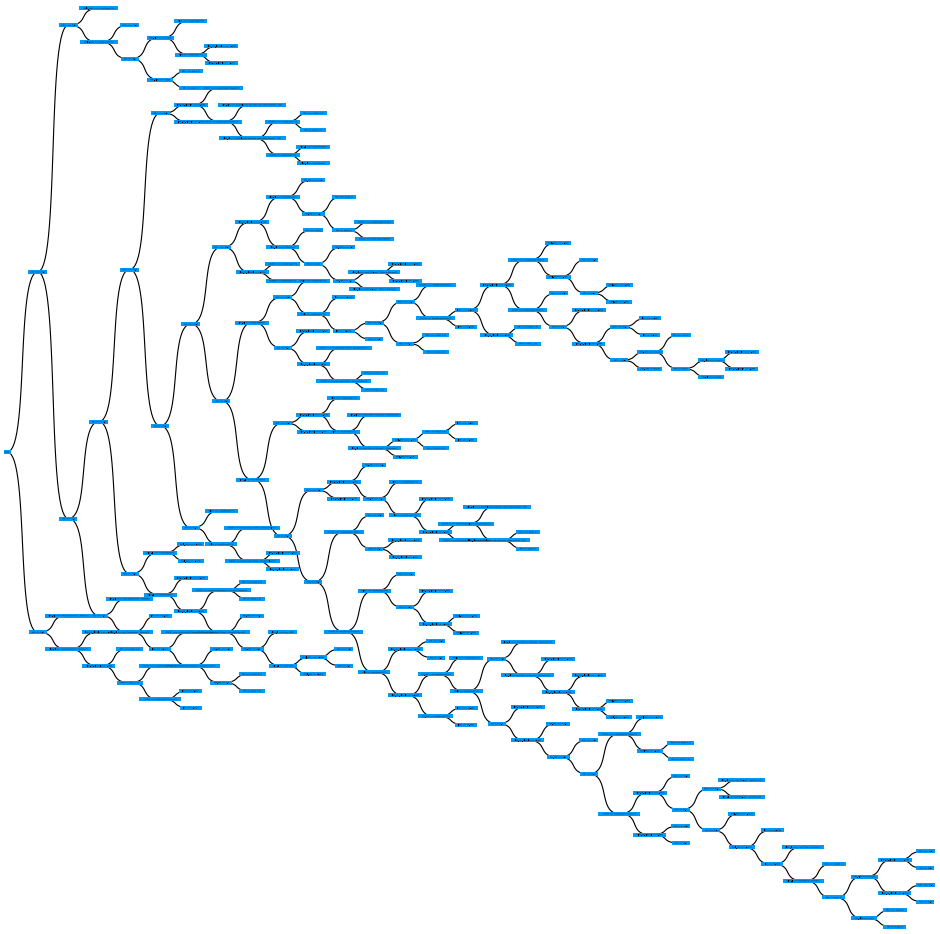


FIGURE A.7 Représentation de l'arbre #2 de la forêt pour prédire le décès du candidat.

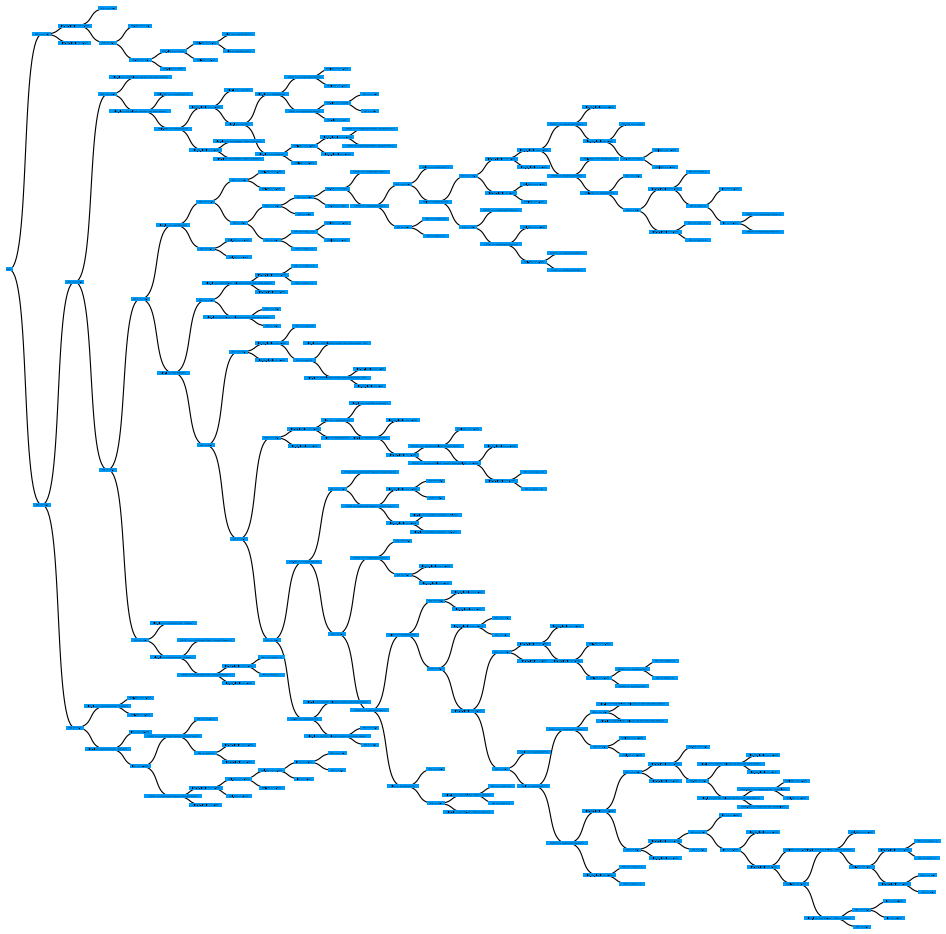


FIGURE A.8 Représentation de l'arbre #3 de la forêt pour prédire le décès du candidat.

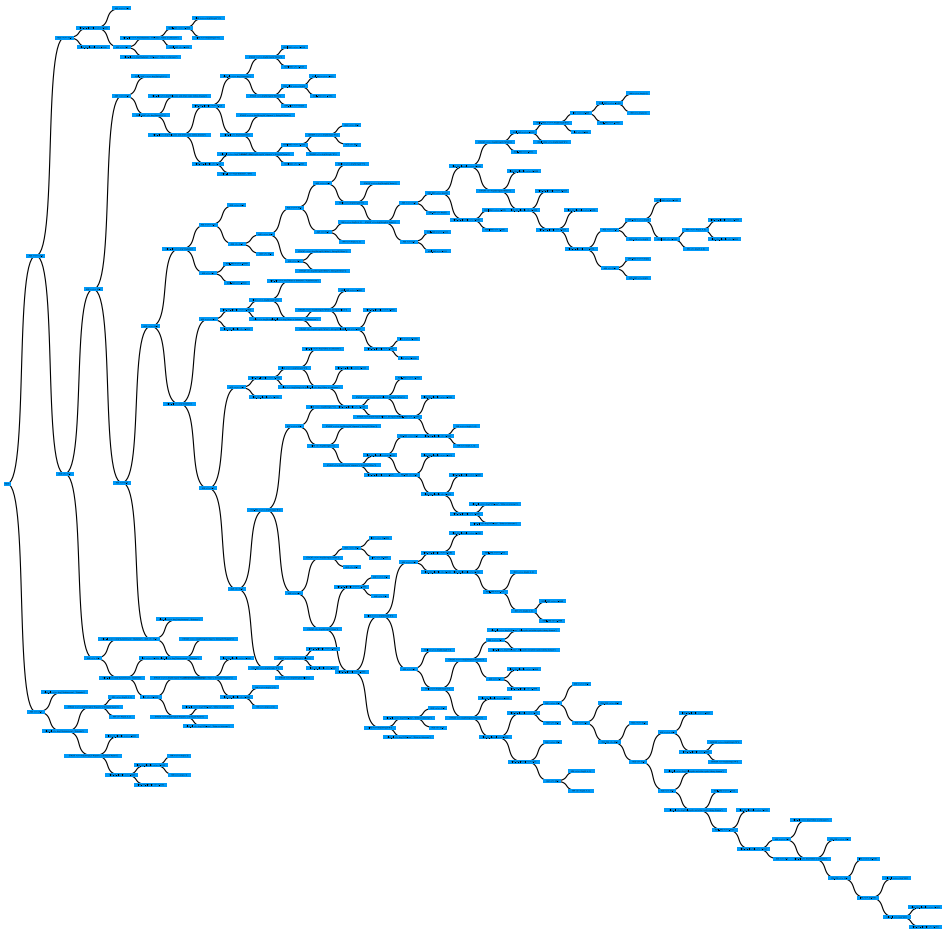


FIGURE A.9 Représentation de l'arbre #4 de la forêt pour prédire le décès du candidat.

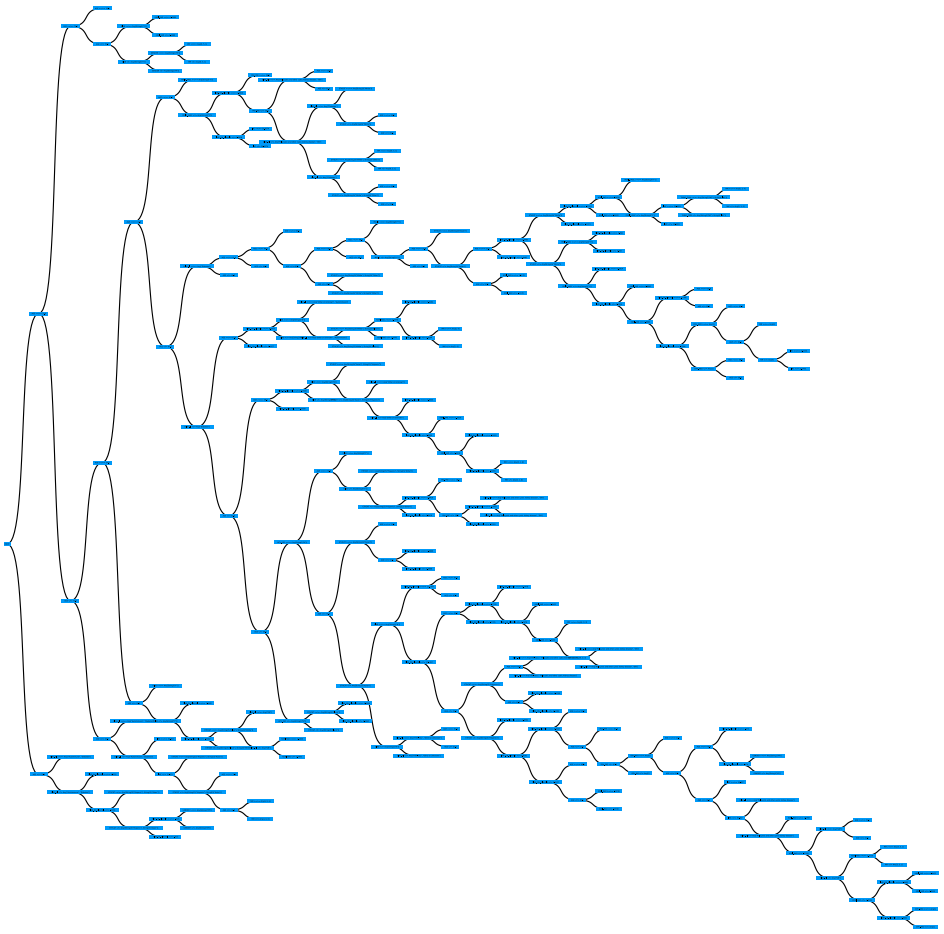


FIGURE A.10 Représentation de l'arbre #5 de la forêt pour prédire le décès du candidat.