ISSN 2607-9917 www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026 https://doi.org/10.25796/bdd.v9i3.87120

Bulletin de la Dialyse à Domicile

Home Dialysis Bulletin (BDD)

Journal international bilingue pour partager les connaissances et l'expérience en dialyse à domicile

(Edition française) (English version available at same address)

Prédiction à l'aide de l'intelligence artificielle du passage en hémodialyse pour des patients en dialyse péritonéale présentant au moins une péritonite

(Predicting the Need for Hemodialysis in Peritoneal Dialysis Patients with at Least One Episode of Peritonitis Using Artificial Intelligence)

Alexander Watson[1], Rémy Lozano[1], Yannick Toussaint[2]

[1]Ecole des Mines (Nancy, France), [2]Laboratoire Lorrain de Recherche en Informatique et ses Applications (Nancy, France)

Pour citer : Watson A, Lozano R, Toussaint Y. Predicting the Need for Hemodialysis in Peritoneal Dialysis Patients with at Least One Episode of Peritonitis Using Artificial Intelligence. Bull Dial Domic [Internet]. [cited 2026 Aug. 29];9(3):115-26. doi: https://doi.org/10.25796/bdd.v9i3.87120

Résumé

Le transfert en hémodialyse est une cause importante d'arrêt définitif de traitement par dialyse péritonéale (DP). L'infection péritonéale est une cause fréquente de l'arrêt de la dialyse péritonéale et du transfert définitif en hémodialyse (HD). Une prédiction correcte de la date approximative de transfert permet de prévoir la préparation de l'abord vasculaire dans des conditions optimales. La prédiction correcte de la date de transfert revêt donc un intérêt clinique. Le but de ce travail est de construire un modèle d'intelligence artificielle qui prédit, étant donné l'historique d'infection péritonéale, si le patient doit être transféré en hémodialyse.

Notre approche repose sur une méthode d'apprentissage à base de réseau neuronal, un réseau neuronal récurrent (RNN) de type Long Short-Term Memory (LSTM) ; l'ensemble d'entraînement, représentant 80% du jeu de données exploité, soit 9 934 séquences patients, a été utilisé pour l'optimisation des paramètres du modèle. L'ensemble de test (10%, soit 1 242 séquences) a été quant à lui maintenu strictement isolé des autres ensembles et réservé exclusivement à l'évaluation finale des performances du modèle.
Le suivi de l'état des patients dans le temps était composé des épisodes d'infections péritonéales. Les réseaux de neurones récurrents se montrent particulièrement adaptés à ce type de prédiction à partir de données évoluant dans le temps. Ce travail a permis de mettre en avant les performances de prédiction d'un tel modèle sur ce type de données. L'introduction d'un biais dans l'apprentissage a permis de limiter les prédictions de faux négatifs.

Nos données, issues de la base de données du Registre de Dialyse Péritonéale de Langue Française (RDPLF), contiennent 15800 patients qui ont présenté au moins une péritonite, dans le but de prédire la possibilité de transfert en hémodialyse à la suite de cette complication. Avec notre modèle le taux de faux négatifs, le taux de patients qui ont été transférés en HD dans les 6 mois alors que le modèle ne le prédisait pas, était 5.3%.

Mots-clés : dialyse péritonéale, intelligence artificielle, réseau de neurones, péritonite, transfert, hémodialyse.

Summary

Transfer to hemodialysis is a major cause of permanent discontinuation of peritoneal dialysis (PD) treatment. Peritoneal infection is a common cause of discontinuation of peritoneal dialysis and permanent transfer to hemodialysis (HD). Accurately predicting the approximate date of transfer allows for the preparation of a vascular access under optimal conditions. Accurately predicting the date of transfer is therefore of clinical importance. The goal of this study is to develop an artificial intelligence model that predicts, based on a patient's history of peritoneal infection, whether the patient should be transferred to hemodialysis.

Our approach is based on a neural network learning method, specifically a Long Short-Term Memory (LSTM) recurrent neural network (RNN); the training set, representing 80% of the dataset—that is, 9,934 patient sequences—was used to optimize the model's parameters. The test set (10%, or 1,242 sequences) was kept strictly separate from the other sets and reserved exclusively for the final evaluation of the model's performance.

The monitoring of patients' conditions over time was based on peritoneal infections. Recurrent neural networks are particularly well-suited to this type of prediction based on time-series data. This study highlighted the predictive performance of such a model on this type of data. Introducing a bias into the training process helped limit false-negative predictions.

Our data, drawn from the French-Language Peritoneal Dialysis Registry (RDPLF) database, included 15,800 patients who had experienced at least one episode of peritonitis, with the aim of predicting the likelihood of transfer to hemodialysis following this complication. With our model, the false-negative rate—the proportion of patients who were transferred to HD within 6 months even though the model did not predict it—was 5.3%.

Keywords: peritoneal dialysis, artificial intelligence, neural network, peritonitis, transfer, hemodialysis.

www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026
https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

Introduction

D'après le RDPLF, 36% des causes d'interruption définitive de traitement par dialyse péritonéale sont liées à un transfert en hémodialyse. Cependant, seuls 8 % des patients sont porteurs d'une fistule artério-veineuse (FAV) en France [1]. Cela nécessite de recourir initialement à l'utilisation d'un cathéter central, dans la majorité des cas, faute d'abord vasculaire préalable [2]. Cette méthode augmente significativement le risque infectieux et le risque ultérieur de perte d'abord vasculaire [3].

La prédiction fiable de la date de passage en hémodialyse revêt un intérêt clinique : elle permettrait d'anticiper le transfert du patient vers l'hémodialyse, dans un délai suffisant pour créer une fistule artério-veineuse utilisable. Le temps de maturation d'une FAV est habituellement de 6 semaines ; néanmoins, en pratique, le temps nécessaire aux examens préalables, la disponibilité du chirurgien et la réservation du bloc opératoire nécessitent d'anticiper davantage. Selon l'avis des néphrologues, la possibilité de prédire le transfert en hémodialyse trois mois à l'avance donnerait les meilleures chance de débuter l'hémodialyse au moyen d'un FAV fonctionnelle et éviterait donc le recours à un cathéter central.

Une approche préventive consistant à poser systématiquement une fistule à l'ensemble des patients dès leur admission en DP pourrait sembler être une solution simple à ce problème. Mais cette option n'est pas cliniquement souhaitable pour deux raisons. D'une part, la fistule artério-veineuse peut avoir une durée de vie limitée en devenant non fonctionnelle au bout de quelques années, sans même avoir été utilisée. D'autre part, mêmes si elles demeurent marginales, des complications infectieuses ou hémodynamiques sont toujours possibles.

La survenue d'une ou plusieurs péritonites est responsable en France de 14 % des transferts en hémodialyse [1], souvent dans un contexte d'urgence. Une première tentative de prédiction du passage en hémodialyse basée sur un modèle de Random Forest avait été infructueuse, il a donc été nécessaire de faire appel à des techniques plus sophistiquées. Les techniques d'apprentissage profond par réseaux de neurones se montrent performants dans la prédiction de complications pour des patients en DP ; en particulier, les réseaux de neurones récurrents (Recurrent Neuronal Networks ou RNNs) sont utilisés pour la prédiction de séries temporelles [4–6]. Notre objectif était donc d'entraîner un RNN avec les données de patients traités par DP et ayant présenté au moins un épisode de péritonite, et d'évaluer ses performances pour la prédiction du risque de transfert en hémodialyse après une péritonite.

Matériel et Méthodes

Les données ont été extraites de la base de données du RDPLF, à partir des centres de France métropolitaine, après anonymisation complète à l'aide d'une clef de hachage générée aléatoirement. Tout le travail a été effectué en langage Python à l'aide des bibliothèques suivantes : *pandas*, *scikit-learn*, *tensorflow*/*keras*, *matplotlib* et *seaborn*.

Nous avons utilisé deux tables du jeu de données du RDPLF : la table *patient* contenait 40 952 patients ; la table *péritonite* représentait un sous-groupe de la table *patient* et concernait 15 800 patients présentant entre 1 et 20 épisodes de péritonite.

journal officiel du Registre de Dialyse Péritonéale de Langue Française RDPLF www.rdplf.org

www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026
https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

Après exclusion des données aberrantes ou manquantes, les variables suivantes renseignées pour chaque instant temporel ont été retenues : le centre de dialyse du patient, le sexe, le groupe néphrologique, la cause du traitement, le traitement précédent la DP, le premier type de DP, le type de DP à 90 jours, le statut d'inscription sur les listes de transplantation, la classe du germe de péritonite, le germe de péritonite, la cause de péritonite, le système et le type de traitement, le nombre de changement de type d'assistance du patient, le nombre total de péritonites du patient (à l'instant considéré), l'indice de Charlson en début de traitement, la taille, le poids, les changements de statut diabétique, le nombre de traitements, l'âge de la première DP et l'âge à l'instant considéré.

Plusieurs variables ont ensuite été ajoutées au jeu de données. Nous avons donc construit les variables suivantes : Durée DP, la durée passée par un patient en DP ; Delta Jours, indiquant le nombre de jours s'étant écoulés depuis le dernier épisode de péritonite ; Densité Infection 6 mois, donnant le nombre d'épisodes de péritonites pour le patient dans les 6 derniers mois passés en dialyse péritonéale.

Ensuite nous avons construit les séquences pour chaque patient de la base de données. Les séquences temporelles des patients n'étant pas directement disponibles dans la base de données, elles ont été construites à partir des dates des épisodes de péritonite, utilisées comme marqueurs temporels. À l'instant initial, l'ensemble des variables issues de la table patient sont renseignées et conservées à chaque instant de la séquence, ces variables étant par nature statiques. L'âge constitue la seule exception : il a été recalculé à chaque instant de la séquence par différence entre la date de l'épisode de péritonite correspondant et la date de naissance du patient. Les variables évoluant au cours du temps sont donc essentiellement celles issues de la table péritonite, ce qui implique que le modèle fonde ses décisions en grande partie sur les caractéristiques spécifiques aux épisodes de péritonite des patients. Le nombre d'épisodes de péritonite variant d'un patient à l'autre, les séquences présentent des longueurs hétérogènes. La longueur maximale observée étant de 20 épisodes, cette valeur a été retenue comme longueur fixe pour l'ensemble des séquences. Un *post-padding* a été appliqué aux séquences de longueur inférieure. La technique de *post-padding* complète les séquences de longueur inférieure à 20 par des valeurs nulles en fin de séquence, qui sont exclues du calcul lors de la phase d'apprentissage via l'indice de masquage réservé à cet effet. Les valeurs ajoutées n'ont pas d'influence sur le calcul de l'erreur durant l'apprentissage mais elles permettent aux séquences de différentes longueurs d'être traitées de la même façon par le modèle.

Les variables catégorielles du jeu de données nécessitaient un encodage spécifique. Les réseaux de neurones récurrents ne pouvant pas traiter directement des variables catégorielles, une représentation vectorielle de ces dernières a été mise en œuvre. Cette étape concerne l'ensemble des variables catégorielles de la base d'entraînement, telles que la variable PreTypDP, qui encode le type de dialyse péritonéale pratiquée initialement (DP mixte, DPCA, DPA quotidienne, DPI). Un encodage catégoriel (label *encoding*) indépendant a préalablement été appliqué à chaque variable catégorielle, assignant à chacune de ses modalités un entier unique *(Tableau I)*, avec un décalage de 1 afin de réserver l'indice 0 au padding, utilisé pour l'alignement des séquences de longueur variable. L'index de chaque valeur de variable a ensuite été converti en représentation vectorielle. L'algorithme Word2Vec, dans sa configuration Skip-gram, a ainsi été entraîné exclusivement sur les séquences issues des données d'entraînement, afin d'éviter toute fuite d'information. Le modèle Skip-gram apprend ainsi, pour chaque token (index), à prédire les tokens présents

journal officiel du Registre de Dialyse Péritonéale de Langue Française RDPLF www.rdplf.org

www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026
https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

dans une fenêtre contextuelle de taille 5 (à l'aide de 5 événements voisins), capturant de ce fait les relations de co-occurrence entre différentes valeurs de variables au sein des trajectoires cliniques des patients. Chaque valeur catégorielle est ainsi représentée par un vecteur dense de dimension 50 *(Tableau II)*. La dimension a été choisie afin de permettre de représenter de façon assez différenciée les catégories présentes dans les données sans les rendre trop volumineuses durant l'entraînement du modèle. Une fois initialisées, les représentations vectorielles des valeurs de variables catégorielles sont injectées sous forme de matrice de poids dans la couche d'entrée du réseau neuronal (*embedding*). Ces poids évoluent au cours de l'entraînement du modèle afin d'ajuster la représentation des catégories (*fine-tuning*).

Tableau I. Exemple d'encodage catégoriel pour la variable PreTypDP

Valeur	Token (Index)
Aucune donnée (*padding*/masque)	0
DP mixte (CC & CA)	1
DPA quotidienne	2
DPCA	3
DPI	4

DPCA : dialyse péritonéale continue ambulatoire, DPA : dialyse péritonéal automatisée quotidienne, DPI : dialyse péritonéale automatisée intermittente, PreTypDP : type de DP initiale (DPCA, DPA ou mixte), DP mixte (CC et CA) : association DPCA et DPA

Tableau II. Dimensions par valeur pour PreTypDP

Valeurs	Dim 1	Dim 2	Dim 3	Dim 4	…	Dim 50
DP mixte (CC & CA)	-0.0593	0.0272	0.0170	0.0284	…	0.0082
DPA quotidienne	-0.0425	0.0217	-0.0088	0.0080	…	0.0034
DPCA	0.0046	-0.0024	0.0120	0.0182	…	0.0103
DPI	0.0156	-0.0190	-0.0004	0.0069	…	-0.0190

DPCA : dialyse péritonéale continue ambulatoire, DPA : dialyse péritonéal automatisée quotidienne, DPI : dialyse péritonéale automatisée intermittente, PreTypDP : type de DP initiale (DPCA, DPA ou mixte), DP mixte (CC et CA) : association DPCA et DPA

Une variable cible binaire a été introduite afin de superviser l'apprentissage du modèle. À chaque instant de la séquence, c'est-à-dire à chaque épisode de péritonite, cette variable indique si le patient a effectué un passage en hémodialyse dans les six mois suivant l'épisode concerné. Elle prend la valeur 1 en cas de passage en HD dans cette fenêtre temporelle, et 0 dans le cas contraire. Ainsi, une prédiction positive émise lors d'un épisode de péritonite permet, dans le meilleur des cas, d'anticiper le passage en HD avec jusqu'à six mois en avance, offrant un délai suffisant pour planifier la pose de la fistule.

Nous avons retenu pour l'architecture de notre modèle un RNN de type LSTM [4,7]. Ce modèle repose sur deux flux d'entrée distincts *(Figure 1)* : les variables numériques préalablement standardisées et les variables catégorielles sous forme de vecteur de dimension 50. Ces deux flux sont concaténés à chaque pas de temps pour former un vecteur d'état, constituant l'entrée de la couche récurrente. La séquence est ensuite traitée par la couche LSTM qui produit un état caché à chaque pas de temps et préserve ainsi l'intégralité de l'information temporelle intermédiaire. Afin de limiter le sur-apprentissage, nous avons appliqué à cette couche une régularisation de type L2, qui introduit la somme des carrés des poids comme terme de pénalité, et un *dropout*, qui fige aléatoirement une proportion fixe des poids des entrées à chaque pas de temps. La couche LSTM est suivie d'une couche dense dont la fonction d'activation est la tangente hyperbolique

journal officiel du Registre de Dialyse Péritonéale de Langue Française RDPLF www.rdplf.org

www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026
https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

(Figure 1). Un *dropout* et un facteur de régularisation de type L2 sont de nouveau appliqués pour cette couche dense *(Figure 1)*. La couche de sortie est une couche dense suivie d'une fonction d'activation sigmoïdale afin d'obtenir une probabilité de passage en HD en sortie pour chaque pas de temps de la séquence d'entrée *(Figure 1)*. Ainsi, chaque prédiction intègre l'ensemble des informations disponibles jusqu'à l'instant considéré, incluant les états cachés propagés par la couche LSTM depuis les pas de temps précédents.

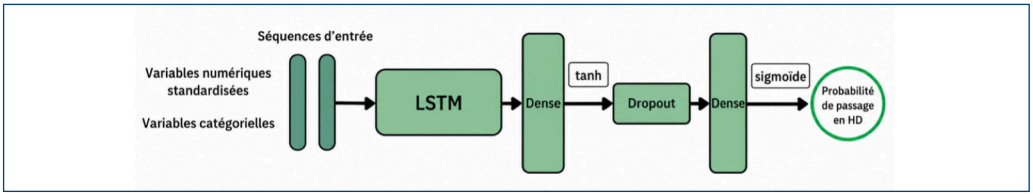

Figure 1. Architecture du modèle

Avant d'entraîner le modèle, nous avons partitionné le jeu de données en trois ensembles de patients distincts, tout en conservant la même proportion de patients transférés en HD dans chaque ensemble. L'ensemble d'entraînement, représentant 80% du jeu de données exploité, soit 9 934 séquences patients, est utilisé pour l'optimisation des paramètres du modèle. L'ensemble de validation (10%, soit 1 242 séquences) permet de surveiller la progression de l'apprentissage. L'ensemble de test (10%, soit 1 242 séquences) est quant à lui maintenu strictement isolé des deux autres ensembles et réservé exclusivement à l'évaluation finale des performances du modèle. Enfin, la partition a été construite de façon à assurer une représentation suffisante des patients ayant effectué un passage en hémodialyse au sein de chacun des trois ensembles.

L'entraînement du modèle a ensuite été réalisé à partir de cette partition du jeu de données. Le nombre d'époques, lorsque l'intégralité de l'ensemble d'entraînement a été parcouru, a été fixé à 50. Lors de chaque époque, les poids du modèle sont optimisés selon l'algorithme Adam de descente du gradient stochastique. Le gradient est calculé à partir d'une fonction de perte qui évalue l'écart entre les prédictions du modèle sur l'ensemble d'entraînement et les vraies valeurs de cet ensemble. Nous avons retenu comme fonction de perte l'entropie-croisée binaire pondérée (ECBP) [8,9]. Elle est basée sur l'entropie-croisée binaire standard (ECB), mais elle adresse le déséquilibre des catégories prédites, car le passage en HD n'est associé qu'à un seul instant de la séquence pour la majorité des patients. Elle pénalise également plus fortement les erreurs de prédiction de passage en HD.

$$\mathrm{ECB} = p_1 \times \log(p_1) - p_0 \times \log(p_0)$$

$$ECBP = ECB \times [\, y_{réel} \times 3.88 + (1 - y_{réel}) \,]$$

Où p_1 représente la probabilité de passage en HD, $p_0 = 1 - p_1$ et $y_{réel}$ représente la valeur réelle de la variable cible. Le terme 3.88 est le ratio entre le nombre d'exemples négatifs et le nombre d'exemples positifs dans l'ensemble d'entraînement.

Pendant l'entraînement le critère d'évaluation était l'aire sous la courbe ROC (AUC, *Area Under the Curve*) [10] qui traduit la discrimination du modèle, c'est-à-dire la capacité du modèle à distinguer les deux classes. Nous avons également ajouté un mécanisme d'*early stopping* qui arrêtait prématurément l'entraînement après 5 époques sans amélioration de l'AUC. Les performances du modèle ont également été évaluées à travers sa calibration de façon graphique.

journal officiel du Registre de Dialyse Péritonéale de Langue Française RDPLF www.rdplf.org

www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026
https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

Le choix des hyperparamètres *(Tableau III)* s'est fait de façon empirique. Nous avons entraîné le modèle plusieurs fois avec différentes valeurs des hyperparamètres tout en moyennant les résultats sur cinq entraînements afin de les rendre fiables. Les valeurs de dropout représentent la proportion des poids figés de la couche associée du modèle *(Tableau III, Figure 1)*. Les facteurs de régularisation se rapportent à la couche LSTM et à la couche dense en sortie de LSTM.

Tableau III. Valeurs retenues des hyperparamètres pour l'entraînement du modèle

Hyperparamètres	**Valeur retenue**
***Dropout* du LSTM**	0.1
Facteur de régularisation L2 du LSTM	0.001
Facteur de régularisation L2 de la couche dense	0.001
***Dropout* après la couche dense**	0.1
***Learning rate* de l'optimiseur Adam**	0.001

LSTM : Long Short-Term Memory

Résultats

L'entraînement a été effectué sur 5 itérations successives indépendantes. Les prédictions finales sur l'ensemble de test ont été obtenues en calculant la moyenne arithmétique des probabilités de sortie du modèle.

Le modèle renvoie pour chaque patient une séquence de probabilités de passage en HD après chaque épisode d'infection péritonéale du patient. Nous avons également évalué les prédictions du modèle en utilisant différents seuils de décision de passage en HD *(Tableau IV)*. Ainsi toute probabilité prédite au-dessus du seuil est comptée comme un passage en HD.

Tableau IV. Valeurs obtenues des différents critères selon la fonction de perte et le seuil choisis

Fonction de perte (*Loss*)	**ECBP**	**ECBP**	**ECB**	**ECB**
Seuil de décision	0.5	Youden (0.35)	0.5	Youden (0.35)
Critères				
AUC	0.946	0.946	0.949	0.949
Sensibilité	0.881	0.947	0.637	0.910
Spécificité	0.842	0.802	0.956	0.856

ECBP : entropie croisée binaire pondérée
ECB : entropie croisée binaire
AUC : Area Under Curve, aire sous la courbe ROC
Youden : seuil de décision optimal au sens de Youden

L'indice de Youden, est un critère dérivé de la courbe ROC *(Figure 2 et Figure 3)* permettant de déterminer le seuil de décision optimal d'un modèle de classification binaire [11]. Il est défini comme la somme de la sensibilité et de la spécificité, diminuée de ***1 : J=sensibilité+spécificité - 1***, ce qui est équivalent à maximiser la différence entre le taux de vrais positifs et le taux de faux positifs. Il varie entre 0 et 1 : une valeur de 0 indique des performances équivalentes à un classifieur aléatoire, tandis qu'une valeur de 1 correspond à une classification parfaite. Le seuil optimal est celui qui maximise cet indice sur la courbe ROC, offrant ainsi le meilleur compromis entre sensibilité et spécificité.

journal officiel du Registre de Dialyse Péritonéale de Langue Française **RDPLF** www.rdplf.org

www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026
https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

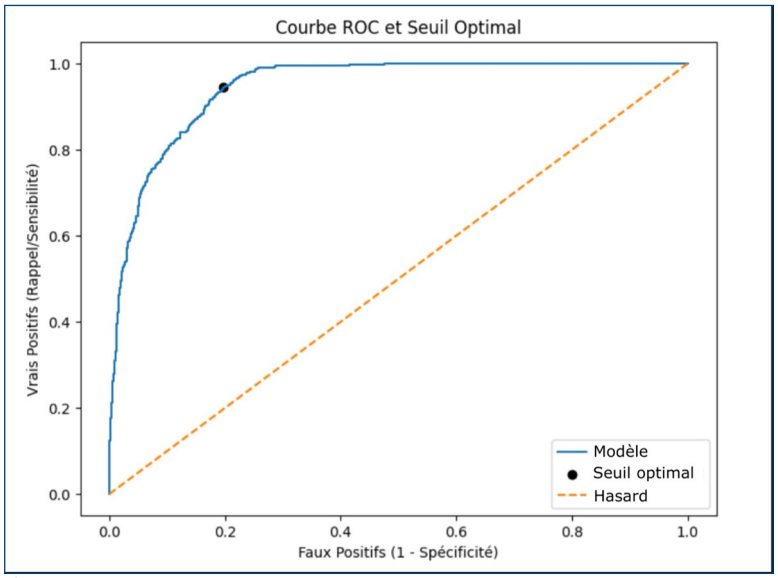

Figure 2. Courbe ROC pour le modèle avec ECBP
Le point sur les courbes ci-dessus indique les taux de vrais positifs et de faux positifs prédits pour le seuil optimal de Youden.

La courbe ROC (*Receiver Operating Characteristic*) *(Figure 2 et Figure 3)* est construite en représentant, pour chaque seuil de décision possible, le taux de vrais positifs (sensibilité) en ordonnée en fonction du taux de faux positifs (1 - *spécificité*) en abscisse. Les seuils testés correspondent à l'ensemble des probabilités de passage en hémodialyse renvoyées par le modèle. Une droite diagonale d'équation $y = x$, passant par l'origine, est tracée afin de représenter les performances d'un classifieur aléatoire, pour lequel le taux de vrais positifs est égal au taux de faux positifs quel que soit le seuil considéré. Une courbe ROC située en dessous de cette diagonale indiquerait des performances inférieures au hasard, le modèle produisant alors davantage de faux positifs que de vrais positifs pour les seuils concernés. L'objectif est donc d'obtenir une courbe ROC la plus éloignée possible de cette diagonale, vers le coin supérieur gauche du graphique.

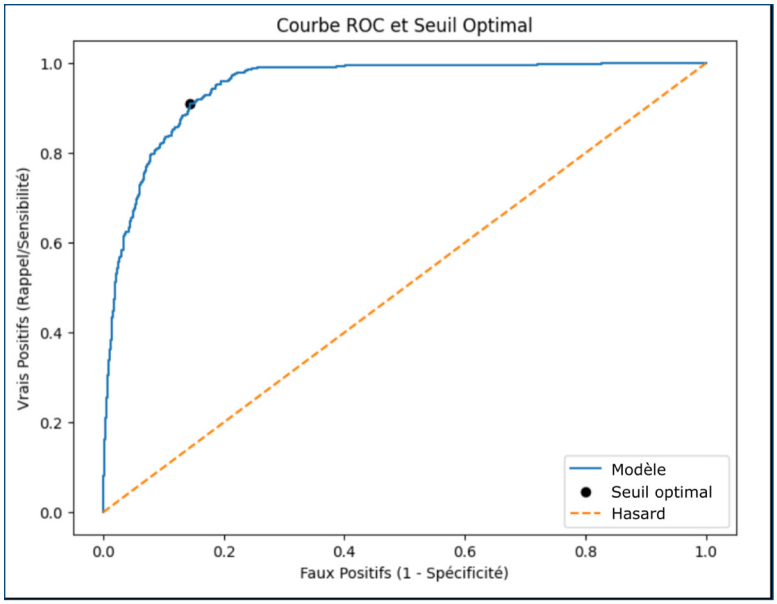

Figure 3. Courbe ROC pour le modèle avec ECB

www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026
https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

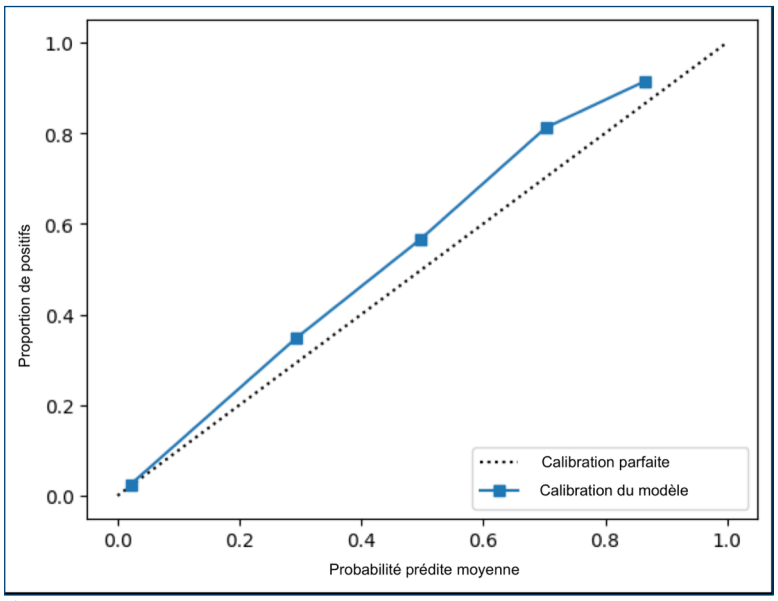

Figure 4. Courbe de calibration pour le modèle avec ECB

La courbe de calibration *(Figure 4 et Figure 5)* une représentation graphique permettant de confronter les probabilités prédites par le modèle aux fréquences de passage en hémodialyse effectivement observées dans l'ensemble de test. Les prédictions sont réparties en cinq quantiles, et pour chacun d'eux, la probabilité prédite moyenne est mise en regard de la fréquence observée correspondante. La diagonale $y = x$ matérialise une calibration parfaite, pour laquelle les probabilités prédites coïncident exactement avec les fréquences observées. Une déviation vers la partie inférieure droite du graphique indique que le modèle surestime le risque de passage en hémodialyse, tandis qu'une déviation vers la partie supérieure gauche traduit une sous-estimation de ce risque.

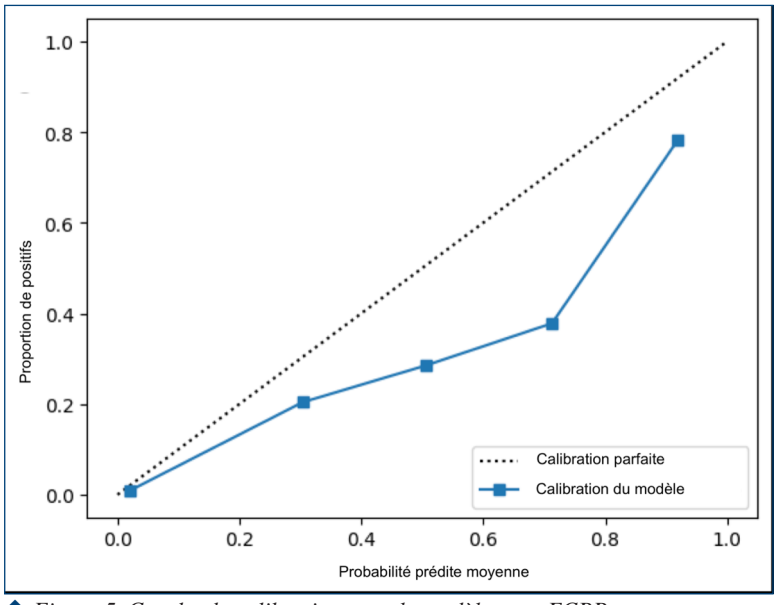

Figure 5. Courbe de calibration pour le modèle avec ECBP

journal officiel du Registre de Dialyse Péritonéale de Langue Française RDPLF www.rdplf.org

www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026
https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

Discussion

La survenue d'une péritonite en dialyse péritonéale fait courir un risque de transfert en hémodialyse important, mais ce dernier dépend à la fois des caractéristiques des patients et de l'infection. Cette étude a permis de montrer qu'il est possible, après entrainement d'un réseau neuronal, de prédire la nécessité de transfert en hémodialyse, après infection péritonéale, avec une fiabilité suffisante pour que son application clinique soit envisageable, et aider la décision médicale.

L'une des priorités au cours de l'entraînement du modèle était de limiter tout risque de sur-apprentissage où les capacités de généralisation de prédiction du modèle seraient dégradées. Nous avons ainsi utilisé des techniques de dropout et *d'early stopping*, pour améliorer la généralisation, et de régularisation L2, pour limiter la sensibilité du modèle au bruit *(Tableau III)*.

Le choix de la fonction de perte ECBP se justifie par la volonté de réduire le nombre de faux négatifs au détriment d'un nombre de faux positifs plus élevé. D'après le *Tableau IV*, la sensibilité augmente avec l'utilisation de l'ECBP à la place de l'ECB, indépendamment du seuil choisi. À l'inverse de la sensibilité, la spécificité baisse lorsque l'ECBP est utilisée *(Tableau IV)*.

Durant l'apprentissage, le seuil de décision retenu était le seuil optimal au sens de Youden lorsque celui-ci était inférieur à 0.5, et 0.5 dans le cas contraire. L'utilisation d'un seuil inférieur à 0.5 revient à considérer qu'une probabilité prédite modérée est suffisante pour déclencher une prédiction de passage en HD, rendant ainsi le modèle plus sensible aux faibles variations de probabilité. Ce choix est cohérent avec l'objectif de minimisation des faux négatifs formulé précédemment : en abaissant le seuil de décision, le modèle favorise la détection des cas positifs à nouveau au détriment d'un taux de faux positifs plus élevé (taux de faux positifs de 15.8% avec un seuil à 0.5 contre 19.8% avec le seuil de Youden, pour l'ECBP, *Tableau IV*).

Les performances d'un modèle de classification binaire peuvent être évaluées selon deux dimensions complémentaires : la discrimination et la calibration [12,13].

L'AUC constitue un critère de sélection de modèle [7,14]. Plus l'AUC est élevée, plus la courbe ROC *(Figure 2 et Figure 3)* s'éloigne de la diagonale du hasard, ce qui traduit de meilleures performances discriminantes du modèle. Formellement, l'AUC est interprétable comme la probabilité que le score prédit pour un exemple positif soit supérieur à celui prédit pour un exemple négatif [15]. La discrimination mesure la capacité du modèle à distinguer efficacement les deux classes. Dans notre cas, l'AUC atteint 0.946 et 0.949 (en fonction de la configuration, *Tableau IV*), des valeurs proches de 1, indiquant une excellente capacité de discrimination.

La calibration, en revanche, évalue l'adéquation entre les probabilités prédites par le modèle et les fréquences observées dans les données réelles [16]. La courbe associée à la configuration avec ECB *(Figure 4)* dévie vers la partie supérieure gauche, indiquant que le modèle sous-estime le risque de passage en HD, ce qui se traduit par un taux élevé de faux négatifs. À l'inverse, la courbe associée à la configuration avec ECBP *(Figure 5)* dévie vers la partie inférieure droite, reflétant une surestimation du risque, qui se manifeste par un taux de faux positifs plus élevé. Cet écart indique que le modèle sur-estime les probabilités de passage en HD par rapport aux observations réelles. Autrement dit, le modèle tend à être trop confiant dans ses prédictions

journal officiel du Registre de Dialyse Péritonéale de Langue Française RDPLF www.rdplf.org

www.bdd.rdplf.org Volume 9, n° 3, Septembre 2026
https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

positives [17], ou à privilégier des erreurs de type faux positifs (patients incorrectement classés comme devant passer en HD) plutôt que des faux négatifs (patients nécessitant une hémodialyse mais non identifiés). Cependant, ce biais permet de minimiser le taux de faux négatifs (11.9% avec ECBP contre 36.3%, lorsque le seuil de décision est fixé à 0.5, *Tableau IV*), une priorité dans notre contexte clinique, où l'omission de patients nécessitant une hémodialyse est plus critique que l'inclusion erronée de patients ne l'exigeant pas.

Des limitations de ce modèle émergent à travers la partition des données. En particulier, les patients sont répartis aléatoirement, indépendamment de leur centre de provenance, entre les 3 ensembles de données utilisés. Cela signifie que le modèle peut apprendre des spécificités des pratiques de chaque centre concernant le traitement en DP et le passage en HD des patients. Il s'agissait d'un choix volontaire dans la mesure où les prédictions s'adaptent aussi aux pratiques usuelles des différents centres.

Conclusion

Les techniques d'intelligence artificielle et d'apprentissage automatique se montrent efficaces dans la prédiction de passage en HD de patients en dialyse péritonéale après survenue d'une ou plusieurs péritonites. La stratégie d'entraînement du modèle visait à limiter le taux de faux négatifs. Néanmoins un équilibre a dû être trouvé entre une prédiction confiante du modèle et la représentation adéquate de la réalité à travers cette prédiction.

Cette approche permet d'envisager le développement d'une application, à tester par une validation externe, afin de déterminer si elle permet de réduire le taux de transfert non programmé et le risque de devoir recourir à un cathéter central.

Contributions des auteurs

Tous les auteurs ont contribué à la conception de l'étude.

AW : rédaction, construction du modèle, analyse des résultats. RL : rédaction, construction du modèle, analyse des résultats. YT : conseils dans la constrution du modèle, l'analyse des résultats et la rédaction.

Considérations éthiques

Cette étude a été menée conformément aux principes éthiques de la Déclaration de Helsinki. Le consentement éclairé n'a pas été requis pour cette étude car ce sont les données rétrospectives d'un registre pour lequel l'autorisation d'inclusion a été demandée aux patients ; les données des patients et celles relatives aux institutions ont été entièrement anonymisées et analysées de manière rétrospective. Les identifiants patients et centres ont été préalablement anonymisés au moyen d'une clef de hachage générée aléatoirement.

Financement

Aucun financement spécifique n'a été reçu pour ce travail.

Remerciements

Les auteurs remercient le Dr Christian Verger (RDPLF) et le Dr Jacques Chanliau (RDPLF) pour leur aide tout au long de ce travail, le Dr Emmanuel Fabre pour l'exportation anonymisée des données à partir de la base de données du RDPLF.

Conflits d'intérêts

Les auteurs déclarent n'avoir aucun conflit d'intérêt avec ce travail.

Disponibilités des données

Les ensembles de données générés et analysés au cours de la présente étude sont disponibles auprès des auteurs et du RDPLF sous demande raisonnable.

https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

Déclaration relative à l'intelligence artificielle

Les auteurs déclarent que ce manuscrit est le fruit d'un travail personnel et original. Aucun outil ou application d'intelligence artificielle n'a été utilisé en dehors de ceux conçus par les auteurs eux mêmes, qui faisait le sujet de ce travail, pour l'analyse des données et la génération des résultats.

ORCID iDs
Alexander Watson : https://orcid.org/0009-0008-5143-7882
Rémy Lozano : https://orcid.org/0009-0005-7138-136X
Yannick Toussaint : https://orcid.org/0009-0005-0507-7204

Références

1. Verger C, Fabre E, Veniez G, Padernoz MC. Synthetic 2018 data report of the French Language Peritoneal Dialysis and Home Hemodialysis Registry (RDPLF) Bull Dial Domic Internet. 2019 Apr 10;2(1):1–10.doi https://doi.org/10.25796/bdd.v2i1.19093

2. Cheikh Hassan HI, Murali KM, Chen JHC, Mullan J. Back-up arteriovenous fistula in peritoneal dialysis patients: a retrospective cohort study of long-term meaningful outcomes. Intern Med J. 2026 May 23;imj.70493. doi: https://doi.org/10.1111/imj.70493.

3. Raskin D, Vachharajani TJ, Partovi S, Khan A, Lyden SP, Kirksey L. Value-Based Model for Vascular Access Management in the End-Stage Kidney Disease Population. Semin Dial. 2026 Jan;39(1–2):15–25. doi: https://doi.org/10.1111/sdi.70024 .

4. Hsu FY, Hwang RH, Tsai MH, Wang JT. Predicting Peritoneal Dialysis Failure Within the Next Three Months Based on Deep Learning and Important Features Analysis. Information. 2024 Dec 5;15(12):776. doi: https://doi.org/10.3390/info15120776

5. Tangri N, Ansell D, Naimark D. Predicting technique survival in peritoneal dialysis patients: comparing artificial neural networks and logistic regression. Nephrol Dial Transplant. 2008 Apr 3;23(9):2972–81. doi: https://doi.org/10.1093/ndt/gfn187.

6. Bai Q, Tang W. Artificial intelligence in peritoneal dialysis: general overview. Ren Fail. 2022 Dec 31;44(1):682–7. doi: https://doi.org/10.1080/0886022X.2022.2064304.

7. Ioannou GN, Tang W, Beste LA, Tincopa MA, Su GL, Van T, et al. Assessment of a Deep Learning Model to Predict Hepatocellular Carcinoma in Patients With Hepatitis C Cirrhosis. JAMA Netw Open. 2020 Sep 1;3(9):e2015626. doi: https://doi.org/10.1001/jamanetworkopen.2020.15626.

8. Sugino T, Kawase T, Onogi S, Kin T, Saito N, Nakajima Y. Loss Weightings for Improving Imbalanced Brain Structure Segmentation Using Fully Convolutional Networks. Healthcare. 2021 Jul 26;9(8):938. doi: https://doi.org/10.3390/healthcare9080938.

9. Yeung M, Sala E, Schönlieb CB, Rundo L. Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation. Comput Med Imaging Graph. 2022 Jan;95:102026. doi: https://doi.org/10.1016/j.compmedimag.2021.102026

10. Fang J, Xu Y, Zhao Y, Yan Y, Liu J, Liu J. Weighing features of lung and heart regions for thoracic disease classification. BMC Med Imaging. 2021 Dec;21(1):99. doi: https://doi.org/10.1186/s12880-021-00627-y.

11. Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden Index and Optimal Cut-Point Estimated from Observations Affected by a Lower Limit of Detection. Biom J. 2008 Jun;50(3):419–30. doi: https://doi.org/10.1002/bimj.200710415

12. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures. Epidemiology. 2010 Jan;21(1):128–38. doi: https://doi.org/10.1097/EDE.0b013e3181c30fb2

13. Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. Stat Med. 2014 Feb 10;33(3):517–35. doi: https://doi.org/10.1002/

journal officiel du Registre de Dialyse Péritonéale de Langue Française **RDPLF** www.rdplf.org

https://doi.org/10.25796/bdd.v9i3.87120
ISSN 2607-9917

sim.5941
14. Whitney HM, Drukker K, Giger ML. Performance metric curve analysis framework to assess impact of the decision variable threshold, disease prevalence, and dataset variability in two-class classification. J Med Imaging. 2022 May 31;9(03). doi: https://doi.org/10.1117/1.JMI.9.3.035502
15. Li J. Area under the ROC Curve has the most consistent evaluation for binary classification. Qin H, editor. PLOS ONE. 2024 Dec 23;19(12):e0316019. doi: https://doi.org/10.1371/journal.pone.0316019
16. Friesacher HR, Engkvist O, Mervin L, Moreau Y, Arany A. Achieving well-informed decision-making in drug discovery: a comprehensive calibration study using neural network-based structure-activity models. J Cheminformatics. 2025 Mar 5;17(1):29. doi: https://doi.org/10.1186/s13321-025-00964-y
17. Zhang F, Dvornek N, Yang J, Chapiro J, Duncan J. Layer Embedding Analysis in Convolutional Neural Networks for Improved Probability Calibration and Classification. IEEE Trans Med Imaging. 2020 Nov;39(11):3331–42. doi: https://doi.org/10.1109/TMI.2020.2990625

*journal officiel du **R**egistre de **D**ialyse **P**éritonéale de **L**angue **F**rançaise* **RDPLF** www.rdplf.org