

Rapport de projet tutoré

Léonie BREUZA, Matthias MAZET, Kento OKADO, Enzo VIGNAUD

2025-04-07

Contents

Contexte	2
Méthodes	2
<i>Données</i>	2
<i>Protocole</i>	2
<i>Traitement des données</i>	2
<i>Analyses statistiques</i>	3
Résultats	6
<i>Statistiques univariées/bivariées</i>	6
<i>Modèles</i>	6
Discussion/Conclusion	6
Impact Environnemental et Sociétal	7
<i>Impact environnemental personnel</i>	7
<i>Impact global du projet</i>	7
<i>Politique de la structure d'accueil</i>	7

Contexte

Les greffes de reins possèdent encore aujourd’hui une certaine probabilité de rejet du greffon par le patient. Cette probabilité augmente fortement ($>80\%$) lorsque ce dernier souffre d’hyperimmunité, c’est-à-dire que le système immunitaire du patient produit une quantité excessive d’anticorps, et notamment d’anticorps anti-HLA responsables du rejet de greffon.

Actuellement, les patients atteints d’hyperimmunité suivent un protocole d’immunoabsorption visant à réduire plus ou moins durablement la quantité d’anticorps anti-HLA dans leur corps, et ainsi à augmenter le taux de réussite de la greffe. En plus d’être lourd à supporter (deux semaines de prélèvements et traitements sanguin), ce protocole ne garantit pas le succès de l’opération même si le patient l’achève entièrement. Il semble donc crucial de pouvoir anticiper les effets des séances d’immunoabsorption sur chaque patient afin de lui prescrire un protocole adapté et si possible allégé.

À l’aide d’un indicateur de fluorescence de la quantité d’anticorps anti-HLA, le Mean Fluorescence Intensity (MFI), cette étude a donc cherché à modéliser l’évolution de cette quantité au cours du protocole d’immunoabsorption. Notamment, l’anticipation du moment où un patient descendrait en dessous d’un seuil déterminant de MFI permettant de décider ou non si les chances de succès d’une greffe seraient suffisantes.

Méthodes

Données

Nous avons travaillé à partir de données longitudinales prélevées sur dix patients hyperimmunisés en attente d’une greffe de rein à Grenoble. Chaque patient présente vingt relevés espacés sur deux semaines, soit deux relevés par séance avec généralement une pause le week-end.

Protocole

Actuellement, le protocole d’immunoabsorption se répartit en dix séances de désimmunisation étalées sur 2 semaines avec généralement une pause de deux jours le week-end. Chaque séance vise à faire baisser le taux d’anticorps anti-HLA chez un patient en prélevant son plasma et en le traitant à l’aide d’une colonne de désimmunisation. Ce taux est contrôlé via une à un indice de fluorescence de la colonne appelé le *Mean Fluorescence Index (MFI)* et le but à la fin des dix séances est de le faire descendre en dessous de 3 000, la norme pour Grenoble. Il est à noter que la norme au niveau national est un indice de MFI de 2 000.

Afin de suivre l’évolution de la quantité de MFI chez un patient, deux relevés sont réalisés à chaque séance, un juste avant le prélèvement et traitement du plasma et un autre à la fin de la séance.

Traitement des données

Afin de pouvoir exploiter les informations de chaque patient, nous avons commencé par traiter l’ensemble du jeu de données. Pour cela, nous avons créé un script R permettant de le mettre en forme de manière automatique par un script R, pour le jeu de données actuel mais aussi qu’il puisse s’adapter si les données venaient à s’alimenter.

La première étape a été de renommer toutes les variables du jeu de données afin de les raccourcir tout en pouvant les comprendre aisément lors de l’interprétation, notamment pouvoir différencier les variables des classes I et II qui sont autrement identiques.

Par la suite, le jeu de données fut nettoyé en commençant par le retrait des variables jugées “inutiles” car elles contenaient des informations répétitives ou des modalités constantes pour tous les individus.

Dans notre base, nous possédions aussi les informations de patients pour lesquels les valeurs de MFI étaient diluées pour cause de précision des prélèvements, nous avons alors fusionné les profils de patients dilués et originaux afin de ne conserver que les informations pertinentes.

Ensuite, un travail d'harmonisation des données a été réalisé afin de faire en sorte que toutes les modalités des variables soient comparables en termes de majuscules, termes utilisés et erreurs de saisies.

Dans la continuité de ces traitements, les valeurs nulles contenues dans le fichier ont dû être prises en charges afin de pouvoir utiliser les variables qui en contiennent dans les calculs de modèles ou les fonctions appelées en R. Quand il était possible, il a été décidé de remplacer les valeurs manquantes par les valeurs moyennes de la variable pour le patient en question. En revanche, il existe des cas où cette solution n'était pas envisageable : dans nos variables d'intérêt, car cela pourrait fausser les calculs ainsi que les conclusions tirées de cette étude. Il a alors été décidé de conserver ces variables telles quelles.

Enfin, nous avons jugé judicieux de calculer certaines variables dans le but d'analyser plus naturellement les données, par exemple en analysant l'âge au lieu de la date de naissance, ou alors analyser le jeu de données selon d'autres axes de lecture, comme avec le nombre de semaine de traitement, le temps passé depuis le début du traitement ainsi que l'évolution des MFI suite à chaque séance (variables que nous appelleront *deltaMFI* par la suite).

Pour terminer, nous avons choisi d'exporter le jeu de données nouvellement créé afin de pouvoir le réutiliser plus aisément à la fois dans des analyses R et dans le logiciel Monolix. Ces deux outils n'ayant pas exactement le même but, un jeu de données fut créé pour chaque utilisation. Le premier, destiné à R, fut exporté directement une fois le traitement terminé tandis que pour le second, à destination de Monolix, nous avons sélectionné les variables qui semblaient les plus intéressantes pour expliquer les variables d'intérêt dans les modèles générés.

Voici, en résumé dans le tableau suivant, les variables sélectionnées pour la sélection du modèle :

Nom	Description	Type/Modalités
Durée de la séance (min)	Durée de la séance = durée du traitement du plasma ? (en min)	Continue
Volume plasma traité	Volume de plasma traité lors de la séance d'immunoabsorption en question. (en mL)	Continue
Poids (kg)	Poids du patient lors de la séance. (en kg)	Continue
Taille (cm)	Taille du patient lors de la séance. (en cm)	Continue
Sexe	Sexe du patient.	Binaire (H/F)
Grossesses	Nombre de grossesses du patient antérieures au protocole.	Discrète (0 à 4 + NC pour les hommes)
Greffe antérieure	Nombre de greffes (de rein ?) effectué sur le patient avant le protocole.	Discrète (0 ; 1 ; 2)
AUTRES ?		
Temps*	Écoulement du temps au fil du protocole, du premier au dernier prélèvement. (en min)	Continue
delta MFI*	Variation de la quantité de MFI entre le début et la fin d'une même séance.	Continue

* : *covariables construites.*

Analyses statistiques

Outils utilisés : R et Monolix.

Analyses univariées

Dans un premier temps, une fois le nouveau jeu de données formé, nous avons effectué des analyses univariées (variable par variable) de chacune des variables avec la distribution des modalités et un graphique correspondant pour les variables qualitatives et le calcul des moyennes, variables et quartiles ainsi qu'un graphique adapté pour les variables quantitatives. Ce travail a été effectué dans le but de comprendre la population cible de notre jeu de données et de voir

les similarités et les différences principales des patients de notre étude.

Analyses bivariées

Une fois les premières analyses effectuées, nous avons essayé de déterminer l'existence de liens entre les variables d'intérêt et les covariables, et peut-être même des covariables entre-elles ce qui nous permettrait d'affirmer qu'elles apportent des informations communes et nous que nous pouvons alors réduire le nombre de variables à étudier.

Pour cela nous avons croisé chacune des variables avec les variables d'intérêt et avons conclu visuellement ou numériquement quand c'était possible et interprétable de l'attrait des variables étudiées entre elles.

Modélisation

Dans ce projet, nous avons utilisé des modèles non linéaires à effets mixtes pour décrire l'évolution longitudinale des quantités de MFI. Ces modèles sont utilisés quand des mesures d'un critère biologique répétées au cours du temps sont relevées sur une population de sujets.

Ils permettent de distinguer deux sources de variabilité : les effets fixes, qui reflètent la tendance moyenne au sein de la population, et les effets aléatoires, qui capturent l'hétérogénéité interindividuelle.

La forme générale du modèle s'écrit :

$$MFI_{i,j} = f(t_{ij}, \phi_i, v_i) + \epsilon_{ij}$$

où :

- $MFI_{i,j}$ est la j -ième mesure pour l'individu i ;
- f est une fonction non linéaire décrivant la dynamique du processus ;
- $\phi_i = A\beta + Bb_i$ sont les paramètres individuels de l'individu i , décomposés en effets fixes β et effets aléatoires b_i ;
- v_i représente les covariables associées à l'individu i ;
- $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ représente l'erreur résiduelle.

Modèle à un compartiment

Nous représentons la dynamique des MFI par un modèle à un compartiment, dans lequel les anticorps sont produits naturellement en continu. Les séances de désimmunisation sont modélisées comme des périodes durant lesquelles les MFI sont éliminés activement.

(Insérer un schéma du modèle compartimental ici)

La quantité de MFI dans le compartiment A suit la dynamique suivante :

$$\begin{aligned} \frac{dA}{dt} &= K_p \cdot A - \mu(t) \cdot K_e \cdot A \\ A(t_0) &= A_0 \end{aligned}$$

où :

- K_p est le taux de production des MFI ;
- K_e est le taux d'élimination pendant les séances ;
- $\mu(t)$ est une fonction indicatrice valant 1 pendant les séances et 0 en dehors ;
- A_0 est la quantité initiale de MFI.

Nous avons également considéré une formulation non linéaire basée sur la cinétique de Michaelis-Menten :

$$K(t) = \frac{V_{\max} \cdot MFI(t)}{K_m + MFI(t)}$$

où V_{\max} est la vitesse maximale de la réaction, et K_m la constante de Michaelis.

Trois variantes sont testées : saturation sur le taux de production, sur le taux d'élimination, ou sur les deux simultanément.

Modèle à deux compartiments

Ce modèle distingue deux compartiments : un compartiment de production (A) et un compartiment d'élimination (B), entre lesquels les MFI circulent.

(Insérer un schéma du modèle à deux compartiments ici)

Le système est modélisé par :

$$\begin{aligned}\frac{dA}{dt} &= K_p \cdot A - K_A \cdot A + K_B \cdot B \\ \frac{dB}{dt} &= K_A \cdot A - K_B \cdot B - \mu(t) \cdot K_e \cdot B \\ A(t_0) &= A_0, \quad B(t_0) = B_0\end{aligned}$$

où K_A et K_B représentent les taux de transfert entre les compartiments A et B.

Modèle *Kinetic-Pharmacodynamic* (K-PD)

Les modèles K-PD permettent de modéliser un effet pharmacodynamique sans mesurer explicitement les concentrations du médicament. Ils reposent sur l'introduction d'un compartiment fictif, représentant la "concentration virtuelle" d'un médicament induit par le traitement.

Chaque séance est modélisée comme une dose virtuelle administrée au début de la séance. Plusieurs schémas de dosage ont été testés :

- dose constante (1) à chaque séance ;
- dose croissante de 1 à 10 ;
- dose proportionnelle au volume de plasma traité ;
- dose proportionnelle à la durée de la séance.

Le système est modélisé comme suit :

$$\begin{aligned}\frac{dA}{dt} &= -K_{DE} \cdot A \\ \frac{dMFI}{dt} &= K_p \cdot \left(1 - \frac{IR}{EDK_{50} + IR}\right) - K_e \cdot MFI, \quad \text{avec } IR = K_{DE} \cdot A\end{aligned}$$

où K_{DE} est le taux d'élimination du compartiment virtuel, EDK_{50} le paramètre de sensibilité, et K_p , K_e les constantes de production et d'élimination.

Une reformulation permet une meilleure interprétation :

$$\begin{aligned}\frac{dA}{dt} &= -K_{DE} \cdot A \\ \frac{dMFI}{dt} &= K_p \cdot \left(1 - \frac{A}{A_{50} + A}\right) - K_e \cdot MFI\end{aligned}$$

où A_{50} est la quantité de médicament nécessaire pour inhiber de moitié la production des MFI.

Modèles d'inhibition de la croissance tumorale

Enfin, nous avons exploré des modèles issus de la modélisation de la croissance tumorale et de son inhibition (TGI), adaptés à notre problématique. Ces modèles présentent une dynamique proche de celle des MFI, en particulier lorsqu'on considère les MFI comme une "masse" pouvant croître ou décroître sous l'effet d'un traitement.

Le modèle général est :

$$\frac{dMFI}{dt} = g(MFI(t), K_p, x_i) - K_e \cdot MFI(t), \quad MFI(t_0) = MFI_0$$

où g représente la fonction de croissance, dépendante de K_p et d'autres paramètres spécifiques x_i , et K_e modélise l'effet d'élimination.

Estimation des paramètres

Les modèles considérés intègrent à la fois des effets fixes et des effets aléatoires. Dans un premier temps, seuls les paramètres des effets fixes seront estimés, afin de permettre la comparaison des modèles. L'estimation des effets aléatoires sera réalisée dans un second temps, uniquement sur le modèle final retenu. L'ensemble des estimations sera effectué à l'aide de l'algorithme **SAEM** (Stochastic Approximation Expectation-Maximization).

Sélection de modèles

Nous comparerons les modèles selon 3 critères :

- Critère d'information : BIC (Bayesian Information Criterion)
- Précision d'estimation des paramètres : Relative Standard Error (RSE%)
- Graphiques : Visual Predictive Check (VPC), Résidus, Individual fits

Résultats

Statistiques univariées/bivariées

- Profils des patients : taille, âge, sexe, etc.
- Boxplots des durée de séance par patient, du volume de plasma traité par patient (Cf présentation intermédiaire).
- Visualisation du nb de patients sous les 3000 MFI à chaque séance (Cf présentation intermédiaire => graphique à améliorer).

Graphiques pertinents par la suite

Modèles

Résultats des modèles (graphiques + indices de qualité/comparaison) :

- modèles K-PD.
- modèles TGI.

Meilleur modèle : lequel et pourquoi.

Discussion/Conclusion

Objectif de l'étude : modéliser l'évolution de la quantité de MFI au cours du protocole afin d'anticiper le moment où une greffe devient possible => réduire la charge du protocole pour les patients.

Principaux résultats :

- Quelle(s) covariable(s) influence(nt) l'évolution du taux de MFI ?
- Meilleur modèle : rappel de ses paramètres d'entrée + fiabilité

Limites de l'étude (hypothèses, a priori...) :

- Pas assez de patients/données ?
- Compréhension du système immunitaire encore trop limité ?

Ouverture : qu'est-ce qu'il pourrait être fait par la suite ?

- Améliorer la précision des prélèvements ?

Impact Environnemental et Sociétal

Cf site et pdf sur Moodle.

Impact environnemental personnel

- Trajets domicile-travail et autres déplacements.
- Consommation des équipements utilisés (ordinateurs fixes ou portables, temps d'utilisation serveur, etc.).
- Autres impacts ?

Impact global du projet

Impact environnemental

Impact sociétal

Politique de la structure d'accueil

Pas concerné pour ce projet. (?)