

Intelligence Artificielle Et Prédiction Précoce De L'insuffisance Rénale Aiguë : Promesse Clinique, Limites Méthodologiques Et Perspectives Pour La Néphrologie

Tebay Nohaa El Mouttaki Abba² Armi Sara² El Hassani Maha² Belyazid Ismail²
Centre Hospitalier Universitaire Cheikh Zayd, Rabat, Service De Néphrologie, Dialyse Et Transplantation Rénale.

Résumé

Introduction : L'insuffisance rénale aiguë (IRA) reste une complication fréquente et grave en milieu hospitalier, associée à une augmentation de risque de mortalité, de la durée de séjour, du recours à l'épuration extrarénale et de progression vers la maladie rénale chronique. Les critères diagnostiques actuels reposent essentiellement sur l'élévation de la créatinine sérique et la baisse de la diurèse, deux marqueurs souvent tardifs. L'intelligence artificielle (IA), en particulier le machine learning et le deep learning, pourrait permettre d'identifier les patients à risque avant l'installation clinique de l'IRA.

Objectif : Décrire les principales applications de l'IA dans la prédiction de l'IRA, résumer les performances rapportées, discuter les limites méthodologiques et proposer les conditions nécessaires à une intégration clinique raisonnable en néphrologie.

Méthodes : Une revue narrative structurée de la littérature récente a été réalisée à partir d'articles publiés entre 2018 et 2026, portant sur les modèles d'IA appliqués à la prédiction de l'IRA en milieu hospitalier, périopératoire, réanimation adulte, réanimation pédiatrique, sepsis, chirurgie cardiaque, exposition aux produits de contraste et transplantation.

Résultats : Les modèles d'IA rapportent globalement de bonnes performances, avec des AUROC souvent comprises entre 0,75 et 0,90 selon les contextes. Une méta-analyse périopératoire incluant 19 études et 304 076 patients a rapporté une sensibilité combinée de 0,77 et une spécificité de 0,75 pour la prédiction de l'IRA par IA (1). Les modèles les plus prometteurs exploitent les dossiers médicaux électroniques, les données biologiques répétées, la diurèse, les constantes vitales et les traitements néphrotoxiques.

Conclusion : L'IA pourrait devenir un outil d'aide à la décision utile pour prévenir l'IRA, mais son intérêt clinique ne peut être déduit de la seule performance statistique. La validation externe, l'interprétabilité, l'évaluation prospective, la réduction des biais et l'intégration dans le flux de soins sont indispensables avant une diffusion large.

Mots-clés : insuffisance rénale aiguë; intelligence artificielle ; machine learning ; deep learning ; néphrologie; prédiction ; dossier médical électronique.

Date of Submission: 11-07-2026

Date of Acceptance: 21-07-2026

I. Introduction

L'insuffisance rénale aiguë est un syndrome fréquent, hétérogène et potentiellement sévère, rencontré dans de nombreux contextes hospitaliers. Elle complique notamment les séjours en réanimation, la chirurgie cardiaque, le sepsis, la transplantation, l'exposition aux produits de contraste et les traitements néphrotoxiques. Les conférences KDIGO ont rappelé que l'IRA demeure un enjeu majeur en raison de son association avec la mortalité, la morbidité et les séquelles rénales à long terme (2). 2

Le diagnostic classique de l'IRA repose sur la créatinine sérique et la diurèse. Ces marqueurs sont accessibles et largement utilisés, mais ils ne permettent pas toujours d'identifier précocement les patients à risque. La créatinine augmente souvent après l'atteinte rénale, tandis que la diurèse peut être influencée par l'état hémodynamique, les traitements diurétiques, la volémie et le contexte clinique.

Cette limite explique l'intérêt croissant pour des outils capables de prévoir l'IRA avant son expression biologique complète. L'IA permet d'analyser de grands volumes de données, d'intégrer des variables répétées dans le temps et d'identifier des relations non linéaires difficiles à détecter par les approches statistiques traditionnelles. Dans une cohorte hospitalière de plus de 121 000 patients, Koyner et al. ont montré qu'un modèle de gradient boosting pouvait prédire une IRA KDIGO stade 2 avec une AUROC de 0,90 à 24 heures et de 0,87 à 48 heures (3). 3

L'enthousiasme pour ces outils doit cependant rester prudent. Les performances publiées proviennent souvent d'études rétrospectives, monocentriques ou fondées sur des bases de données spécifiques. La transposition d'un modèle d'un hôpital à un autre expose à une perte de performance liée aux différences de population, de pratiques, de codage et de qualité des données. Kelly et al. soulignent que la réussite clinique de l'IA dépend non seulement de l'exactitude du modèle, mais aussi de son évaluation clinique, de sa régulation, de son intégration dans les soins et de sa capacité à éviter les biais (4). 4

L'objectif de cet article est de proposer une synthèse structurée et originale des applications actuelles de l'IA dans la prédiction précoce de l'IRA, en insistant sur les résultats cliniquement pertinents, les limites et les perspectives pour la pratique néphrologique.

II. Méthodes

Il s'agit d'une revue narrative structurée de la littérature récente. Les articles retenus concernaient l'utilisation de l'IA, du machine learning ou du deep learning pour prédire l'IRA, l'IRA sévère, le recours à l'épuration extrarénale ou le pronostic des patients après IRA.

Les travaux sélectionnés couvraient plusieurs contextes : hospitalisation générale, réanimation adulte, réanimation pédiatrique, chirurgie cardiaque, période périopératoire, sepsis, produit de contraste, transplantation hépatique et cardiologie interventionnelle. Les principaux éléments analysés étaient la population étudiée, le type d'algorithme, la source des données, l'horizon de prédiction, la performance du modèle, l'existence d'une validation externe et les limites méthodologiques.

Cette revue n'est pas présentée comme une revue systématique PRISMA, afin de ne pas revendiquer une méthodologie qui nécessiterait un protocole préenregistré, une stratégie de recherche exhaustive, un diagramme de flux et une sélection indépendante en double lecture.

III. Résultats

Performance globale des modèles d'IA

Les modèles d'IA appliqués à la prédiction de l'IRA rapportent globalement des performances encourageantes. Dans la méta-analyse de Zhang et al., 19 études portant sur 304 076 patients ont été incluses, avec une sensibilité combinée de 0,77, une spécificité de 0,75 et un diagnostic odds ratio de 10,7 pour la prédiction périopératoire de l'IRA (1). 1 Les auteurs soulignaient cependant que seules quelques études comportaient une validation externe, ce qui limite la généralisation des résultats (1). 1

Une revue systématique récente centrée sur la réanimation a également montré que les modèles d'IA pour l'IRA présentent un potentiel réel, mais que leur applicabilité clinique reste limitée par le risque de biais, l'insuffisance de validation externe, le manque d'interprétabilité et l'hétérogénéité des méthodes (5). 5

Modèles fondés sur les dossiers médicaux électroniques

Les dossiers médicaux électroniques constituent la source principale des modèles prédictifs modernes. Ils permettent d'exploiter les données démographiques, biologiques, hémodynamiques, médicamenteuses et évolutives disponibles en routine.

Koyner et al. ont développé un modèle de machine learning chez 121 158 patients hospitalisés adultes, sans insuffisance rénale terminale initiale. Le modèle prédisait l'IRA KDIGO stade 2 avec une AUROC de 0,90 dans les 24 heures et de 0,87 dans les 48 heures, et prédisait le recours à l'épuration extrarénale avec une AUROC de 0,96 à 48 heures (3). 3

Mohamadlou et al. ont évalué un algorithme de machine learning sur plus de 300 000 séjours hospitaliers. L'AUROC était de 0,800 pour la prédiction de l'IRA 12 heures avant son apparition, 0,795 à 24 heures, 0,761 à 48 heures et 0,728 à 72 heures (6). 6 Cette baisse progressive de performance lorsque l'horizon de prédiction s'allonge illustre la difficulté de prédire très précocement un événement multifactoriel.

Deep learning et prédiction dynamique

Les modèles de deep learning cherchent à dépasser l'estimation statique du risque pour produire une prédiction dynamique. Tomašev et al. ont proposé une approche de prédiction continue de l'IRA future à partir de données longitudinales de dossiers médicaux électroniques, dans une étude publiée dans Nature (7). 7 Ce travail a contribué à populariser l'idée d'une surveillance automatisée et évolutive du risque rénal en milieu hospitalier.

Rank et al. ont développé un modèle de deep learning en temps réel pour prédire l'IRA après chirurgie cardiothoracique. Leur modèle dépassait la performance prédictive humaine dans leur cohorte, suggérant un potentiel d'intégration aux systèmes de surveillance périopératoire (8). 8 Toutefois, une supériorité statistique sur la prédiction humaine ne démontre pas automatiquement une amélioration des résultats cliniques.

Chirurgie cardiaque et période périopératoire

La chirurgie cardiaque est l'un des domaines les plus étudiés, car l'IRA postopératoire y est fréquente et associée à une morbidité importante. Tseng et al. ont comparé plusieurs modèles de machine learning chez 671 patients opérés de chirurgie cardiaque. Le modèle de forêt aléatoire obtenait une AUROC de 0,839, et le modèle d'ensemble associant forêt aléatoire et XGBoost atteignait une AUROC de 0,843 (9). 9

Dans cette étude, les variables les plus influentes étaient la diurèse peropératoire, la transfusion de culots globulaires et le taux d'hémoglobine préopératoire (9). 9 Ces variables sont cliniquement plausibles et montrent l'intérêt des modèles explicables : le clinicien peut comprendre les facteurs qui contribuent au risque et adapter sa surveillance.

La transplantation hépatique représente un autre contexte périopératoire à haut risque. Lee et al. ont comparé plusieurs modèles de machine learning à la régression logistique pour prédire l'IRA après transplantation hépatique ; le gradient boosting machine obtenait une AUROC de 0,90, contre 0,61 pour la régression logistique (10). 10

Sepsis et réanimation adulte

L'IRA associée au sepsis est fréquente, sévère et souvent multifactorielle. Yue et al. ont développé des modèles de machine learning à partir de la base MIMIC-III chez 3 176 patients septiques. L'IRA survenait chez 75,5% des patients, et le modèle XGBoost obtenait la meilleure performance, avec une AUROC de 0,821 (11). 11

Zhang et al. ont étudié une question plus thérapeutique : prédire la réponse volumique chez les patients avec IRA oligurique en réanimation. Le modèle XGBoost obtenait une AUROC de 0,860, contre 0,728 pour la régression logistique (12). 12 Cette approche est particulièrement intéressante, car elle rapproche l'IA d'une décision clinique concrète : faut-il poursuivre le remplissage ou éviter une surcharge hydrosodée ?

Réanimation pédiatrique

L'IA peut également être pertinente en néphrologie pédiatrique. Dong et al. ont développé un modèle multicentrique chez 16 863 patients de réanimation pédiatrique âgés de 1 mois à 21 ans. Le modèle prédisait l'IRA modérée à sévère avec une AUROC de 0,89 et un délai médian d'anticipation de 30 heures (13). 13

Le même modèle identifiait 70% des épisodes nécessitant une épuration extrarénale, 58% des IRA stade 2 ou 3 et 41% de toutes les IRA (13). 13 Parmi les patients correctement identifiés comme à risque, 79% recevaient un médicament potentiellement néphrotoxique pendant la fenêtre de prédiction, ce qui suggère une possibilité d'intervention préventive (13). 13

Produit de contraste et cardiologie interventionnelle

L'exposition aux produits de contraste est un autre champ d'application. Yun et al. ont utilisé un modèle de deep learning pour prédire l'IRA après scanner injecté chez 14 185 patients. Le modèle de réseau neuronal récurrent obtenait une AUROC de 0,755, supérieure aux autres modèles évalués (14). 14

En cardiologie interventionnelle, Huang et al. ont comparé un modèle de référence de prédiction de l'IRA après angioplastie coronaire à plusieurs modèles utilisant des techniques de machine learning. Le meilleur modèle, fondé sur le gradient boosting, obtenait une AUROC de 0,752 contre 0,711 pour le modèle de référence (15). 15

Interprétabilité des modèles

L'interprétabilité est un enjeu majeur pour la néphrologie. Un modèle qui prédit un risque sans expliquer les facteurs contributifs risque d'être peu accepté par les cliniciens. Les méthodes d'explication comme SHAP et LIME permettent d'identifier les variables qui influencent le plus la prédiction individuelle.

Tseng et al. ont utilisé SHAP pour expliquer les prédictions de leur modèle d'IRA après chirurgie cardiaque, en mettant en évidence le rôle de la diurèse, de la transfusion et de l'hémoglobine (9). 9 Hu et al. ont également utilisé SHAP et LIME pour interpréter un modèle de prédiction de mortalité chez des patients ayant une IRA en réanimation (16). 16

IV. Discussion

Apport clinique potentiel

L'apport principal de l'IA dans l'IRA est la possibilité d'anticiper le risque avant l'apparition des critères diagnostiques classiques. Cette anticipation pourrait permettre une revue plus précoce des médicaments néphrotoxiques, une optimisation hémodynamique, une adaptation des doses, une surveillance biologique renforcée, une discussion néphrologique précoce et une meilleure planification de l'épuration extrarénale.

L'IA est particulièrement adaptée à l'IRA parce que celle-ci est dynamique et multifactorielle. Les modèles peuvent intégrer des variations de créatinine, de diurèse, de pression artérielle, de lactates,

d'hémoglobine, de médicaments, de ventilation et de paramètres infectieux, alors qu'un score classique ne retient souvent qu'un nombre limité de variables.

De la prédiction à l'action

La question centrale n'est pas seulement : « le modèle prédit-il l'IRA ? », mais plutôt : « que fait-on de cette prédiction ? ». Un modèle utile doit être associé à une conduite à tenir claire. Par exemple, une alerte pourrait déclencher une révision des néphrotoxiques, une évaluation de la volémie, une adaptation des antibiotiques, une surveillance de la diurèse ou une consultation spécialisée.

L'étude pédiatrique de Dong et al. illustre bien cette logique, car elle montre qu'une proportion importante des patients identifiés comme à risque recevait des médicaments potentiellement néphrotoxiques pendant la fenêtre de prédiction (13). 13 Cela suggère que l'alerte algorithmique pourrait être transformée en intervention clinique concrète.

Limites méthodologiques

La majorité des études disponibles sont rétrospectives. Cette caractéristique limite l'interprétation causale et ne permet pas de savoir si l'utilisation du modèle améliore réellement le devenir des patients.

La validation externe reste insuffisante. Rockenschaub et al. ont montré que la performance de modèles de deep learning pouvait diminuer lorsqu'ils étaient appliqués à des hôpitaux non vus pendant l'entraînement, avec des baisses d'AUROC pouvant atteindre 0,200 (17). 17 Cette observation rappelle qu'un modèle performant localement peut devenir moins fiable dans un autre système de soins.

La calibration locale est également importante. Sun et al. ont évalué des modèles en conditions réelles dans plusieurs hôpitaux et ont montré que les performances étaient meilleures lorsque les modèles étaient calibrés avec les données du site de déploiement (18). 18

Risques éthiques

L'IA peut reproduire ou amplifier les inégalités présentes dans les données. Si certaines populations sont sous-représentées, les modèles peuvent être moins fiables chez elles. Ibrahim et al. ont décrit ce phénomène comme une « pauvreté en données de santé », susceptible d'aggraver les inégalités d'accès aux bénéfices de la santé numérique (19). 19

Un autre risque est la fatigue d'alerte. Un modèle trop sensible peut produire de nombreux faux positifs, ce qui expose à une perte de confiance des soignants. À l'inverse, un modèle trop spécifique peut manquer des patients réellement à risque. Le seuil d'alerte doit donc être adapté au contexte et au type d'intervention disponible.

Place du néphrologue

L'IA ne remplace pas le raisonnement néphrologique. Elle doit être comprise comme un outil d'aide à la décision. Le néphrologue reste indispensable pour interpréter le contexte clinique, distinguer les mécanismes fonctionnels, obstructifs ou organiques, adapter les traitements, décider du moment de l'épuration extrarénale et intégrer les préférences du patient.

L'avenir le plus réaliste n'est pas une IA autonome, mais une IA intégrée à une équipe clinique. Les modèles les plus utiles seront probablement ceux qui associent une prédiction fiable, une explication compréhensible et une proposition d'action.

V. Conclusion

L'intelligence artificielle représente une avancée importante pour la prédiction précoce de l'insuffisance rénale aiguë. Les études récentes montrent des performances prometteuses dans plusieurs contextes, notamment la réanimation, la chirurgie cardiaque, le sepsis, la pédiatrie, la transplantation et l'exposition aux produits de contraste.

Cependant, l'IA ne doit pas être évaluée uniquement par l'AUROC. Pour devenir utile en pratique, elle doit être validée en externe, interprétable, intégrée dans le flux de soins, acceptable pour les cliniciens et associée à des interventions concrètes. Les futures recherches devront démontrer non seulement que l'IA prédit l'IRA, mais surtout qu'elle permet d'en réduire la fréquence, la sévérité ou les conséquences.

Bibliographie Vancouver

- [1]. Zhang H, Wang Y, Wu S, Ngo J, Feng Y, He X, Et Al. Artificial Intelligence For The Prediction Of Acute Kidney Injury During The Perioperative Period: Systematic Review And Meta-Analysis Of Diagnostic Test Accuracy. BMC Nephrol. 2022;23:405. 1
- [2]. Ostermann M, Bellomo R, Burdmann EA, Doi K, Endre Z, Goldstein SL, Et Al. Controversies In Acute Kidney Injury: Conclusions From A Kidney Disease: Improving Global Outcomes Conference. Kidney Int. 2020;98(2):294-309. 2
- [3]. Koyner JL, Carey KA, Edelson DP, Churpek MM. The Development Of A Machine Learning Inpatient Acute Kidney Injury Prediction Model. Crit Care Med. 2018;46(7):1070-1077. 3

- [4]. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado GS, King D. Key Challenges For Delivering Clinical Impact With Artificial Intelligence. *BMC Med.* 2019;17:195. 4
- [5]. Shi T, Lin Y, Zhao H, Kong G. Artificial Intelligence Models For Predicting Acute Kidney Injury In The Intensive Care Unit: A Systematic Review Of Modeling Methods, Data Utilization, And Clinical Applicability. *JAMIA Open.* 2025;8(4):Ooaf065. 5
- [6]. Mohamadlou H, Lynn-Palevsky A, Barton C, Chettipally UK, Shieh L, Calvert J, Et Al. Prediction Of Acute Kidney Injury With A Machine Learning Algorithm Using Electronic Health Record Data. *Can J Kidney Health Dis.* 2018;5:2054358118776326. 6
- [7]. Tomašev N, Glorot X, Rae JW, Zieliński M, Askham H, Saraiva A, Et Al. A Clinically Applicable Approach To Continuous Prediction Of Future Acute Kidney Injury. *Nature.* 2019;572(7767):116-119. 7
- [8]. Rank N, Pfahringer B, Kempfert J, Stamm C, Kühne T, Schoenrath F, Et Al. Deep-Learning-Based Real-Time Prediction Of Acute Kidney Injury Outperforms Human Predictive Performance. *NPJ Digit Med.* 2020;3:139. 8
- [9]. Tseng PY, Chen YT, Wang CH, Chiu KM, Peng YS, Hsu SP, Et Al. Prediction Of The Development Of Acute Kidney Injury Following Cardiac Surgery By Machine Learning. *Crit Care.* 2020;24:478. 9
- [10]. Lee HC, Yoon SB, Yang SM, Kim WH, Ryu HG, Jung CW, Et Al. Prediction Of Acute Kidney Injury After Liver Transplantation: Machine Learning Approaches Vs. Logistic Regression Model. *J Clin Med.* 2018;7(11):428. 10
- [11]. Yue S, Li S, Huang X, Liu J, Hou X, Zhao Y, Et Al. Machine Learning For The Prediction Of Acute Kidney Injury In Patients With Sepsis. *J Transl Med.* 2022;20:215. 11
- [12]. Zhang Z, Ho KM, Hong Y. Machine Learning For The Prediction Of Volume Responsiveness In Patients With Oliguric Acute Kidney Injury In Critical Care. *Crit Care.* 2019;23:112. 12
- [13]. Dong J, Feng T, Thapa-Chhetry B, Cho BG, Shum T, Inwald D, Et Al. Machine Learning Model For Early Prediction Of Acute Kidney Injury In Pediatric Critical Care. *Crit Care.* 2021;25:288. 13
- [14]. Yun D, Cho S, Kim YC, Kim DK, Oh KH, Joo KW, Et Al. Use Of Deep Learning To Predict Acute Kidney Injury After Intravenous Contrast Media Administration: Prediction Model Development Study. *JMIR Med Inform.* 2021;9(10):E27177. 14
- [15]. Huang C, Murugiah K, Mahajan S, Li SX, Dhruva SS, Haimovich JS, Et Al. Enhancing The Prediction Of Acute Kidney Injury Risk After Percutaneous Coronary Intervention Using Machine Learning Techniques: A Retrospective Cohort Study. *Plos Med.* 2018;15(11):E1002703. 15
- [16]. Hu C, Tan Q, Zhang Q, Li Y, Wang F, Zou X, Et Al. Application Of Interpretable Machine Learning For Early Prediction Of Prognosis In Acute Kidney Injury. *Comput Struct Biotechnol J.* 2022;20:2861-2870. 16
- [17]. Rockenschaub P, Hilbert A, Kossen T, Elbers P, Von Dincklage F, Madai VI, Et Al. The Impact Of Multi-Institution Datasets On The Generalizability Of Machine Learning Prediction Models In The ICU. *Crit Care Med.* 2024;52(9):E456-E466. 17
- [18]. Sun H, Depraetere K, Meesseman L, Silva PC, Szymanowsky R, Fliegenschmidt J, Et Al. Machine Learning-Based Prediction Models For Different Clinical Risks In Different Hospitals: Evaluation Of Live Performance. *J Med Internet Res.* 2022;24(6):E34295. 18
- [19]. Ibrahim H, Liu X, Zariffa N, Morris AD, Denniston AK. Health Data Poverty: An Assailable Barrier To Equitable Digital Health Care. *Lancet Digit Health.* 2021;3(4):E260-E265. 19