



École doctorale 393 Pierre Louis de santé publique

THÈSE DE DOCTORAT

Discipline : Biostatistiques

présentée par

François GROLLEAU

**Causal inference methods for personalized medicine:
an application to the timing of renal replacement
therapy initiation**

dirigée par Raphaël PORCHER

Soutenue le *X* novembre 2023 devant le jury composé de :

M. Antoine CHAMBAZ	Université Paris Cité	Examinateur
M. Yohann FOUCHER	Université de Poitiers	Rapporteur
M ^{me} Julie JOSSE	INRIA Sophia-Antipolis	Rapporteure
M. Romain PIRRACCHIO	University of California San Francisco	Examinateur
M ^{me} Laura RICHERT	Université de Bordeaux	Examinaterice
M. Raphaël PORCHER	Université Paris Cité	Directeur de thèse

Centre de recherche en épidémiologie et
statistiques. INSERM-UMR 1153.
Équipe METHODS
Hôpital Hôtel-Dieu
1, place du Parvis Notre-Dame
75 004 Paris

Université Paris Cité.
École doctorale 393 Pierre Louis de
santé publique: épidémiologie et sci-
ences de l'information biomédicale.
21, rue de l'école de medecine
75 006 Paris

Remerciements

Mes remerciements vont en premier à Raphaël Porcher qui a su me guider si intelligemment tout au long de ce projet.

Je remercie mon ami François Petit pour m'avoir patiemment indiqué le chemin de la rigueur et de la précision.

Je remercie les Professeurs Yohann Foucher et Julie Josse pour avoir accepté de rapporter ce travail de thèse.

Je remercie les Professeurs Antoine Chambaz, Romain Pirracchio et Laura Richert pour avoir accepté d'examiner critiquement la qualité de ce manuscrit.

Je remercie le Professeur Philippe Ravaud qui, pour sa vision et ses idées scientifiques intransigeantes, a depuis longtemps été une source d'inspiration pour moi.

Je remercie le Professeure Isabelle Boutron pour m'avoir fait confiance en tant que chef de clinique dans une discipline, l'épidémiologie et les biostatistiques, qui ne correspondait pas à ma formation initiale.

Je remercie le Professeur Viet-Thi Tran pour ses conseils clairvoyants et pour nos longues discussions qui m'ont permis de structurer ma pensée en ce qui concerne le futur de la médecine à l'heure de la révolution du numérique.

Je remercie le Professeur Steve Oudot pour m'avoir laissé la chance de participer au projet "*Precision medicine using topology*".

Je remercie les Professeurs Didier Dreyfuss, Stéphane Gaudry et Jean-Pierre Quenot pour la qualité de nos interactions scientifiques, et pour m'avoir fait partager leur expertise exceptionnelle dans le domaine de l'insuffisance rénale aiguë et de la réanimation en général.

Je remercie les post-doctorants et doctorants en statistiques avec lesquels j'ai eu la chance de travailler et aux côtés desquels j'ai beaucoup appris. Je remercie en particulier à Alan Balendran, Céline Béji, Florie Bouvier, Bénédicte Colnet et Étienne Peyrot.

Je remercie mes collègues de bureau Stéphanie Chardon, Audrey Conjaud, Élise Diard, Amel Ghizia et Léa Sicard pour leur gentillesse et pour leur soutien constant.

Enfin, je souhaite souligner que l'idée d'une autonomie complète ou indépendance pour atteindre un objectif prédéfini (de thèse ou non) n'est qu'une fantaisie trop répandue. J'ai une dette de gratitude immense envers tous mes enseignants, tous mes mentors et plus largement tous les individus non mentionnés ci-dessus qui m'ont aidé tout au long de mon parcours.

À ma mère Pascale,

À mon père Jean-Louis,

À mes grands-parents Éliane, Pierrot, Danielle et Bobby,

À mon oncle Jean-Pierre,

À ma sœur Chloé et son époux Thomas,

À Cynthia

Valorisation scientifique

Publications issues de cette thèse

- S. Gaudry, **F. Grolleau**, S. Barbar et al. Continuous renal replacement therapy versus intermittent hemodialysis as first modality for renal replacement therapy in severe acute kidney injury: a secondary analysis of AKIKI and IDEAL-ICU studies. *Critical Care* **26**, 93 (2022).
- **F. Grolleau**, R. Porcher, S. Barbar et al. Personalization of renal replacement therapy initiation: a secondary analysis of the AKIKI and IDEAL-ICU trials. *Critical Care* **26**, 1–10 (2022).
- **F. Grolleau**, F. Petit, R. Porcher. A comprehensive framework for the evaluation of individual treatment rules from observational data. *arXiv:2207.06275, en révision*.
- **F. Grolleau**, F. Petit, S. Gaudry et al. Personalizing renal replacement therapy initiation in the intensive care unit: a reinforcement learning-based strategy. *medRxiv, 2023.06. 13.23291349, en révision*.

Publications annexes

- N. Clementy*, L. Labombarda*, **F. Grolleau*** et al. Electrocardiogram Versus Invasive Electrophysiological Study in the Prediction of Major Conduction Delays in Myotonic Dystrophy Type 1. *soumis*. *contribution égale des auteurs.
- J. Puntonet, B. Nodot, **F. Grolleau** et al. Artificial Intelligence for Radiographic Detection of CT-scan confirmed Fractures of the Extremities. *soumis*.
- D. Nezam, R. Porcher, **F. Grolleau** et al. Kidney histopathology can predict kidney function in ANCA-associated vasculitides with acute kidney injury treated with plasma exchanges. *Journal of the American Society of Nephrology* **33**, 628–637 (2022).
- D. Nezam, R. Porcher, **F. Grolleau** et al. Authors' reply: Kidney histopathology can predict kidney function in ANCA-associated vasculitides with acute kidney injury treated with plasma exchanges. *Journal of the American Society of Nephrology* **33**, 1224 (2022).
- R. Bey, R. Goussault, **F. Grolleau** et al. Fold-stratified cross-validation for unbiased and privacy-preserving federated learning. *Journal of the American Medical Informatics Association* **27**, 1244–1251 (2020)

Communications orales et conférences

Statistiques

- AI for science. Science for AI, **AISSA/CNRS seminar**, Causality in practice 2023 (Paris)
- Session “Utilisation de l’IA en recherche clinique”, **EPICLIN/JSCCLCC 2023** (Nancy)
- International society for clinical biostatistics, **ISCB 2022** (Newcastle, UK)
- American causal inference conference, **ACIC 2022** (UC Berkeley, CA)
- Paris artificial intelligence research institute day, **PRAIRIE day 2022** (Paris)
- **Causal Inference in Medicine**, 2022 seminar (Ghent university, BEL)
- Session “Modélisation avancée et analyse de biomarqueurs”, **EPICLIN/JSCCLCC 2022** (Paris)
- Dynamic Treatment Regime talk No. 2, **INRIA/CRESS workshop** 2021 (En ligne)

Médecine

International symposium on intensive care and emergency medicine, **ISICEM 2022** (Bruxelles)

Packages et libraraires

R

- **Mestim** disponible sur *CRAN*
- **CACEmix** (bientôt disponible sur *CRAN*)

Python

- **Speedboot** disponible sur *Pypi*
- **ClusterITE** (disponible sur <https://github.com/fcgrolleau/ClusterITE>)

Applications web

- <http://dynamic-rrt.eu>
- <http://rrt-personalization.eu>

Abréviations

- AFT:** *Accelerated Failure Time*
- AIPW:** *Augmented Inverse Probability of Treatment Weigthing*
- AKI:** *Acute Kidney Injury*
- ATE:** *Average Treatment Effect*
- CACE:** *Complier Average Causal Effect*
- CABG:** *Coronary Artery Bypass Graft*
- CATE:** *Conditional Average Treatment Effect*
- CI:** *Confidence Interval*
- CRRT:** *Continuous Renal Replacement Therapy*
- dWOLS:** *dynamic Weighted Ordinary Least Squares*
- EM:** *Expectation-Maximization*
- HTE:** *Heterogeneous Treatment Effects*
- ICU:** *Intensive Care Unit*
- IHD:** *Intermittent Hemodialysis*
- IPW:** *Inverse Probabilty of Treatment Weigthing*
- IRLS:** *Iterated Reweighted Least Squares*
- ITR:** *Individualized Treatment Rule*
- LOOCV:** *Leave One Out Cross Validation*
- MILP:** *Mixed-Integer Linear Programming*
- HR:** *Hazard Ratio*
- KDIGO:** *Kidney Disease Improving Global Outcomes*
- PATH:** *Predictive Approaches to Treatment effect Heterogeneity*
- RCT:** *Randomized Controlled Trial*
- RMSE:** *Root Mean Square Error*
- RRT:** *Renal Replacement Therapy*
- RMST:** *Restricted Mean Survival Time*
- SMART:** *Sequential Multiple Assignment Randomized Trial*
- SOFA:** *Sepsis-related Organ Failure Assessment*
- TMLE:** *Targeted Maximum Likelihood Estimation*
- VC:** *Vapnik–Chervonenkis*

Sommaire

1 Contexte médical et scientifique	13
1.1 Insuffisance rénale aiguë en réanimation	13
1.1.1 Insuffisance rénale aiguë	13
1.1.2 Épuration extra-rénale	14
1.2 Quand débuter l'épuration extra-rénale?	17
1.3 Objectifs de cette thèse	20
1.3.1 Limites des approches fondées sur la biologie et la physiologie	20
1.3.2 Limites de l'évaluation d'effets traitements à l'échelle populationnelle à un temps donné	21
1.3.3 Objectifs de cette thèse	22
2 Revue de certains outils théoriques utilisés dans cette thèse	23
2.1 Inférence causale	23
2.1.1 Notations et hypothèses d'identifiabilité	23
2.1.2 Estimands causaux	24
2.1.3 Estimateurs de l'effet traitement moyen	25
2.2 M-estimation	28
2.2.1 Théorie	28

2.2.2	Illustration du principe d'empilement d'équations d'estimation	29
2.2.3	Implémentation avec le package Mestim	31
2.3	Hetérogénéité de l'effet du traitement	32
2.3.1	Modélisation de l'effet du traitement par niveaux de risque prédit	32
2.3.2	Modélisation directe de l'effet du traitement	35
2.4	Mélange d'experts	40
2.4.1	Modèle probabiliste	40
2.4.2	Reformulation et représentation graphique	41
2.4.3	Vraisemblances sur données complètes et incomplètes	41
2.4.4	Algorithme d'espérance-maximisation (EM)	42
2.4.5	Identifiabilité	44
2.4.6	Propriétés asymptotiques	44
2.5	Apprentissage par renforcement	45
2.5.1	Formalisation du problème et notations	45
2.5.2	Estimation d'une politique optimale par maximisation de valeur	47
2.5.3	Politiques de type “Quand débuter le traitement ?”	49
2.5.4	Estimation d'une politique optimale par modélisations récursives	50
2.5.5	Note concernant les critères de jugement censurés	51
3	Continuous renal replacement therapy versus intermittent hemodialysis as first modality for renal replacement therapy in severe acute kidney injury: a secondary analysis of AKIKI and IDEAL-ICU studies	53
4	Personalization of renal replacement therapy initiation: a secondary analysis of the AKIKI and IDEAL-ICU trials	73

SOMMAIRE	11
5 A comprehensive framework for the evaluation of individual treatment rules from observational data	93
6 Personalizing renal replacement therapy initiation in the intensive care unit: a statistical reinforcement learning-based strategy with external validation on the AKIKI randomized controlled trials	127
7 Discussion et perspectives	167
7.1 La causalité est-elle fondamentale ou simplement utile ?	167
7.2 Systèmes de médecine personnalisée plus intelligents	168
7.3 Impasses et perspectives pour l'IA appliquée à la santé	170
Conclusion	173
Bibliographie	175
Annexe A : Un modèle causal général pour l'estimation de CACE sans hypothèses de monotocité ni d'exclusion restriction	183
Annexe B : Calculs et définitions	205

Chapter 1

Contexte médical et scientifique

One who has a “why” to live for can endure almost any “how.”

Friedrich Nietzsche, *Twilight of the Idols*

1.1 Insuffisance rénale aiguë et épuration extra-rénale en réanimation

L’insuffisance rénale aiguë affecte à ce jour environ un patient de réanimation sur deux. Sa survenue est associée à une mortalité élevée et des séquelles à long terme incluant des événements cardiovasculaires et une insuffisance rénale chronique. L’épuration extra-rénale, introduite au début des années soixante, fut une avancée importante pour mitiger ce risque et améliorer la survie des patients de réanimation. Dans cette section, nous rappellerons d’abord certaines caractéristiques importantes de l’insuffisance rénale aiguë et développerons ensuite plusieurs points concernant l’épuration extra-rénale.

1.1.1 Insuffisance rénale aiguë

L’insuffisance rénale aiguë est définie par l’occurrence d’un déclin brutal de la fonction excrétrice rénale telle que mesurée par le débit de filtration glomérulaire [1]. Du fait de l’impossibilité de mesurer cette grandeur physiologique en routine clinique, le diagnostic positif de l’insuffisance rénale aiguë est réalisé en présence d’une augmentation de la créatinine plasmatique (un déchet azoté habituellement excrété par le rein dans les urines) et/ou d’une réduction du débit urinaire (celui-ci reflète la fonction excrétrice hydrique rénale). La sévérité de l’insuffisance rénale aiguë est évaluée par l’amplitude et la durée de ces changements. Les recommandations d’experts *Kidney Disease Improving Global Outcomes* (KDIGO) [2] pour le diagnostic positif et de sévérité de l’insuffisance rénale aiguë s’appuient sur ces notions ; celles-ci sont énoncées en détail dans le tableau 1. L’insuffisance rénale aiguë est souvent caractérisée plus finement selon son mécanisme. Celui-ci peut être i) prérénal (en cas d’altération hémodynamique), ii) postrénal (par obstruction des voies urinaires), ou iii) parenchymateux (par atteinte intrinsèque d’une ou plusieurs structures rénales : glomérulaire, tubulaire, vasculaire ou interstitielle). En réanimation, l’insuffisance

Tableau 1: Référence KDIGO pour le diagnostic et l'évaluation de la sévérité de l'insuffisance rénale aiguë. La présence d'un critère suffit pour le diagnostic et la classification par stade de sévérité (le critère le plus défavorable est retenu). Un stade plus élevé indique une sévérité plus importante.

Stade	Créatinine plasmatique	Débit urinaire
	1.5-1.9 fois la valeur de base en 48h	
1	OU augmentation $\geq 0.3\text{mg/dL}$	$\leq 0.5\text{ml/kg/min}$ pendant plus de 6h
2	2.0-2.9 fois la valeur de base	$\leq 0.5\text{ml/kg/min}$ pendant plus de 12h
	plus de 3 fois la valeur de base	$\leq 0.3\text{ml/kg/min}$ pendant plus de 24h
3	OU valeur $\geq 4.0\text{mg/dL}$	$\leq 0.1\text{ml/kg/min}$ pendant plus de 12h

rénale aiguë étant souvent associée à la survenue d'un sepsis, (et/ou) un état de choc, (et/ou) à une exposition à des néphrotoxiques, le mécanisme est souvent attribué à une nécrose tubulaire aiguë. L'intrication de plusieurs mécanismes est cependant fréquente et la caractérisation fine des mécanismes est difficile en pratique.

Au cours de l'insuffisance rénale aiguë, les dosages répétés quotidiens de la créatinine et de l'urée permettent de suivre la progression de l'atteinte rénale. L'augmentation de la créatinine et de l'urée reflètent l'augmentation parallèle d'autres métabolites toxiques dont plusieurs sont non dosés.¹ La fonction excrétrice étant altérée au cours de l'insuffisance rénale aiguë, l'excrétion de potassium et de protons diminue. Ces éléments s'accumulent alors dans le plasma et on parle alors respectivement d'hyperkaliémie et d'acidose métabolique. La sévérité du déséquilibre de la physiologie de l'eau et des électrolytes est corrélée positivement à i) l'importance de l'atteinte rénale et ii) au taux de production de métabolites toxiques (catabolisme). Si les déséquilibres hydro-électrolytiques et urémiques sont sévères, le patient est à risque de décès rapide (celui-ci fait souvent suite à un trouble conductif cardiaque d'évolution défavorable). L'objectif principal de l'épuration extra-rénale est de corriger les déséquilibres hydro-électrolytiques et urémiques menaçants pour prévenir le décès du patient.

1.1.2 Épuration extra-rénale

Les deux modalités d'épuration extra-rénale les plus souvent utilisées dans le contexte d'insuffisance rénale aiguë sont i) l'hémodialyse intermittente et ii) l'hémofiltration continue.² Ces techniques ont des points communs. Elles requièrent toutes deux un accès vasculaire veineux par cathéter idéalement en position jugulaire ou fémorale pour la mise en place d'une circulation extra-corporelle. Sur le circuit, le sang veineux est mis en mouvement à l'aide d'une pompe à galet ; il perfuse une membrane semi-perméable où sont échangés de l'eau, des solutés et des électrolytes (voir Figure 1 pour un schéma des deux modalités d'épuration extra-rénale).

¹Ces métabolites sont vus comme des médiateurs de l'effet toxique de l'urée ; quand ces métabolites indosés augmentent dans le plasma, nous parlons de désordre (ou déséquilibre) "urémique".

²D'autres modalités d'épuration extra-rénale sont possibles pour le traitement de l'insuffisance rénale aiguë (hémodialyse continue, hémodiafiltration continue, dialyse péritonéale, etc.). Pour une description exhaustive des ces techniques voir Gaudry et al. [3] et Teitelbaum [4].

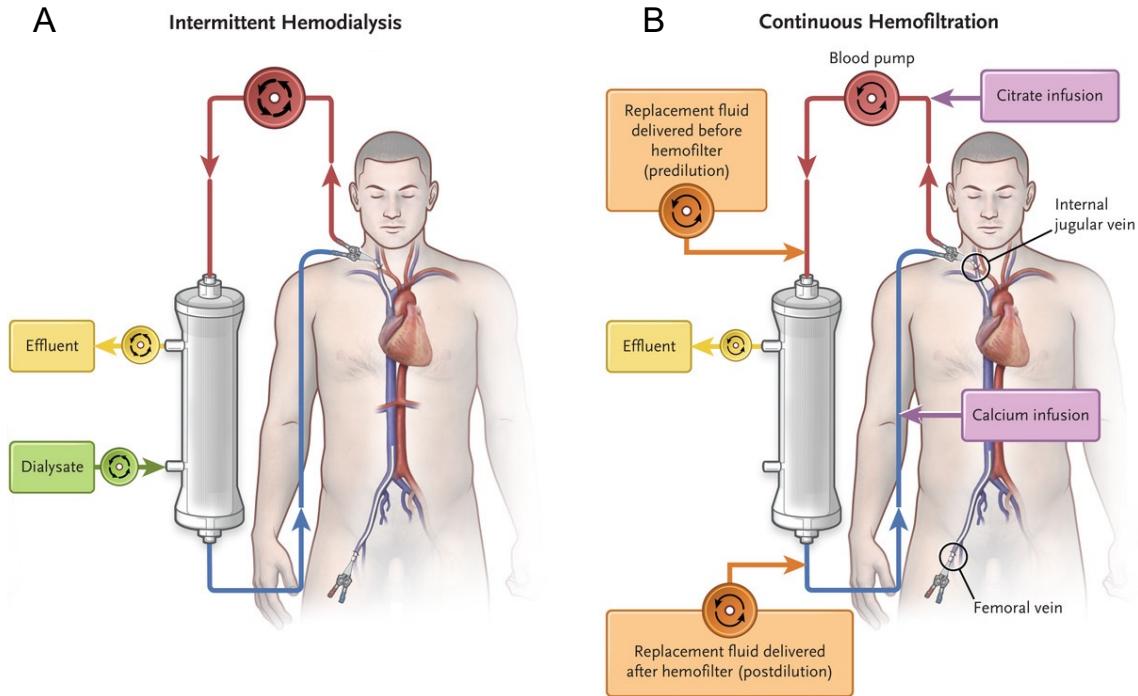


Figure 1: Circulation extra-corporelle pour l'hémodialyse intermittente (Panneau A) et l'hémofiltration continue (Panneau B) (Gaudry et al. [3]). À noter, dans l'hémofiltration continue, une anticoagualtion régionale du circuit par calcium-citrate est nécessaire afin de prévenir les thromboses sur la membrane semi-perméable. Un soluté (prédilution) est délivré en amont du filtre pour prévenir ces thromboses. Un soluté (postdilution) est également délivré en aval du filtre pour compenser la filtration (effluent) en excès.

L'hémodialyse intermittente, s'appuie sur la méthode de diffusion pour corriger les déséquilibres hydro-électrolytiques et urémiques (voir figure 2A et sa légende pour une explication de la méthode de diffusion). Avec cette modalité la circulation extra-corporelle est branchée au cathéter et mise en marche 3 à 7 fois par semaine pour des séances durant 3 à 6 heures. Ces temps de traitements étant courts pour corriger les déséquilibres hydro-électrolytiques et urémiques, l'hémodialyse intermittente requiert l'utilisation de débits sanguins et de dialysat élevés dans la membrane semi-perméable. La correction rapide des désordres peut être un avantage en cas de situations menaçantes (par exemple une hyperkaliémie importante). La durée courte du traitement a également l'avantage de fréquemment libérer les patients de la circulation extra-corporelle ce qui facilite la réalisation de soins infirmiers et la mobilisation précoce pour la rééducation (kinésithérapie). Chez les patients hémodynamiquement instables, l'hémodialyse intermittente est à risque d'hypotensions peridialytiques. Certains réglages permettent cependant de limiter ce risque (allongement des séances jusqu'à 16 heures, réduction du débit d'ultrafiltration, composition modifiée et température abaissée du dialysat). En cas de sepsis, l'adaptation des doses d'antibiotiques est nécessaire chez les patients traités par hémodialyse intermittente ; cette adaptation des doses est parfois délicate. La normalisation rapide de l'urée pouvant induire des troubles neurologiques, l'hémodialyse intermittente est déconseillée chez les patients ayant une hypertension intracrânienne ou un œdème cérébral (patients traumatisés

crâniens ou ayant une hépatite fulminante aiguë par exemple).

L'hémofiltration s'appuie essentiellement sur la méthode de convection pour corriger les déséquilibres hydro-électrolytiques et urémiques (voir figure 2B et sa légende pour une explication de la méthode de convection). L'administration “continue” de cette modalité d'épuration extra-rénale consiste en la mise en marche de la circulation extra-corporelle 24 heures sur 24 dans le but de limiter les hypotensions peridialytiques chez des patients hémodynamiquement instables (en hémofiltration continue le débit d'ultrafiltration peut être abaissé tout en maintenant une clairance correcte ; les débits sanguins requis sont par ailleurs plus faibles qu'en hémodialyse intermittente). Chez les patients sous vasopresseurs, l'hémofiltration continue est considérée comme causant un stress hémodynamique plus faible que l'hémodialyse intermittente. Ce bénéfice n'a cependant jamais été démontré rigoureusement [5]. L'utilisation continue de l'hémofiltration permet, en revanche, une élimination des antibiotiques plus stable dans le temps qu'avec l'hémodialyse intermittente ce qui simplifie parfois l'adaptation des doses. La correction des déséquilibres hydro-électrolytique est lente avec l'hémofiltration continue ; cela peut être un inconvénient dans certaines situation cliniques (hyperkaliémie importante) et un avantage dans d'autres (patients traumatisés crâniens ou ayant une hépatite fulminante aiguë).

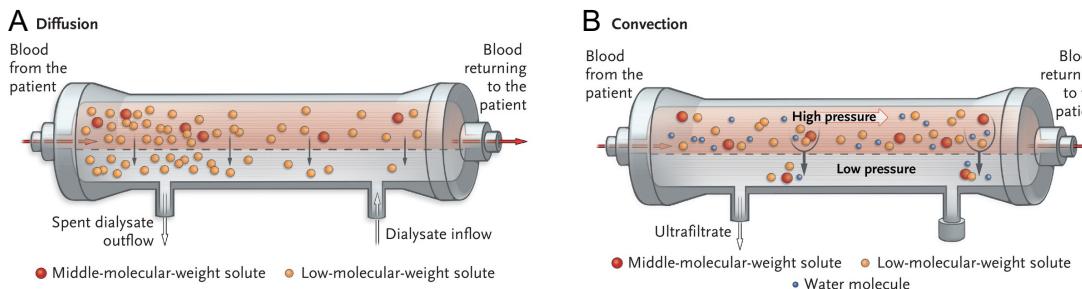


Figure 2: Mécanismes de correction des déséquilibres hydro-électrolytiques et urémiques en hémodialyse intermittente (Panneau A: méthode de diffusion) et en hémofiltration continue (Panneau B: méthode de convection) (Gaudry et al. [3]). Dans la méthode de diffusion, les composés de bas poids moléculaire sont épurés par diffusion à travers la membrane dans le sens de leur gradient osmotique. Un petit gradient de pression hydrostatique (flèches verticales fines) est également exercé à travers la membrane pour générer un débit d'ultrafiltration éliminant eau et sodium dans le dialysat. Dans la méthode de convection, un gradient de pression hydrostatique important est imposé à travers la membrane (flèches verticales larges). L'eau et les électrolytes sont épurés en excès. Un soluté (postdilution) est délivré en aval du filtre pour compenser la filtration en excès (voir Figure 1B). À noter, la méthode de convection permet l'épuration de composés de poids moléculaire plus importants que la méthode de diffusion. Cela représente un avantage théorique qui ne se traduit pas par une amélioration de critères de jugements pertinents pour les patients [6].

1.2 Quand débuter l'épuration extra-rénale ? Essais randomisés et conclusions à l'échelle populationnelle

L'épuration extra-rénale est un traitement invasif qui, bien que parfois indispensable, expose certains patients en état critique à des risques non négligeables, en particulier des thromboses, saignements, infections liées au cathéter et une récupération rénale altérée par l'épuration extra-rénale [7]. Du fait de ces risques, les indications de début d'épuration extra-rénale ont été débattues depuis l'introduction de l'épuration extra-rénale il y a plus de soixante ans. Ces débats scientifiques sont légitimes, ils reflètent le souhait de trouver un compromis optimal entre i) le bénéfice de la restauration de l'équilibre hydro-électrolytique et urémique, et ii) le risque iatrogène induit par le recours à l'épuration extra-rénale.

Pour la population de patients venant de développer une insuffisance rénale aiguë sévère (stade 3 KDIGO) compatible avec une nécrose tubulaire aiguë, jusqu'à récemment, deux stratégies étaient envisagées :

- **Stratégie précoce (préemptive) :** Débuter l'épuration extra-rénale dès que possible (en pratique dans les 6 heures suivant le diagnostic d'insuffisance rénale aiguë stade 3 KDIGO)
- **Stratégie tardive :** Surveiller régulièrement l'évolution des désordres induits par l'insuffisance rénale aiguë et débuter l'épuration extra-rénale uniquement si les déséquilibres urémiques/hydro-électrolytiques se dégradent au-delà de seuils prédéfinis, par exemple :
 - urée plasmatique $\geq 40\text{mmol/L}$,
 - débit urinaire $\leq 0.5/\text{ml/kg/min}$ pendant plus de 72h,
 - kaliémie $\geq 6\text{mmol/L}$ (ou $\geq 5.5\text{mmol/L}$ et persistante malgré traitement médical par bicarbonate et/ou insuline-glucose),
 - pH $\leq 7.15\text{mmol/L}$ dans un contexte d'acidose métabolique pure ou d'acidose métabolique mixte sans possibilité d'augmenter la ventilation alvéolaire,
 - œdème aigu du poumon hydrostatique persistant sous traitement diurétique et responsable d'hypoxémie sévère (débit d'oxygène $\geq 5\text{L/min}$ pour maintenir une SpO₂ $\geq 95\%$ ou FiO₂ $\geq 50\%$ en ventilation invasive ou en ventilation non invasive).

Au cours de la dernière décennie, trois grands essais randomisés contrôlés multicentriques, AKIKI [8], IDEAL-ICU [9] et STARRT-AKI [10], ont été conduits pour comparer stratégies précoce et tardive. Dans l'essai AKIKI, étaient exclus les patients dont les déséquilibres urémiques/hydro-électrolytiques étaient déjà au-delà des seuils listés ci-dessus au moment du diagnostic d'insuffisance rénale aiguë sévère. Aucun des essais AKIKI, IDEAL-ICU et STARRT-AKI, n'a montré de différence significative de survie entre les deux stratégies (critère de jugement principal). Une mété-analyse sur données individuelles incluant neuf essais randomisés publiés entre 2010 et 2020, dont certains monocentriques, comparant stratégies précoce et tardive n'a pas non plus mis en évidence de différence significative de survie entre les groupes [11]. En revanche, les essais montrent que l'utilisation d'une stratégie tardive permettait de s'affranchir d'épuration extra-rénale chez 40% des patients (moyenne sur les essais AKIKI, IDEAL-ICU et STARRT-AKI). Les caractéristiques et particularités des trois plus grands essais randomisés contrôlés sont rapportées dans le

Tableau 2. S'appuyant sur le résultat de ces essais, les recommandations internationales de la *Surviving Sepsis Campaign* [12] recommandent désormais d'utiliser une stratégie tardive pour la population de patients considérée dans cette section. Pour affiner le timing optimal d'initiation de l'épuration extra-rénale, un autre essai randomisé multicentrique, nommé AKIKI 2, a été conduit [13]. Celui-ci cherchait à comparer l'utilisation d'une stratégie tardive à une stratégie très tardive, celles-ci étaient définies par :

- **Stratégie tardive** : identique à la stratégie tardive présentée précédemment,
- **Stratégie très tardive** : Surveiller régulièrement l'évolution des désordres induits par l'insuffisance rénale aiguë et débuter l'épuration extra-rénale uniquement si les déséquilibres urémiques/hydro-électrolytiques se dégradent au-delà des seuils suivants :
 - urée plasmatique $\geq 50\text{mmol/L}$,
 - kaliémie $\geq 6\text{mmol/L}$ (ou $\geq 5.5\text{mmol/L}$ et persistante malgré traitement médical par bicarbonate et/ou insuline-glucose),
 - pH $\leq 7.15\text{mmol/L}$ dans un contexte d'acidose métabolique pure ou d'acidose métabolique mixte sans possibilité d'augmenter la ventilation alvéolaire,
 - œdème aiguë du poumon hydrostatique persistant sous traitement diurétique et responsable d'hypoxémie sévère (débit d'oxygène $\geq 5\text{L/min}$ pour maintenir une SpO₂ $\geq 95\%$ ou FiO₂ $\geq 50\%$ en ventilation invasive ou en ventilation non invasive).

Dans l'essai AKIKI 2, les patients étaient inclus dans une cohorte dès qu'ils développaient une insuffisance rénale aiguë de stade 3 KDIGO (essai "niché dans une cohorte"). Les désordres induits par l'insuffisance rénale aiguë étaient surveillés régulièrement pour la survenue d'une dégradation au-delà d'un des seuils définissant la stratégie tardive (voir ci-dessus). Les patients de la cohorte rencontrant un de ces critères étaient randomisés selon un ratio 1:1 pour recevoir soit une stratégie tardive (impliquant un début de l'épuration extra-rénale rapidement après la randomisation) soit une stratégie très tardive (impliquant une surveillance régulière). L'essai AKIKI 2 reposait sur l'hypothèse qu'à l'échelle populationnelle, une stratégie très tardive réduirait le recours à l'épuration extra-rénale sans altérer la survie des patients. Cependant, les analyses de survie de l'essai ont montré que la stratégie très tardive pouvait être délétère pour la survie des patients (HR non ajusté 1.34 IC 95% [0.96-1.89], HR ajusté 1.65 IC 95% [1.09-2.50]). Cette stratégie n'est donc pas recommandée à l'échelle populationnelle en pratique clinique.

Tableau 2: Caractéristiques et particularités de trois grands essais randomisés contrôlés multicentriques comparant stratégies précoce et tardive dans l'insuffisance rénale aigüe.

	AKIKI	IDEAL-ICU	STARRT-AKI
Patients	Patients de réanimation recevant ventilation mécanique et/ou vasopresseurs Stade 3 KDIGO	Patients de réanimation à la phase aigüe d'un choc septique Stade F RIFLE	Patients de réanimation Stade 2 ou 3 KDIGO
Stratégie précoce (préemptive)	-Urée plasmatique $\geq 40 \text{ mmol/L}$ -Débit urinaire $\leq 0.5 \text{ ml/kg/min}$ pendant plus de 72h. -Kaliémie $\geq 6 \text{ mmol/L}$ -pH ≤ 7.15 avec contexte d'acidose métabolique ou d'acidose mixte -Edème pulmonaire aigüe hydrostatique persistant sous traitement diurétique et responsable d'hypoxémie sévère	-Absence de récupération rénale après 48h -pH ≤ 7.15 avec contexte d'acidose métabolique -Kaliémie $\geq 6.5 \text{ mmol/L}$ -Surcharge hydrosodée réfractaire aux diurétiques et responsable d'œdème aigüe du poumon	-Kaliémie $\geq 6 \text{ mmol/L}$ -Insuffisance rénale aigüe persistant plus de 72h après la randomisation -pH ≤ 7.20 ou bicarbonates $\leq 12 \text{ mmol/L}$ -Défaillance respiratoire sévère et perception clinique de surcharge hydrosodée
Stratégie tardive	619 (311+308)	488 (246+242)	2927 (1463+1462)
Patients (n) (stratégie précoce + stratégie tardive)	52 heures	44 heures	25 heures
Différence de médianes entre les bras stratégie précoce et stratégie tardive pour le temps ayant initiation de l'épuration extra-rénale			
Proportion de patients du bras stratégie tardive n'ayant pas initié l'épuration extra-rénale à la fin du suivi	49%	38%	38%
Critère de jugement principal (stratégie précoce vs stratégie tardive)	Mortalité à J60 150/311 vs 153/308; p=0.79	Mortalité à J90 138/246 vs 28/242; p=0.38	Mortalité à J90 343/1465 vs 639/1462; p=0.92
Survie par estimation de Kaplan-Meier à J60 (stratégie précoce vs stratégie tardive)	51.5% vs 50.3%	43% vs 47%	57% vs 58%

1.3 Limites des approches actuelles et objectifs de cette thèse

Au cours des dix dernières années, trois grands essais randomisés comparant stratégies d'initiation précoce et tardive de l'épuration extra-rénale ont donc été conduits en réanimation [8–10]. Cependant les analyses et méta-analyses de ces essais ne sont pas parvenues à démontrer une différence d'impact de ces stratégies sur des critères pertinents pour les malades [11]. En réanimation, les grands essais randomisés bien conduits sont fréquemment négatifs, y compris pour des interventions ne concernant pas l'insuffisance rénale aiguë [14]. Pour cette raison, dans cette discipline, l'identification d'effets traitements individuels – et non populationnels – est désormais considérée comme une priorité [15].

1.3.1 Limites des approches fondées sur la biologie et la physiologie

L'objectif du clinicien, comme du chercheur en médecine, de proposer au temps opportun, des traitements adaptés aux caractéristiques de chaque patient n'est pas nouveau. Les origines exactes de la médecine “personnalisée” sont anciennes et imprécises [16]. Dans certaines maladies comme le cancer du sein [17], le cancer colorectal [18] ou la leucémie myeloïde chronique [19] l'utilisation de connaissances biologiques a permis des avancées majeures dans le domaine de la médecine personnalisée.

L'insuffisance rénale aiguë est un syndrome hétérogène pour lequel la compréhension fine des mécanismes est probablement plus difficile à caractériser. De nombreux biomarqueurs (NGAL, KIM1, CCL14, etc.) ont été étudiés dans l'insuffisance rénale aiguë. Ceux-ci n'ont cependant pas démontré leur utilité clinique à ce jour [20]. Concernant l'approche physiologique, les modèles actuels cherchent à distinguer précisément la part lésionnelle (agression structurale du rein), de la part fonctionnelle (baisse du débit de filtration gloméru-laire) dans l'insuffisance rénale aiguë [1, 21]. L'objectif ultime de ces approches serait de caractériser en temps réel au lit du malade la proportion de néphrons³ avec :

- a. Ni atteinte lésionnelle, ni atteinte fonctionnelle (tissu rénal “sain”)
- b. Une atteinte fonctionnelle sans atteinte lésionnelle (tissu rénal soumis à une atteinte hémodynamique “pure”)
- c. Une atteinte lésionnelle sans atteinte fonctionnelle (tissu rénal soumis à une agression “directe” inflammatoire, infectieuse, toxique etc.)
- d. Une atteinte lésionnelle et fonctionnelle (tissu rénal sclérosé, à faible potentiel de récupération)

Bien qu'intéressante conceptuellement, cette cartographie dynamique précise de l'atteinte rénale semble extrêmement difficile à réaliser en pratique. En particulier, nous remarquons que même une mesure parfaite en continu de l'atteinte lésionnelle par de futurs biomarqueurs serait insuffisante pour atteindre cet objectif conceptuel. Les mécanismes lésionnels et fonctionnels étant intriqués, en l'absence de biomarqueurs spécifiques pour b. et c. (ci-dessus), une quantité non négligeable d'information pertinente est perdue.

³Le néphron est l'unité structurale et fonctionnelle du rein. Il permet la formation d'urine. Chacun des deux reins humains en contient environ 1 000 000.

1.3.2 Limites de l'évaluation d'effets traitements à l'échelle populationnelle à un temps donné

Dans le paradigme classique de la médecine fondée sur les faits (*evidence-based medicine*), les résultats d'essais randomisés (simples ou combinés dans une méta-analyse) sont le plus haut niveau de preuve scientifique pour prendre la décision de traiter ou non un individu donné [22]. Avec cette approche les résultats d'essais randomisés ont souvent conduit à la prescription de traitements identiques pour toute la population de patients compatibles avec les critères d'inclusion de l'essai considéré. Cette attitude laisse peut laisser penser, à tort, que le bénéfice d'un traitement est identique pour tous les patients éligibles à l'essai. Les auteurs des récentes recommandations PATH (*Predictive Approaches to Treatment effect Heterogeneity*) ont rappelé l'invraisemblance de cette hypothèse [23] ; ils argumentent que si dans une population un effet traitement est constant sur une échelle relative (risque relatif ou hazard ratio), il est forcément variable d'un individu à l'autre sur une échelle absolue (réduction absolue du risque). L'inverse est également vrai : si dans une population un effet traitement est constant sur l'échelle absolue, il est forcément variable sur l'échelle relative.⁴

En réanimation, où les traitements évalués sont souvent invasifs, les résultats négatifs d'essais randomisés peuvent être expliqués par le mélange d'individus qui bénéficieraient du traitement et d'individus pour lesquels le traitement serait délétère. D'autre part dans cette discipline où l'état des patients change rapidement, il est légitime de penser que pour un individu donné, l'effet d'un traitement change sensiblement au cours du temps. Plusieurs experts ont par exemple articulé le fait que chez un patient très instable à la phase aiguë de l'insuffisance rénale aiguë l'épuration extra-rénale puisse être délétère ; mais ce traitement devient parfois nécessaire (bénéfique) au deuxième ou troisième jour — une fois le cap critique passé [24, 25]. Les analyses actuelles des essais randomisés n'apportent aucune information vis à vis de ce type de questions. Pour y répondre rigoureusement, il faudrait réaliser un essai dans lequel les patients seraient randomisés chaque jour pendant trois jours pour initier ou non l'épuration extra rénale. Ce type d'essai appellé SMART (*Sequential Multiple Assignment Randomized Trial*) [26] n'a jamais été conduit en réanimation pour des raisons de coût, de temps et de contraintes logistiques.

Il est important de se rappeler que les stratégies évaluées dans les essais (voir section 1.2) conduisent les cliniciens à initier l'épuration extra rénale à des temps différents pour des patients différents.⁵ En un certain sens, il s'agit donc déjà de stratégies "personnalisées". A propos de la stratégie tardive désormais reprise par les recommandations internationales, il est important de remarquer que :

1. La personnalisation a été conçue avant le début des essais. Les variables sélectionnées (urée, kaliémie, œdème aiguë du poumon, etc.) et leurs seuils reposent sur la compréhension (forcément imparfaite) de mécanismes biologiques et physiologiques.
2. La stratégie tardive est "stationnaire" dans le temps : pour la décision d'initiation, les variables retenues et leurs seuils sont identiques à la phase initiale (défaillances

⁴Pour que l'effet d'un traitement soit constant sur les deux échelles pour tous les patients éligibles à un essai, il est nécessaire — mais pas suffisant — que tous les patients de l'essai aient exactement le même risque (sévérité) à la randomisation. Cette condition étant improbable en pratique, dans un essai l'hétérogénéité de réponse au traitement existe quasiment toujours.

⁵Le temps 0 est défini par convention comme le temps auquel survient le stade 3 (KDIGO) de l'insuffisance rénale aiguë.

souvent intenses) et à la phase retardée (défaillances d'organes souvent stabilisées) de l'insuffisance rénale aiguë.

3. L'optimalité n'est pas certaine : le fait que la stratégie tardive ait été retenue comme supérieure aux stratégies précoce et très tardive n'implique pas qu'aucune stratégie ne soit supérieure à la stratégie tardive.

1.3.3 Objectifs de cette thèse

L'objectif principal de cette thèse était d'illustrer l'utilisation d'outils statistiques avancés dans le but de développer une stratégie d'initiation de l'épuration extra-rénale en réanimation qui soit i) personnalisée par l'analyse rigoureuse de données clinique, ii) non-stationnaire dans le temps, iii) supérieure à la meilleure stratégie actuellement retenue. Un objectif secondaire de cette thèse était d'illustrer comment les applications concrètes en biostatistiques nourrissent les développements théoriques en inférence causale — et vice versa.

Dans le chapitre suivant, nous décrivons certains outils théoriques utilisés dans cette thèse. Dans le chapitre 3 nous illustrons le principe des essais émulés en réanimation. Dans le chapitre 4 nous présentons un travail évaluant l'hétérogénéité de l'effet d'un traitement par niveaux de risque prédit. Dans le chapitre 5 nous introduisons une méthode pour l'évaluation de règles de médecine personnalisée dans différentes situations à partir de données observationnelles. Dans le chapitre 6, nous décrivons le développement et la validation d'une stratégie dynamique de l'épuration extra-rénale en réanimation. Le chapitre 7 présente une interprétation personnelle visant à mettre en perspective les résultats importants de cette thèse. Enfin, en Annexe A nous allons au-delà de la médecine personnalisée en réanimation et introduisons un modèle causal faisant le lien entre différents outils théoriques utilisés dans cette thèse.

Chapter 2

Revue de certains outils théoriques utilisés dans cette thèse

We see the world in terms of our theories.

Thomas Kuhn, *The Structure of Scientific Revolutions*

2.1 Inférence causale

Le point de départ de la théorie de l’inférence causale est de remarquer que dans des données observationnelles, les associations statistiques entre les variables ne sont pas nécessairement de nature causale. L’objectif de l’inférence causale est d’établir un cadre conceptuel pour décrire des relations *causales* à partir de données observées. En d’autres termes, la théorie vise à annuler le biais de confusion pour décrire précisément l’effet d’interventions — ou actions — sur des résultats futurs.

2.1.1 Notations et hypothèses d’identifiabilité

Dans sa version la plus simple, le cadre conceptuel de l’inférence causale suppose l’existence de deux variables aléatoires latentes $\{Y^{a=0}, Y^{a=1}\}$ dits “résultats potentiels” (*potential outcomes*, voir modèle causal de Neyman et Rubin [27–29]). Ces variables représentent le résultat d’un individu s’il venait à recevoir respectivement l’intervention $A = 0$ ou l’intervention $A = 1$. Dans le cas usuel que nous présentons dans cette section, l’identification des effets causaux d’intérêt repose, sur les trois hypothèses suivantes.

Hypothèse 1 (Cohérence et non interférence ou en anglais “Stable Unit Treatment Value Assumption”). *Il n’existe qu’une version de l’intervention $A = 1$ et qu’une version de l’intervention $A = 0$. Par ailleurs, aucune intervention sur un individu j ne peut affecter le résultat Y d’un individu $i \neq j$. Classiquement on note cette hypothèse*

$$Y_i = A_i Y_i^{a=1} + (1 - A_i) Y_i^{a=0}, \quad i = 1, \dots, n.$$

Hypothèse 2 (Absence de facteurs de confusion non mesurés ou en anglais “Ignorability”). *Toutes les variables causant à la fois l’intervention A et la survenue de*

Y sont mesurées. Cette hypothèse se note

$$\{Y^{a=0}, Y^{a=1}\} \perp\!\!\!\perp A|X$$

où X désigne un ensemble de variables mesurées avant l'intervention A .

Hypothèse 3 (Positivité ou en anglais “Overlap”). Pour tous les niveaux réalistes de X , les individus pouvaient recevoir, avec une probabilité non nulle, chacune des deux interventions. En notant le score de propension $e(x) = \mathbb{E}[A|X = x]$, cette hypothèse s'écrit de la manière suivante. Il existe une constante η telle que

$$\forall x \in \mathcal{X}, \quad 0 < \eta < e(x) < 1 - \eta < 1.$$

2.1.2 Estimands causaux

L'estimand représente la quantité cible à estimer. En inférence causale, celui-ci peut être vu comme un paramètre caractérisant certains aspects de la distribution jointe $p_{X,A,\dots,Y^{a=1},Y^{a=0}}$.¹ Ci-dessous nous présentons quelques exemples d'estimands causaux courants ainsi que leur interprétation.

- **Effet traitement moyen** (ou *average treatment effect*: ATE)

$$\Delta \stackrel{\text{def}}{=} \mathbb{E}(Y^{a=1} - Y^{a=0})$$

Cet estimand représente l'effet à l'échelle populationnelle de l'intervention $A = 1$ sur le résultat Y dans un essai randomisé de taille infinie où l'intervention serait comparée au contrôle $A = 0$.

- **Effet traitement moyen chez les traités** (ou *ATE in the treated*: ATT)

$$\Delta^{A1} \stackrel{\text{def}}{=} \mathbb{E}(Y^{a=1} - Y^{a=0}|A = 1)$$

Cet estimand représente l'effet de l'intervention $A = 1$ sur le résultat Y dans un essai randomisé de taille infinie où l'intervention serait comparée au contrôle $A = 0$ avec une analyse en sous groupe des patients qui, en dehors de l'essai, auraient spontanément reçu l'intervention.

- **Effet traitement moyen chez les non traités** (ou *ATE in the untreated*: ATU)

$$\Delta^{A0} \stackrel{\text{def}}{=} \mathbb{E}(Y^{a=1} - Y^{a=0}|A = 0)$$

Cet estimand représente l'effet de l'intervention $A = 1$ sur le résultat Y dans un essai randomisé de taille infinie où l'intervention serait comparée au contrôle $A = 0$ avec une analyse en sous groupe des patients qui, en dehors de l'essai, auraient spontanément reçu le contrôle.

- **Effet traitement chez les observants** (ou *complier average causal effect*: CACE)

$$\Delta^{C1} \stackrel{\text{def}}{=} \mathbb{E}(Y^{a=1} - Y^{a=0}|C = 1)$$

¹Alternativement un estimand peut être vu comme une fonctionnelle de la distribution jointe. Afin de simplifier les notations, nous considérons dans ce chapitre un estimand comme un paramètre.

Cet estimand représente l'effet de l'intervention $A = 1$ sur le résultat Y dans un essai randomisé de taille infinie où l'intervention serait comparée au contrôle $A = 0$ avec une analyse en sous groupe des patients qui auraient toujours reçu intervention/contrôle tel que le demande la randomisation. Ici on considère la variable latente C qui indique le statut observant ou non d'un patient. De nouveaux estimateurs de CACE sont donnés dans l'Annexe A.

Dans la section suivante nous nous concentrerons sur un estimand particulier : l'effet traitement moyen Δ .

2.1.3 Estimateurs de l'effet traitement moyen

Considérant les données $(X_i, A_i, Y_i)_{1 \leq i \leq n} \stackrel{\text{iid}}{\sim} \mathcal{P}$, observationnelles ou issues d'un essai randomisé, on s'intéresse à l'estimation de l'effet traitement moyen Δ . Pour cela, il est nécessaire de poser un modèle (paramétrique ou non) pour le score de propension $\mathbb{E}[A|X]$ et/ou un modèle pronostique (paramétrique ou non) pour $\mathbb{E}[Y|A, X]$. Dans cette section, nous notons respectivement ces modèles $e(X; \beta)$ et $q(A, X; \gamma)$ et considérons que ceux-ci sont correctement spécifiés. Pour clarifier l'exposition, dans le reste de cette sous-section, nous supposerons que l'estimation des paramètres de ces modèles est donnée par un oracle, c'est à dire $\hat{\beta} = \beta$ et $\hat{\gamma} = \gamma$.

- **Estimateur IPW (inverse probability of treatment weighting)**

$$\hat{\Delta}_{IPW} = n^{-1} \sum_{i=1}^n \frac{A_i Y_i}{e(X_i; \hat{\beta})} - \frac{(1 - A_i) Y_i}{1 - e(X_i; \hat{\beta})}$$

A taille d'échantillon finie, l'estimateur $\hat{\Delta}_{IPW}$ est sans biais [30].

- **Estimateur IPW normalisé**

$$\hat{\Delta}_{IPW^*} = \sum_{i=1}^n \frac{\frac{A_i Y_i}{e(X_i; \hat{\beta})}}{\sum_{i=1}^n \frac{A_i}{e(X_i; \hat{\beta})}} - \frac{\frac{(1 - A_i) Y_i}{1 - e(X_i; \hat{\beta})}}{\sum_{i=1}^n \frac{1 - A_i}{1 - e(X_i; \hat{\beta})}}$$

A taille d'échantillon finie, l'estimateur $\hat{\Delta}_{IPW^*}$ est biaisé (voir *weighted importance sampling* dans Rubinstein et Kroese, page 149 [31]). La variance de cet estimateur étant inférieure à celle de $\hat{\Delta}_{IPW}$, l'utilisation de $\hat{\Delta}_{IPW^*}$ est cependant préférable en pratique.

- **Estimateur par G-computation**

$$\hat{\Delta}_G = n^{-1} \sum_{i=1}^n q(1, X_i; \hat{\gamma}) - q(0, X_i; \hat{\gamma})$$

A taille d'échantillon finie, l'estimateur $\hat{\Delta}_G$ est sans biais. La variance de $\hat{\Delta}_G$ est inférieure à celle de $\hat{\Delta}_{IPW^*}$.

- **Estimateur AIPW** (*augmented inverse probability of treatment weighting*)

$$\begin{aligned}\hat{\Delta}_{AIPW} = n^{-1} \sum_{i=1}^n & \left[\frac{A_i(Y_i - q(1, X_i; \hat{\gamma}))}{e(X_i; \hat{\beta})} + q(1, X_i; \hat{\gamma}) \right] \\ & - \left[\frac{(1 - A_i)(Y_i - q(0, X_i; \hat{\gamma}))}{1 - e(X_i; \hat{\beta})} + q(0, X_i; \hat{\gamma}) \right]\end{aligned}$$

A taille d'échantillon finie, l'estimateur $\hat{\Delta}_{AIPW}$ est sans biais si l'un ou l'autre des modèles score de propension ou pronostique est sans biais (c'est à dire si $\hat{\beta} = \beta$ ou $\hat{\gamma} = \gamma$). Pour cette raison, l'estimateur $\hat{\Delta}_{AIPW}$ est dit "double-robuste". Quand l'un ou l'autre des modèles score de propension ou pronostique est bien spécifié, la variance asymptotique de $\hat{\Delta}_{AIPW}$ est inférieure à celle de $\hat{\Delta}_{IPW}$ [32]. Parmi la classe d'estimateurs de type *IPW* décrite par Robins et al., l'estimateur $\hat{\Delta}_{AIPW}$ est l'estimateur avec la plus petite variance asymptotique [33]. Cette propriété d'efficience rend l'utilisation de $\hat{\Delta}_{AIPW}$ attractive en pratique.

- **Estimateur TMLE** (*targeted maximum likelihood estimation*)

Cet estimateur est un estimateur par substitution ; son expression simplifiée est différente en fonction de la nature binaire ou continue de Y . Dans le cas où Y est continu l'expression de l'estimateur TMLE est :

$$\hat{\Delta}_{TMLE} = n^{-1} \sum_{i=1}^n q(1, X_i; \hat{\gamma}) - q(0, X_i; \hat{\gamma}) + \hat{\epsilon} \left\{ H(1, X_i; \hat{\beta}) - H(0, X_i; \hat{\beta}) \right\}$$

où

$$H(A_i, X_i; \hat{\beta}) = \frac{A_i}{e(X_i; \hat{\beta})} - \frac{1 - A_i}{1 - e(X_i; \hat{\beta})}$$

désigne la "variable astucieuse" (*clever covariate*) et ϵ est estimé lors d'une l'étape dite de "ciblage" par

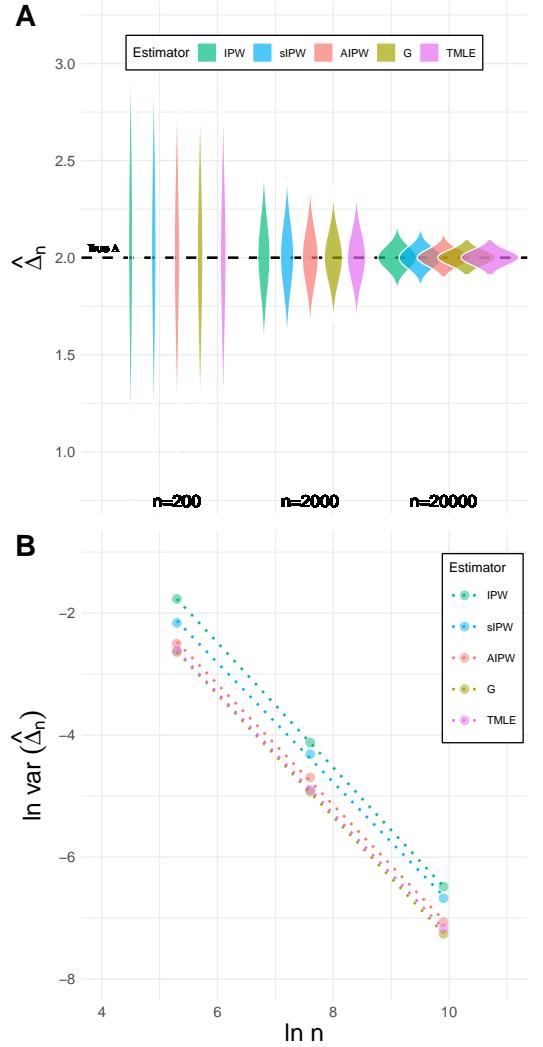
$$\hat{\epsilon} = \arg \min_{\epsilon} \sum_{i=1}^n \left\{ Y_i - q(A_i, X_i; \hat{\gamma}) - \epsilon H(A_i, X_i; \hat{\beta}) \right\}^2.$$

L'expression de l'estimateur TMLE dans le cas où Y est binaire est donnée dans l'annexe B1. A taille d'échantillon finie, l'estimateur $\hat{\Delta}_{TMLE}$ est sans biais si l'un ou l'autre des modèles score de propension ou pronostique est sans biais. Pour cette raison, l'estimateur $\hat{\Delta}_{TMLE}$ est dit "double-robuste". Les estimateurs $\hat{\Delta}_{TMLE}$ et $\hat{\Delta}_G$, contrairement aux autres estimateurs présentés dans cette section, n'appartiennent pas à la classe d'estimateurs de type *IPW* décrite par Robins et al. L'estimateur $\hat{\Delta}_{TMLE}$ est généralement considéré plus efficient que $\hat{\Delta}_{AIPW}$ et moins sensible aux valeurs aberrantes [34]. La variance asymptotique de $\hat{\Delta}_{TMLE}$ est inférieure à celle de $\hat{\Delta}_G$.

Les cinq estimateurs de l'effet traitement mentionnés ci-dessus sont asymptotiquement normaux ; ils sont convergents à la vitesse $O_p(n^{-1/2})$ (vitesse de convergence maximale) si $e(x; \hat{\beta})$ et/ou $q(a, x; \hat{\gamma})$ convergent à une vitesse supérieure à $O_p(n^{-1/4})$ [35]. Cette dernière propriété rend légitime l'utilisation de modèles non-paramétriques (boosting, forêts aléatoires, réseaux de neurones profonds, etc.) pour estimer les fonctions $e(x; \beta)$ et/ou $q(a, x; \gamma)$.

La variance asymptotique des cinq estimateurs mentionnés ci-dessus est cependant différente. Ce point est illustré ci-dessous dans des simulations à visée pédagogique où les modèles pour $e(x; \beta)$ et $q(a, x; \gamma)$ sont paramétriques et bien spécifiés (Figure 3).²

Figure 3: Variance et vitesse de convergence des estimateurs de l'effet traitement moyen : simulation de 1000 échantillons de taille $n = 200, 2000$ et 20000 . Panneau A: Diagrammes en violon décrivant la distribution des 1000 estimations pour les estimateurs $\hat{\Delta}_{IPW}$, $\hat{\Delta}_{IPW^*}$, $\hat{\Delta}_G$, $\hat{\Delta}_{AIPW}$ et $\hat{\Delta}_{TMLE}$. Notez la forme des violons, dès $n = 200$, celle-ci décrit bien la normalité asymptotique des estimateurs. Toutes les distributions semblent centrées sur la vraie valeur Δ : cela illustre le caractère sans biais de $\hat{\Delta}_{IPW}$, $\hat{\Delta}_G$, $\hat{\Delta}_{AIPW}$ et $\hat{\Delta}_{TMLE}$ (l'estimateur $\hat{\Delta}_{IPW^*}$ étant convergent, dès $n = 200$ son caractère biaisé est indétectable numériquement). Notez que pour une taille d'échantillon donnée, la distribution des estimations est de variance différente d'un estimateur à l'autre : cela illustre la différence d'efficience des estimateurs. Notez enfin que la variance diminue quand la taille de l'échantillon augmente. Panneau B: Vitesse de convergence commune $n^{1/2}$ et efficience différente des cinq estimateurs. Pour tout estimateur $\hat{\Delta}_n$ asymptotiquement uniformément intégrable vérifiant $n^\alpha(\hat{\Delta}_n - \Delta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma)$ l'égalité $\ln \text{var}(\hat{\Delta}_n) = \ln \Sigma - 2\alpha \ln n + o_p(1)$ est vérifiée. Cela suggère, quand n est grand, de modéliser $\ln \text{var}(\hat{\Delta}_n)$ en fonction de $\ln n$ par régression linéaire. La pente -2α décrit la vitesse de convergence et l'ordonnée à l'origine $\ln \Sigma$ décrit la variance asymptotique de $\hat{\Delta}_n$. Dans cette simulation on observe que les cinq estimateurs ont la même pente et une ordonnée à l'origine différente. Par ailleurs, dans cette simulation on observe que la pente $-2\hat{\alpha} \approx -1$ semble identique pour les cinq estimateurs. Cela suggère que tous les estimateurs ont le même paramètre $\alpha = 1/2$ décrivant leur vitesse de convergence.



²Le code R reproduisant ces simulations est accessible sur <https://github.com/fcgrolleau/ate-simulations>.

2.2 M-estimation

La théorie de la M-estimation permet de déduire les propriétés asymptotiques d'une classe d'estimateurs fréquemment rencontrée en pratique. Dans cette section nous adaptons l'exposition proposée par Tsiatis et al. [36] pages 32 à 38, et Boos & Stefanski pages 297 à 337 [37] pour formuler un rappel succinct de la théorie en nous focalisant sur les M-estimateurs de type ψ .³ Nous illustrerons ensuite la mise en pratique de la M-estimation dans un exemple. L'implémentation du calcul d'intervalles de confiance est également développée à l'aide du package **Mestim** qui est un développement annexe de cette thèse.

2.2.1 Théorie

En voyant d'abord l'estimand d'intérêt comme un paramètre (possiblement multidimensionnel) d'une distribution de probabilité, un M-estimateur de type ψ est défini de la manière suivante.

Définition 1 (M-estimateur de type ψ – notation paramétrique). *En considérant $Z_1, \dots, Z_n \stackrel{iid}{\sim} F_\theta$, l'estimateur $\hat{\theta}_n$ est un M-estimateur de type ψ s'il existe une fonction $\psi(\cdot, \cdot)$ telle que $\hat{\theta}_n$ soit la solution (supposant qu'elle existe et soit bien définie) d'une équation de la forme*

$$\sum_{i=1}^n \psi(Z_i, \hat{\theta}_n) = \mathbf{0}.$$

Par convention, la fonction ψ d'un M-estimateur est appelée “fonction d'estimation sans biais”.⁴ La définition 1 est équivalente à la définition ci-dessous où un estimateur est vu comme une fonctionnelle d'une distribution de probabilité.

Définition 2 (M-estimateur de type ψ – notation fonctionnelle). *En considérant $Z_1, \dots, Z_n \stackrel{iid}{\sim} F$, la fonctionnelle $T(\cdot)$ est un M-estimateur de type ψ si il existe une fonction $\psi(\cdot, \cdot)$ telle que*

$$\int_Z \psi(z, T(F_n)) dF_n(z) = \mathbf{0}.$$

Dans le reste de cette section, afin de simplifier les notations nous utilisons la notation paramétrique donnée dans la définition 1. Les résultats de cette section sont donnés avec la notation fonctionnelle en annexe B2. Un résultat asymptotique important de la théorie de la M-estimation est le suivant.

Théorème 1 (Variance asymptotique sandwich). *Si $\hat{\theta}_n$ est un M-estimateur de type ψ avec $\psi(z, \delta)$ dérivable selon δ , z –(presque partout) alors $\hat{\theta}_n$ est \sqrt{n} –convergent et*

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma)$$

où la matrice de variance-covariance Σ a pour expression

$$\Sigma = \left[\mathbb{E} \left\{ \frac{\partial \psi(Z, \delta)}{\partial \delta^T} \Big|_{\delta=\theta} \right\} \right]^{-1} \mathbb{E} \left\{ \psi(Z, \theta) \psi^T(Z, \theta) \right\} \left[\left[\mathbb{E} \left\{ \frac{\partial \psi(Z, \delta)}{\partial \delta^T} \Big|_{\delta=\theta} \right\} \right]^{-1} \right]^T.$$

³Des M-estimateurs de type ρ ont également été décrits. Ces estimateurs étant moins fréquemment rencontrés en pratique nous ne les aborderons pas dans cette section.

⁴Notez que certains estimateurs biasés satisfont la définition 1. L'estimateur $\hat{\Delta}_{IPW*}$ est un exemple : bien que biaisé, il a une fonction d'estimation sans bias ; c'est donc un M-estimateur de type ψ .

Ce résultat découle de l'application successive du théorème central limite et du théorème de Slutsky. Une preuve est donnée dans Boos et Stefansky, page 300 [37].

Corollaire 1 (Variance asymptotique sandwich empirique). *Si $\hat{\theta}_n$ est un M-estimateur de type ψ avec $\psi(z, \delta)$ dérivable selon δ , z-(presque partout) alors la distribution asymptotique de $\hat{\theta}_n$ peut être estimée par*

$$\hat{\theta}_n \stackrel{d}{\sim} \mathcal{N}(0, n^{-1}\hat{\Sigma})$$

où

$$\hat{\Sigma} = A^{-1}B(A^{-1})^T$$

avec

$$A = n^{-1} \sum_{i=1}^n \left\{ \frac{\partial \psi(Z_i, \delta)}{\partial \delta^T} \Big|_{\delta=\hat{\theta}_n} \right\},$$

$$B = n^{-1} \sum_{i=1}^n \left\{ \psi(Z_i, \hat{\theta}_n) \psi^T(Z_i, \hat{\theta}_n) \right\}.$$

2.2.2 Illustration du principe d'empilement d'équations d'estimation

En guise d'exemple, nous proposons d'estimer la variance asymptotique de l'estimateur $\hat{\Delta}_G$ de l'effet traitement moyen. Pour calculer $\hat{\Delta}_G$, nous proposons dans cet exemple d'estimer distinctement un modèle pronostique pour les individus recevant l'intervention (traitement) et pour les individus ne le recevant pas (configuration parfois appelée *Two-learners*). L'objectif est d'illustrer comment la théorie de la M-estimation permet dans ces conditions d'obtenir un intervalle de confiance asymptotique pour $\hat{\Delta}_G$ tenant compte de l'incertitude sur les deux modèles pronostiques estimés. Pour cela nous utiliserons la technique d'empilement d'équations d'estimations. Avec la configuration dite *Two-learners*, on utilise les représentations $q_1(X) \stackrel{\text{def}}{=} \mathbb{E}[Y|X, A = 1]$ et $q_0(X) \stackrel{\text{def}}{=} \mathbb{E}[Y|X, A = 0]$ pour reformuler

$$\Delta \stackrel{\text{def}}{=} \mathbb{E}[Y^{a=1} - Y^{a=0}] = \mathbb{E}[q_1(X) - q_0(X)].$$

On suppose que pour $q_1(x)$ et $q_0(x)$ les modèles $q_1(x; \gamma_1)$ et $q_0(x; \gamma_0)$ sont correctement spécifiés et on note par ailleurs $\gamma = (\gamma_0^T, \gamma_1^T)^T$. En considérant les données $\mathbf{Z}_i = (\mathbf{X}_i^T, A_i, Y_i)^T$, $i = 1, \dots, n$, l'expression de l'estimateur de l'effet traitement moyen par G-computation *Two-learners* est :

$$\hat{\Delta}_G = n^{-1} \sum_{i=1}^n q_1(\mathbf{X}_i; \hat{\gamma}_1) - q_0(\mathbf{X}_i; \hat{\gamma}_0). \quad (2.1)$$

Dans cet exemple minimal à visée pédagogique nous considérons le résultat Y continu et utilisons des modèles linéaires pour $q_1(\cdot; \gamma_1)$ et $q_0(\cdot; \gamma_0)$. Dans ces conditions, $\hat{\gamma}_0$ est solution de l'équation d'annulation du gradient des moindres carrés dans le sous-groupe $\{A = 0\}$:

$$\sum_{i=1}^n \mathbb{1}(A_i = 0)(Y_i - \hat{\gamma}_0^T \mathbf{X}_i) \mathbf{X}_i = \mathbf{0}. \quad (2.2)$$

De la même manière, $\hat{\gamma}_1$ est solution de l'équation d'annulation du gradient des moin-

dres carrés dans le sous-groupe $\{A = 1\}$:

$$\sum_{i=1}^n \mathbb{1}(A_i = 1)(Y_i - \hat{\gamma}_1^T \mathbf{X}_i) \mathbf{X}_i = \mathbf{0}. \quad (2.3)$$

En notant le vecteur “empilement” des paramètres estimés $\hat{\boldsymbol{\theta}} = (\hat{\gamma}_0^T, \hat{\gamma}_1^T, \hat{\Delta}_G)^T$,⁵ les représentations des équations (2.2) et (2.3) permettent d’identifier respectivement

$$\begin{aligned} \psi_0(\mathbf{z}, \boldsymbol{\theta}) &= \mathbb{1}(a = 0)(y - \gamma_0^T \mathbf{x}) \mathbf{x} \quad \text{et} \\ \psi_1(\mathbf{z}, \boldsymbol{\theta}) &= \mathbb{1}(a = 1)(y - \gamma_1^T \mathbf{x}) \mathbf{x}. \end{aligned}$$

Ces identifications impliquent que $\hat{\gamma}_0$ et $\hat{\gamma}_1$, satisfaisant la Définition 1, sont tous deux des M-estimateurs de type ψ . On remarque que l’équation (2.1) peut être réarrangée de façon à ce que $\hat{\Delta}_G$ apparaisse comme solution de l’équation

$$\sum_{i=1}^n \left\{ q_1(\mathbf{X}_i; \hat{\gamma}_1) - q_0(\mathbf{X}_i; \hat{\gamma}_0) - \hat{\Delta}_G \right\} = 0.$$

Cette représentation permet d’identifier

$$\begin{aligned} \psi_2(\mathbf{z}, \boldsymbol{\theta}) &= q_1(\mathbf{x}; \gamma_1) - q_0(\mathbf{x}; \gamma_0) - \Delta_G \\ &= (\gamma_1 - \gamma_0)^T \mathbf{x} - \Delta_G. \end{aligned}$$

Cette identification implique que $\hat{\Delta}_G$ est un M-estimateur de type ψ si $\hat{\gamma}_0$ et $\hat{\gamma}_1$ sont eux-mêmes des M-estimateurs de type ψ .⁶ Dans ces conditions, on dit parfois que l’estimateur $\hat{\Delta}_G$ est un M-estimateur partiel. On peut désormais identifier la fonction d’estimation sans biais ψ comme un “empilement” de fonctions d’estimation sans biais :

$$\psi(\mathbf{z}, \boldsymbol{\theta}) = \{\psi_0(\mathbf{z}, \boldsymbol{\theta}), \psi_1(\mathbf{z}, \boldsymbol{\theta}), \psi_2(\mathbf{z}, \boldsymbol{\theta})\}^T.$$

En appliquant le résultat du Corollaire 1 on a

$$\hat{\boldsymbol{\theta}} \stackrel{\sim}{\sim} \mathcal{N}(0, n^{-1} \hat{\Sigma})$$

où la matrice $\hat{\Sigma}$ est obtenue via la formule de l’estimateur empirique de la variance asymptotique sandwich. Plus spécifiquement pour $\hat{\Delta}_G$, la distribution asymptotique estimée est

$$\hat{\Delta}_G \stackrel{\sim}{\sim} \mathcal{N}(0, n^{-1} \hat{\Sigma}_{der})$$

où $\hat{\Sigma}_{der}$ désigne le dernier élément de $\text{diag}(\hat{\Sigma})$. Un intervalle de confiance asymptotique pour $\hat{\Delta}_G$ peut donc être calculé par

$$\hat{IC}_{1-\alpha}(\hat{\Delta}_G) = \hat{\Delta}_G \pm z_{1-\alpha/2} \sqrt{\hat{\Sigma}_{der}}$$

où z_ν désigne le quantile ν d’une loi normale centrée réduite.

⁵Notez qu’avec l’astuce “d’empilement”, l’estimateur $\hat{\Delta}_G$ est ici représenté comme un élément (le dernier) du vecteur de paramètres estimés.

⁶Cette dernière condition a été vérifiée ci-dessus avec l’identification des fonctions ψ_0 et ψ_1 .

2.2.3 Implémentation avec le package `Mestim`

Le package `Mestim`, disponible sur CRAN, est un développement annexe de cette thèse. Il permet d'utiliser la méthode d'empilement des équations d'estimation pour obtenir dans R la distribution asymptotique de M-estimateurs de type ψ . Nous décrivons rapidement le fonctionnement du package avant d'illustrer son implémentation (Figure 4).

Le programme `Mestim::get_vcov` prend en entrée des données (`df` désigne $(Z_i)_{1 \leq i \leq n}$), un “empilement” des paramètres estimés (`thetas_hat` désigne $\hat{\theta}$) et un “empilement” des fonctions d'estimation sans biais correspondantes (`psi` désigne ψ). Ce programme réalise – en parallèle – une dérivation algorithmique formelle de ψ afin d'obtenir la matrice jacobienne $\frac{\partial\psi(z,\delta)}{\partial\delta^T}$. L'estimateur sandwich empirique est ensuite calculé tel qu'indiqué dans le Corollaire et la matrice de variance-covariance $n^{-1}\hat{\Sigma}$ est retournée à l'utilisateur pour permettre le calcul d'intervalle de confiance asymptotiques.

Figure 4: Exemple minimal d'empilement d'équations d'estimation dans `Mestim` : calcul de la variance asymptotique de l'estimateur $\hat{\Delta}_G$ dans la configuration dite *Two-learners* avec X, Y continus et unidimensionnels.

```
thetas_hat <- list(gamma_0 = gamma_0_hat,
                     gamma_1 = gamma_1_hat,
                     delta = delta_hat)

psi_0 <- expression( (1-A) * (Y - gamma_0 * X) * X )
psi_1 <- expression( A * (Y - gamma_1 * X) * X )
psi_2 <- expression( (gamma_1 - gamma_0) * X - delta )

psi <- list(psi_0, psi_1, psi_2)

res <- Mestim::get_vcov(data = df, thetas = thetas_hat, M = psi)
var_hat_delta_hat <- tail(diag(res$vcov), 1)
```

L'utilisation du programme `base::parse` permettant la constitution d'expressions à partir de chaînes de caractères, la constitution des fonctions `psi_j` est facile à automatiser dans le cas où X est multidimensionnel. À terme, le package `Mestim` a un triple objectif :

- **Pédagogique** : Le principe d'empilement d'équations d'estimation est illustré avec plusieurs exemples dans une vignette disponible sur <https://fcgrolleau.github.io/Mestim/vignette/vignette.html>.
- **Aide à la réalisation de simulations de Monte-Carlo** : `Mestim` a été utilisé dans des simulations de Monte-Carlo évaluant la couverture d'intervalles de confiances dans le travail d'autres étudiants en thèse.⁷
- **Pièce élémentaire pour la construction d'autres packages** : Le package `Mestim` est appelé dans le package `CACEmix` afin de calculer des intervalles de confiance asymptotiques pour l'effet traitement chez les compliants (CACE).

⁷`Mestim` ne s'applique que pour des M-estimateurs de type ψ . Pour les estimateurs d'autres classes, nous avons développé la librairie Python `Speedboot` qui repose sur le bootstrap. Celle-ci accessible sur Pypi et à l'adresse <https://github.com/fcgrolleau/Speedboot>.

2.3 Hetérogénéité de l'effet du traitement

Dans cette section nous suivons la structure d'exposition des recommandations PATH (*Predictive Approaches to Treatment effect Heterogeneity*) [23] et distinguons deux approches pour évaluer l'hétérogénéité de réponse au traitement : i) modélisation de l'effet du traitement par niveaux de risque prédit (*risk modeling approach*) et ii) modélisation directe de l'effet du traitement (*treatment effect modeling approach*).

2.3.1 Modélisation de l'effet du traitement par niveaux de risque prédit

L'argumentation développée dans cette section se focalise sur l'utilisation de données d'essais randomisés.⁸ L'approche de modélisation par niveaux de risque prédit s'appuie sur le constat que l'hétérogénéité de "l'effet du traitement" dépend de l'échelle choisie pour mesurer cet effet [38]. L'effet du traitement peut être mesuré sur une échelle absolue (réduction de risque) ou relative (ratio de risque ou hazard ratio en cas de données censurées). L'approche présentée dans cette section repose sur l'utilisation d'un modèle validé pour prédire un résultat d'intérêt Y^* binaire (Y^* peut être choisi égal à Y ou Y^* peut être choisi comme étant un résultat différent de Y). Ci-dessous, nous présentons la procédure en cinq étapes en nous focalisant sur le cas fréquent d'un essai randomisé avec critère de jugement principal censuré (par exemple temps avant décès).

1. Prédire le Risque (probabilité postérieure) d'un événement d'intérêt à partir des covariables pré-randomisation X , en utilisant un modèle existant, validé :

$$\tilde{R}(X) \leftarrow \hat{\mathbb{P}}(Y^* = 1|X).$$

En l'absence de modèle existant, un modèle peut être développé et validé à partir des données de l'essai à analyser.

2. Modéliser le critère de jugement principal en utilisant une interaction entre l'intervention A et une transformation du risque prédit $\tilde{R}(X)$. Si le critère de jugement principal est censuré, on peut par exemple utiliser le modèle de Cox suivant :

$$\lambda(t|A, X; \beta) = \lambda_0(t) \exp \{ \beta_0 A + \beta_1^T g^{(d)}(\tilde{R}(X)) + \beta_2^T A \times g^{(d)}(\tilde{R}(X)) \}$$

où t désigne le temps de suivi post-randomisation, $g^{(d)} : \mathbb{R} \rightarrow \mathbb{R}^d$ désigne une fonction deux fois différentiable (par exemple spline cubique naturelle), les paramètres $(\beta_0, \beta_1^T, \beta_2^T)^T = \beta$ sont estimés par méthode de maximisation de vraisemblance partielle.⁹

3. Prédire l'effet traitement conditionnel à $\tilde{R}(X)$ sur les échelles relatives et absolues. Avec le modèle mentionné ci-dessus pour un critère de jugement principal censuré on obtient les effets traitement conditionnels suivants.

⁸Grâce à l'emploi d'outils d'inférence causale non détaillés ici, la méthode présentée dans cette section peut être étendue à l'utilisation de données observationnelles.

⁹La fonction $g^{(d)}$ peut être définie à priori (e.g., spline cubique naturelle à d nœuds) où estimée dans un modèle additif généralisé avec un paramètre de lissage déterminé par LOOCV.

- Sur une échelle relative, le hazard ratio conditionnel estimé est

$$\hat{HR}(X) = \frac{\lambda(t|1, X; \hat{\beta})}{\lambda(t|0, X; \hat{\beta})} = \exp \{ \hat{\beta}_0 + \hat{\beta}_2^T g^{(d)}(\tilde{R}(X)) \}.$$

Des intervalles de confiance peuvent être calculés en utilisant la formule de l'estimation asymptotique de l'erreur type pour $\ln \hat{HR}(X)$:

$$\hat{se}\{ \ln \hat{HR}(X) \} = \sqrt{d(X)^T \text{cov}(\hat{\beta}) d(X)}$$

où

$$d(X) = \begin{bmatrix} 1 \\ 0 \\ g^{(d)}(\tilde{R}(X))^T \end{bmatrix}.$$

- Sur une échelle absolue, pour le temps de suivi post-randomisation t , la différence de risque conditionnelle estimée est

$$\hat{RD}(X, t) = \hat{S}(t|1, X) - \hat{S}(t|0, X)$$

où

$$\hat{S}(t|A, X) = \exp\{ -\hat{\Lambda}_0(t) \}^{\exp \hat{\beta}_0 A + \hat{\beta}_1^T g^{(d)}(\tilde{R}(X)) + \hat{\beta}_2^T A \times g^{(d)}(\tilde{R}(X))}$$

désigne la fonction de survie conditionnelle et $\Lambda_0(t)$ désigne la fonction de base du hasard cumulé. Cette dernière fonction peut être estimée par l'estimateur non-paramétrique de Breslow [39] et des intervalles de confiance peuvent être calculés pour $\hat{RD}(X, t)$ en utilisant un ré-échantillonnage bootstrap.

4. Test statistique de l'hétérogénéité de l'effet du traitement en comparant deux modèles non emboîtés. Dans le cas d'un critère de jugement censuré, on peut réaliser un test basé sur le rapport de vraisemblance partielle [40] pour comparer les deux modèles de Cox suivants

$$\begin{aligned} \mathcal{M}_1 : \quad & \lambda(t|A, X; \beta) = \lambda_0(t) \exp \{ \beta_0 A + \beta_1 \tilde{R}(X) + \beta_2 A \times \tilde{R}(X) \}, \\ \mathcal{M}_2 : \quad & \lambda(t|A, X; \beta) = \lambda_0(t) \exp \{ \beta_0 A + \beta_1^T g^{(d)}(\tilde{R}(X)) + \beta_2^T A \times g^{(d)}(\tilde{R}(X)) \}. \end{aligned}$$

La P -valeur teste l'hypothèse nulle que le modèle hétérogène souple \mathcal{M}_2 décrit les données aussi bien que le modèle hétérogène linéaire \mathcal{M}_1 (avec \mathcal{M}_1 on a : $\ln HR(X) = \beta_0 + \beta_2 \tilde{R}(X)$). Dans ces conditions, rejeter l'hypothèse nulle est conservateur pour tester l'hétérogénéité de l'effet du traitement. On remarque en particulier que le test est plus conservateur qu'un test comparant \mathcal{M}_2 à un modèle supposant un effet constant du traitement sur l'échelle relative (c'est-à-dire un modèle \mathcal{M}_0 impliquant $\ln HR(X) = \beta_0$).

5. Présentation de l'effet du traitement conditionnel au risque prédit sur les échelles relative et absolue. L'effet du traitement est présenté par quarts ou cinquièmes du risque prédit accompagné du lissage et de la P -valeur pour l'hétérogénéité tous deux résultats de la modélisation (voir Figure 5).

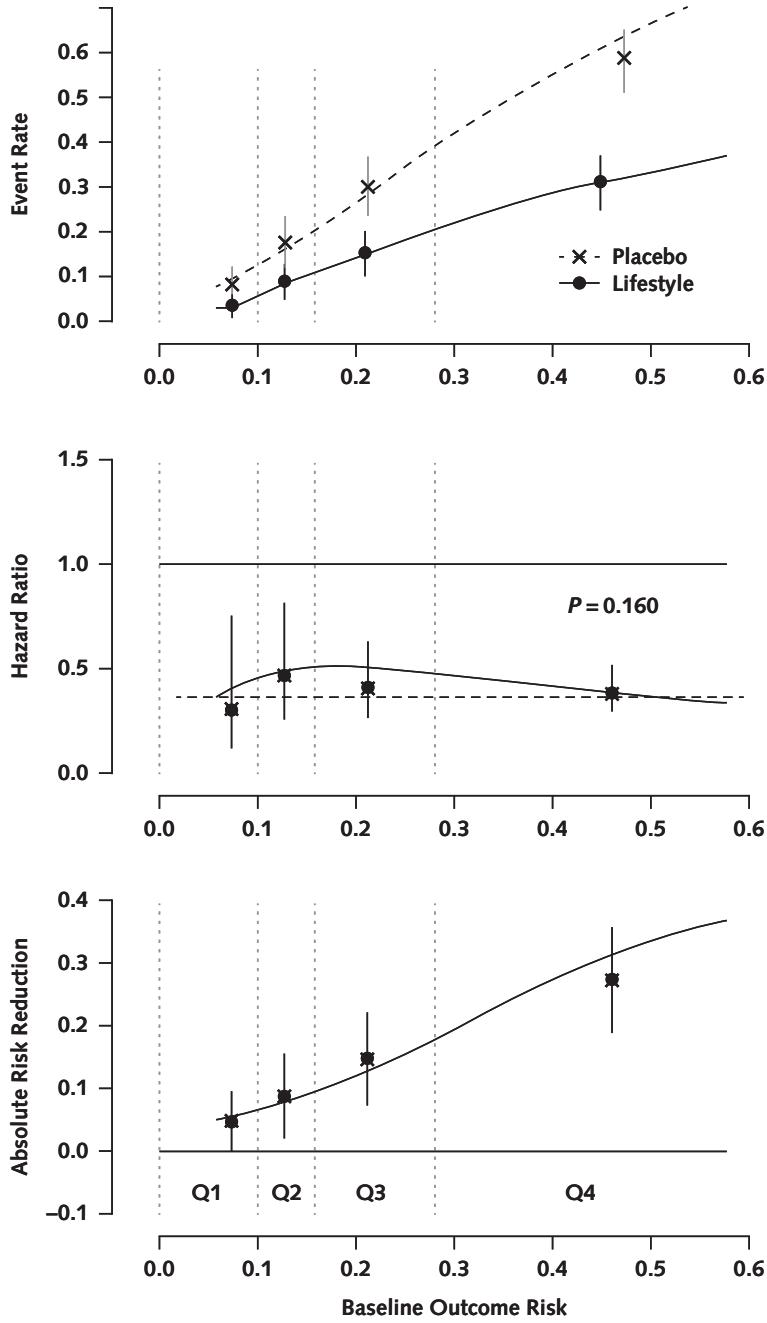


Figure 5: Exemple de résultats d'hétérogénéité de l'effet du traitement par niveaux de risque prédit $\tilde{R}(X)$, d'après Kent et al. Cette présentation des résultats est suggérée dans les recommandations PATH [23]. Les lignes verticales pointillées indiquent les quantiles de risque prédit. La ligne horizontale pointillée indique l'effet traitement moyen observé dans l'essai. Dans le cadre de la personnalisation du recours à l'épuration extra-rénale en réanimation, nous reprenons cette présentation (voir Chapitre 4).

2.3.2 Modélisation directe de l'effet du traitement

Une autre approche pour la modélisation de l'hétérogénéité de l'effet du traitement est l'estimation de la fonction d'effet traitement moyen conditionnel (*conditional average treatment effect ou CATE*) notée

$$\tau(X) \stackrel{\text{def}}{=} \mathbb{E}(Y^{a=1} - Y^{a=0}|X).$$

Comme pour les tâches de prédictions, après avoir été estimée, la fonction d'effet traitement moyen conditionnel doit être évaluée (idéalement sur un jeu de donnée indépendant de celui ayant servi à l'entraînement). Nous distinguons donc deux sous-sections abordant successivement l'estimation et l'évaluation de ces fonctions visant à modéliser directement l'effet du traitement. On note qu'en supposant correctes les hypothèses (1–3) de la section 2.1, les méthodes exposées ci-dessous sont utilisables pour des données observationnelles – comme pour des données d'essai randomisé.

A. Estimation de la fonction d'effet traitement moyen conditionnel

Le développement de méthodes robustes et efficientes pour l'estimation de l'effet traitement moyen conditionnel est un champ de recherche actif. Une liste des principales méthodes disponibles à ce jour est présentée dans le tableau 3.

Tableau 3: Méthodes d'estimation de l'effet traitement moyen conditionnel (CATE).

Catégorie	Méthode	Référence
Méta-apprentissage	R-learner	Nie et Wager (2021) [41]
	DR-learner	Kennedy (2020) [42]
	X-learner	Künzel et al. (2019) [43]
	OW-learner	Zhao et al. (2012) [44]
	S-learner	Hill (2011) [45]
	A-learner	Blatt et al. (2004) [46]
	T-learner	Hansotia et Rukstales (2002) [47]
Méthodes d'apprentissage automatique modifiées	Causal BART	Hahn et al. (2020) [48]
	Causal Forest	Athey et al. (2019) [49]
	Causal Boosting	Powers et al. (2018) [50]

Dans cette section, nous nous focalisons sur R-learner, une méthode récente qui possède de bonnes propriétés théoriques. Le point de départ de R-learner est de ré-exprimer la fonction $\tau(\cdot)$ comme la régression des résidus de la fonction pronostique $q(X) \stackrel{\text{def}}{=} \mathbb{E}[Y|X]$ sur les résidus du score de propension $e(\cdot)$:

$$\tau(\cdot) = \arg \min_{\tau} \left\{ \mathbb{E} \left(\left[\{Y - q(X)\} - \{A - e(X)\} \tau(X) \right]^2 \right) \right\}.$$

Cette expression dite de résidualisation formulée initialement par Robinson [51] donne son nom à R-learner. En s'appuyant sur cette formulation, Nie et Wager [41] proposent d'estimer $\tau(\cdot)$ en deux étapes :

1. Découper les données $(X_i^T, A_i, Y_i)_{i \leq i \leq n}$ en K blocs (typiquement 5 ou 10). Obtenir

K fonctions estimées ($\hat{q}^{(1)}, \dots, \hat{q}^{(K)}$) pour $q(\cdot)$ et K fonctions estimées ($\hat{e}^{(1)}, \dots, \hat{e}^{(K)}$) pour $e(\cdot)$ par apprentissage supervisé sur les K regroupements de $(K - 1)$ blocs (pour chaque modèle, le bloc restant peut être utilisé pour le réglage d'hyperparamètres). Cette étape est parfois appelée étape de “cross-fitting” [52].

2. Estimer $\tau(\cdot)$ en optimisant l'objectif de perte pénalisée suivant

$$\begin{aligned}\hat{\tau}(\cdot) &= \arg \min_{\tau} \left[\hat{L}_n\{\tau(\cdot)\} + \Lambda_n\{\tau(\cdot)\} \right], \\ \hat{L}_n\{\tau(\cdot)\} &= n^{-1} \sum_{i=1}^n \left[\{Y_i - \hat{q}^{(-k(i))}(X_i)\} - \{A_i - \hat{e}^{(-k(i))}(X_i)\} \tau(X_i) \right]^2\end{aligned}$$

où $\hat{q}^{(-k(i))}(X_i), \hat{e}^{(-k(i))}(X_i)$ indiquent la prédiction d'un modèle entraîné sur le regroupement des $(K - 1)$ blocs ne comprenant pas l'observation i , $\Lambda_n\{\tau(\cdot)\}$ indique un terme de régularisation de la complexité de $\tau(\cdot)$. En pratique, l'étape 2 peut être réalisée par apprentissage supervisé avec les entrées $(X_i^T)_{i \leq i \leq n}$ et les sorties

$$\left(\frac{Y_i - \hat{q}^{(-k(i))}(X_i)}{A_i - \hat{e}^{(-k(i))}(X_i)} \right)_{i \leq i \leq n}$$

où est appliquée une régularisation explicite (par exemple ridge ou lasso), ou implicite (par exemple réseau de neurone profond avec descente de gradient stochastique arrêtée précocement).

Nie et Wager montrent que si les estimations de $q(\cdot)$ et $e(\cdot)$ convergent chacune à une vitesse de $O_p(n^{-1/4})$ pour l'erreur quadratique moyenne, alors la vitesse de convergence de $\hat{\tau}(\cdot)$ est beaucoup plus rapide que $O_p(n^{-1/4})$ et dépendante de la complexité de la fonction inconnue $\tau(\cdot)$.

On note que malgré ses propriétés asymptotiques intéressantes, la fonction $\hat{\tau}(\cdot)$ estimée avec R-learner peut avoir une erreur quadratique moyenne non négligeable avec l'utilisation d'un échantillon – typiquement de petite taille en recherche clinique. Cette limitation affecte également toutes les autres méthodes visant à modéliser directement l'effet du traitement [53]. Pour cette raison, il apparaît impératif d'évaluer et valider la performance des fonctions estimées d'effet traitement moyen conditionnel.

B. Évaluation d'une fonction estimée de l'effet traitement moyen conditionnel

Pour évaluer la discrimination de la fonction estimée $\hat{\tau}(\cdot)$ on peut calculer la *C-statistic for benefit* [54] dans un jeu de donnée indépendant de celui ayant servi à l'entraînement. Cette métrique est calculable pour des résultats Y binaires ou continus. La méthodologie suppose la création de paires de patients avec ($A = 1$) et sans intervention ($A = 0$) – de caractéristiques X similaires (où X est supposé contenir toutes les variables pronostiques et toutes les variables influençant la réponse à l'intervention).¹⁰ Toutes les paires de paires sont ensuite évaluées et classées en i) paire de paire concordante, ii) paire de paire disconcordante, iii) paire de paire avec égalité du bénéfice prédit, iv) paire de paire avec égalité

¹⁰Avec des données d'essai randomisé, un appariement sur la fonction de l'effet traitement moyen conditionnel est généralement utilisé (s'il persiste des différences de caractéristiques entre individus appariés celles-ci sont dues au hasard et l'estimateur résultant pour *C-statistic for benefit* est sans biais).

du bénéfice observé. La procédure de calcul de *C-statistic for benefit* est détaillée dans l'algorithme 1. L'objectif est d'estimer

$$C_{\text{for-benefit}} \stackrel{\text{def}}{=} \mathbb{P}(\hat{\tau}(X_V) > \hat{\tau}(X_W) \mid \tau(X_V) > \tau(X_W))$$

où

$$\tau(X_V) \stackrel{\text{def}}{=} Y_{V:A=1} - Y_{V:A=0}$$

représente le bénéfice observé pour la paire V ,

$$\hat{\tau}(X_V) \stackrel{\text{def}}{=} (\hat{\tau}(X_{V:A=1}) + \hat{\tau}(X_{V:A=0}))/2$$

représente le bénéfice prédit pour la paire V ; l'indice $(-)_{V:A=1}$ indiquant la variable aléatoire du patient de la paire V recevant l'intervention $A = 1$, etc. Une valeur de $C_{\text{for-benefit}}$ inférieure à 0.5 indique que les recommandations de traitement faites par la fonction estimée $\hat{\tau}(\cdot)$ sont incorrectes pour plus de la moitié des patients. La discrimination augmente quand $C_{\text{for-benefit}}$ tend vers 1. On remarque cependant que même si on utilisait un oracle pour $\tau(\cdot)$ et un appariement parfait des patients, la $C_{\text{for-benefit}}$ ne pourrait, en principe, jamais atteindre 1. Cela est dû au fait qu'il existe une incertitude fondamentale sur le bénéfice réel c'est-à-dire

$$Y^{a=1} - Y^{a=0} | X \sim f_{Y^{a=1} - Y^{a=0} | X}$$

où $f_{Y^{a=1} - Y^{a=0} | X}$ décrit une loi de probabilité d'espérance $\tau(X)$ (oracle) mais de variance non nulle. Dans ces conditions, il est difficile de distinguer si une $C_{\text{for-benefit}}$ donnée est faible du fait de l'erreur d'estimation sur $\tau(\cdot)$ ou du fait d'une variance importante pour la loi de probabilité $f_{Y^{a=1} - Y^{a=0} | X}$ (incertitude sur le bénéfice réel). Pour autant, il ne faut pas conclure que $C_{\text{for-benefit}}$ est inutile. Comme pour un modèle de prédiction conventionnel non causal, il existe pour $\hat{\tau}(\cdot)$ un compromis entre la discrimination et la calibration [55]. En parallèle de l'estimation de $C_{\text{for-benefit}}$, il est donc important de rapporter la calibration de $\hat{\tau}(\cdot)$. Un exemple de diagramme de calibration pour une fonction estimée de l'effet traitement moyen conditionnel est présenté dans la Figure 6.

Nous terminons cette section en remarquant que le développement de méthodes pour l'évaluation de règles individuelles de traitement est aussi un champ de recherche actif. Xia et al. [56] critiquent $C_{\text{for-benefit}}$ en argumentant que ce métrique n'est défini qu'algorithmiquement, et son estimation est dépendante de la technique d'appariement. Plusieurs autres métriques ont été proposés : Imai et Li ont développé *Population Average Prescriptive Effect* (PAPE), un métrique évaluant la valeur d'une règle individuelle de traitement en tenant compte du nombre de patients traités [57]. Par ailleurs, Sadatsafavi et al. proposaient *Concentration of Benefit index* [58] alors que, plus tôt, Janes et al. introduisaient un cadre pour évaluer le caractère prédictif de la réponse au traitement de biomarqueurs [59].

Algorithme 1 Procédure d'estimation non paramétrique de $C_{\text{for-benefit}}$.**Entrée:** Données de validation $(X_i, A_i, Y_i)_{1 \leq i \leq n}$ et la fonction $\hat{\tau}(\cdot)$ à évaluer.

1. Calculer les effets traitements moyens conditionnels prédits $\hat{\tau}_i \leftarrow \hat{\tau}(X_i)$.
2. Apparier 1 : 1 les patients avec ($A = 1$) et sans ($A = 0$) intervention sur leurs caractéristiques X pour obtenir $m \leq n/2$ paires de patients. Différentes techniques sont possibles. Avec des données d'essai randomisé par exemple : éliminer au hasard certains patients du groupe $\{A = 1\}$ ou $\{A = 0\}$ le plus grand pour réaliser ensuite un *matching* 1:1 par rang de $\hat{\tau}_i$.
3. Pour chaque paire $p = 1, \dots, m$ calculer :
 - Le bénéfice observé pour la paire $\tau_p \leftarrow Y_{i:A_i=1} - Y_{i:A_i=0}$,
 - Le bénéfice prédit pour la paire $\hat{\tau}_p \leftarrow ((\hat{\tau}_i)_{i:A_i=1} + (\hat{\tau}_i)_{i:A_i=0})/2$.
4. Évaluer chacune des $\binom{m}{2}$ paires de paires avec les règles suivantes :
 - Paire de paire avec égalité du bénéfice observé

$$\tau_v = \tau_w$$

- Paire de paire avec égalité du bénéfice prédit

$$\hat{\tau}_v = \hat{\tau}_w \quad \text{et} \quad \tau_v \neq \tau_w$$

- Paire de paire concordante

$$(\hat{\tau}_v > \hat{\tau}_w \quad \text{et} \quad \tau_v > \tau_w) \quad \text{ou} \quad (\hat{\tau}_v < \hat{\tau}_w \quad \text{et} \quad \tau_v < \tau_w)$$

- Paire de paire discordante sinon

où pour chaque paire de paire, v et w indiquent distinctement une des paires.

Sortie: $\hat{C}_{\text{for-benefit}} = \frac{N_c + 0.5N_{ebp}}{N_c + N_d + N_{ebp}}$, où N_c , N_d , N_{ebp} indiquent respectivement le nombre de paires de paires concordantes, discordantes et avec égalité du bénéfice prédit.

Le code R et Python implémentant l'estimation de $C_{\text{for-benefit}}$ est accessible sur <https://github.com/fcgrolleau/building-blocks>.

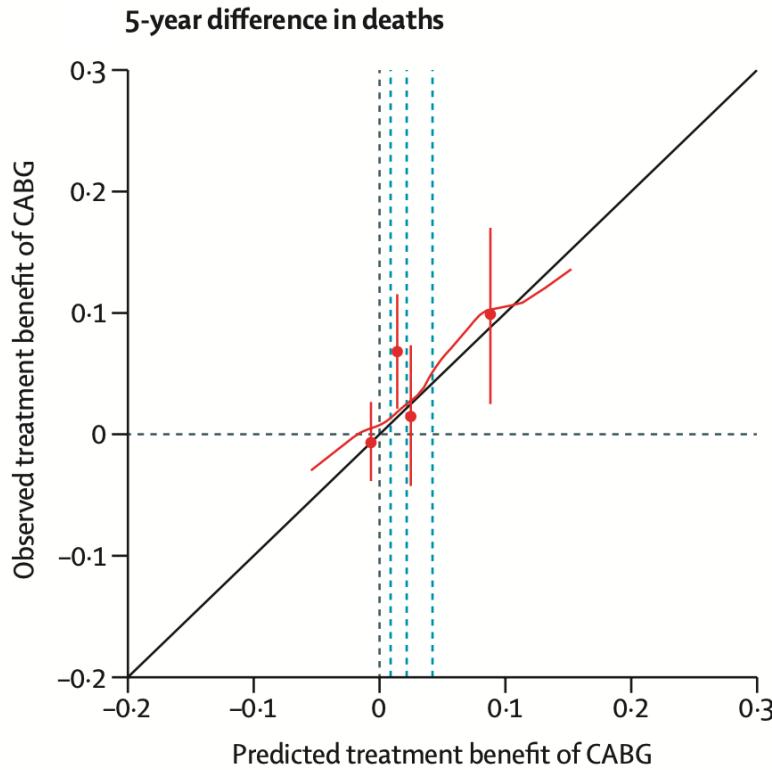


Figure 6: Diagramme de calibration pour une fonction estimée de l'effet traitement moyen conditionnel évaluant chirurgie cardiaque *vs* cardiologie interventionnelle pour la survie à 5 ans de patients avec une coronaropathie complexe (validation externe d'après Takahashi et al. [60]). Les effets traitements moyens conditionnels prédits sont regroupés par quarts (lignes pointillées verticales). La moyenne de chaque quart est présentée sur l'axe des abscisses. L'effet traitement moyen est calculé pour chaque quart et présentée en ordonnée. Ici, les données de validation proviennent d'un essai randomisé 1:1 et aucun appariement n'est nécessaire ; un lissage des points $(\hat{\tau}(X_i), 2Y_i - 1)_{1 \leq i \leq n}$ peut être réalisé par exemple par splines de lissage avec un paramètre de pénalisation déterminé par LOOCV [61]. Avec des données observationnelles, l'appariement de patients traités par chirurgie cardiaque et traités par cardiologie interventionnelle aurait été nécessaire pour évaluer la calibration. CABG: *Coronary Artery Bypass Graft*.

2.4 Mélange d'experts et algorithme d'espérance-maximisation

Les mélanges d'experts sont des modèles statistiques utilisés dans le cadre de problèmes d'apprentissage supervisé. Ces modèles s'appuient sur deux idées importantes : i) la division souple de l'espace des entrées (approche dite *divide and conquer*) et ii) le mélange des modèles de régressions obtenus sur chaque sous-espace des entrées. Afin de clarifier l'exposition dans cette section, nous considérons un problème d'apprentissage supervisé en utilisant un mélange de deux experts. La généralisation à des mélanges d'experts à plus de deux composantes est sans difficulté ; celle-ci est exposée dans [62, 63].

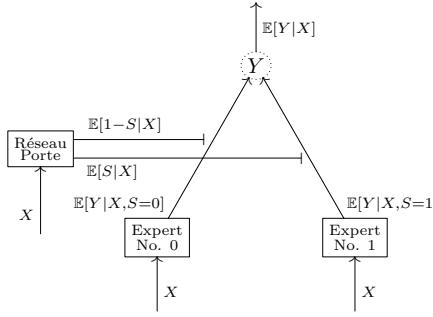


Figure 7: Mélange d'expert à deux composantes. Représentation graphique de l'équation 2.4 détaillant deux réseaux experts et un réseau porte.

2.4.1 Modèle probabiliste

Le modèle de mélange de deux experts suppose le mécanisme génératif des données suivant. On considère que la variable X suit une distribution ayant une densité inconnue $X \sim p_X$ et on suppose l'existence d'une variable aléatoire latente binaire S . On suppose que les densités conditionnelles $p_{Y|X,S}(y|x, 0; \alpha_0)$ et $p_{Y|X,S}(y|x, 1; \alpha_1)$ correspondent à une loi de la famille exponentielle avec pour paramètres respectifs α_0 et α_1 (ceux-ci incluent un paramètre de poids et possiblement un paramètre de dispersion). La variable S étant binaire, la densité conditionnelle $p_{Y|X,S}$ peut être notée

$$p_{Y|X,S}(y|x, s; \alpha) = p_{Y|X,S}(y|x, 1; \alpha_1)^s p_{Y|X,S}(y|x, 0; \alpha_0)^{1-s}$$

où on note $\alpha = (\alpha_0^T, \alpha_1^T)^T$ pour simplifier les notations. Le modèle de mélange paramétrique de deux experts auquel on s'intéresse suppose que $p_{Y,S|X}$ a pour densité

$$\begin{aligned} p_{Y,S|X}(y, s|x; \theta) &= \left[p_{Y|X,S}(y|x, 1; \alpha_1) \text{expit}(\beta^T x) \right]^s \left[p_{Y|X,S}(y|x, 0; \alpha_0) \{1 - \text{expit}(\beta^T x)\} \right]^{1-s} \\ &= p_{Y|X,S}(y|x, s; \alpha) \text{expit}(\beta^T x)^s \{1 - \text{expit}(\beta^T x)\}^{1-s} \end{aligned}$$

où on note $\theta = (\alpha^T, \beta^T)^T$. Dans l'annexe B3 on montre que cette représentation pour $p_{Y,S|X}$ induit bien une densité de probabilité.

2.4.2 Reformulation et représentation graphique

Le modèle probabiliste présenté ci-dessus impose une contrainte sur la forme des lois conditionnelles $p_{S|X}$ et $p_{Y|X}$. En particulier, en marginalisant contre Y , on remarque que

$$\begin{aligned} p_{S|X}(s|x; \theta) &= \int_Y p_{Y,S|X}(y, s|x; \theta) dy \\ &= \text{expit}(\beta^T x)^s \{1 - \text{expit}(\beta^T x)\}^{1-s} \int_Y p_{Y|X,S}(y|x, s; \alpha) dy \\ &= \text{expit}(\beta^T x)^s \{1 - \text{expit}(\beta^T x)\}^{1-s}. \end{aligned}$$

L'égalité ci-dessus permet d'exprimer l'espérance conditionnelle de S sachant X comme

$$\begin{aligned} \mathbb{E}_\theta[S|X = x] &= \sum_{s=0}^1 s p_{S|X}(s|x; \theta) = \sum_{s=0}^1 s \text{expit}(\beta^T x)^s \{1 - \text{expit}(\beta^T x)\}^{1-s} \\ &= \text{expit}(\beta^T x). \end{aligned}$$

Par ailleurs, en marginalisant maintenant la densité $p_{Y,S|X}$ contre S on obtient

$$\begin{aligned} p_{Y|X}(y|x; \theta) &= \sum_{s=0}^1 p_{Y,S|X}(y, s|x; \theta) \\ &= \text{expit}(\beta^T x) p_{Y|X,S}(y|x, 1; \alpha_1) + \{1 - \text{expit}(\beta^T x)\} p_{Y|X,S}(y|x, 0; \alpha_0). \end{aligned}$$

L'égalité ci-dessus permet d'exprimer l'espérance conditionnelle de Y sachant X comme

$$\begin{aligned} \mathbb{E}_\theta[Y|X = x] &= \int_Y y p_{Y|X}(y|x; \theta) dy \\ &= \text{expit}(\beta^T x) \int_Y y p_{Y|X,S}(y|x, 1; \alpha_1) dy + \{1 - \text{expit}(\beta^T x)\} \int_Y y p_{Y|X,S}(y|x, 0; \alpha_0) dy \\ &= \mathbb{E}_\theta[S|X = x] \mathbb{E}_\theta[Y|X = x, S = 1] + \mathbb{E}_\theta[1 - S|X = x] \mathbb{E}_\theta[Y|X = x, S = 0]. \quad (2.4) \end{aligned}$$

En interprétant la variable latente S comme indiquant “l'expert” associée à chaque observation de telle sorte que $S = 1$ indique une observation associée à l'expert No. 1 et $S = 0$ indique une observation associée à l'expert No. 0, une représentation graphique de l'équation (2.4) est fréquemment formulée ; celle-ci est donnée dans la Figure 7.

2.4.3 Vraisemblances sur données complètes et incomplètes

On suppose l'existence de données incomplètes (observées) $\mathcal{D} = (X_i, Y_i)$, $i = 1, \dots, n$ et de données complètes (non observées) $\mathcal{C} = (X_i, S_i, Y_i)$, $i = 1, \dots, n$ toutes deux i.i.d. selon le modèle probabiliste décrit ci-dessus. La log-vraisemblance sur données incomplètes est

$$\begin{aligned} l(\theta; \mathcal{D}) &= \ln \prod_{i=1}^n p_{Y|X}(Y_i|X_i; \theta) \\ &= \sum_{i=1}^n \ln \left\{ \text{expit}(\beta^T X_i) p_{Y|X,S}(Y_i|X_i, 1; \alpha_1) + \{1 - \text{expit}(\beta^T X_i)\} p_{Y|X,S}(Y_i|X_i, 0; \alpha_0) \right\}. \end{aligned}$$

La log-vraisemblance sur données complètes est

$$\begin{aligned}
l_c(\theta; \mathcal{C}) &= \ln \prod_{i=1}^n p_{Y,S|X}(Y_i, S_i | X_i; \theta) \\
&= \sum_{i=1}^n S_i \ln p_{Y|X,S}(Y_i | X_i, 1; \alpha_1) \\
&\quad + \sum_{i=1}^n (1 - S_i) \ln p_{Y|X,S}(Y_i | X_i, 0; \alpha_0) \\
&\quad + \sum_{i=1}^n S_i \ln \{ \text{expit}(\beta^T X_i) \} + (1 - S_i) \ln \{ 1 - \text{expit}(\beta^T X_i) \}.
\end{aligned}$$

La dernière égalité indique que la maximisation de $l_c(\theta; \mathcal{C})$ en θ peut être envisagée comme trois problèmes de maximisation distincts : maximisation de la première somme en α_1 , maximisation de la seconde somme en α_0 et maximisation de la troisième somme en β . Par comparaison, la maximisation de $l(\theta; \mathcal{D})$ semble plus complexe ; l'expression de $l(\theta; \mathcal{D})$ contenant le logarithme d'une somme, l'équation ne se simplifie pas.

2.4.4 Algorithme d'espérance-maximisation (EM)

Plutôt que de maximiser directement $l(\theta; \mathcal{D})$ en θ , l'algorithme EM propose de maximiser $l_c(\theta; \mathcal{C})$ en θ . Une difficulté est liée au fait que S_i n'étant pas observée, $l_c(\theta; \mathcal{C})$ est une variable aléatoire. L'algorithme d'EM procède en deux étapes.

- **Étape d'espérance (*E step*)**. L'espérance de $l_c(\theta; \mathcal{C})$ est calculée conditionnellement aux données observées et aux valeurs estimées pour θ à l'itération (p) afin d'obtenir la fonction Q déterministe (n'impliquant pas la variable aléatoire S_i) :

$$\begin{aligned}
Q(\theta, \theta^{(p)}) &\stackrel{\text{def}}{=} \mathbb{E}_{\theta^{(p)}}[l_c(\theta; \mathcal{C}) | \mathcal{D}] \\
&= \sum_{i=1}^n h_i \ln p_{Y|X,S}(Y_i | X_i, 1; \alpha_1) \\
&\quad + \sum_{i=1}^n (1 - h_i) \ln p_{Y|X,S}(Y_i | X_i, 0; \alpha_0) \\
&\quad + \sum_{i=1}^n h_i \ln \{ \text{expit}(\beta^T X_i) \} + (1 - h_i) \ln \{ 1 - \text{expit}(\beta^T X_i) \} \quad (2.5)
\end{aligned}$$

où on utilise le raccourci

$$h_i = \frac{\text{expit}(\beta^{(p)T} X_i) p_{Y|X,S}(Y_i | X_i, 1; \alpha_1^{(p)})}{\text{expit}(\beta^{(p)T} X_i) p_{Y|X,S}(Y_i | X_i, 1; \alpha_1^{(p)}) + \{1 - \text{expit}(\beta^{(p)T} X_i)\} p_{Y|X,S}(Y_i | X_i, 0; \alpha_0^{(p)})}.$$

Preuve.

$$Q(\theta, \theta^{(p)}) \stackrel{\text{def}}{=} \mathbb{E}_{\theta^{(p)}}[l_c(\theta; \mathcal{C}) | \mathcal{D}]$$

$$\begin{aligned}
&= \sum_{i=1}^n \mathbb{E}_{\theta^{(p)}}[S_i|\mathcal{D}] \ln p_{Y|X,S}(Y_i|X_i, 1; \alpha_1) \\
&\quad + \sum_{i=1}^n \mathbb{E}_{\theta^{(p)}}[1 - S_i|\mathcal{D}] \ln p_{Y|X,S}(Y_i|X_i, 0; \alpha_0) \\
&\quad + \sum_{i=1}^n \mathbb{E}_{\theta^{(p)}}[S_i|\mathcal{D}] \ln \{\text{expit}(\beta^T X_i)\} + \mathbb{E}[1 - S_i|\mathcal{D}] \ln \{1 - \text{expit}(\beta^T X_i)\}
\end{aligned}$$

Par ailleurs, en utilisant l'indépendance des observations et en appliquant le théorème de Bayes on remarque :

$$\begin{aligned}
\mathbb{E}_{\theta^{(p)}}[S_i|\mathcal{D}] &= \mathbb{P}(S_i = 1|Y_1, X_1, \dots, Y_i, X_i, \dots, Y_n, X_n; \theta^{(p)}) \\
&= \mathbb{P}(S_i = 1|Y_i, X_i; \theta^{(p)}) \\
&= \frac{\mathbb{P}(S_i = 1|X_i; \theta^{(p)}) \mathbb{P}(Y_i|S_i = 1, X_i; \theta^{(p)})}{\mathbb{P}(Y_i|X_i; \theta^{(p)})} \\
&= \frac{\mathbb{P}(S_i = 1|X_i; \theta^{(p)}) \mathbb{P}(Y_i|S_i = 1, X_i; \theta^{(p)})}{\mathbb{P}(S_i = 1|X_i; \theta^{(p)}) \mathbb{P}(Y_i|S_i = 1, X_i; \theta^{(p)}) + \mathbb{P}(S_i = 0|X_i; \theta^{(p)}) \mathbb{P}(Y_i|S_i = 0, X_i; \theta^{(p)})} \\
&= \frac{\text{expit}(\beta^{(p)T} X_i) p_{Y|X,S}(Y_i|X_i, 1; \alpha_1^{(p)})}{\text{expit}(\beta^{(p)T} X_i) p_{Y|X,S}(Y_i|X_i, 1; \alpha_1^{(p)}) + \{1 - \text{expit}(\beta^{(p)T} X_i)\} p_{Y|X,S}(Y_i|X_i, 0; \alpha_0^{(p)})} \\
&= h_i.
\end{aligned}$$

La substitution de ce résultat dans le résultat précédent prouve la formule pour $Q(\theta, \theta^{(p)})$ énoncée dans l'équation (2.5). \square

En pratique, à l'itération $(p+1)$ l'étape d'espérance consiste à estimer $(h_i)_{1 \leq i \leq n}$ en utilisant la formule donnée dans le raccourci ci-dessus.

- **Étape de maximisation (*M step*)**. Cette étape vise à maximiser Q en θ pour obtenir la nouvelle estimation des paramètres $\theta^{(p+1)}$:

$$\theta^{(p+1)} = \arg \max_{\theta} Q(\theta, \theta^{(p)}).$$

En observant l'équation (2.5), on constate que les paramètres α_1 , α_0 et β n'influencent la fonction Q qu'à travers la première, seconde et troisième somme respectivement. Par conséquent, l'étape de maximisation se réduit aux problèmes de maximisation distincts suivants :

$$\begin{aligned}
\alpha_1^{(p+1)} &= \arg \max_{\alpha_1} \sum_{i=1}^n h_i \ln p_{Y|X,S}(Y_i|X_i, 1; \alpha_1) \\
\alpha_0^{(p+1)} &= \arg \max_{\alpha_0} \sum_{i=1}^n (1 - h_i) \ln p_{Y|X,S}(Y_i|X_i, 0; \alpha_0) \\
\beta^{(p+1)} &= \arg \max_{\beta} \sum_{i=1}^n h_i \ln \{\text{expit}(\beta^T X_i)\} + (1 - h_i) \ln \{1 - \text{expit}(\beta^T X_i)\}.
\end{aligned}$$

On remarque qu'à chacun de ces problèmes de maximisation correspond un problème de maximisation de log-vraisemblance. Étant donné notre paramétrisation utilisant des lois exponentielles, on reconnaît ces problèmes de maximisation comme des problèmes de maximisation de log-vraisemblance pondérées de modèles linéaires généralisés. Ces problèmes peuvent être facilement résolus par différents algorithmes d'optimisation convexe (montée de gradient, IRLS, etc.).

En pratique, à l'itération $(p+1)$, l'étape de maximisation consiste à i) estimer $\alpha_1^{(p+1)}$ en régressant (avec un modèle linéaire généralisé adapté au modèle génératif) Y_i sur X_i avec les poids h_i , ii) estimer $\alpha_0^{(p+1)}$ en régressant (avec un modèle linéaire généralisé adapté au modèle génératif) Y_i sur X_i avec les poids $1 - h_i$, iii) estimer $\beta^{(p+1)}$ en régressant h_i sur X_i avec un modèle logistique.

Dempser et al. [64] ont prouvé, en s'appuyant sur l'inégalité de Jensen, que l'augmentation de Q conduite à l'étape de maximisation implique une augmentation de la vraisemblance sur données incomplètes :

$$l(\theta^{(p+1)}; \mathcal{D}) \geq l(\theta^{(p)}; \mathcal{D}).$$

Par conséquent la vraisemblance sur données incomplètes croît à chaque itération de l'algorithme d'EM. En pratique cela implique une convergence vers un maximum local de la vraisemblance sur données incomplètes.

2.4.5 Identifiabilité

Le modèle de mélange de deux experts est dit identifiable si l'application $\Theta \mapsto p_{Y|X}(\cdot|.; \theta)$ est injective c'est à dire précisément si

$$\forall (\theta, \theta')^T \in \Theta^2, \quad \left(p_{Y|X}(\cdot|.; \theta) \equiv p_{Y|X}(\cdot|.; \theta') \quad \Rightarrow \quad \theta = \theta' \right).$$

Dans le cas général du fait d'une permutation possible des experts, les mélanges d'experts ne sont pas identifiables. Ci-dessous, nous illustrons ce point par un contre exemple pour un mélange de deux experts : en choisissant $\alpha'_0 = \alpha_1$, $\alpha'_1 = \alpha_0$, et $\beta' = -\beta$ on a $\theta' \neq \theta$ et

$$\begin{aligned} & p_{Y|X}(y|x; \theta') \\ &= \text{expit}(-\beta^T x) p_{Y|X,S}(y|x, 1; \alpha_0) + \{1 - \text{expit}(-\beta^T x)\} p_{Y|X,S}(y|x, 0; \alpha_1) \\ &= \{1 - \text{expit}(\beta^T x)\} p_{Y|X,S}(y|x, 1; \alpha_0) + \text{expit}(\beta^T x) p_{Y|X,S}(y|x, 0; \alpha_1) \\ &= \text{expit}(\beta^T x) p_{Y|X,S}(y|x, 1; \alpha_1) + \{1 - \text{expit}(\beta^T x)\} p_{Y|X,S}(y|x, 0; \alpha_0) \\ &= p_{Y|X}(y|x; \theta). \end{aligned}$$

Jiang et Tanner ont cependant montré qu'en ordonnant les experts, les mélanges d'experts de lois exponentielles normale, binomiale, poisson et gamma, sont identifiables sous certaines conditions [65].

2.4.6 Propriétés asymptotiques

En supposant que, sous une contrainte d'ordre à préciser, l'estimateur $\hat{\theta}$ est l'unique solution d'une équation du maximum de vraisemblance sur données incomplètes, les propriétés asymptotiques de $\hat{\theta}$ peuvent être déduites en appliquant les résultats de la M-

estimation (section 2.2). Par exemple, dans le cas d'un mélange de deux experts binomiaux $p_{Y|X,S}(Y_i|X_i, 0; \alpha_0)$ et $p_{Y|X,S}(Y_i|X_i, 1; \alpha_1)$, on peut supposer que sous une contrainte d'ordre $\alpha_0 \prec \alpha_1$ (à expliciter), l'estimateur $\hat{\theta} = (\hat{\alpha}_0^T, \hat{\alpha}_1^T, \hat{\beta}^T)^T$ est l'unique solution de l'équation

$$\sum_{i=1}^n \frac{\partial}{\partial \theta} \ln \left\{ \expit(\beta^T X_i) \expit(\alpha_1^T X_i)^{Y_i} \{1 - \expit(\alpha_1^T X_i)\}^{1-Y_i} + \{1 - \expit(\beta^T X_i)\} \expit(\alpha_0^T X_i)^{Y_i} \{1 - \expit(\alpha_0^T X_i)\}^{1-Y_i} \right\} \Big|_{\theta=\hat{\theta}} = \mathbf{0}.$$

On remarque que cette équation est de la forme

$$\sum_{i=1}^n \psi(X_i, Y_i; \hat{\theta}) = \mathbf{0}$$

où étant donné notre paramétrisation utilisant des lois exponentielles, la fonction $\psi(x, y; \delta)$ est dérivable selon δ , (x, y) —(presque partout). Reconnaissant ainsi, sous les hypothèses mentionées ci-dessus, $\hat{\theta}$ comme un M-estimateur de type ψ , le théorème énoncé dans la section 2.2.1 permet de conclure que $\hat{\theta}$ est \sqrt{n} —convergent, asymptotiquement normal et de variance asymptotique calculable par l'estimateur sandwich empirique.

2.5 Apprentissage par renforcement de stratégies dynamiques optimales

Dans cette section, nous considérons les prises de décisions séquentielles en situation d'incertitude. En ce qui concerne les applications médicales, nous nous intéressons aux situations où, observant les caractéristiques changeantes d'un patient donné, on ne connaît pas, aux différentes étapes d'évolution de sa maladie, les meilleures décisions de traitement à prendre. Dans ce contexte, notre objectif est d'utiliser des données observationnelles pour estimer une stratégie de traitement dynamique optimale pour la maximisation d'un critère de jugement prédéfini. Les méthodes existantes pour résoudre ce problème s'inscrivent dans le champ de l'apprentissage par renforcement hors ligne (*offline reinforcement learning*) [36, 66]. Après introduction du cadre conceptuel, nous décrivons brièvement certaines méthodes pour estimer une stratégie de traitement (appelée “politique” par la suite) dynamique optimale dans ce contexte.

2.5.1 Formalisation du problème et notations

Nous considérons l'observation d'un ensemble de $i = 1, \dots, n$ trajectoires indépendantes et identiquement distribués selon une distribution \mathbb{P} décrivant l'évolution d'individus pendant K étapes. Pour un individu i , à l'étape $k \in \{1, \dots, K\}$ on observe les variables $S_k^{(i)} \in \mathcal{S}_k$, et $A_k^{(i)} \in \mathcal{A}_k \supset \{0\}$. Ces variables décrivent respectivement l'état (*State*) de l'individu i à l'étape k et l'action (ou action référence notée 0) prise par l'individu i à l'étape k .¹¹ A l'étape $K+1$, on observe le résultat final $Y^{(i)}$ pour l'individu i . Pour simplifier les notations

¹¹Dans le contexte médical l'action est souvent prise non pas par l'individu i mais par son médecin (décision de traitement par exemple).

par la suite, on note $S_{k1:k2}^{(i)} \stackrel{\text{def}}{=} (S_{k1}^{(i)T}, \dots, S_{k2}^{(i)T})^T$, $A_{k1:k2}^{(i)} \stackrel{\text{def}}{=} (A_{k1}^{(i)T}, \dots, A_{k2}^{(i)T})^T$ et on définit le vecteur histoire comme $H_k^{(i)} \stackrel{\text{def}}{=} (S_{1:k}^{(i)T}, A_{1:(k-1)}^{(i)T})^T$. On définit la généralisation appropriée du score de propension par

$$e_{k,a}(H_k) \stackrel{\text{def}}{=} \mathbb{P}(A_k = a | H_k).$$

Nous supposons pour tout $a_k \in \mathcal{A}_k$, $k = 1, \dots, K$, l'existence de résultats potentiels $Y^{(i)}(a_{1:K})$ et $S_k^{(i)}(a_{1:(k-1)})$. Le résultat potentiel $Y^{(i)}(a_{1:K})$ s'interprète comme le résultat (*reward*) final qui aurait été observé pour l'individu i si celui-ci avait suivi la séquence d'actions $a_{1:K}$. Sans perte de généralité, dans cette section, on suppose qu'il est désirable d'obtenir un résultat final élevé. Le résultat potentiel $S_k^{(i)}(a_{1:(k-1)})$ s'interprète comme l'état à l'étape k qui aurait été observé pour l'individu i si celui-ci avait suivi la séquence d'actions $a_{1:(k-1)}$.

Pour l'identification d'effets causaux dans le contexte de prise de décisions séquentielles, nous nous appuyons sur les généralisations appropriées des hypothèses de cohérence, ignorabilité (absence de facteurs de confusion non mesurés), et positivité.

Hypothèse 4 (Cohérence séquentielle). *Les observations sont cohérentes avec les résultats potentiels dans le sens où*

$$Y^{(i)} = Y^{(i)}(A_{1:K}^{(i)}), \quad S_k^{(i)} = S_t^{(i)}(A_{1:(k-1)}^{(i)}), \quad i = 1, \dots, n.$$

Hypothèse 5 (Ignorabilité séquentielle). *Les actions ne sont pas prises en réponse aux informations futures, mais seulement en réponse aux informations mesurées c'est à dire : Pour tout $a_{1:K}$,*

$$\left\{ (S_{k'}(A_{1:(k-1)}, a_{k:(k'-1)})_{k'=k+1}^K, Y(A_{1:(k-1)}, a_{k:K})) \right\} \perp\!\!\!\perp A_k | H_k, \quad \forall k = 1, \dots, K.$$

Hypothèse 6 (Positivité séquentielle). *Il existe des constantes $0 < \eta$ et $\eta_0 < 1$ telles que pour tout $h_k \in \mathcal{H}_k$, $k = 1, \dots, K$, les relations suivantes sont vérifiées :*

$$\forall a \in \mathcal{A}_k \setminus \{0\}, \quad e_{k,a}(h_k) > \eta, \quad \text{et} \quad e_{k,0}(h_k) > 1 - \eta_0.$$

Une politique déterministe π est définie comme une fonction qui, pour chaque étape $k = 1, \dots, K$, assigne une action à une histoire, c'est à dire :

$$\pi_k : \mathcal{H}_k \rightarrow \mathcal{A}_k \quad \text{et} \quad \pi \stackrel{\text{def}}{=} \{\pi_k\}_{k=1}^K.$$

Dans le cas particulier où $\pi_k : \mathcal{S}_k \rightarrow \mathcal{A}_k$ et $\pi_1 = \pi_2 = \dots = \pi_K$, on parle de politique déterministe stationnaire. En reprenant les notations de Murphy [67], adaptées par Chakraborty et Moodie [68] (page 36) nous définissons $f_k(S_k | H_{k-1}, A_{k-1})$ la densité conditionnelle de transition vers l'état S_k , sachant les observations H_{k-1} et la prise d'action A_{k-1} . La fonction de densité des trajectoires observées (h_K, a_K, y) peut ainsi être représentée par

$$f(h_K, a_K, y) = f(s_1) \mathbb{P}(a_1 | s_1) \left[\prod_{k=2}^K f_k(s_k | h_{k-1}, a_{k-1}) \mathbb{P}(a_k | h_k) \right] \mathbb{P}(y | h_K, a_K),$$

et nous notons \mathbb{E} l'espérance sous la distribution associée à cette densité. Par ailleurs, en utilisant l'hypothèse 5, la fonction de densité des trajectoires (h_K, a_K, y) sous une politique

déterministe alternative π est

$$f(h_K, a_K, y; \pi) = f(s_1) \mathbb{1}\{a_1 = \pi_1(s_1)\} \left[\prod_{k=2}^K f_k(s_k | h_{k-1}, a_{k-1}) \mathbb{1}\{a_k = \pi_k(h_k)\} \right] \mathbb{P}(y | h_K, a_K),$$

et nous notons \mathbb{E}_π l'espérance sous la distribution associée à cette densité. La valeur d'une politique déterministe π est définie par

$$V(\pi) \stackrel{\text{def}}{=} \mathbb{E}_\pi[Y] = \mathbb{E}[Y(\pi_1(H_1), \dots, \pi_K(H_K))],$$

elle s'interprète comme le résultat final moyen attendu en réalisant les actions A_1, \dots, A_K toujours tel que le recommande la politique déterministe π . On définit également la fonction de valeur conditionnelle

$$q_{\pi,k}(H_k) \stackrel{\text{def}}{=} \mathbb{E}_\pi[Y | H_k] = \mathbb{E}[Y(A_{1:(k-1)}, \pi_k(H_k), \dots, \pi_K(H_K)) | H_k].$$

Pour une classe Π de politiques déterministes, une politique déterministe optimale π^* (en supposant qu'elle soit identifiable) est définie par

$$\pi^* \stackrel{\text{def}}{=} \arg \max_{\pi \in \Pi} V(\pi).$$

L'objectif est d'estimer π^* pour une classe prédéfinie Π de politiques déterministes. On remarque que l'optimalité n'est pas ici définie en un sens universel mais bien au sein d'une classe de politique prédéfinie Π – qui est probablement de grande dimension VC.

2.5.2 Estimation d'une politique optimale par maximisation de valeur

Une approche classique pour estimer π^* est d'estimer la fonction de valeur $V(\cdot)$ puis de résoudre

$$\hat{\pi}^* = \arg \max_{\pi \in \Pi} \hat{V}(\pi)$$

où Π est une classe prédéfinie de politiques déterministes rendant le problème d'optimisation soluble numériquement.¹² Un approche directe est d'utiliser l'estimateur IPW généralisé au cas de décision séquentielles [69] :

$$\hat{V}_{IPW}(\pi) = n^{-1} \sum_{i=1}^n \gamma_K^{(i)}(\pi) Y^{(i)}$$

où la fonction $\gamma_k^{(i)}(\cdot)$ est définie récursivement par

$$\gamma_k^{(i)}(\pi) = \frac{\gamma_{k-1}^{(i)}(\pi) \mathbb{1}\{A_k^{(i)} = \pi_k(H_k^{(i)})\}}{\sum_{a \in \mathcal{A}_k} \mathbb{1}\{a_k = \pi_k(H_k^{(i)})\} \hat{e}_{k,a}(H_k^{(i)})}, \quad \gamma_0^{(i)}(\pi) = 1.$$

L'estimateur \hat{V}_{IPW} repose sur la méthode d'échantillonage préférentiel (*importance sampling*) ; il est sans biais si pour tout k et pour tout $a \in \mathcal{A}_k$, $\hat{e}_{k,a}(\cdot)$ est sans biais. En

¹²L'algorithme de recherche MILP est par exemple utilisable pour résoudre ce type de problèmes quand la dimension VC de Π n'est pas trop grande.

pratique, il est conseillé d'utiliser une version normalisée de l'estimateur :

$$\hat{V}_{IPW*}(\pi) = \frac{\sum_{i=1}^n \gamma_K^{(i)}(\pi) Y^{(i)}}{\sum_{i=1}^n \gamma_K^{(i)}}.$$

Bien que biaisé, l'estimateur \hat{V}_{IPW*} est convergent et de variance inférieure à celle de \hat{V}_{IPW} [70]. Il existe cependant deux limitations principales à l'utilisation des estimateurs \hat{V}_{IPW*} et \hat{V}_{IPW} : i) ils requièrent pour tout k et pour tout a une spécification correcte des modèles $e_{k,a}(\cdot)$ pour être sans biais (argument de robustesse faible) et ii) ils n'exploitent que les trajectoires observées qui sont entièrement compatibles avec les actions recommandées par la politique π (argument d'efficience faible). Une possible solution à ces limitations est d'utiliser l'estimateur AIPW généralisé au cas de décision séquentielles [71, 72] :

$$\hat{V}_{AIPW}(\pi) = n^{-1} \sum_{i=1}^n \left\{ \gamma_K^{(i)}(\pi) Y^{(i)} - \sum_{k=1}^K \{ \gamma_k^{(i)}(\pi) - \gamma_{k-1}^{(i)}(\pi) \} \hat{q}_{\pi,k}(H_k^{(i)}) \right\}$$

où $\hat{q}_{\pi,k}(\cdot)$ est un estimateur pour la fonction de valeur conditionnelle $q_{\pi,k}(\cdot)$ telle que définie précédemment. L'estimateur \hat{V}_{AIPW} présente un avantage théorique de double robustesse. En pratique, il a essentiellement l'avantage d'exploiter les portions de trajectoires observées qui sont compatibles avec les actions recommandées par la politique π jusqu'à l'apparition d'une incompatibilité (argument d'efficience augmentée). Une limitation de l'estimateur \hat{V}_{AIPW} tient au fait que $q_{\pi,k}(\cdot)$ est un paramètre de nuisance dépendant de π , et par conséquent l'estimation de $q_{\pi,k}(\cdot)$ doit être répétée pour chaque politique $\pi \in \Pi$ à évaluer. Du point de vue de l'optimisation, cette caractéristique rend le problème $\hat{\pi}^* = \arg \max_{\pi \in \Pi} \hat{V}_{AIPW}(\pi)$ difficilement tractable pour des classes de politiques Π non-triviales. Pour résoudre ce problème computationnel plusieurs heuristiques ont été proposées ; voir par exemple Zhang et al. [71] équation (7) et Tsiatis et al. [36] section 7.4.3 page 282 pour une discussion à ce sujet.

Remarque (Gestion des individus entrants dans un état terminal). *Dans le cas où un individu i entre dans un état terminal à l'étape l , par convention, on note pour tout $k = l, \dots, K$, $S_k^{(i)} = \Phi$, $H_k^{(i)} = \Phi$ et $A_k^{(i)} = 0$. Par exemple, un patient qui décède entre les étapes $(l-1)$ et l est considéré en état terminal aux étapes $k = l, \dots, K$: pour ce patient il n'est plus nécessaire de prendre une action aux étapes $k = l, \dots, K$. En présence d'individus entrant en état terminal, il est possible d'utiliser les estimateurs \hat{V}_{IPW} et \hat{V}_{IPW*} en imposant pour tout k et pour tout $a \in \mathcal{A}_k \setminus \{0\}$, $\hat{e}_{k,a}(\Phi) = 0$, $\hat{e}_{k,0}(\Phi) = 1$, et $\pi_k(\Phi) = 0$. Dans le cas où le résultat final est connu dès l'entrée en état terminal d'un individu, l'estimateur \hat{V}_{AIPW} est également utilisable. Par exemple, si un patient décède entre les étapes $(l-1)$ et l et le résultat final Y est "nombre de jours passés en dehors de l'hôpital dans les 60 jours suivant l'étape 1," ce résultat final est connu dès l'étape l . Formellement, il existe un ensemble connu de fonctions $\{g_k\}_{k=1}^K$ tel que quand $S_k = \Phi$, on a $Y = g_{k-1}(S_{1:(k-1)})$. Dans ces conditions, on peut utiliser \hat{V}_{AIPW} en imposant pour tout k et pour tout π*

$$\hat{q}_{\pi,k}(H_k^{(i)}) = g_{k-1}(S_{1:(k-1)}^{(i)}) \quad \text{quand} \quad H_k^{(i)} = \Phi.$$

Note : Pour conduire cette gestion des individus entrants dans un état terminal, il est nécessaire d'adapter légèrement l'hypothèse 7. L'hypothèse de positivité séquentielle modifiée requise dans ce contexte est énoncée dans l'annexe B4.

2.5.3 Cas particulier de l'estimation des politiques de type “Quand débuter le traitement ?”

Dans cette section nous proposons une brève introduction à la méthode décrite par Nie et al [72] pour estimer une politique optimale de type “Quand débuter le traitement ?”. Le formalisme et la procédure d'estimation sont d'abord énoncée sous une forme simplifiée, une interprétation est donnée ensuite.

Une politique π de type “Quand débuter le traitement ?” est définie par une variable aléatoire “étape” $\tau_\pi \in \{1, \dots, K, K+1\}$ tel que pour tout $k = \{1, \dots, K\}$, si $A_{k-1} = 1$, $\pi_k(H_k) = 1$, sinon si $k \geq \tau_\pi$, $\pi_k(H_k) = 1$, sinon $\pi_k(H_k) = 0$. Ce formalisme impose la structure suivante : i) si le traitement a déjà été débuté, alors la politique doit recommander de le continuer ; et par ailleurs, ii) lorsque la politique recommande de débuter le traitement mais cette recommandation n'est pas suivie, aux étapes suivantes, la politique doit persister dans sa recommandation de débuter le traitement. Pour estimer une politique optimale de type “Quand débuter le traitement ?”, Nie et al. proposent de résoudre le problème

$$\hat{\pi}^* = \arg \max_{\pi \in \Pi} \hat{\Delta}(\pi, \mathbf{0})$$

où

$$\hat{\Delta}(\pi, \mathbf{0}) = n^{-1} \sum_{i=1}^n \sum_{k=1}^K \mathbb{1}\{k \geq \tau_\pi^{(i)}\} \frac{\mathbb{1}\{A_{1:(k-1)}^{(i)} = 0\}}{\prod_{k'=1}^{k-1} \hat{e}_{k',0}(H_{1:k'}^{(i)})} \hat{\Psi}_k(S_{1:k}^{(i)})$$

avec

$$\begin{aligned} \hat{\Psi}_k(S_{1:k}^{(i)}) &= \hat{q}_{now}(S_{1:k}^{(i)}, k) - \hat{q}_{next}(S_{1:k}^{(i)}, k) \\ &\quad + \frac{\mathbb{1}\{A_k^{(i)} = 1\}}{\hat{e}_{k,1}(H_{1:k}^{(i)})} \{Y^{(i)} - \hat{q}_{now}(S_{1:k}^{(i)}, k)\} \\ &\quad - \frac{\mathbb{1}\{A_k^{(i)} = 0\} \mathbb{1}\{A_{k+1}^{(i)} = 1\}}{\hat{e}_{k,0}(H_{1:k}^{(i)}) \hat{e}_{k+1,1}(H_{1:(k+1)}^{(i)})} \{Y^{(i)} - \hat{q}_{next}(S_{1:k}^{(i)}, k)\} \end{aligned}$$

et

$$\begin{aligned} \hat{q}_{now}(S_{1:k}^{(i)}, k) &= \hat{\mathbb{E}}[Y | S_{1:k} = S_{1:k}^{(i)}, A_{1:(k-1)} = 0, A_k = 1] \\ \hat{q}_{next}(S_{1:k}^{(i)}, k) &= \hat{\mathbb{E}}\left[\frac{\mathbb{1}\{A_{k+1} = 1\}}{\hat{e}_{k+1,1}(H_{1:(k+1)})} Y | S_{1:k} = S_{1:k}^{(i)}, A_{1:(k-1)} = 0, A_k = 1\right]. \end{aligned}$$

Nie et al. généralisent cette procédure au cas de politiques de type “Quand débuter un traitement parmi un set de traitements prédéfinis ?”.

Dans la procédure ci-dessus $\hat{\Delta}(\pi, \mathbf{0})$ s'interprète comme la différence entre la valeur d'une politique π et la valeur d'une politique dénotée $\mathbf{0}$ qui ne recommande jamais de débuter le traitement (avec cette politique, on a $\tau_0 \equiv K+1$). La variable $\tau_\pi^{(i)}$ désigne la première étape à laquelle la politique π recommande d'initier le traitement. Par exemple $\tau_\pi^{(i)} = 3$ indique que pour le patient i , la politique π recommande l'administration du traitement aux étapes $k = 3, \dots, K$, mais pas aux étapes précédentes. Les “scores” $\hat{\Psi}_k(S_{1:k}^{(i)})$ s'interprètent comme l'estimation double robuste de l'effet traitement individuel pour : débuter le traitement à l'étape k versus débuter le traitement à l'étape $k+1$. Cette formulation de l'estimateur $\hat{\Delta}(\pi, \mathbf{0})$ est la conséquence d'une décomposition de $\Delta(\pi, \mathbf{0})$ en

une somme d'avantages locaux tel que décrit par Murphy [67, Lemma 1]. En ce qui concerne l'implémentation, Nie et al. proposent trois modifications à la formulation simplifiée de $\hat{\Delta}(\pi, \mathbf{0})$ présentée ci-dessus : i) la normalisation de chaque terme à estimer dans l'esprit de l'estimateur \hat{V}_{IPW*} , (première équation section 5), ii) la reformulation de q_{next} pour une estimation de cette fonction par régression pondérée non paramétrique (remarque 3) et iii) l'estimation de tous les paramètres de nuisance par méthode de *cross-fitting*. Sur le plan théorique, Nie et al. proposent une borne supérieure probabiliste asymptotique pour le regret $V(\pi^*) - V(\hat{\pi})$ [72, Theorem 7]. En pratique, cette borne peut aider à la décision de i) arrêter la procédure d'optimisation $\hat{\pi}^* = \arg \max_{\pi \in \Pi} \hat{\Delta}(\pi, \mathbf{0})$ (quand la borne supérieure du regret est faible), ii) augmenter la dimension VC de la classe de politique Π (quand la borne supérieure du regret est faible et la politique optimale estimée à une valeur très peu supérieure à la politique actuellement implémentée ; c'est à dire si “ $\hat{\Delta}(\pi, e)$ est faible”). La gestion des individus entrant en état terminal n'étant pas naturelle avec l'estimateur $\hat{\Delta}(\pi, \mathbf{0})$, les auteurs proposent explicitement un estimateur modifié pour gérer ces situations [72, Algorithm 2].

2.5.4 Estimation d'une politique optimale par modélisations récursives

Une limitation des méthodes de maximisation de valeur est liée aux problèmes d'optimisation $\hat{\pi}^* = \arg \max_{\pi \in \Pi} \hat{V}(\pi)$ ou $\hat{\pi}^* = \arg \max_{\pi \in \Pi} \hat{\Delta}(\pi, \mathbf{0})$. Ces problèmes d'optimisation étant non lisses et non continus, ils sont potentiellement difficiles à résoudre numériquement. Une alternative classique pour s'affranchir de ces difficultés computationnelles est d'utiliser l'algorithme *Q-learning* qui ne requiert pas de définir explicitement à l'avance une classe de politique Π . Cet algorithme consiste à estimer récursivement pour toutes les étapes $k = K, \dots, 1$, la fonction de valeur optimale conditionnelle aux états et aux actions :

$$\begin{aligned} Q_k^*(H_k, A_k) &\stackrel{\text{def}}{=} \mathbb{E}_{\pi^*}[Y|H_k, A_k] \\ &= \mathbb{E}[Y(A_{1:(k-1)}, A_k, \pi_{k+1}^*(H_{k+1}), \dots, \pi_K^*(H_K))|H_k, A_k]. \end{aligned}$$

La procédure d'estimation de $\{Q_k^*\}_{k=1}^K$ est décrite dans l'Algorithme 2. Une fois ces fonctions estimées, la politique optimale peut être directement obtenue en prenant

$$\hat{\pi}_k^*(H_k) = \arg \max_{a \in \mathcal{A}_k} \hat{Q}_k^*(H_k, a), \quad k = 1, \dots, K.$$

Avec cette approche, la classe de politiques déterministes Π est induite par le choix des modèles d'apprentissage supervisé pour $\{Q_k^*\}_{k=1}^K$; celle-ci est d'autant plus expressive que les modèles choisis sont souples. Par ailleurs, du fait de la nature récursive de la procédure *Q-learning*, l'erreur de spécification des modèles s'amplifie de $k = K$ vers $k = 1$. Pour cette raison, l'utilisation de modèles non paramétriques souples est souvent recommandée pour l'estimation de $\{Q_k^*\}_{k=1}^K$ [66]. Dans le cas où le résultat final Y est continu, Wallace et Moodie [73], proposent une adaptation de l'algorithme *Q-learning* rendant la procédure d'estimation double robuste et efficiente. Cette procédure appelée *dWOLS* repose sur l'utilisation de modèles linéaires pondérés pour la production de politiques interprétables et permettant la construction d'intervales de confiance à l'échelle individuelle. Une description de la procédure *dWOLS* est donnée dans le Chapitre 6 (voir page 4 du supplément). Dans le contexte de la médecine personnalisée, une revue récente des méthodes d'estimation d'une politique optimale a été conduite par Li et al. [74].

2.5.5 Note concernant les critères de jugement censurés

En médecine, de nombreux résultats finaux d'intérêt sont censurés : temps restant avant décès, temps restant en vie sans rechute (*progression-free survival*), etc. Plusieurs méthodes d'apprentissage par renforcement ont récemment été développées pour estimer une stratégie de traitement dynamique optimale pour des résultats censurés. Par exemple, Simoneau et al. [75] proposent de spécifier des modèles paramétriques AFT (*Accelerated Failure Time*) et d'utiliser d'une procédure récursive de résolution d'équations d'estimation généralisées pondérées (*weighted generalized estimating equations*). Cho et al. [76] proposent d'assouplir ces hypothèses paramétriques en utilisant des forêts aléatoires ; leur méthode permet par ailleurs d'estimer une politique optimale au sein de classes de politiques plus flexibles (non linéaires).

Une alternative à l'utilisation de méthodes gérant explicitement les résultats censurés est de pré-traiter ces résultats censurés pour obtenir des pseudo-observations pour chaque individu [77]. Les pseudo-observations peuvent ensuite être utilisées comme résultat final (scalaire) à optimiser par n'importe quelle méthode ne gérant pas explicitement les résultats censurés (les méthodes décrites précédemment dans cette section sont utilisables). La technique de pré-traitement par pseudo-observations permet de gérer les situations de censures dépendantes des temps avant événements ainsi que les situations de risques compétitifs [78].

Algorithme 2 Procédure récursive d'estimation non paramétrique de $\{Q_k^*\}_{k=1}^K$.

Entrée: Données longitudinales au format élargi : $(S_1^{(i)}, A_1^{(i)}, \dots, S_K^{(i)}, A_K^{(i)}, Y^{(i)})_{1 \leq i \leq n}$.

- Étape $k = K$. Estimer Q_K^* par méthode d'apprentissage supervisé avec les labels Y et les observations $(H_K^{(i)}, A_K^{(i)})$.[†]
- Étape $k = K - 1$
 - Produire les pseudo-résultats[‡] $\tilde{Y}_K^{(i)} \leftarrow \max_{a \in \mathcal{A}_K} \hat{Q}_K^*(H_K^{(i)}, a)$
 - Estimer Q_{K-1}^* par méthode d'apprentissage supervisé avec les labels $\tilde{Y}_K^{(i)}$ et les observations $(H_{K-1}^{(i)}, A_{K-1}^{(i)})$.
- Étape $k = \dots$
 - ...
 - ...
- Étape $k = 1$
 - Produire les pseudo-résultats $\tilde{Y}_2^{(i)} \leftarrow \max_{a \in \mathcal{A}_2} \hat{Q}_2^*(H_2^{(i)}, a)$
 - Estimer Q_1^* par méthode d'apprentissage supervisé avec les labels $\tilde{Y}_2^{(i)}$ et les observations $(H_1^{(i)}, A_1^{(i)})$.

Sortie: $\{\hat{Q}_k^*\}_{k=1}^K$

[†]Le réglage d'hyperparamètres peut être réalisé par validation croisée. S'il est nécessaire de gérer des individus entrant en état terminal on utilisera le set $\{(H_K^{(i)}, A_K^{(i)}), Y^{(i)}\}_{i: H_K^{(i)} \neq \Phi}$ et des adaptations similaires doivent être reconduites aux étapes $k = K - 1, \dots, 1$.

[‡]S'il est nécessaire de gérer des individus entrant en état terminal cette ligne doit être modifiée pour $\tilde{Y}_K^{(i)} \leftarrow \mathbb{1}\{H_K^{(i)} \neq \Phi\} \max_{a \in \mathcal{A}_K} \hat{Q}_K^*(H_K^{(i)}, a) + \mathbb{1}\{H_K^{(i)} = \Phi\} g_{K-1}(S_{1:(k-1)}^{(i)})$ et des adaptations similaires doivent être reconduites aux étapes $k = K - 2, \dots, 1$.

Chapter 3

Continuous renal replacement therapy versus intermittent hemodialysis as first modality for renal replacement therapy in severe acute kidney injury: a secondary analysis of AKIKI and IDEAL-ICU studies

You are smarter than your data. Data do not understand causes and effects; humans do.

Judea Pearl, *The book of why*

La musique associée à ce chapitre est ‘Into My Arms’ de Nick Cave.

RESEARCH

Open Access



Continuous renal replacement therapy versus intermittent hemodialysis as first modality for renal replacement therapy in severe acute kidney injury: a secondary analysis of AKIKI and IDEAL-ICU studies

Stéphane Gaudry^{1,2,3,4*}, François Grolleau⁵, Saber Barbar⁶, Laurent Martin-Lefevre⁷, Bertrand Pons⁸, Éric Boulet⁹, Alexandre Boyer¹⁰, Guillaume Chevrel¹¹, Florent Montini¹², Julien Bohe¹³, Julio Badie¹⁴, Jean-Philippe Rigaud¹⁵, Christophe Vinsonneau¹⁶, Raphaël Porcher⁵, Jean-Pierre Quenot^{17,18,19†} and Didier Dreyfuss^{3,20†}

Abstract

Background: Intermittent hemodialysis (IHD) and continuous renal replacement therapy (CRRT) are the two main RRT modalities in patients with severe acute kidney injury (AKI). Meta-analyses conducted more than 10 years ago did not show survival difference between these two modalities. As the quality of RRT delivery has improved since then, we aimed to reassess whether the choice of IHD or CRRT as first modality affects survival of patients with severe AKI.

Methods: This is a secondary analysis of two multicenter randomized controlled trials (AKIKI and IDEAL-ICU) that compared an early RRT initiation strategy with a delayed one. We included patients allocated to the early strategy in order to emulate a trial where patients would have been randomized to receive either IHD or CRRT within twelve hours after the documentation of severe AKI. We determined each patient's modality group as the first RRT modality they received. The primary outcome was 60-day overall survival. We used two propensity score methods to balance the differences in baseline characteristics between groups and the primary analysis relied on inverse probability of treatment weighting.

Results: A total of 543 patients were included. Continuous RRT was the first modality in 269 patients and IHD in 274. Patients receiving CRRT had higher cardiovascular and total-SOFA scores. Inverse probability weighting allowed to adequately balance groups on all predefined confounders. The weighted Kaplan–Meier death rate at day 60 was 54.4% in the CRRT group and 46.5% in the IHD group (weighted HR 1.26, 95% CI 1.01–1.60). In a complementary analysis of less severely ill patients (SOFA score: 3–10), receiving IHD was associated with better day 60 survival

*Correspondence: stephanegaudry@gmail.com

†Jean-Pierre Quenot and Didier Dreyfuss contributed equally to this work

³ Present Address: Common and Rare Kidney Diseases, INSERM, UMR-S 1155, Hôpital Tenon, Sorbonne Université, 4 rue de la Chine, 75020 Paris, France

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

compared to CRRT (weighted HR 1.82, 95% CI 1.01–3.28; $p < 0.01$). We found no evidence of a survival difference between the two RRT modalities in more severe patients.

Conclusion: Compared to IHD, CRRT as first modality seemed to convey no benefit in terms of survival or of kidney recovery and might even have been associated with less favorable outcome in patients with lesser severity of disease. A prospective randomized non-inferiority trial should be implemented to solve the persistent conundrum of the optimal RRT technique.

Keywords: Renal replacement therapy, Acute kidney injury, Critical care

Background

Acute kidney injury (AKI) complicates the course of many critically ill patients and is associated with both increased morbidity and mortality [1, 2]. The treatment of AKI is based on both conservative measures and timely use of renal replacement therapy (RRT). Intermittent hemodialysis (IHD) and continuous RRT (CRRT) with hemofiltration are the two main modalities in the intensive care unit (ICU). Controversy exists as to which is the optimal one in this setting [3–5]. Indeed, CRRT allows for slower fluid removal which may ensure better hemodynamic stability and slower control of solute concentration than IHD (thereby minimizing fluid shift). This is the reason why many critical care physicians favor CRRT, at least at the initial stage of ICU stay [6].

However, the COVID crisis has highlighted the shortage of RRT devices [7]. To overcome this problem, some authorities have suggested the use of CRRT for short periods (i.e., 12 h a day) in order to increase machine availability and the use of the same machine in two different patients in the same day [8]. A simpler alternative would be to favor IHD as it allows treating more than two patients (usually three) with the same machine in one day. Indeed, IHD may be well tolerated owing to simple interventions including isovolemic initiation, reduced dialysate temperature, preferential use of bicarbonate buffer, sodium profiling (i.e., dialysate $[Na^+] > 145 \text{ mmol/L}$) and conservative initial ultrafiltration [9].

Improvement in critical care delivery resulted in marked improvement of patient prognosis in recent years [10, 11]. For instance, mortality was substantially higher in both groups of a large RCT dating back 15 years that compared IHD and CRRT [12] than in recent RCTs on RRT timing where both modalities were used [13–15]. This justifies the reassessment of the impact of RRT modality on survival. New data can therefore be generated which will in turn generate new research questions and rekindle research interest in this area.

Our teams recently conducted two independent large multicenter RCTs on the timing of RRT initiation (AKIKI [13] and IDEAL-ICU [15] trials). The choice of RRT technique was left at clinician discretion at each study

site, in accordance with French guidelines [16]. The present study merged the two datasets in order to evaluate whether the choice of IHD or CRRT as first modality affects survival of critically ill patients with severe AKI.

Methods

Study design and patients

This study is a secondary analysis of two open pragmatic multicenter RCTs (the AKIKI [13] and IDEAL-ICU [15] trials) that compared an early RRT initiation strategy with a delayed one. Both were multicenter studies involving critically ill patients with severe AKI among other organ failures.

Patients allocated to the early strategy received RRT less than 12 h after documentation of severe AKI. This allows for studying this patient population at a time when baseline characteristics are unlikely to have changed noticeably, resulting in measurable confounding we can account for.

Following from the principles of an ideal target RCT, we emulated a trial where patients would have been randomized to receive either IHD or CRRT within twelve hours after the documentation of severe AKI. In a bid to mimic the intention to treat analysis from a randomized trial, we determined each patient's modality group as the first RRT modality they received. We then used propensity score-based analyses to correct for confounding.

The AKIKI trial and the IDEAL-ICU trial received approval for all participating centers from competent French legal authority (Comité de Protection des Personnes d'Île de France VI, ID RCB 2013-A00765-40, NCT01932190 for AKIKI and Comité de Protection des Personnes Est I ID RCB 2012-A00519-34 for IDEAL-ICU), and consent of patient or relatives was obtained before inclusion (except in emergencies where deferred consent was allowed by the Institutional Review Board).

Interventions

In both trials, the choice of RRT modality was left at clinician discretion and a switch to either modality was allowed over the course of a patient ICU stay. No patient was treated with hybrid machine such as sustained low efficiency dialysis (SLED). In the present study, we

analyzed outcomes according to the first modality used. Investigators were encouraged to follow current guidelines [16–18]. All study centers used dialyzers with biocompatible membranes for IHD and CRRT. The recommended buffer in dialysate and replacement fluid was bicarbonate. Recommendations included the delivery of a Kt/V of 3.9 per week during IHD and an effluent volume of 20–25 mL/kg/h for CRRT.

Outcomes

The primary outcome was 60-day survival measured from the date of initiation of RRT until death or day 60. Secondary outcomes were the status at ICU and hospital discharge, kidney recovery (defined as RRT discontinuation and spontaneous urine output higher than 1000 mL per 24 h in the absence of diuretic therapy or higher than 2000 mL per 24 h with diuretics) before day 28, survival with no need for RRT (RRT dependence) at day 28, length of stay in ICU, number of days free of RRT, mechanical ventilation and vasopressor at 28 days.

Statistical analyses

The main analysis relied on inverse probability of treatment weighting (IPTW) [19, 20] to balance the differences in baseline characteristics between treatment groups. A propensity score model was estimated using logistic regression, with the first RRT modality as dependent variable and the following pre-randomization characteristics as covariates: age, gender, weight, comorbidities (heart failure, hypertension, diabetes, cirrhosis, respiratory disease, cancer, AIDS, organ graft), treatments with immunosuppressive drugs, severity (respiratory SOFA, cardiovascular SOFA, bilirubin SOFA, platelet SOFA, Glasgow SOFA, renal SOFA, global SOFA), urea plasma concentration, pH, creatinine plasma concentration before ICU admission, and trial (AKIKI or IDEAL-ICU). These variables were specified before outcome analyses. Polynomials were used to handle potentially nonlinear effects of continuous variables. Nonlinear terms were removed if they did not improve balance of groups after weighting. Standardized mean differences (SMD) were examined to assess balance between groups before and after weighting [21, 22], and a value below 10% was considered as indicating clinically meaningful balance of a covariate [23]. Causal treatment effects were assessed for the whole weighted population (i.e. average treatment effect) in terms of hazard ratio (HR), risk ratio (RR), absolute risk difference (ARD) or mean difference (MD). We estimated survival through a weighted Kaplan–Meier estimator and used restricted mean survival time (RMST) to measure the average survival time for each group over the 60 days follow-up period [24].

One difficulty stemmed from the fact that some centers only provided one of the two RRT modalities, possibly violating the positivity assumption, which is crucial for propensity score-based methods. We, therefore, performed a sensitivity analysis estimating the treatment effects by overlap weighting (OW) [25] through a second propensity score model which accounted for centers. Specifically, we used a mixed-effects logistic regression model with the same aforementioned covariates as fixed effects and a random center effect.

The residual confounding effect that would be due to centers in the IPTW analysis was investigated through additional prognostic modelling and showed that very little confounding due to centers was expected (Additional file 1: Table S1 and Figure S2). The positivity assumption and between-group balance were further evaluated for each weighting technique by plotting the distribution of the propensity score in both groups (Additional file 1: Figures S3 and S4). Ninety-five percent confidence intervals (95% CI) were estimated by bootstrapping. Missing data were handled through multiple imputations by chained equations using outcomes as well as all confounders mentioned above in the imputation model. Because roughly 5% of patients had one or more missing predictors (Additional file 1: Figure S5), 5 independent imputed data sets were generated and analyzed separately. Estimates were then pooled using Rubin's rules.

We assessed treatment effect heterogeneity, using the same methodology stratified by thirds of baseline risk as evaluated by SOFA scores.

All analyses were performed using the R statistical software version 4.0.0 or later (R Foundation for Statistical Computing, <https://www.R-project.org/>).

Results

Patients

A total of 543 critically ill patients with severe AKI who received RRT in the early strategy of our two previous studies were included in the present one (304 from the AKIKI trial and 239 from the IDEAL-ICU trial). The median time between random allocation to the early strategy and RRT initiation was 2 h (IQR, 1 to 3) and 3 h (IQR, 2 to 4) in AKIKI and IDEAL-ICU, respectively.

Two hundred sixty-nine patients received CRRT and 274 IHD as first RRT modality (Fig. 1). Patient characteristics at baseline are depicted in Table 1. Patients receiving IHD were more frequently hypertensive, had higher serum creatinine levels and higher coagulation-SOFA scores at baseline. On the opposite, patients receiving CRRT had higher cardiovascular and total-SOFA scores.

Of the 269 patients who initially received CRRT, 56 (20.8%) switched to IHD on average at day 6.4. Of the 274

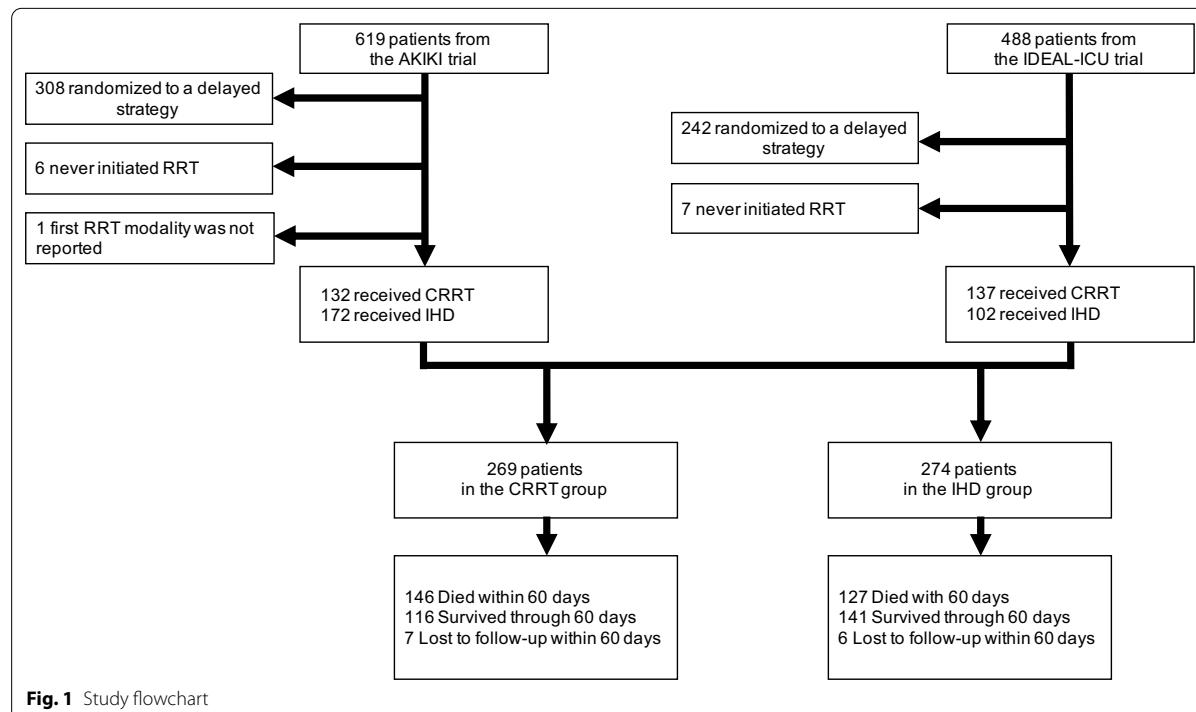


Fig. 1 Study flowchart

patients who initially received IHD, 48 (17.5%) switched to CRRT on average at day 5.7.

Unweighted analyses

Through the 60-day follow-up period, 273 patients (50%) died (146 in the CRRT group and 127 in the IHD group). Thirteen patients (2%) were lost to follow-up. The Kaplan–Meier death rate at day 60 was 54.7% in the CRRT group and 46.5% in the IHD group (Fig. 2A) (HR 1.27, 95% CI 1.00 to 1.61) The ARD was 8.1%, 95% CI – 0.5 to 16.7%. The RMST was 33.6 days in the CRRT group and 37.4 days in the IHD group (RMST difference 3.8 days, 95% CI 0.9 to 8.0).

Weighted analyses

Groups were modestly imbalanced at baseline for hypertension, immunosuppressive drugs, SOFA scores, and creatinine levels. However, inverse probability weighting allowed to adequately balance the groups on all predefined confounders (Table 1 and Additional file 1: Figure S6). The weighted Kaplan–Meier death rate at day 60 was 54.4% in the CRRT group and 46.5% in the IHD group (Fig. 2B) (HR 1.26, 95% CI 1.01 to 1.60). The weighted ARD was 7.9%, 95% CI – 0.7 to 16.1%). The weighted RMST was 33.9 days in the CRRT group and 37.4 days in the IHD group (weighted RMST difference 3.5 days, 95% CI 0.9 to 8.0).

In further IPTW analyses, ICU and hospital mortality did not differ between groups. Same holds true for survival with no RRT dependence at day 28, kidney function recovery within 28 days, RRT-free days at day 28, ICU-free days at day 28, and length of ICU stay. However, the number of ventilator-free days and vasopressor-free days at day 28 were significantly higher in the IHD group than in CRRT group (Table 2).

The overlap weighting (OW) analysis of the primary endpoint is presented in the supplementary appendix (Additional file 1: Figure S1). This sensitivity analysis did not change the magnitude of the treatment effect.

Complementary analysis

In an analysis stratified by thirds of SOFA scores, we found that IHD was associated with greater 60-day survival as compared to CRRT in less severely ill patients (weighted HR 1.82, 95% CI 1.01 to 3.28; $p < 0.01$ for patients with SOFA 3–10). In contrast, we found no evidence of a survival difference in more severely ill patients (Fig. 3).

Discussion

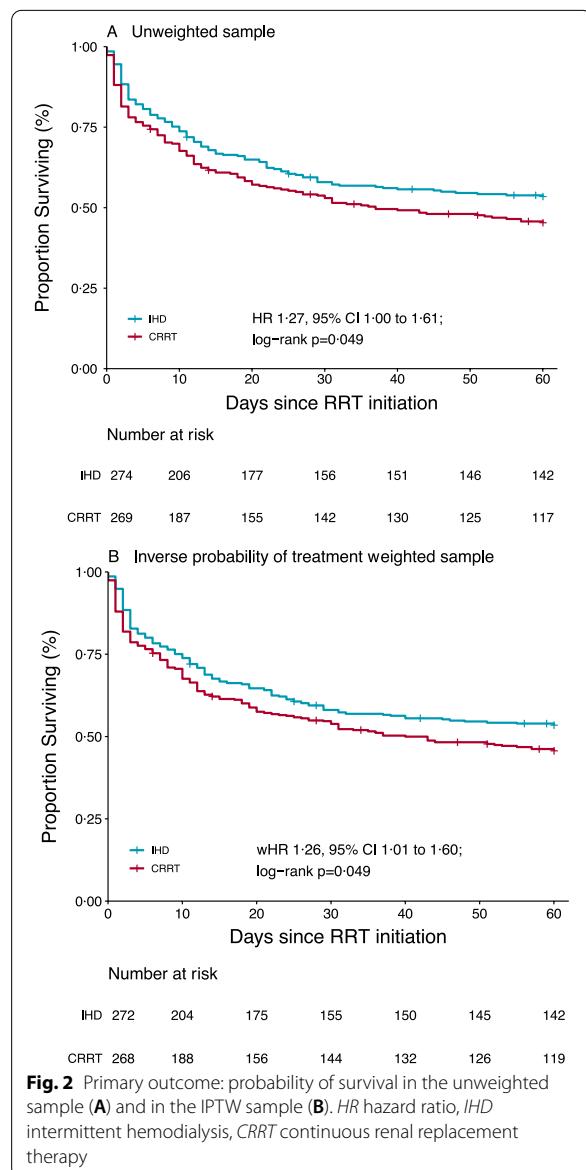
This study compared continuous and intermittent RRT for the first session in critically ill patients with severe AKI. Continuous renal replacement therapy was associated with shorter survival within 60 days and with

Table 1 Patient characteristics at baseline. All characteristics reported in the table were determined at inclusion in the AKIKI or IDEAL-ICU trial, or before initiation of renal replacement therapy*

Characteristic	Before weighting				After inverse probability weighting			
	Total (n = 543)	CRRT group (n = 269)	IHD group (n = 274)	SMD (%)	CRRT group (n = 268.2)	IHD group (n = 271.9)	SMD (%)	p value
Age, years	66.5 (13.3)	66.2 (13.4)	66.8 (13.3)	4.4	66.5 (13.4)	66.7 (13.2)	1.5	0.87
Female sex	200 (36.8%)	101 (37.6%)	99 (36.1%)	2.9	97.4 (36.3%)	101.2 (37.2%)	1.9	0.83
Serum creatinine before ICU admission, µmol/L*	86.2 (34.1)	84.6 (36.7)	87.8 (31.3)	9.4	86.6 (33.8)	86.8 (32.1)	0.6	0.95
Coexisting conditions								
Chronic kidney disease	54 (9.9%)	23 (8.6%)	31 (11.3%)	9.3	24.3 (9.1%)	30.6 (11.3%)	7.2	0.43
Chronic hypertension	301 (55.4%)	136 (50.6%)	165 (60.2%)	19.5	148.4 (55.4%)	153.3 (56.4%)	2.0	0.82
Diabetes mellitus	109 (20.1%)	55 (20.4%)	54 (19.7%)	1.8	53.7 (20.0%)	53.5 (19.7%)	0.9	0.92
Congestive heart failure	43 (7.9%)	22 (8.2%)	21 (7.7%)	1.9	21.6 (8.1%)	22.2 (8.2%)	0.4	0.97
Cirrhosis	53 (9.8%)	27 (10.0%)	26 (9.5%)	1.8	26.3 (9.8%)	25.1 (9.2%)	2.0	0.82
Respiratory disease	62 (11.4%)	30 (11.2%)	32 (11.7%)	1.7	30.4 (11.3%)	31.3 (11.5%)	0.5	0.95
Cancer	89 (16.4%)	48 (17.8%)	41 (15.0%)	7.8	43.2 (16.1%)	43.2 (15.9%)	0.6	0.95
AIDS	5 (0.9%)	3 (1.1%)	2 (0.7%)	4.0	3.1 (1.1%)	4.0 (1.5%)	2.8	0.79
Immunosuppressive drugs	32 (5.9%)	21 (7.8%)	11 (4.0%)	16.1	15.3 (5.7%)	14.0 (5.2%)	2.4	0.79
Organ transplantation	5 (0.9%)	4 (1.5%)	1 (0.4%)	11.7	2.6 (1.0%)	2.6 (1.0%)	<0.1	0.99
SOFA score at inclusion (0 to 24)	11.4 (3.2)	11.7 (3.1)	11.1 (3.3)	18.6	11.5 (3.2)	11.4 (3.2)	1.8	0.84
Renal SOFA (1 to 5)	3.7 (1.1)	3.6 (1.1)	2.9 (1.2)	4.4	3.7 (1.1)	3.7 (1.1)	0.4	0.97
Cardiovascular SOFA (1 to 5)	4.4 (1.3)	4.6 (1.1)	4.3 (1.4)	24.6	4.5 (1.2)	4.4 (1.3)	2.5	0.78
Liver SOFA (1 to 5)	1.8 (1.1)	1.8 (1.1)	1.8 (1.1)	4.0	1.8 (1.1)	1.8 (1.1)	3.1	0.73
Neurologic SOFA (1 to 5)	2.2 (1.5)	2.2 (1.5)	2.3 (1.5)	6.8	2.2 (1.5)	2.2 (1.5)	0.4	0.96
Coagulation SOFA (1 to 5)	3.2 (1.6)	3.0 (1.6)	3.3 (1.6)	17.3	3.2 (1.6)	3.1 (1.6)	2.1	0.81
Body weight, kg	82.0 (21.9)	81.1 (21.9)	83.0 (21.9)	8.8	83.0 (23.5)	82.3 (21.9)	3.3	0.73
Laboratory values								
Serum creatinine, µmol/L	286.6 (126.9)	273.7 (122.5)	299.3 (130.0)	20.3	287.5 (129.1)	290.2 (124.0)	2.1	0.82
Serum urea, mmol/L	19.8 (9.1)	19.4 (9.0)	20.3 (9.2)	10.1	19.8 (9.0)	20.0 (9.0)	1.5	0.87
Serum potassium, mmol/L	4.38 (0.77)	4.41 (0.78)	4.35 (0.76)	8.3	4.38 (0.77)	4.38 (0.79)	0.5	0.96
Arterial blood pH	7.30 (0.10)	7.30 (0.10)	7.30 (0.09)	4.3	7.30 (0.10)	7.30 (0.10)	1.3	0.89

Data are mean (SD) or n (%). SMD standardized mean difference, expressed as a percentage; SOFA score Sequential Organ Failure Assessment score

*The serum creatinine concentration before ICU admission was either determined with the use of values measured in the 12 months preceding the ICU stay or was estimated using the Modification of Diet in Renal Disease Study Group formula



needed to solve this problem [8]. Intermittent hemodialysis allows faster removal of uremic toxins and control of electrolyte and acid–base disturbance than CRRT. This makes the same IHD machine available for several patients in a same day, particularly in countries where IHD is provided by usual ICU nurses, as is the case in France. In these conditions, it is of utmost importance to reassess the actual safety of IHD in critically ill patients. We took advantage of the fact that both studies included in the present analysis mandated very precise settings in order to maximize tolerance of IHD [9, 13, 15, 26]. In the present study, the choice of modality of RRT did not affect survival outcomes in the most severely ill patients. This result contradicts some common opinion [3] and guidelines [17] that promote CRRT for these patients. An explanation for this finding may be that any effect of RRT modality on prognosis was likely to be obscured by the severity of illness. In contrast, the less severely ill patients (according to the SOFA score) benefited significantly ($p < 0.01$) from IHD. This benefit of IHD in these patients may be explained by the short duration of RRT sessions which permits patients to be mobilized for nursing, early mobilization and medical procedure (such as CT scan or operating theater).

The KDIGO guidelines give a weak recommendation in favor of CRRT [17]. In contrast, French recommendations stipulate the equivalence of the two techniques [16]. Large multicenter randomized studies [12, 27] and a meta-analysis of the Cochrane Collaboration involving 15 RCTs and a total of 1550 patients concluded that CRRT did not differ from IHD with respect to mortality and kidney recovery [28]. Worthy of note is the fact that one randomized study [29] reported a higher mortality with CRRT than with IHD. Moreover, an individual patient data meta-analysis reported that time to cessation of RRT through 28 days was longer when CRRT was used as the initial modality of RRT [30]. Despite this, some key opinion leaders continue to favor CRRT over IHD in particular because of the alleged although unproven superiority of the former in hemodynamically unstable patients. This contention is however not supported by the conclusion of the Cochrane meta-analysis [28] which showed no difference in the occurrence of hemodynamic instability or hypotension or need for escalation of vasoressor therapy between RRT modalities. Then, there is no evidence-based data to support the fear of IHD in critically ill patients. Worthy of note is the fact that studies included in the Cochrane meta-analysis were published more than ten years ago whereas the prognosis of critically ill patients markedly improved in recent years [10, 11] and RRT modalities and indications were refined. For instance, mortality rate was more than 65% in both arms of the largest multicenter RCT comparing CRRT

longer duration of mechanical ventilation and vasopressor support than IHD. This association persisted after adjustment for confounders through propensity scores methods. Moreover, in a complementary analysis, we showed that this difference was mainly the result of the poorer outcome of the less severely ill patients allocated to CRRT as compared to those allocated to IHD. These results may inform the debate on RRT modalities in critically ill patients.

The present COVID crisis is associated with a shortage of RRT devices [7], and innovative solutions are

Table 2 Average effect of CRRT versus IHD on secondary outcomes in the original and inverse probability of treatment weighted samples

Outcome	CRRT group (n = 269)	IHD group (n = 274)	RR or Difference	(95% CI)	Weighted RR or weighted difference	(95% CI)
Hospital mortality	146 (54.2%)	132 (48.3%)	1.12	(0.95 to 1.32)	1.13	(0.94 to 1.34)
ICU mortality	130 (48.3%)	114 (41.6%)	1.16	(0.96 to 1.40)	1.14	(0.93 to 1.39)
Survival with no RRT dependence at day 28	117 (43.5%)	132 (48.0%)	0.91	(0.75 to 1.10)	0.92	(0.75 to 1.13)
Kidney recovery at day 28	123 (45.7%)	134 (48.9%)	0.93	(0.78 to 1.12)	1.00	(0.83 to 1.21)
RRT-free days at day 28	11.5 (12.0)	13.4 (12.6)	-1.9	(-3.9 to 0.2)	-1.6	(-3.9 to 0.4)
Ventilator-free days at day 28	8.0 (10.0)	9.9 (10.8)	-1.9	(-3.7 to -0.2)	-1.9	(-3.4 to -0.1)
Vasopressor-free days at day 28	11.6 (11.7)	14.1 (12.3)	-2.5	(-4.5 to -0.5)	-2.1	(-4.1 to -0.1)
ICU-free days at day 28	6.0 (8.6)	7.7 (9.5)	-1.6	(-3.2 to -0.1)	-1.4	(-3.1 to 0.2)
Length of ICU stay, days	14.2 (15.4)	14.4 (20.8)	-0.2	(-3.8 to 2.5)	0.1	(-3.0 to 2.6)

Data are mean (SD) or n (%). All results are pooled over the imputed datasets. RRT renal replacement therapy, CRRT continuous renal replacement therapy, IHD intermittent hemodialysis, ICU intensive care unit, RR risk ratio, CI confidence interval

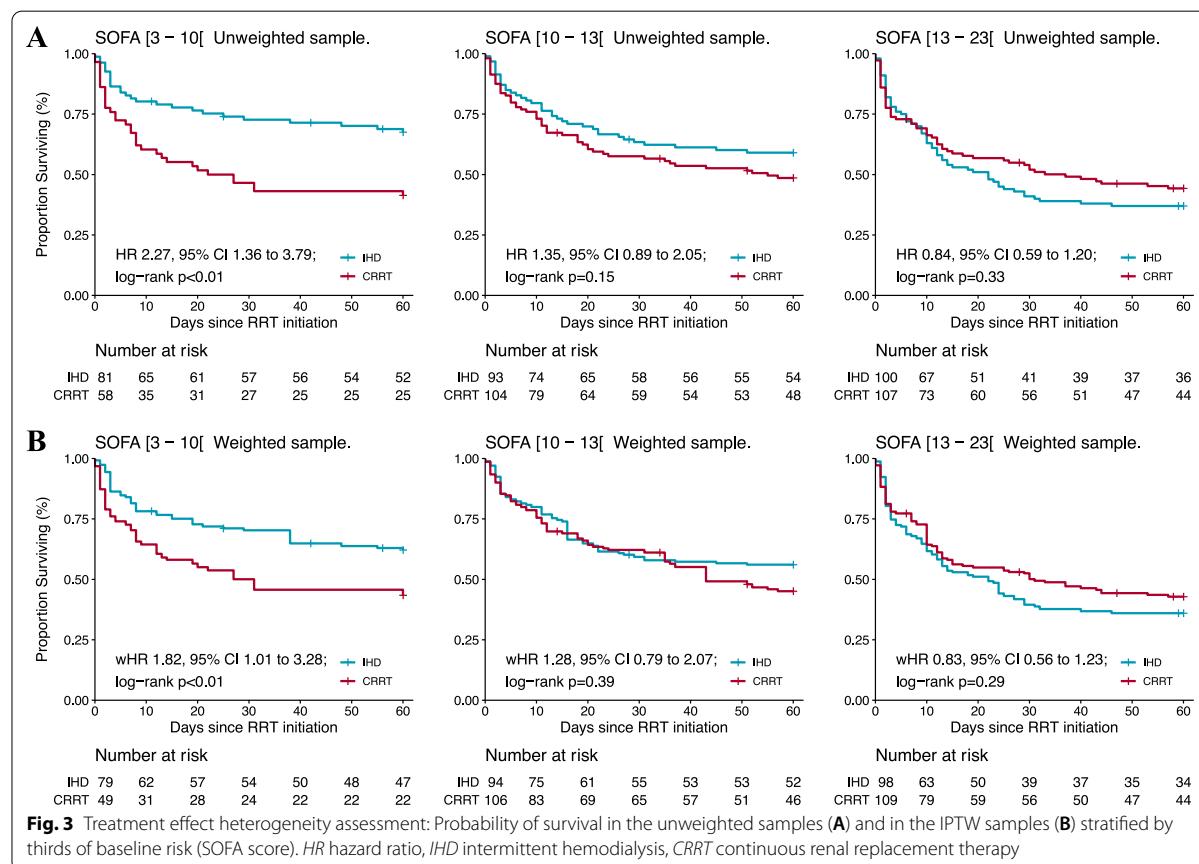


Fig. 3 Treatment effect heterogeneity assessment: Probability of survival in the unweighted samples (A) and in the IPTW samples (B) stratified by thirds of baseline risk (SOFA score). HR hazard ratio, IHD intermittent hemodialysis, CRRT continuous renal replacement therapy

and IHD that was published in 2006 [12]. In contrast, recent RCTs on RRT indication (in which both methods were used) found that mortality was comprised between 43 and 58% [13–15].

This simple finding would justify a reappraisal of the respective merits and dangers of RRT techniques and the performance of new RCT comparing RRT modalities in a large cohort of patients with pre-specified RRT indications. This is all the most reasonable as the present study shows that compared to CRRT, mortality was not increased in a large population of critically ill patients when IHD was used as the first RRT session. The first session is frequently hampered by hemodynamic instability since patients may not be stabilized at the time of initiation of the technique. There is no reason to believe that subsequent RRT sessions would be less tolerated than the first one. In such condition, focus on the first modality initiated should have favored CRRT. Finding an association of opposite direction in fact strengthens our conclusions.

Our study included 543 critically ill patients (a number superior to the largest RCT included in the meta-analysis of the Cochrane [28]) who received RRT within twelve hours after the occurrence of severe AKI. The use of strict criteria for RRT initiation and the substantial overlap of the propensity score distribution (see Additional file 1: Figure S3) led to mimic the intention to treat analysis from an RCT. Indeed, the moderate imbalance of baseline characteristics was adequately corrected by inverse probability weighting (all SMDs less than 10%).

Our study has limitations. First, even though the present study used high-quality data from two large RCTs, it was not a RCT itself and bias could have arisen because of unknown or unmeasured confounders. However, after we applied robust methods to draw inferences from observational data, all the known and clinically relevant prognostic variables were balanced between groups (see Additional file 1: Figure S3 and S6). Second, we only analyzed patients allocated to the early group of AKIKI and IDEAL-ICU studies. However, time from ICU admission to RRT initiation was short and similar to those in previous RCTs, as well as reason for initiation [12].

The advantage of considering the early arm only stems from the fact that RRT was started just after randomization occurred (2–3 h). Then, baseline characteristics were unlikely to change noticeably. This would not be the case if we also included the delayed arm where RRT was initiated 48 to 52 h later than in the early arm. In such conditions, it would not be possible to eliminate major confounding. Third, patients receiving CRRT had higher cardiovascular and total-SOFA scores at baseline. Even if we used high-quality methodology with a propensity

score analysis, confounding factors may persist. For instance, fluid overload which is difficult to accurately assess may have played a role.

Forth, even though the HR and RMSTD pointed to large treatment effect, these results may be fragile as statistical significance may depend on only a few events. However, a recent systematic review on critical care trials showed that for the trials with a statistically significant mortality difference, the fragility index was 4 (IQR 1–20) [31]. Fifth, when we considered 60-days survival through ARD, as opposed to HR and RMSTD, the significance of the association between CRRT and higher mortality was lost. Yet, this may just be a consequence of a lack of power in the ARD analyses: ARD assesses survival at a particular time point, whereas RMSTD and HR take advantage of all the information accrued from baseline through day 60. Sixth, we did not assess immediate adverse event such as hypotension, arrhythmia and increase in vasopressor during RRT session. Seventh, we did not assess a potential advantage on fluid control of CRRT over IHD. However, patient-centered outcomes such as duration of ICU and hospital stay may be affected by fluid balance but did not differ according to RRT modality. Eighth, a significant proportion of patients (almost 20%) were switched from a modality to another. However, the objective of the study was to assess the impact of the first RRT modality on prognosis and this first RRT modality was continued for a long duration in both groups (almost 6 days).

It should be noted that the clinical decision for choosing the optimal RRT mode is not only relying on outcome data but also on other factors such as practicability, availability of machines and the existence of trained nurse for IHD, which is the case in French ICUs.

Conclusion

In this secondary analysis of two large multicenter randomized controlled trials on RRT timing, CRRT as first modality did not provide any benefit in term of survival within 60 days or kidney recovery and might be associated with less favorable outcome in patients with lesser severity of disease. Our results should merely be viewed as generating an unexpected hypothesis and an impetus for performing new large RCTs comparing RRT modalities.

Abbreviations

RRT: Renal replacement therapy; AKI: Acute kidney injury; IHD: Intermittent hemodialysis; CRRT: Continuous renal replacement therapy; RCT: Randomized controlled trial; ICU: Intensive care unit; AKIKI: Artificial Kidney Initiation in Kidney Injury; IDEAL-ICU: Initiation of dialysis early versus delayed in intensive care unit; KDIGO: Kidney Disease Improving Global Outcomes; SOFA: Sepsis related

Organ Failure Assessment; HR: Hazard ratio; IPTW: Inverse probability of treatment weighting; SMD: Standardized mean differences; RR: Risk ratio; ARD: Absolute risk difference; MD: Mean difference (MD); RMST: Restricted mean survival time.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13054-022-03955-9>.

Additional file 1. Table S1. Balance of prognostic score and centre effects before and after inverse probability of treatment weighting.

Table S2. Baseline characteristics of patients in the two groups after overlap weighting. **Figure S1.** Probability of survival in the overlap weighted sample. **Figure S2.** Bubble plot evaluating the association between centre-specific random effect for treatment allocation and centre-specific random effect for prognosis. **Figure S3.** Propensity score distribution in the original (A) and inverse probability of treatment weighted (B) samples.

Figure S4. Propensity score (including centres as random effects) distribution in the original (A) and overlap weighted (B) samples. **Figure S5.** Fraction of missing data for each variable in the original sample. **Figure S6.** Standardised mean difference for each potential confounder in the unweighted, IPTW and OW samples.

Kidney Diseases, INSERM, UMR-S 1155, Hôpital Tenon, Sorbonne Université, 4 rue de la Chine, 75020 Paris, France. ⁴Investigation Network Initiative—Cardiovascular and Renal Clinical Trialists, Bobigny, France. ⁵Centre of Research in Epidemiology and Statistics (CRESS), Université de Paris, French Institute of Health and Medical Research (INSERM), National Institute of Agricultural Research (INRA), Paris, France. ⁶Hôpital Caremeau, Nîmes, France. ⁷Réanimation Polyvalente, CHR départementale La Roche Sur Yon, La Roche sur Yon, France. ⁸CHU Pointe-À-Pitre/Abymes, Pointe-a-Pitre, France. ⁹Réanimation Et USC, GH Carnelle Portes de l'Oise, 95260 Beaumont sur Oise, France. ¹⁰CHU de Bordeaux, Service de Réanimation Médicale, 33000 Bordeaux, France. ¹¹Réanimation Polyvalente, CH Sud Francilien, Corbeil Essones, France. ¹²Réanimation Polyvalente, Centre Hospitalier d'Avignon, Avignon, France. ¹³Anesthésie Réanimation Médicale et Chirurgicale, CH Lyon Sud, Pierre Benite, France. ¹⁴Réanimation Polyvalente, Hôpital Nord Franche-Comté CH Belfort, Belfort, France. ¹⁵Réanimation Polyvalente, CH de Dieppe, Dieppe, France. ¹⁶Médecine Intensive Réanimation, CH Bethune Beuvry – Germont et Gauthier, Bethune, France. ¹⁷Department of Intensive Care, François Mitterrand University Hospital, Dijon, France. ¹⁸Lipness Team, INSERM Research Center LNC-UMR1231 and LabExLiSTIC, University of Burgundy, Dijon, France. ¹⁹INSERM CIC 1432, Clinical Epidemiology, University of Burgundy, Dijon, France. ²⁰Médecine Intensive-Réanimation, APHP, Hôpital Louis Mourier, Université de Paris, Colombes, France.

Received: 10 February 2022 Accepted: 3 March 2022

Published online: 04 April 2022

Acknowledgements

We thank patients and their surrogates. We thank all medical and nurses teams from all study sites.

Authors' contributions

SG, FG, SB CV, RP, JPQ and DD were responsible for the design, analyzing and writing the manuscript. LML, BP, AB, GC, FM, JB, J-PR were responsible for recruitment and clinical care of the patients. All authors read and approved the final manuscript.

Funding

AKIKI trial was supported by a grant from the Programme Hospitalier de Recherche Clinique National, 2012 (AOM12456), funded by the French Ministry of Health. IDEAL-ICU was Supported by a grant (A00519-34) from the Programme Hospitalier de Recherche Clinique National, funded by the French Ministry of Health.

Availability of data and materials

The datasets used during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The AKIKI trial and the IDEAL-ICU trial received approval for all participating centers from competent French legal authority (Comité de Protection des Personnes d'Ile de France VI, ID RCB 2013-A00765-40, NCT01932190 for AKIKI and Comité de Protection des Personnes Est I ID RCB 2012-A00519-34 for IDEAL-ICU), and consent of patient or relatives was obtained before inclusion (except in emergencies where deferred consent was allowed by the Institutional Review Board).

Consent for publication

Patients or their surrogates were informed both verbally and with a written document by the investigators. They could refuse to participate at any time, and their decision was recorded in patient files.

Competing interests

None of the authors have competing interests in the manuscript.

Author details

¹Département de Réanimation Médico-Chirurgicale, APHP Hôpital Avicenne, Bobigny, France. ²Health Care Simulation Center, UFR SMBH, Université Sorbonne Paris Nord, Bobigny, France. ³Present Address: Common and Rare

References

1. Hoste EAJ, Bagshaw SM, Bellomo R, Cely CM, Colman R, Cruz DN, et al. Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study. *Intensive Care Med.* 2015;41(8):1411–23.
2. Ronco C, Bellomo R, Kellum JA. Acute kidney injury. *Lancet.* 2019;394(10212):1949–64.
3. Bagshaw SM, Darmon M, Ostermann M, Finkelstein FO, Wald R, Tolwani AJ, et al. Current state of the art for renal replacement therapy in critically ill patients with acute kidney injury. *Intensive Care Med.* 2017;43(6):841–54.
4. Truché A-S, Darmon M, Baily S, Clech C, Dupuis C, Misset B, et al. Continuous renal replacement therapy versus intermittent hemodialysis in intensive care patients: impact on mortality and renal recovery. *Intensive Care Med.* 2016;42(9):1408–17.
5. Schneider AG, Bellomo R, Bagshaw SM, Glassford NJ, Lo S, Jun M, et al. Choice of renal replacement therapy modality and dialysis dependence after acute kidney injury: a systematic review and meta-analysis. *Intensive Care Med.* 2013;39(6):987–97.
6. Legrand M, Darmon M, Joannidis M, Payen D. Management of renal replacement therapy in ICU patients: an international survey. *Intensive Care Med.* 2013;39(1):101–8.
7. Goldfarb DS, Bernstein JA, Zhdanova O, Hammer E, Block CA, Caplin NJ, et al. Impending shortages of kidney replacement therapy for COVID-19 patients. *Clin J Am Soc Nephrol.* 2020;15:880–2.
8. Wald R, Bagshaw SM. COVID-19-associated acute kidney injury: learning from the first wave. *J Am Soc Nephrol.* 2021;32(1):4–6.
9. Schortgen F, Soubrier N, Delclaux C, Thuong M, Girou E, Brun-Buisson C, et al. Hemodynamic tolerance of intermittent hemodialysis in critically ill patients: usefulness of practice guidelines. *Am J Respir Crit Care Med.* 2000;162(1):197–202.
10. Zhang Z, Spieth PM, Chiumello D, Goyal H, Torres A, Laffey JG, et al. Declining mortality in patients with acute respiratory distress syndrome: an analysis of the acute respiratory distress syndrome network trials. *Crit Care Med.* 2019;47(3):315–23.
11. Zimmerman JE, Kramer AA, Knaus WA. Changes in hospital mortality for United States intensive care unit admissions from 1988 to 2012. *Crit Care.* 2013;17(2):R81.
12. Vinsonneau C, Camus C, Combes A, Costa de Beauregard MA, Klouche K, Boulain T, et al. Continuous venovenous haemodiafiltration versus intermittent haemodialysis for acute renal failure in patients with multiple-organ dysfunction syndrome: a multicentre randomised trial. *Lancet.* 2006;368(9533):379–85.

13. Gaudry S, Hajage D, Schortgen F, Martin-Lefevre L, Pons B, Boulet E, et al. Initiation strategies for renal-replacement therapy in the intensive care unit. *N Engl J Med.* 2016;375(2):122–33.
14. STARRT-AKI Investigators, Canadian Critical Care Trials Group, Australian and New Zealand Intensive Care Society Clinical Trials Group, United Kingdom Critical Care Research Group, Canadian Nephrology Trials Network, Irish Critical Care Trials Group, et al. Timing of initiation of renal-replacement therapy in acute kidney injury. *N Engl J Med.* 2020;383(3):240–51.
15. Barbar SD, Clerc-Jehl R, Bourredjem A, Hernu R, Montini F, Bruyère R, et al. Timing of renal-replacement therapy in patients with acute kidney injury and sepsis. *N Engl J Med.* 2018;379(15):1431–42.
16. Vinsonneau C, Allain-Launay E, Blayau C, Darmon M, Ducheyron D, Gaillot T, et al. Renal replacement therapy in adult and pediatric intensive care: recommendations by an expert panel from the French Intensive Care Society (SRLF) with the French Society of Anesthesia Intensive Care (SFAR) French Group for Pediatric Intensive Care Emergencies (GFRUP) the French Dialysis Society (SFD). *Ann Intensive Care.* 2015;5(1):58.
17. ACUTE KIDNEY INJURY | KDIGO [Internet]. [cité 4 mars 2014]. Disponible sur: <http://kdigo.org/home/guidelines/acute-kidney-injury/>.
18. Brochard L, Abroug F, Brenner M, Broccard AF, Danner RL, Ferrer M, et al. An Official ATS/ERS/ESICM/SCCM/SRLF statement: prevention and management of acute renal failure in the ICU patient: an international consensus conference in intensive care medicine. *Am J Respir Crit Care Med.* 2010;181(10):1128–55.
19. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Stat Med.* 2004;23(19):2937–60.
20. Austin PC, Stuart EA. Estimating the effect of treatment on binary outcomes using full matching on the propensity score. *Stat Methods Med Res.* 2017;26(6):2505–25.
21. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med.* 2009;28(25):3083–107.
22. Franklin JM, Rassen JA, Ackermann D, Bartels DB, Schneeweiss S. Metrics for covariate balance in cohort studies of causal effects. *Stat Med.* 2014;33(10):1685–99.
23. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med.* 2015;34(28):3661–79.
24. Conner SC, Sullivan LM, Benjamin EJ, LaValley MP, Galea S, Trinquart L. Adjusted restricted mean survival times in observational studies. *Stat Med.* 2019;38(20):3832–60.
25. Li F, Thomas LE, Li F. Addressing extreme propensity scores via the overlap weights. *Am J Epidemiol.* 2019;188(1):250–7.
26. Gaudry S, Hajage D, Martin-Lefevre L, Louis G, Moschietto S, Titeca-Beauport D, et al. The Artificial Kidney Initiation in Kidney Injury 2 (AKIKI2): study protocol for a randomized controlled trial. *Trials.* 2019;20:726.
27. Lins RL, Elseviers MM, Van der Niepen P, Hoste E, Malbrain ML, Damas P, et al. Intermittent versus continuous renal replacement therapy for acute kidney injury patients admitted to the intensive care unit: results of a randomized clinical trial. *Nephrol Dial Transplant.* 2009;24(2):512–8.
28. Rabindranath K, Adams J, Macleod AM, Muirhead N. Intermittent versus continuous renal replacement therapy for acute renal failure in adults. *Cochrane Database Syst Rev.* 2007;3:CD003773.
29. Mehta RL, McDonald B, Gabai FB, Pahl M, Pascual MT, Farkas A, et al. A randomized clinical trial of continuous versus intermittent dialysis for acute renal failure. *Kidney Int.* 2001;60(3):1154–63.
30. Wang Y, Gallagher M, Li Q, Lo S, Cass A, Finfer S, et al. Renal replacement therapy intensity for acute kidney injury and recovery to dialysis independence: a systematic review and individual patient data meta-analysis. *Nephrol Dial Transplant.* 2018;33(6):1017–24.
31. Messika J, Gaudry S, Tubach F, Guillo S, Dreyfuss D, Hajage D, et al. Under-reporting of end-of-life decisions in critical care trials: a call to modify CONSORT statement. *Am J Respir Crit Care Med.* 2017;197:263–6.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Supplementary Appendix

Continuous renal replacement therapy versus intermittent haemodialysis as first modality for renal replacement therapy in severe acute kidney injury in intensive care unit

Table of Contents

Tables.....	2
Table S1. Balance of prognostic score and centre effects before and after inverse probability of treatment weighting.	2
Table S2. Baseline characteristics of patients in the two groups after overlap weighting.	3
Figures	4
Figure S1. Probability of survival in the overlap weighted sample.....	4
Figure S2. Bubble plot evaluating the association between centre-specific random effect for treatment allocation and centre-specific random effect for prognosis.	5
Figure S3. Propensity score distribution in the original (A) and inverse probability of treatment weighted (B) samples.	6
Figure S4. Propensity score (including centres as random effects) distribution in the original (A) and overlap weighted (B) samples.	6
Figure S5. Fraction of missing data for each variable in the original sample.	7
Figure S6. Standardised mean difference for each potential confounder in the unweighted, IPTW and OW samples.	8

Tables

Table S1. Balance of prognostic score and centre effects before and after inverse probability of treatment weighting.

Mixed effect Cox proportional hazard regression was used to derive a multivariable prediction model for 60-days survival (primary outcome). Centres were used as the covariate for random effects while all confounding variables (see main text) were used for fixed effects. We then allocated each patient their prognostic score and corresponding centre-specific random effects (best linear unbiased prediction, BLUP). In the unweighted and inverse probability of treatment weighting (IPTW) pseudo populations, we assessed the balance of the prognostic score and centre effects between the two groups.

	CRRT group	IHD group	SMD (%)	P value
Unweighted sample				
Number of patients	269	274		
Prognostic score, mean (SD)	2.71 (0.59)	2.68 (0.62)	4.8	0.56
Centre effects, mean (SD)	0.02 (0.14)	0.01 (0.14)	5.2	0.41
IPTW sample				
Number of patients (sum of weights)	268.2	271.9		
Prognostic score, mean (SD)	2.69 (0.59)	2.70 (0.61)	1.1	0.91
Centre effects, mean (SD)	0.01 (0.14)	0.01 (0.14)	0.1	0.99

CCRT=Continuous Renal Replacement Therapy.

IHD=Intermittent Haemodialysis.

SMD=Standardised Mean Difference, expressed as a percentage.

IPTW=Inverse Probability of Treatment Weighting.

Table S2. Baseline characteristics of patients in the two groups after overlap weighting.
Numbers of patients reported are the sum of stabilised weights.

Characteristic	CRRT group (n=69.4)	IHD group (n=70.2)	SMD (%)	P value
Age, years	66.9 (12.9)	66.8 (12.9)	0.5	0.96
Female sex	23.9 (34.5%)	24.3 (34.6%)	0.3	0.98
Serum creatinine, µmol/L	86.7 (29.0)	86.3 (28.8)	1.5	0.89
Coexisting conditions				
Chronic kidney disease	6.5 (9.3%)	7.7 (11.0%)	5.3	0.66
Hypertension	38.2 (55.0%)	38.2 (54.4%)	1.1	0.92
Diabetes mellitus	15.1 (21.7%)	15.2 (21.6%)	0.1	0.99
Congestive heart failure	6.3 (9.1%)	6.4 (9.1%)	<0.001	>0.99
Cirrhosis	7.5 (10.8%)	7.7 (11.0%)	0.5	0.96
Respiratory disease	8.2 (11.9%)	8.2 (11.6%)	0.7	0.95
Cancer	11.5 (16.6%)	11.7 (16.6%)	0.2	0.99
AIDS	0.8 (1.1%)	0.8 (1.1%)	0.3	0.98
Immunosuppressive drugs	3.0 (4.3%)	3.2 (4.6%)	1.5	0.89
Organ transplantation	0.4 (0.6%)	0.4 (0.6%)	0.1	>0.99
SOFA score at inclusion (0 to 24)	11.5 (3.2)	11.5 (3.1)	2.1	0.85
Renal SOFA (1 to 5)	3.7 (1.1)	3.7 (1.1)	2.4	0.83
Haemodynamic SOFA (1 to 5)	4.4 (1.3)	4.4 (1.3)	2.1	0.86
Liver SOFA (1 to 5)	1.8 (1.1)	1.8 (1.1)	4.0	0.73
Neurologic SOFA (1 to 5)	2.4 (1.6)	2.4 (1.5)	0.5	0.96
Coagulation SOFA (1 to 5)	3.2 (1.7)	3.1 (1.6)	2.9	0.80
Body weight, kg	86.7 (29.0)	86.3 (28.8)	1.5	0.89
Laboratory values				
Serum creatinine, µmol/L	282.0 (124.4)	289.4 (119.3)	6.1	0.59
Serum urea, mmol/L	19.9 (9.4)	19.7 (8.9)	1.7	0.88
Serum potassium, mmol/L	4.41 (0.78)	4.41 (0.77)	0.1	0.99
Arterial blood pH	7.30 (0.09)	7.30 (0.09)	<0.1	>0.99

Data are n (%) or mean (SD).

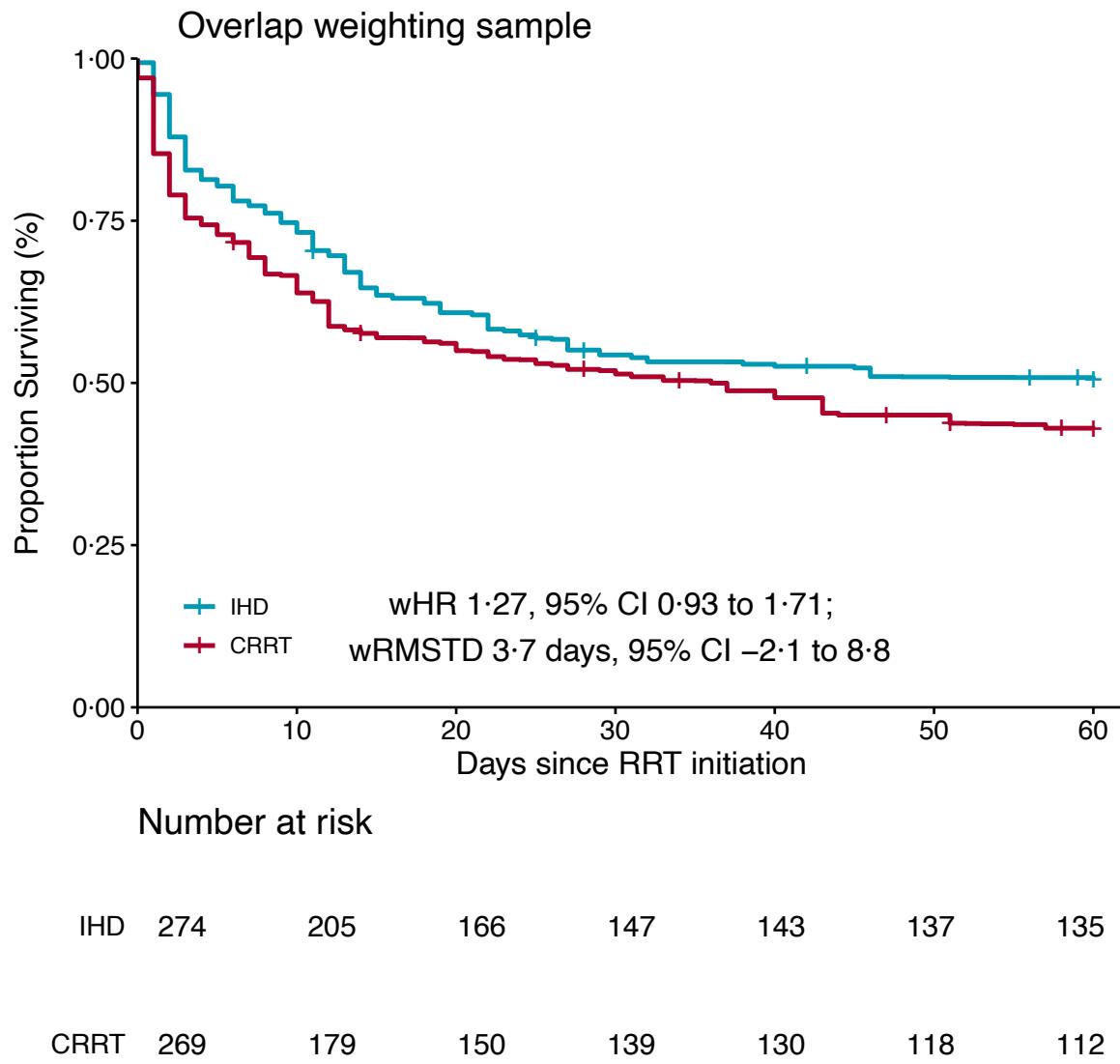
SMD=Standardised Mean Difference, expressed as a percentage.

SOFA score=Sequential Organ Failure Assessment score.

Figures

Figure S1. Probability of survival in the overlap weighted sample.

The weighted Kaplan-Meier death rate at day 60 was 57·1% in the CRRT group and 49·5% in the IHD group (overlap weighted ARD 7·6%, 95% CI -2·4% to 18·0%).



wHR=weighted Hazard Ratio.

wRMSTD=weighted Restricted Mean Survival Time.

CRRT=Continuous Renal Replacement Therapy.

IHD=Intermittent Haemodialysis.

Figure S2. Bubble plot evaluating the association between centre-specific random effect for treatment allocation and centre-specific random effect for prognosis.

Each centre was allocated its centre-specific random effect on treatment allocation and its centre-specific random effect on prognosis from the propensity score and prognostic score models respectively. The propensity score model was fitted through mixed-effects logistic regression with all confounding variables as fixed effects (see main text) and a random centre effect. The prognostic model predicted 60-days survival and was fitted through mixed effect Cox proportional hazard regression with treatment allocation and the same confounding variables as fixed effects (see main text) as well as a random centre effect. We found no associational pattern between centre-specific random effect for treatment allocation and centre-specific random effect for prognosis, thus suggesting that centres were not confounding variables and omission of centres from the propensity score model did not result in a violation of the ignorability assumption.

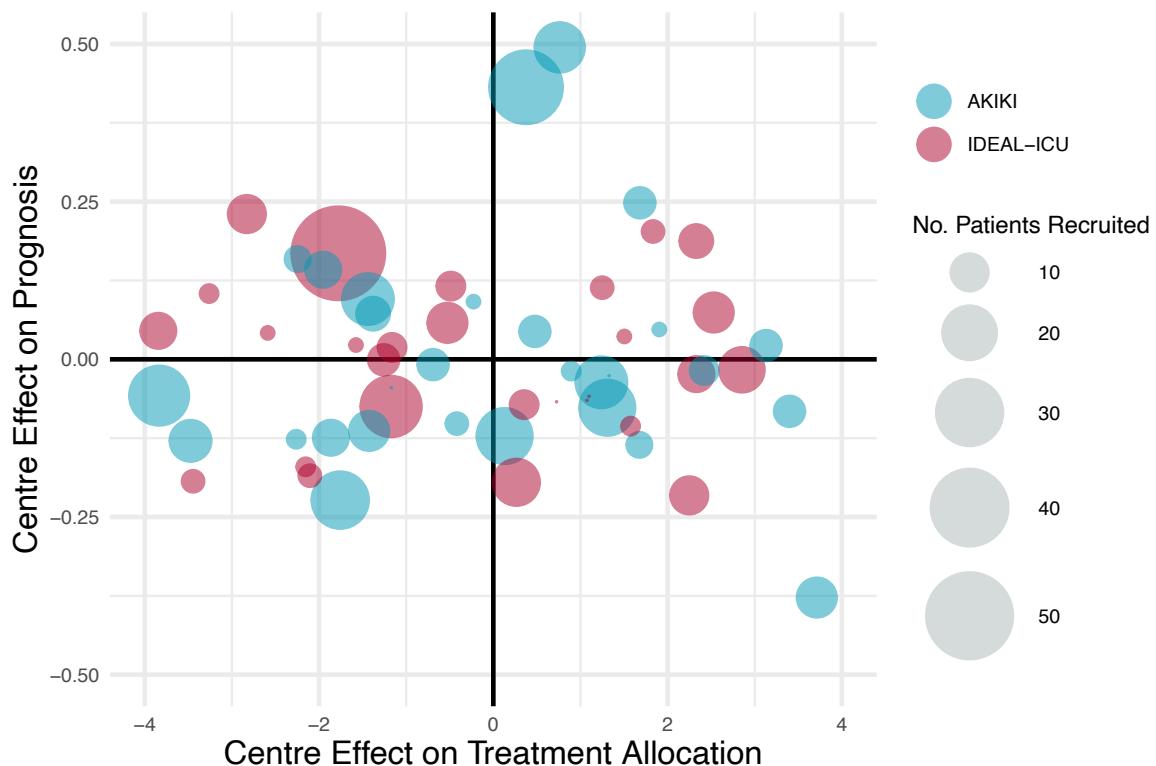


Figure S3. Propensity score distribution in the original (A) and inverse probability of treatment weighted (B) samples.

The overlap of the distributions allows to assess the positivity assumption.

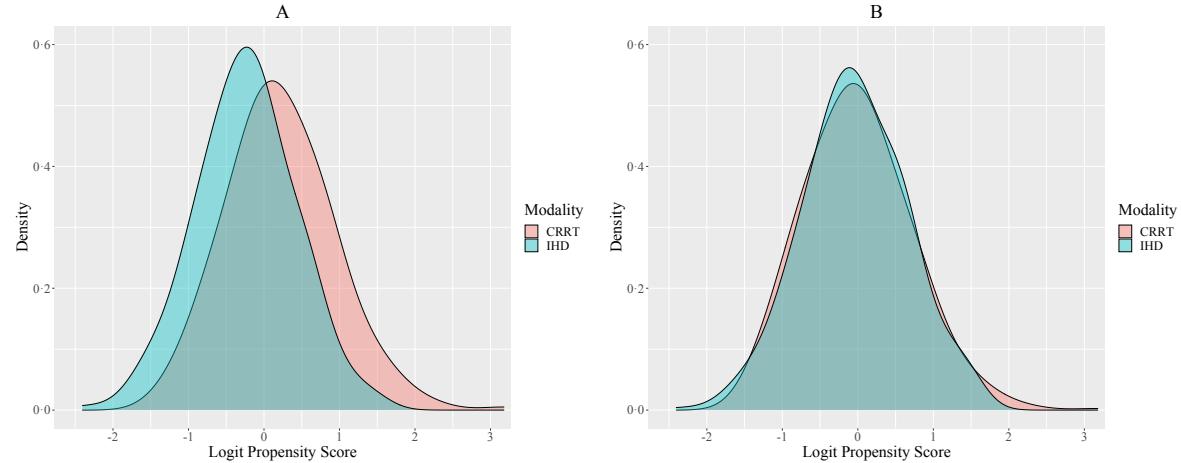
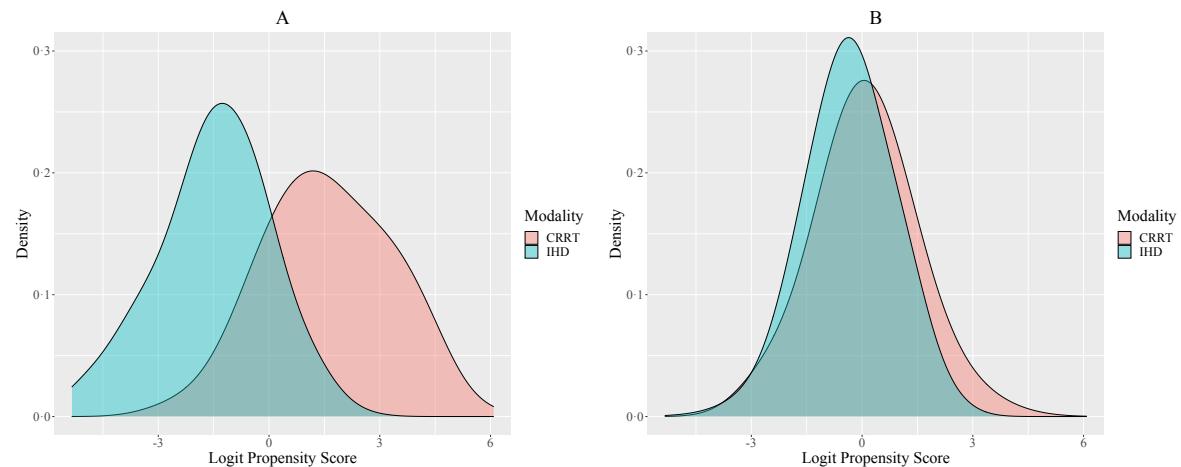
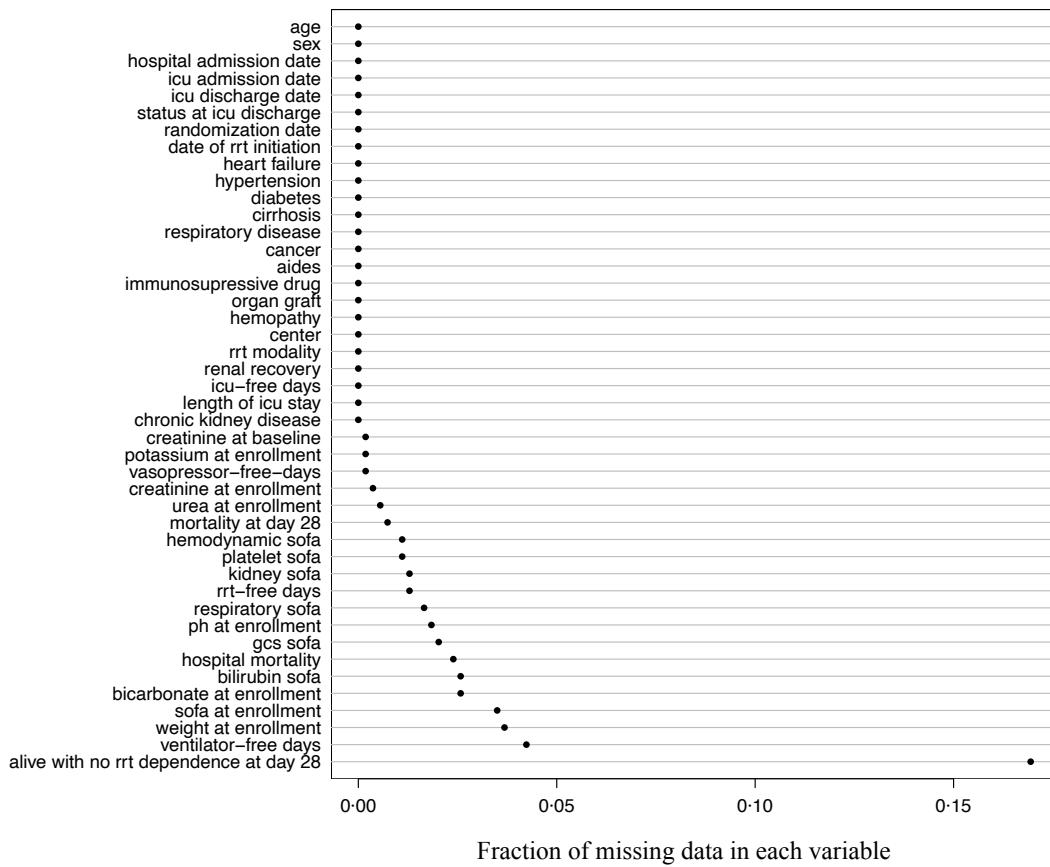


Figure S4. Propensity score (including centres as random effects) distribution in the original (A) and overlap weighted (B) samples.

The overlap of the distributions allows to assess the positivity assumption.

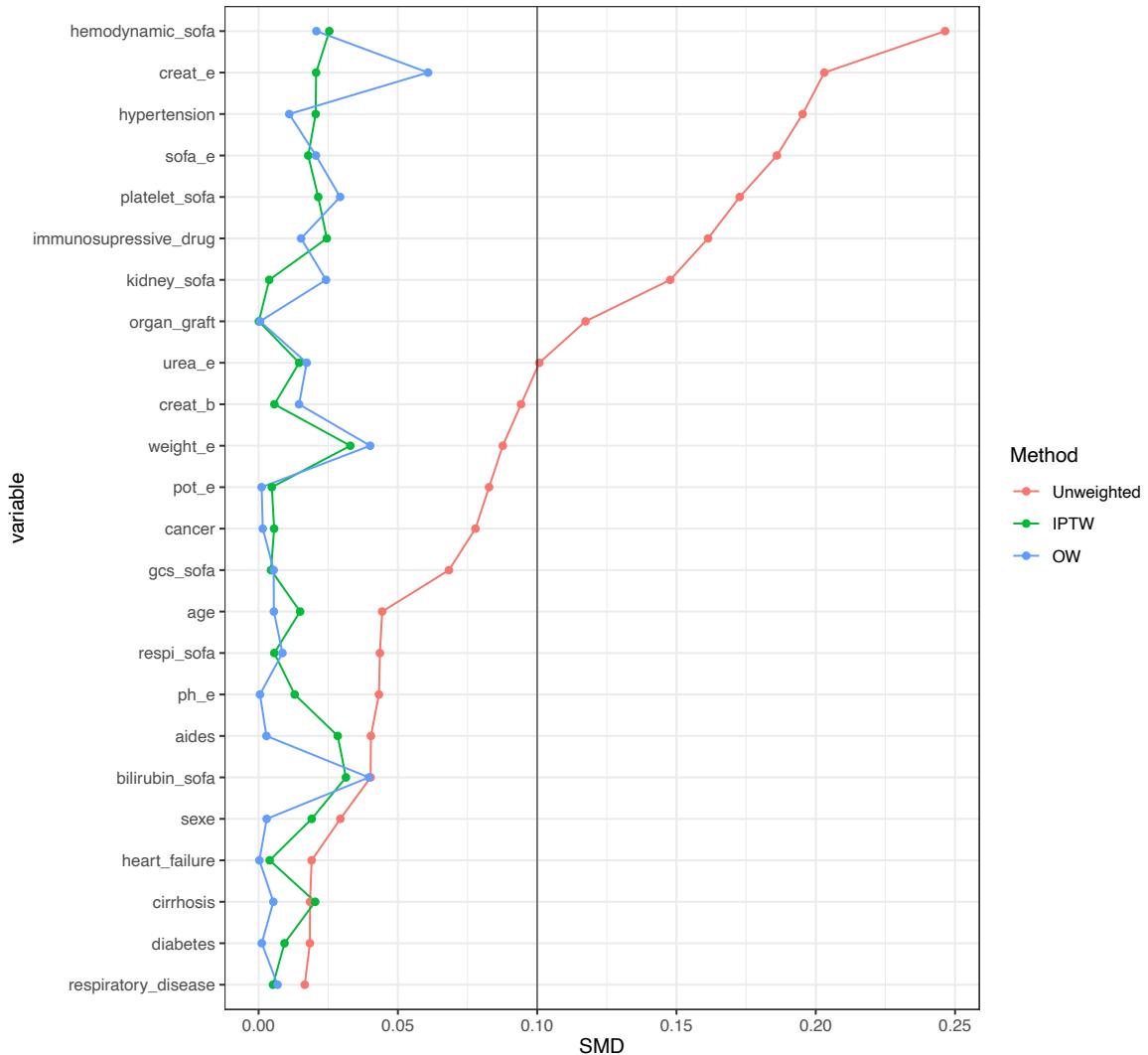


CCRT=Continuous Renal Replacement Therapy.
IHD=Intermittent Haemodialysis.

Figure S5. Fraction of missing data for each variable in the original sample.

RRT=Renal Replacement Therapy.

Figure S6. Standardised mean difference for each potential confounder in the unweighted, IPTW and OW samples.



IPTW=Inverse Probability of Treatment Weighting.

OW=Overlap Weighting.

SMD=Standardised Mean Difference.

Chapter 4

Personalization of renal replacement therapy initiation: a secondary analysis of the AKIKI and IDEAL-ICU trials

It is wise to take admissions of uncertainty seriously, but declarations of high confidence mainly tell you that an individual has constructed a coherent story in his mind, not necessarily that the story is true.

Daniel Kahneman, *Thinking fast and slow*

La musique associée à ce chapitre est ‘Happiness Does Not Wait’ de Ólafur Arnalds.

RESEARCH

Open Access



Personalization of renal replacement therapy initiation: a secondary analysis of the AKIKI and IDEAL-ICU trials

François Grolleau^{1*}, Raphaël Porcher¹, Saber Barbar², David Hajage³, Abderrahmane Bourredjem⁴, Jean-Pierre Quenot^{5†}, Didier Dreyfuss^{6†} and Stéphane Gaudry^{7†}

Abstract

Background: Trials comparing early and delayed strategies of renal replacement therapy in patients with severe acute kidney injury may have missed differences in survival as a result of mixing together patients at heterogeneous levels of risks. Our aim was to evaluate the heterogeneity of treatment effect on 60-day mortality from an early vs a delayed strategy across levels of risk for renal replacement therapy initiation under a delayed strategy.

Methods: We used data from the AKIKI, and IDEAL-ICU randomized controlled trials to develop a multivariable logistic regression model for renal replacement therapy initiation within 48 h after allocation to a delayed strategy. We then used an interaction with spline terms in a Cox model to estimate treatment effects across the predicted risks of RRT initiation.

Results: We analyzed data from 1107 patients (619 and 488 in the AKIKI and IDEAL-ICU trial respectively). In the pooled sample, we found evidence for heterogeneous treatment effects ($P=0.023$). Patients at an intermediate-high risk of renal replacement therapy initiation within 48 h may have benefited from an early strategy (absolute risk difference, -14% ; 95% confidence interval, -27% to -1%). For other patients, we found no evidence of benefit from an early strategy of renal replacement therapy initiation but a trend for harm (absolute risk difference, 8% ; 95% confidence interval, -5% to 21% in patients at intermediate-low risk).

Conclusions: We have identified a clinically sound heterogeneity of treatment effect of an early vs a delayed strategy of renal replacement therapy initiation that may reflect varying degrees of kidney demand-capacity mismatch.

Keywords: Acute kidney injury, Renal replacement therapy, Heterogeneity of treatment effect, Personalized medicine

Introduction

Acute kidney injury (AKI) affects approximately half of critically ill patients and is associated with high mortality and long-term sequelae [1]. Since its introduction in intensive care units (ICU) in the 1960s [2], renal replacement therapy (RRT) has proved to be a major breakthrough for the treatment of AKI, saving countless lives. However, the optimal timing for RRT initiation in patients with severe AKI has been controversial. This is illustrated by opposite hypotheses regarding which of an

*Correspondence: francois.grolleau@aphp.fr

†Jean-Pierre Quenot, Didier Dreyfuss, and Stéphane Gaudry have contributed equally as senior authors.

¹Centre of Research in Epidemiology and Statistics (CRESS), Université de Paris, French Institute of Health and Medical Research (INSERM U1153), French National Research Institute for Agriculture, Food, and Environment (INRAE), Paris, France

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

early or a delayed RRT initiation strategy would be superior to the other in the sample size calculation of recent multicenter randomized controlled trials (RCTs) [3–5]. Moreover, three trials—the largest on the topic—did not demonstrate any survival benefit from either strategy over the other. Likewise, recent meta-analyses concluded that, in the absence of life-threatening condition, the timing of RRT initiation did not affect survival [6, 7].

One suggested reason for the lack of conclusive findings lies in the heterogeneous baseline characteristics of patients included in these trials [8]. Meaningful differences in survival may have been missed as a result of mixing together patients with potential benefit and potential harm from a given initiation strategy. For instance, one may hypothesize that an early RRT initiation strategy is harmful to the patients who would never start it under a delayed strategy. In fact, when a delayed strategy is implemented, we observed that between a third and half of the patients never met the criteria mandating RRT initiation. Conversely, experts have speculated that the patients who would be susceptible to benefit from an early initiation strategy are those who would initiate RRT within 48 h under a delayed strategy [9].

Patient management further tailored to individual's characteristics is much anticipated in critical care medicine [10] and AKI [11]. In that respect, the conventional subgroup analyses performed “one variable at a time” fail to convey meaningful results as they cannot fully capture all the relevant heterogeneity in patient characteristics [12]. Conversely, approaches using multivariable models have the potential to address the challenge of heterogeneous treatment effects (HTE) [13].

The concept of kidney demand-capacity mismatch may be useful to the personalization of RRT initiation, but it has not been evaluated on robust clinical data [14]. In this study, we wished to test if estimating the degree of demand-capacity mismatch could guide RRT initiation strategies. We hypothesized that an early RRT initiation strategy is unnecessary or harmful to the patients at low risk of RRT initiation under a delayed strategy; and beneficial to the patients at a higher risk. Accordingly, we used data from two large multicenter RCTs on RRT timing to develop a risk prediction model for RRT initiation within 48 h after allocation to a delayed strategy and then estimated treatment effects within levels of predicted risks.

Methods

Ethical approval and research transparency

The AKIKI and the IDEAL-ICU trials received approval for all participating centers from competent French legal authority (Comité de Protection des Personnes d'Île de France VI, ID RCB 2013-A00765-40, NCT01932190 for AKIKI and Comité de Protection des Personnes Est I ID

RCB 2012-A00519-34 for IDEAL-ICU), and consent of patient or relatives was obtained before inclusion (except in emergencies where deferred consent was allowed by the Institutional Review Board). We transparently reported our analysis following the PATH [15] and TRIP-POD [16] statements.

Source of data

The study sample included participants from the AKIKI and IDEAL-ICU, two multicenter RCTs conducted in France. The AKIKI trial was conducted at 31 ICUs from September 2013 through January 2016 and recruited 619 patients with severe AKI who required mechanical ventilation, catecholamine infusion, or both (the vast majority with septic shock). The IDEAL-ICU trial recruited in 29 ICUs from July 2012 through October 2016 and included 488 patients with severe AKI and septic shock. Both trials randomly assigned (1:1) patients to either an early or a delayed strategy of RRT initiation. None of these trials showed a significant difference between the two strategies on 60-day mortality. The delayed strategy averted the need for RRT in 49% and 38% of patients in the AKIKI and IDEAL-ICU trials, respectively.

Outcomes

The primary outcome of this study was death at day 60. Secondary outcomes included mean differences in number of days free of RRT, mechanical ventilation and intensive care at 28 days [17] across the same levels of risk.

Prediction model development

We developed a risk prediction model for RRT initiation within 48 h after allocation to a delayed strategy. The derivation sample consisted of the 550 patients allocated to the delayed arms of the AKIKI ($n=308$) and IDEAL-ICU ($n=242$) trials (Table 1). We fit a logistic regression model, using predefined 14 predictors to predict the occurrence of RRT initiation within 48 h after the start of the delayed strategy. Candidate predictor variables were taken from the pre-randomization eligibility screening or clinical examination prior to randomization to the delayed strategy of RRT initiation and included age (years), gender (male vs female), potassium level (mmol/L), blood urea nitrogen level (mmol/L), pH (unitless), the ratio of creatinine at enrollment over creatinine at baseline (unitless), urine output (< 200 ml/day vs ≥ 200 ml/day, as was already categorized in the data), SOFA score at enrollment (unitless), weight (kg), heart failure (yes vs no), hypertension (yes vs no), diabetes mellitus (yes vs no), cirrhosis (yes vs no), non-corticosteroid immunosuppressive drug (yes vs no). Missing data were handled through multiple imputations by chained equations using outcomes as well as all aforementioned

Table 1 Characteristics of the patients at randomization

Characteristic	Delayed strategy <i>n</i> =550	Early strategy <i>n</i> =557
<i>Study</i>		
AKIKI	308 (56.0)	311 (55.8)
IDEAL-ICU	242 (44.0)	246 (44.2)
Age—year	67.7 (13.2)	66.5 (13.3)
Weight—kg	81.6 (22.2)	82.4 (22.2)
Male sex	352 (64.0)	351 (63.0)
<i>Pre-existing conditions</i>		
Heart failure	52 (9.5)	44 (7.9)
Hypertension	304 (55.3)	306 (54.9)
Diabetes mellitus	92 (16.7)	112 (20.1)
Cirrhosis	54 (9.8)	54 (9.7)
Respiratory disease	54 (9.8)	62 (11.1)
Cancer	100 (18.2)	91 (16.3)
Hemopathy	27 (4.9)	34 (6.1)
AIDS	2 (0.4)	5 (0.9)
Non-corticosteroid immunosuppressive drug	36 (6.5)	32 (5.7)
Organ graft	17 (3.1)	5 (0.9)
<i>Severity at enrollment</i>		
SOFA score (0 to 24)	11.5 (3.1)	11.4 (3.2)
Respiratory SOFA (0 to 4)	2.1 (1.1)	1.9 (1.1)
Hemodynamic SOFA (0 to 4)	3.5 (1.2)	3.5 (1.2)
Liver SOFA (0 to 4)	0.8 (1.1)	0.8 (1.1)
Coagulation SOFA (0 to 4)	2.1 (1.6)	2.2 (1.6)
Neurologic SOFA (0 to 4)	1.3 (1.5)	1.2 (1.5)
<i>Laboratory values</i>		
Baseline creatinine (IQR), $\mu\text{mol/L}^*$	88 (71–97)	84 (71–97)
Creatinine at enrollment (IQR), $\mu\text{mol/L}$	268 (211–343)	267 (198–352)
Blood urea nitrogen at enrollment (IQR), mmol/L	19 (14–26)	19 (13–26)
Potassium at enrollment, mmol/L	4.4 (0.8)	4.4 (0.8)
Bicarbonate at enrollment, mmol/L	18 (5)	18 (5)
Arterial blood pH at enrollment	7.30 (0.10)	7.30 (0.10)

All characteristics reported in the table were determined at inclusion in the AKIKI or IDEAL-ICU trial, before initiation of renal replacement therapy

Data are mean (SD), median (IQR) or *n* (%). AIDS = Acquired Immunodeficiency Syndrome. IQR = Interquartile range. SOFA score = Sequential Organ Failure Assessment score

To convert the values for creatinine to milligrams per deciliter, divide by 88.4

*The serum creatinine concentration before ICU admission was either determined with the use of values measured in the 12 months preceding the ICU stay or was estimated

predictors in the imputation models [18]. Five independent imputed data sets were generated and analyzed separately. The nonlinearity of each continuous variable was assessed through penalized spline regression. All continuous variables appeared roughly linearly associated with the logit of the outcome probability; hence, no non-linear terms were used.

Two strategies were used to select predictors with the imputed data [19]. First, we used Wald tests for the pooled regression coefficients to simplify the model with

a backward selection procedure, with *P*-value cut-offs mimicking the use of Akaike information criterion (e.g., a cut-off of 0.157 for variables with 1 *df*). We then used a conventional backward elimination procedure in each imputed data set and retained the model comprising the variables selected in most imputed data sets. Both strategies selected the same variables. Two-by-two interactions between each of the selected variables were then examined using Wald tests for the pooled regression coefficients. No higher-order interactions were considered.

Regression coefficients estimates and their variances were then pooled across imputed data sets [20].

To evaluate the predictive ability of the model, we first calculated the apparent discrimination (c-statistic) and calibration (categorization by fifth of predicted risk) in the derivation sample. The c-statistic measures how well the model discriminates between the patients who initiated RRT within 48 h after allocation a delayed strategy and those who did not. The calibration curve, estimated using local regression [21]], contrasts observed vs predicted probabilities of events and evaluates the accuracy of the predictions. Internal validation of the model was performed by bootstrapping, which allows to correct regression coefficients and model performance for optimism [22]. The variable selection strategy was repeated in 200 bootstrap samples, and performance of models fit in each sample was evaluated in these samples and in the original sample. The differences between these two performances were averaged and taken as a measure of overoptimism. The c-statistic as well as the calibration intercept and slope were corrected for bias by subtracting measures of overoptimism to the apparent performance metrics.

Risk categorization

In the AKIKI ($n=619$), IDEAL-ICU ($n=488$) and pooled ($n=1107$) samples, we categorized patients by fifths of the risk predicted by our final model. In each fifth of risk, we compared early vs delayed strategy of RRT initiation on primary and secondary outcomes. To account for censoring, death at day 60 was calculated from the Kaplan–Meier estimator. As HTE are fundamentally a scale dependent concept [15], we evaluated treatment effects on the absolute risk difference and the hazard ratio scales. For each scale we computed a smooth curve of the treatment effect across levels of risks by using an interaction term between treatment arm and a two knots natural spline transform [23] of the predicted risk in a Cox model. We assessed the evidence for heterogeneous treatment effect by testing the null hypothesis that a Cox model using a linear interaction between treatment arm and the predicted risk fits data equally well as a Cox model using a similar interaction with a spline transform of the predicted risk [24]. Ninety-five percent confidence intervals (95% CI) were calculated by bootstrapping (1000 iterations). All analyses were performed using the R statistical software version 4.0.5 (The R Foundation). More precisely, we used the rms package for model building and internal validation, the survival package for survival analyses, the mgcv package for heterogeneous treatment effects assessment, the boot package for bootstrap, and the mice package for multiple imputation. For transparency and reproducibility, the computer code

used in this study is available as an Additional file 1 at the Journal's website.

Results

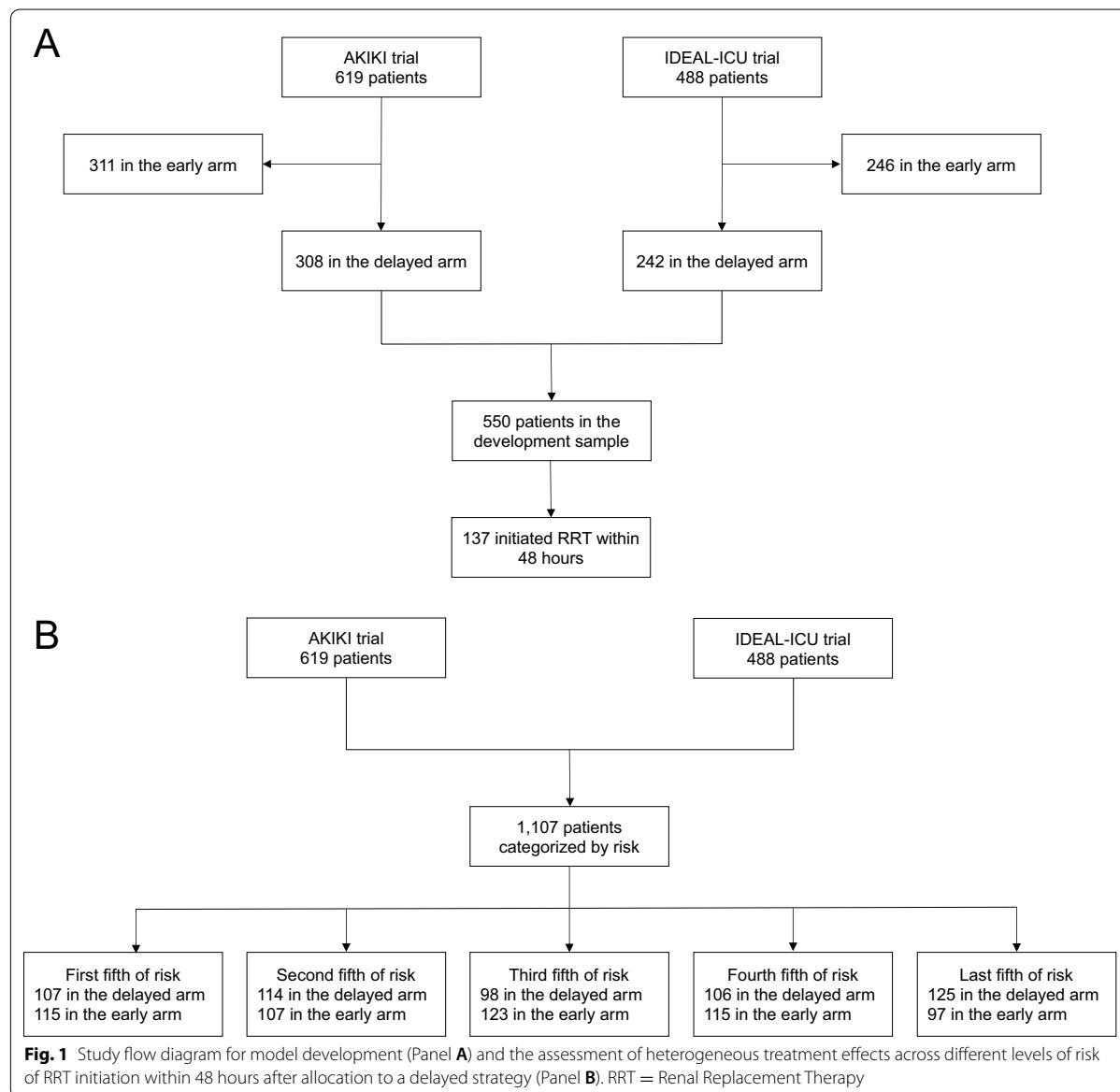
Prediction model for RRT initiation

Of the 550 patients included for model derivation (see Fig. 1, Panel A), 137 patients (25%) initiated RRT within 48 h after allocation to a delayed strategy (62 [20%] and 75 patients [31%] in AKIKI and IDEAL-ICU, respectively). 91% of patients had complete data for all candidate predictors (see Additional file 2: Figure S1); there were no missing data for the event of RRT initiation. The final prediction model included potassium, blood urea nitrogen, pH, non-corticosteroid immunosuppressive drug, SOFA and weight. No two-by-two interaction between variables was added as none showed statistical significance or seemed clinically informative.

The full and final models are presented in Table 2. The apparent and bias-corrected c-statistic were 0.73 (95% CI: 0.70 to 0.80) and 0.70 (95% CI: 0.67 to 0.77), respectively. The predictive performance of the final model was good, as measured by discrimination and calibration (Fig. 2).

Heterogeneity of treatment effect

For the pooled AKIKI and IDEAL-ICU samples ($n=1,107$ see Fig. 1, Panel B), baseline characteristics at randomization are presented in Table 1. In all fifth of risk predicted by our model, patients' characteristics appeared balanced between the randomization arms (see Additional file 2: Table S1). Patients' characteristics by fifth of risk predicted by our model are provided in the Additional file 2: Table S2. Heterogeneity of treatment effect is presented by fifth of risk in Fig. 3. There was no evidence of benefit from an early RRT initiation strategy for individuals within the lowest fifth of RRT initiation risk (absolute risk difference [ARD], 1%; 95% CI – 12% to 14%). However, patients in the fourth fifth of risk, may have benefited from an early strategy of RRT initiation (ARD, – 14%; 95% CI – 27% to – 1%). For patients with the highest risk (last fifth of risk), we found no evidence of benefit from an early initiation strategy (ARD, 7%; 95% CI – 6% to 20%). On both the absolute (i.e., ARD) and relative (i.e., event rate and hazard ratio) scales, the smooth curve suggested that an early RRT initiation strategy may be harmful in patients at an intermediate-low risk (second fifth of risk), while it may be beneficial in patients at an intermediate-high risk (fourth fifth of risk). This pattern was consistent in both the AKIKI and IDEAL-ICU trials when analyzed separately (see Additional file 2: Figure S2). Kaplan–Meier survival for each fifth of risk are given in Fig. 4 and in Additional file 2: Figure S3. No



difference in secondary outcomes was found between early and delayed RRT initiation strategy in any fifth of predicted risk (see Additional file 2: Figure S4).

An implementation of our model has been made available via a user-friendly web interface at <http://rrt-personalization.eu/>. With this web application, clinicians and researchers can obtain the predicted probability of RRT initiation within 48 h after allocation to a delayed strategy in patients with severe AKI. The individual treatment effect of an early vs delayed strategy is then computed and returned with 95% CIs.

Discussion

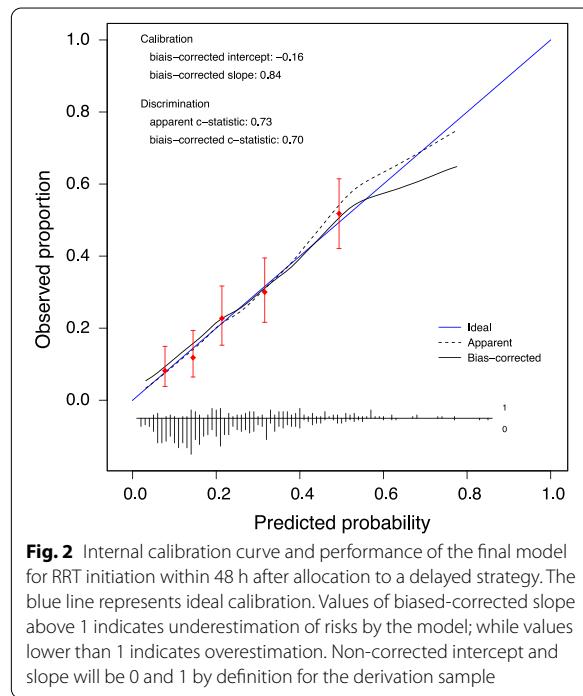
Summary of findings

In this study, we developed a prediction model for the initiation of RRT within 48 h after allocation to a delayed strategy in patients with severe AKI in the ICU. We subsequently used the predictions from this model to identify subgroups (i.e., fifths) of patients at similar risk. We then assessed if the treatment effect of an early vs delayed strategy of RRT initiation was heterogeneous between these subgroups.

Table 2 Univariable analysis, full and final models for RRT initiation within 48 h after allocation to a delayed strategy

Variable	Univariable analysis			Full model			Final model		
	OR	(95% CI)	P	OR	(95% CI)	P	OR	(95% CI)	P
Age (year)	0.993	0.979	1.008	0.361	1.004	0.986	1.022	0.694	—
Sex (male vs female)	1.257	0.834	1.894	0.276	1.188	0.755	1.868	0.456	—
Potassium at enrollment (mmol/L)	1.728	1.343	2.224	<.001	1.381	1.048	1.821	0.022	1.391
Creatinine at enrollment over creatinine at baseline (unitless)	1.181	1.058	1.318	0.003	1.068	0.931	1.225	0.350	—
Urine output (<200 ml/day vs >= 200 ml/day)	1.792	1.183	2.714	0.006	1.129	0.699	1.823	0.621	—
SOFA at enrollment (unitless)	1.151	1.079	1.227	<.001	1.123	1.041	1.210	0.003	1.139
Weight at enrollment (kg)	1.012	1.003	1.021	0.006	1.013	1.003	1.023	0.010	1.013
Heart failure (yes vs no)	0.792	0.395	1.588	0.512	0.845	0.399	1.788	0.659	—
Hypertension (yes vs no)	0.831	0.564	1.224	0.350	0.847	0.541	1.326	0.468	—
Diabetes mellitus (yes vs no)	1.153	0.695	1.913	0.582	0.985	0.558	1.741	0.960	—
Cirrhosis (yes vs no)	1.179	0.628	2.213	0.608	1.135	0.549	2.346	0.733	—
Non-corticosteroid immunosuppressive drug (yes vs no)	2.023	1.005	4.074	0.049	1.971	0.928	4.185	0.078	1.973
Blood urea nitrogen at enrollment (mmol/L)	1.029	1.009	1.050	0.004	1.022	0.998	1.047	0.076	1.026
pH at enrollment (OR for an increase of 0.01)	0.945	0.925	0.966	<.001	0.954	0.932	0.978	<.001	0.954
									<.001

The intercepts were 1.35×10^{12} and 2.51×10^{12} for the full and final models respectively



We stress that although causal understanding of model predictions is always inappropriate, in the case of the present HTE, this interpretation is proper as all variables included in our model were measured prior to randomization. In our main analysis, we found substantial HTE across levels of predicted risks. Except for the upper boundary (i.e., highest levels of risks), the directions of the HTE were aligned with our prespecified hypothesis.

From a clinical standpoint, the predicted risk from our model may be viewed as a proxy for the severity of kidney demand-capacity mismatch of the patients included in the trials. Through this lens, our results seem to indicate that for the most severe patients, an invasive strategy i.e., early RRT was unnecessary and/or harmful (ARD in the last fifth of predicted risk, 7%; 95% CI, - 6% to 20%). This seemed true also of mildly severe patients (ARD in the second fifth of predicted risk, 8%; 95% CI, - 5% to 21%). The only patients who seemed to have benefited from early RRT are those at a high but nonextreme risk (ARD in the fourth fifth of predicted risk, - 14%; 95% CI, - 27% to - 1%). An interpretation for these findings is that starting RRT early could harm the lesser severe patients because they often have no need for such invasive treatment. On the other hand, early RRT could be unnecessary to the most severe patients as their prognosis may outweigh potential benefits; or early RRT could even harm them through the destabilization of a weak equilibrium.

Hitherto, the concept of demand-capacity and personalization of RRT initiation did not rely on the analysis of robust clinical data. The 2021 Surviving Sepsis Campaign guidelines argues for a pragmatic approach: propose a wait-and-see strategy for all patients with severe AKI and no life-threatening complications in the intensive care unit [25].

Strength and limitations

We acknowledge that given large enough sample sizes, more advanced machine learning techniques could potentially yield a more precise estimation of HTEs. These techniques, often referred to as effect-modelling approaches, aim to estimate HTE through direct modelling of the treatment effect [26]. Of note, they are also vulnerable to misspecification and overfitting, and therefore require huge sample sizes [27]. In contrast, we chose to implement a risk-modelling approach and relied on the PATH guidelines for personalized medicine [15]. On the upside, this allowed us to evaluate a clinically sound, a priori-specified hypothesis [9]. Compared to black-box algorithms, we believe the transparency of our parametric modelling methodology offers researchers a window for interpretability.

Despite the good performance of our prediction model as evaluated on biased-corrected metrics, the absence of external validation for our prediction model is a limitation. However, in our methodology, the model predictions are merely a mean for a downstream purpose namely, the assessment of HTEs. A poorly performing model would have limited our ability to find evidence of HTE when treatment effects are in fact truly heterogeneous.

Last, in contrast with other instances where predictions from developed models cannot be readily calculated by clinicians or researchers, we have implemented a user-friendly web interface for our approach. We trust this will help further disseminate, replicate, or refine our findings. We purposely chose to emphasize uncertainty for the individualized treatment effects by providing all metrics along with their 95% CI. We believe that as decision tools have not been evaluated in controlled settings, clinical judgment should however prevail.

Implications for future research

Precision medicine is an active field of research with limited clinical applications so far [28]. Data-driven decision support tools have been made available in cardiology [29], while in critical care HTE were documented for crystalloid fluids [30] or ventilation strategies [31]. In fact, as negative trial findings are widespread, disentangling HTE were judged a research priority in critical care [32]. The identification of HTE may also inform the

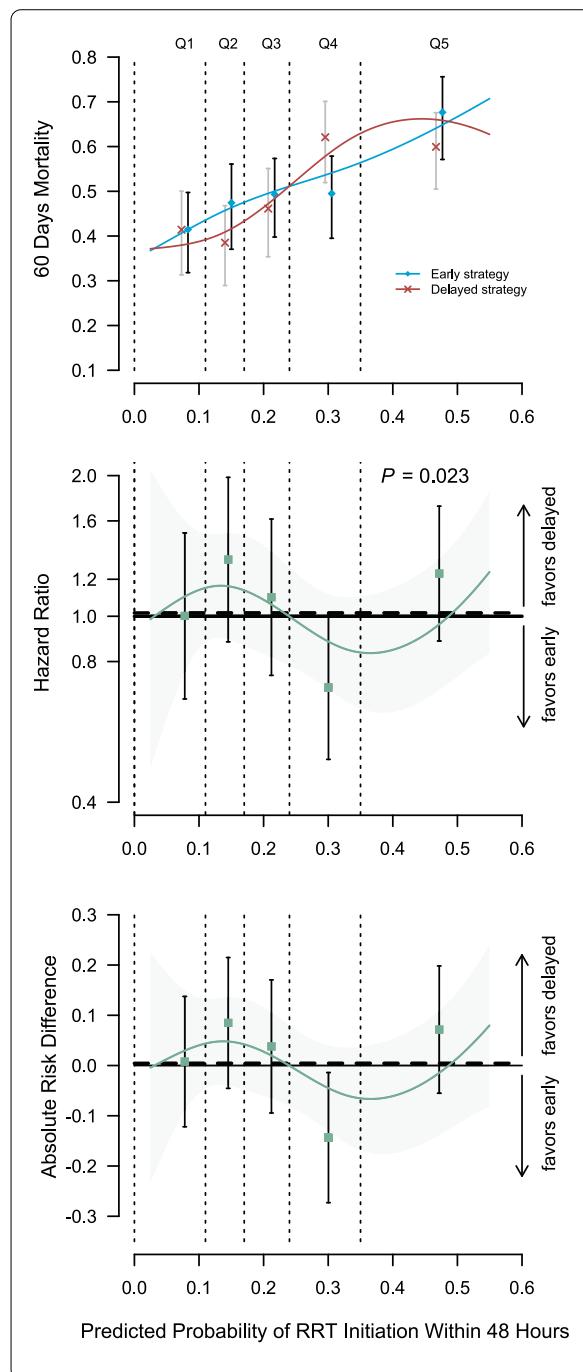


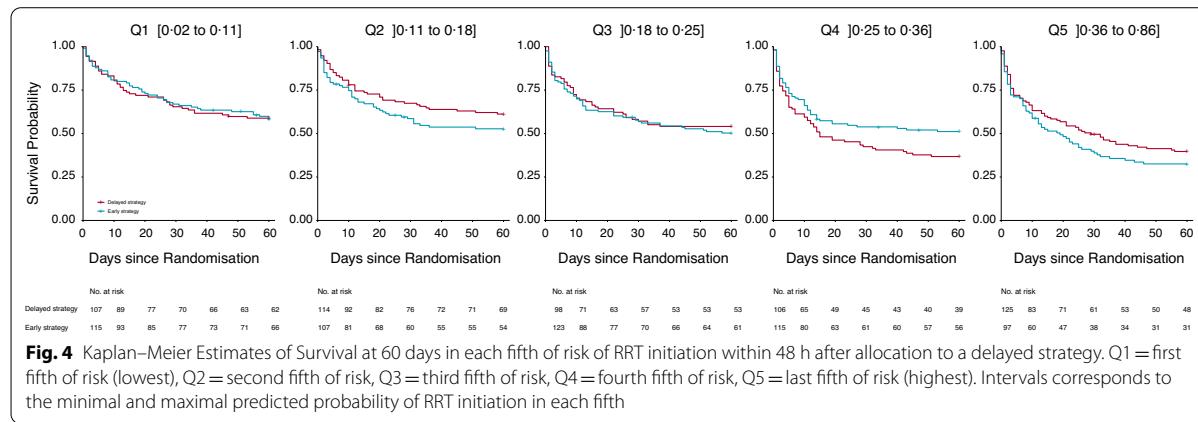
Fig. 3 Heterogeneity of treatment effect (early vs delayed strategy) across different levels of risk of RRT initiation within 48 h after allocation to a delayed strategy. This figure presents heterogeneous treatment effect of an early vs a delayed strategy of RRT initiation as a function of the baseline risk of RRT initiation within 48 h after allocation to a delayed strategy in the pooled AKIKI and IDEAL-ICU sample. The horizontal dashed lines indicate the average treatment effect. P value for a constant effect along the predicted risk (test of heterogeneity of the treatment effect). Q1 = first fifth of risk (lowest), Q2 = second fifth of risk, Q3 = third fifth of risk, Q4 = fourth fifth of risk, Q5 = last fifth of risk (highest)

We believe the risk-modelling methodology presented in our study is transportable to treatments as diverse as corticosteroids for sepsis [35], proton pump inhibitors for gastrointestinal bleeding prevention [36], or extracorporeal membrane oxygenation for acute respiratory distress syndrome [37].

As for RRT initiation strategies, our findings will require further replication using other data sources and methodologies. The way in which this can happen is twofold. First, as in the present study, researchers can consider the static case of an early vs delayed strategy of RRT initiation and use either other RCT data or observational data coupled with robust statistical methods. Second, researchers may also account for the fundamentally dynamic nature of the question. On the one hand, AKI staging systems inaccurately reflect the timing of the underlying pathology [38]; on the other hand definition of the criteria mandating RRT initiation under a delayed strategy ought to be refined [39, 40]. While the latter problem can be addressed with advanced causal inference techniques [41], the former can be tackled through cutting-edge pathophysiological studies. These two approaches are, in our view, complementary and we believe researchers should strive to dig from both ends.

In this secondary analysis of the AKIKI and IDEAL-ICU trials, we have provided proof-of-concept for the HTE of early vs delayed strategy across levels of baseline risk of RRT initiation within 48 h after a delayed strategy. Though consistent between the two trials, our results will require replication and refinement before they can be implemented in practice. We believe that the risk-modelling methodology we described can help move the precision medicine agenda forward as it may be applicable to a wide variety of treatments in critical care.

design of adaptive trials [33]. For instance, enrichment trials recruiting only the patients most likely to benefit from an early RRT initiation strategy could yield larger treatment effect sizes [34].



Abbreviations

AKI: Acute Kidney Injury; ARD: Absolute Risk Difference; CI: Confidence Interval; HTE: Heterogeneous Treatment Effects; HR: Hazard Ratio; ICU: Intensive Care Unit; PATH: Predictive Approaches to Treatment effect Heterogeneity; RCT: Randomized Controlled Trial; RRT: Renal Replacement Therapy; SOFA: Sequential Organ Failure Assessment; TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13054-022-03936-y>.

Additional file 1. Computer code: analysis.R.

Additional file 2. Table S1: Characteristics of the patients at randomization in each arm by fifth of risk of RRT initiation within 48 hours after the start of a delayed strategy. **Table S2:** Characteristics of the patients at randomization by fifth of risk of RRT initiation within 48 hours after the start of a delayed strategy. **Figure S1:** Missing data in the delayed (Panel A) and early (Panel B) strategy arms. **Figure S2:** Heterogeneity of treatment effect (early vs delayed strategy) across different levels of risk of RRT initiation within 48 hours after allocation a delayed strategy in the AKIKI and IDEAL-ICU samples. **Figure S3:** Kaplan-Meier estimates of survival at 60 days in each fifth of risk for the AKIKI, IDEAL-ICU and pooled samples. **Figure S4:** Results on secondary outcomes on the mean difference scale in each fifth of risk for the AKIKI, IDEAL-ICU and pooled samples.

Acknowledgements

We thank all patients included in the AKIKI and IDEAL-ICU trials as well as their surrogates. We express our gratitude to the medical and nursing teams that participated in these trials.

Authors' contributions

Study concept and design: FG, RP, SG, J-PQ, DD. Acquisition of data: SG, SB, J-PQ, DD. Analysis and interpretation of data: FG, RP. Drafting of the manuscript: FG. Critical revision of the manuscript for important intellectual content: FG, SG, SB, DH, AB, RP, J-PQ, DD. Dr. FG and Prof. RP had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors read and approved the final manuscript.

Funding

This study received funding from Assistance-Publique — Hôpitaux de Paris. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Availability of data and materials

Anonymous participant data is available under specific conditions. Proposals will be reviewed and approved by the sponsor, scientific committee, and staff on the basis of scientific merit and absence of competing interests. Once the proposal has been approved, data can be transferred through a secure online platform after the signing of a data access agreement and a confidentiality agreement.

Declarations

Ethics approval and consent to participate

The AKIKI and the IDEAL-ICU trials received approval for all participating centers from competent French legal authority and consent of patient or relatives was obtained before inclusion.

Consent for publication

All authors have consented to the publication of the present manuscript.

Competing interests

The authors have disclosed that they do not have any conflicts of interest.

Author details

¹Centre of Research in Epidemiology and Statistics (CRESS), Université de Paris, French Institute of Health and Medical Research (INSERM U1153), French National Research Institute for Agriculture, Food, and Environment (INRAE), Paris, France. ²Intensive Care Department, Nîmes University Hospital, University of Montpellier, Nîmes, France. ³INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, AP-HP, Hôpital Pitie-Salpêtrière, Département de Santé Publique, Centre de Pharmacocépidémiologie, Sorbonne Université, Paris, France. ⁴Clinical Epidemiology Unit, INSERM CIC1432, Dijon, and Clinical Investigation Center, Clinical Epidemiology/Clinical Trials Unit, Dijon Bourgogne University Hospital, Dijon, France. ⁵Department of Intensive Care, François Mitterrand University Hospital, Lipness Team, INSERM Research Center, LNC-UMR1231 and LabEx LipSTIC, and INSERM CIC 1432, Clinical Epidemiology, University of Burgundy, Dijon, France. ⁶Université de Paris, Service de Médecine Intensive-Réanimation, Hôpital Louis Mourier, AP-HP and INSERM, UMR S1155 "Common and Rare Kidney Diseases: From Molecular Events To Precision Medicine", Sorbonne Université, Paris, France. ⁷Service de Réanimation Médico-Chirurgicale, Hôpital Avicenne, APHP, UFR SMBH, Université Sorbonne Paris Nord, Bobigny, French National Institute

of Health and Medical Research (INSERM), Common and Rare kidney Diseases (CORAKID), Hôpital Tenon, Paris, France.

Received: 7 January 2022 Accepted: 22 February 2022

Published online: 21 March 2022

References

- Chawla LS, Eggers PW, Star RA, Kimmel PL. Acute kidney injury and chronic kidney disease as interconnected syndromes. *N Engl J Med*. 2014;371(1):58–66.
- Parsons FM, Hobson S, Blagg CR, McCracken BH. Optimum time for dialysis in acute reversible renal failure. Description and value of an improved dialyser with large surface area. *Lancet*. 1961;1(7169):129–34.
- Barbar SD, Clerc-Jehl R, Bourredjem A, Hernu R, Montini F, Bruyère R, et al. Timing of renal-replacement therapy in patients with acute kidney injury and sepsis. *N Engl J Med*. 2018;379(15):1431–42.
- Gaudry S, Hajage D, Schortgen F, Martin-Lefevre L, Pons B, Boulet E, et al. Initiation strategies for renal-replacement therapy in the intensive care unit. *N Engl J Med*. 2016;375(2):122–33.
- STARRT-AKI Investigators, Canadian Critical Care Trials Group, Australian and New Zealand Intensive Care Society Clinical Trials Group, United Kingdom Critical Care Research Group, Canadian Nephrology Trials Network, Irish Critical Care Trials Group, et al. Timing of Initiation of Renal-Replacement Therapy in Acute Kidney Injury. *N Engl J Med*. 2020;383(3):240–51.
- Fayad All, Buamscha DG, Ciapponi A. Timing of renal replacement therapy initiation for acute kidney injury. *Cochrane Database Syst Rev*. 2018. <https://doi.org/10.1002/14651858.CD010612.pub2/full>.
- Gaudry S, Hajage D, Benichou N, Chaïbi K, Barbar S, Zarbock A, et al. Delayed versus early initiation of renal replacement therapy for severe acute kidney injury: a systematic review and individual patient data meta-analysis of randomised clinical trials. *Lancet*. 2020;395(10235):1506–15.
- Iwashyna TJ, Burke JF, Sussman JB, Prescott HC, Hayward RA, Angus DC. Implications of heterogeneity of treatment effect for reporting and analysis of randomized trials in critical care. *Am J Respir Crit Care Med*. 2015;192(9):1045–51.
- Barbar SD, Dargent A, Quenot J-P. Timing of renal-replacement therapy in acute kidney injury and sepsis. *N Engl J Med*. 2019;380(4):399.
- Shah FA, Meyer NJ, Angus DC, Awdish R, Azoulay É, Calfee CS, et al. A research agenda for precision medicine in sepsis and acute respiratory distress syndrome: an official American Thoracic Society Research Statement. *Am J Respir Crit Care Med*. 2021;204(8):891–901.
- Schaub JA, Heung M. Precision medicine in acute kidney injury: a promising future? *Am J Respir Crit Care Med*. 2019;199(7):814–6.
- Gaudry S, Hajage D, Schortgen F, Martin-Lefevre L, Verney C, Pons B, et al. Timing of renal support and outcome of septic shock and acute respiratory distress syndrome. A post hoc analysis of the AKIKI Randomized clinical trial. *Am J Respir Crit Care Med*. 2018;198(1):58–66.
- Hamburg MA, Collins FS. The Path to Personalized Medicine. *N Engl J Med*. 2010;363(4):301–4.
- Bouchard J, Mehta RL. Timing of kidney support therapy in acute kidney injury: what are we waiting for? *Am J Kidney Dis*. 2022;79:417–26.
- van Klaveren D, Varadhan R, Kent DM. The Predictive Approaches to Treatment effect Heterogeneity (PATH) statement. *Ann Intern Med*. 2020;172(11):776.
- Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594.
- Schoenfeld DA, Bernard GR, ARDS Network. Statistical evaluation of ventilator-free days as an efficacy measure in clinical trials of treatments for acute respiratory distress syndrome. *Crit Care Med*. 2002;30(8):1772–7.
- van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res*. 2007;16(3):219–42.
- Vergouwe Y, Royston P, Moons KGM, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. *J Clin Epidemiol*. 2010;63(2):205–14.
- Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med*. 1991;10(4):585–98.
- Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Stat Med*. 2014;33(3):517–35.
- Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15(4):361–87.
- Collins GS, Ogundimu EO, Cook JA, Manach YL, Altman DG. Quantifying the impact of different approaches for handling continuous predictors on the performance of a prognostic model. *Stat Med*. 2016;35(23):4124–35.
- Wood SN. On p-values for smooth components of an extended generalized additive model. *Biometrika*. 2013;100(1):221–8.
- Evans L, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Crit Care Med*. 2021;49(11):e1063–143.
- Künzel SR, Sekhon JS, Bickel PJ, Yu B. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proc Natl Acad Sci U S A*. 2019;116(10):4156–65.
- van Klaveren D, Balan TA, Steyerberg EW, Kent DM. Models with interactions overestimated heterogeneity of treatment effects and were prone to treatment mistargeting. *J Clin Epidemiol*. 2019;114:72–83.
- Cutler DM. Early returns from the era of precision medicine. *JAMA*. 2020;323(2):109–10.
- Takahashi K, Serruys PW, Fuster V, Farkouh ME, Spertus JA, Cohen DJ, et al. Redesign and validation of the SYNTAX score II to individualise decision making between percutaneous and surgical revascularisation in patients with complex coronary artery disease: secondary analysis of the multicentre randomised controlled SYNTAXES trial with external cohort validation. *Lancet*. 2020;396(10260):1399–412.
- McKown AC, Huerta LE, Rice TW, Semler MW. Heterogeneity of treatment effect by baseline risk in a trial of balanced crystalloids versus saline. *Am J Respir Crit Care Med*. 2018;198(6):810–3.
- Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA, et al. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *Lancet Respir Med*. 2014;2(8):611–20.
- Semler MW, Bernard GR, Aaron SD, Angus DC, Biros MH, Brower RG, et al. Identifying clinical research priorities in adult pulmonary and critical care: NHLBI working group report. *Am J Respir Crit Care Med*. 2020;202(4):511–23.
- Gasparini M, Chevret S. Intensive care medicine in 2050: clinical trials designs. *Intensive Care Med*. 2019;45(5):668–70.
- Kellum JA, Fuhrman DY. The handwriting is on the wall: there will soon be a drug for AKI. *Nat Rev Nephrol*. 2019;15(2):65–6.
- Stanski NL, Wong HR. Prognostic and predictive enrichment in sepsis. *Nat Rev Nephrol*. 2020;16(1):20–31.
- Granholm A, Marker S, Krag M, Zampieri FG, Thorsen-Meyer H-C, Kaas-Hansen BS, et al. Heterogeneity of treatment effect of prophylactic pantoprazole in adult ICU patients: a post hoc analysis of the SUP-ICU trial. *Intensive Care Med*. 2020;46(4):717–26.
- Zochios V, Brodie D, Parhar KK. Toward precision delivery of ECMO in COVID-19 cardiorespiratory failure. *ASAIO J*. 2020;66(7):731–3.
- Barasch J, Zager R, Bonventre JV. Acute kidney injury: a problem of definition. *Lancet*. 2017;389(10071):779–81.
- Gaudry S, Hajage D, Martin-Lefevre L, Lebbah S, Louis G, Moschietto S, et al. Comparison of two delayed strategies for renal replacement therapy initiation for severe acute kidney injury (AKIKI 2): a multicentre, open-label, randomised, controlled trial. *Lancet*. 2021;397(10281):1293–300.
- Oermann M, Lumertgul N. Wait and see for acute dialysis: but for how long? *Lancet*. 2021;397(10281):1241–3.
- Nie X, Brunskill E, Wager S. Learning when-to-treat policies. *J Am Stat Assoc*. 2021;116(533):392–409.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Personalization of renal replacement therapy initiation: a secondary analysis of the AKIKI and IDEAL-ICU trials

François GROLLEAU,¹ Raphaël PORCHER,² Saber BARBAR,³ David HAJAGE,⁴ Abderrahmane BOURREDJEM,⁵ Jean-Pierre QUENOT*,⁶ Didier DREYFUSS*,⁷ Stéphane GAUDRY*.⁸

* These authors contributed equally as senior authors.

ADITIONAL FILE 1

Table of Contents

Supplementary Results.....	3
Table S1. Characteristics of the patients at randomization in each arm by fifth of risk of RRT initiation within 48 hours after the start of a delayed strategy.....	3
Table S2. Characteristics of the patients at randomization by fifth of risk of RRT initiation within 48 hours after the start of a delayed strategy	4
Figure S1. Missing data in the delayed (Panel A) and early (Panel B) strategy arms.	5
Figure S2. Heterogeneity of treatment effect (early vs delayed strategy) across different levels of risk of RRT initiation within 48 hours after allocation a delayed strategy in the AKIKI and IDEAL-ICU samples.	6
Figure S3. Kaplan-Meier estimates of survival at 60 days in each fifth of risk for the AKIKI, IDEAL-ICU and pooled samples.....	7
Figure S4. Results on secondary outcomes on the mean difference scale in each fifth of risk for the AKIKI, IDEAL-ICU and pooled samples.....	8

Supplementary Results

Table S1. Characteristics of the patients at randomization in each arm by fifth of risk of RRT initiation within 48 hours after the start of a delayed strategy.

All characteristics reported in the table were determined at inclusion in the AKIKI or IDEAL-ICU trial, before initiation of renal replacement therapy. Intervals corresponds to the minimal and maximal predicted probability of RRT initiation in each fifth. *P* values are not adjusted for multiplicity.

Characteristic	Q1 [0.0211,0.114]			Q2 (0.114,0.177]			Q3 (0.177,0.249]			Q4 (0.249,0.355]			Q5 (0.355,0.862]		
	Delayed strategy n=107	Early strategy n=115	<i>P</i>	Delayed strategy n=114	Early strategy n=107	<i>P</i>	Delayed strategy n=98	Early strategy n=123	<i>P</i>	Delayed strategy n=106	Early strategy n=115	<i>P</i>	Delayed strategy n=125	Early strategy n=97	<i>P</i>
Study															
IDEAL-ICU	32 (29.9)	39 (33.9)	0.620	44 (38.6)	40 (37.4)	0.962	40 (40.8)	64 (52.0)	0.128	57 (53.8)	51 (44.3)	0.206	69 (55.2)	52 (53.6)	0.920
Age — year	70.40 (12.53)	65.66 (14.76)	0.011	65.91 (15.09)	66.03 (14.33)	0.955	69.55 (12.11)	67.68 (12.22)	0.257	67.21 (13.75)	67.96 (11.81)	0.662	65.84 (11.80)	65.02 (13.44)	0.627
Weight — kg	72.29 (17.73)	73.02 (14.39)	0.735	77.24 (17.27)	78.01 (16.31)	0.735	82.77 (18.58)	83.12 (21.29)	0.896	83.15 (21.13)	84.77 (20.35)	0.560	91.88 (28.31)	94.67 (30.90)	0.485
Male sex	65 (60.7)	62 (53.9)	0.372	65 (57.0)	71 (66.4)	0.198	64 (65.3)	73 (59.3)	0.443	75 (70.8)	81 (70.4)	1.000	83 (66.4)	64 (66.0)	1.000
Pre-existing conditions															
Heart failure	11 (10.3)	10 (8.7)	0.862	10 (8.8)	7 (6.5)	0.712	8 (8.2)	6 (4.9)	0.473	12 (11.3)	5 (4.3)	0.091	11 (8.8)	16 (16.5)	0.125
Hypertension	64 (59.8)	56 (48.7)	0.127	62 (54.4)	57 (53.3)	0.975	57 (58.2)	74 (60.2)	0.871	56 (52.8)	60 (52.2)	1.000	65 (52.0)	59 (60.8)	0.239
Diabetes mellitus	13 (12.0)	16 (13.9)	1.000	13 (12.5)	23 (17.5)	0.261	11 (11.2)	15 (13.5)	0.510	20 (18.9)	22 (17.1)	1.000	29 (27.2)	27 (27.0)	0.345
Cirrhosis	10 (9.3)	9 (7.8)	0.869	10 (8.8)	6 (5.6)	0.517	7 (7.1)	13 (10.6)	0.518	12 (11.3)	10 (8.7)	0.670	15 (12.0)	16 (16.5)	0.445
Respiratory Disease	9 (8.4)	12 (10.4)	0.775	14 (12.3)	16 (15.0)	0.702	14 (14.3)	10 (8.1)	0.214	7 (6.6)	11 (9.6)	0.577	10 (8.0)	13 (13.4)	0.277
Cancer	22 (20.6)	22 (19.1)	0.921	19 (16.7)	15 (14.0)	0.720	18 (18.4)	17 (13.8)	0.463	18 (17.0)	20 (17.4)	1.000	23 (18.4)	17 (17.5)	1.000
Hepatitis	7 (6.3)	3 (2.6)	1.000	3 (2.6)	5 (4.7)	0.473	4 (4.0)	1 (0.8)	1.000	3 (2.6)	10 (9.7)	0.118	10 (9.0)	5 (5.2)	0.570
AIDS	0 (0.0)	1 (0.9)	1.000	2 (1.8)	1 (0.9)	1.000	0 (0.0)	1 (0.8)	1.000	0 (0.0)	0 (0.0)	1.000	0 (0.0)	2 (2.1)	0.370
Non-corticosteroid immunosuppressive drug	2 (1.9)	1 (0.9)	0.950	1 (0.9)	1 (0.9)	1.000	6 (6.1)	6 (4.9)	0.915	6 (5.7)	12 (10.4)	0.294	21 (16.8)	12 (12.4)	0.465
Organ graft	1 (0.9)	0 (0.0)	0.971	1 (0.9)	1 (0.9)	1.000	3 (3.1)	1 (0.8)	0.461	2 (1.9)	2 (1.7)	1.000	10 (8.0)	1 (1.0)	0.039
Severity at enrollment															
SOFA score (0 to 24)	8.59 (2.39)	8.50 (2.62)	0.783	10.68 (2.46)	10.78 (2.61)	0.789	11.11 (2.28)	11.69 (2.58)	0.083	13.12 (2.60)	12.58 (2.65)	0.128	13.56 (2.86)	13.76 (2.82)	0.598
Respiratory SOFA (0 to 4)	2.18 (1.17)	1.88 (1.09)	0.050	2.03 (1.03)	1.96 (1.16)	0.665	2.05 (1.06)	1.85 (1.16)	0.194	1.95 (1.22)	1.94 (1.18)	0.933	2.18 (1.25)	2.00 (1.19)	0.289
to 4)															
 Hemodynamic															
SOFA (0 to 4)	2.82 (1.65)	2.63 (1.71)	0.407	3.28 (1.40)	3.51 (1.14)	0.179	3.49 (1.17)	3.63 (1.01)	0.326	3.76 (0.76)	3.69 (0.99)	0.518	3.78 (0.78)	3.88 (0.60)	0.335
 Liver SOFA (0 to 4)															
Liver SOFA (0 to 4)	0.56 (0.87)	0.54 (0.88)	0.854	0.74 (0.93)	0.64 (0.99)	0.478	0.67 (0.93)	0.87 (1.09)	0.158	0.94 (1.09)	0.77 (1.10)	0.253	1.18 (1.26)	1.14 (1.30)	0.819
 Coagulation SOFA (0 to 4)															
Coagulation SOFA (0 to 4)	2.57 (1.55)	2.43 (1.74)	0.547	2.19 (1.61)	2.50 (1.54)	0.156	2.10 (1.66)	1.91 (1.60)	0.386	1.88 (1.54)	2.05 (1.47)	0.438	1.82 (1.53)	1.99 (1.50)	0.398
to 4)															
 Neurologic SOFA (0 to 4)															
Neurologic SOFA (0 to 4)	0.74 (1.27)	0.92 (1.32)	0.293	1.19 (1.44)	1.14 (1.51)	0.791	1.22 (1.51)	1.14 (1.39)	0.659	1.69 (1.61)	1.41 (1.59)	0.195	1.57 (1.50)	1.65 (1.61)	0.698
Laboratory values															
 Baseline creatinine, $\mu\text{mol/L}$*															
Baseline creatinine, $\mu\text{mol/L}$ *	82.68 (24.92)	75.44 (22.30)	0.023	83.33 (24.43)	83.98 (26.91)	0.850	89.28 (30.07)	84.43 (26.64)	0.206	96.72 (44.31)	85.04 (32.90)	0.026	93.69 (45.76)	103.85 (51.00)	0.120
 Creatinine at enrollment, $\mu\text{mol/L}$															
Creatinine at enrollment, $\mu\text{mol/L}$	235.10	228.40	0.620	269.88	269.40	0.973	284.28	295.82	0.473	311.57	319.94	0.651	342.49	317.32	0.201
 Blood urea nitrogen at enrollment, mmol/L															
Blood urea nitrogen at enrollment, mmol/L	15.99 (6.92)	15.81 (7.29)	0.846	18.73 (9.41)	18.29 (7.37)	0.697	20.51 (7.44)	20.28 (9.04)	0.844	22.28 (9.89)	21.76 (10.25)	0.701	25.30 (10.81)	23.41 (9.24)	0.171
 Arterial blood pH at enrollment															
Arterial blood pH at enrollment	7.39 (0.07)	7.39 (0.07)	0.467	7.33 (0.07)	7.34 (0.08)	0.320	7.29 (0.07)	7.31 (0.07)	0.161	7.27 (0.07)	7.27 (0.07)	0.958	7.22 (0.10)	7.20 (0.08)	0.233

Data are mean (SD) or n (%); AIDS=Acquired Immunodeficiency Syndrome; SOFA score=Sequential Organ Failure Assessment score.

* The serum creatinine concentration before ICU admission was either determined with the use of values measured in the 12 months preceding the ICU stay or was estimated. To convert the values for creatinine to milligrams per deciliter, divide by 88.4.

Table S2. Characteristics of the patients at randomization by fifth of risk of RRT initiation within 48 hours after the start of a delayed strategy.

All characteristics reported in the table were determined at inclusion in the AKIKI or IDEAL-ICU trial, before initiation of renal replacement therapy. Intervals corresponds to the minimal and maximal predicted probability of RRT initiation in each fifth. *P* values are not adjusted for multiplicity.

Characteristic	Q1 [0.0211,0.114] n=222	Q2 (0.114,0.177] n=221	Q3 (0.177,0.249] n=221	Q4 (0.249,0.355] n=221	Q5 (0.355,0.862] n=222	<i>P</i>
Study						
AKIKI	151 (68.0)	137 (62.0)	117 (52.9)	113 (51.1)	101 (45.5)	
IDEAL-ICU	71 (32.0)	84 (38.0)	104 (47.1)	108 (48.9)	121 (54.5)	<0.001
Age — year	67.9 (13.9)	65.9 (14.7)	68.5 (12.2)	67.6 (12.7)	65.5 (12.5)	0.072
Weight — kg	72.7 (16.0)	77.6 (16.8)	82.9 (20.1)	84.0 (20.7)	93.1 (29.5)	<0.001
Male sex	127 (57.2)	136 (61.5)	137 (62.0)	156 (70.6)	147 (66.2)	0.043
Pre-existing conditions						
Heart failure	21 (9.5)	17 (7.7)	14 (6.3)	17 (7.7)	27 (12.2)	0.227
Hypertension	120 (54.1)	119 (53.8)	131 (59.3)	116 (52.5)	124 (55.9)	0.652
Diabetes mellitus	31 (14.0)	41 (18.6)	39 (17.6)	42 (19.0)	51 (23.0)	0.189
Cirrhosis	19 (8.6)	16 (7.2)	20 (9.0)	22 (10.0)	31 (14.0)	0.162
Respiratory Disease	21 (9.5)	30 (13.6)	24 (10.9)	18 (8.1)	23 (10.4)	0.430
Cancer	44 (19.8)	34 (15.4)	35 (15.8)	38 (17.2)	40 (18.0)	0.742
Hemopathy	15 (6.8)	8 (3.6)	10 (4.5)	13 (5.9)	15 (6.8)	0.507
AIDS	1 (0.5)	3 (1.4)	1 (0.5)	0 (0.0)	2 (0.9)	0.442
Non-corticosteroid immunosuppressive drug	3 (1.4)	2 (0.9)	12 (5.4)	18 (8.1)	33 (14.9)	<0.001
Organ graft	1 (0.5)	2 (0.9)	4 (1.8)	4 (1.8)	11 (5.0)	0.007
Severity at enrollment						
SOFA score (0 to 24)	8.5 (2.5)	10.7 (2.5)	11.4 (2.5)	12.8 (2.6)	13.6 (2.8)	<0.001
Respiratory SOFA (0 to 4)	2.0 (1.1)	2.0 (1.1)	1.9 (1.1)	1.9 (1.2)	2.1 (1.2)	0.595
Hemodynamic SOFA (0 to 4)	2.7 (1.7)	3.4 (1.3)	3.6 (1.1)	3.7 (0.9)	3.8 (0.7)	<0.001
Liver SOFA (0 to 4)	0.5 (0.9)	0.7 (1.0)	0.8 (1.0)	0.9 (1.1)	1.2 (1.3)	<0.001
Coagulation SOFA (0 to 4)	2.5 (1.7)	2.3 (1.6)	2.0 (1.6)	1.9 (1.5)	1.8 (1.5)	<0.001
Neurologic SOFA (0 to 4)	0.8 (1.3)	1.2 (1.5)	1.2 (1.4)	1.5 (1.6)	1.6 (1.5)	<0.001
Laboratory values						
Baseline creatinine, µmol/L*	79 (23)	84 (26)	85 (28)	90 (39)	98 (48)	<0.001
Creatinine at enrollment, µmol/L	232 (100)	269 (104)	291 (118)	316 (137)	331 (145)	<0.001
Blood urea nitrogen at enrollment, mmol/L	16 (7)	19 (8)	20 (8)	22 (10)	24 (10)	<0.001
Potassium at enrollment, mmol/L	3.9 (0.6)	4.2 (0.7)	4.4 (0.6)	4.5 (0.7)	5.0 (0.8)	<0.001
Bicarbonate at enrollment, mmol/L	21 (5)	19 (5)	18 (5)	17 (4)	17 (5)	<0.001
Arterial blood pH at enrollment	7.39 (0.07)	7.34 (0.07)	7.30 (0.07)	7.27 (0.07)	7.21 (0.09)	<0.001

Data are mean (SD) or n (%); AIDS=Acquired Immunodeficiency Syndrome; SOFA score=Sequential Organ Failure Assessment score.

* The serum creatinine concentration before ICU admission was either determined with the use of values measured in the 12 months preceding the ICU stay or was estimated. To convert the values for creatinine to milligrams per deciliter, divide by 88.4.

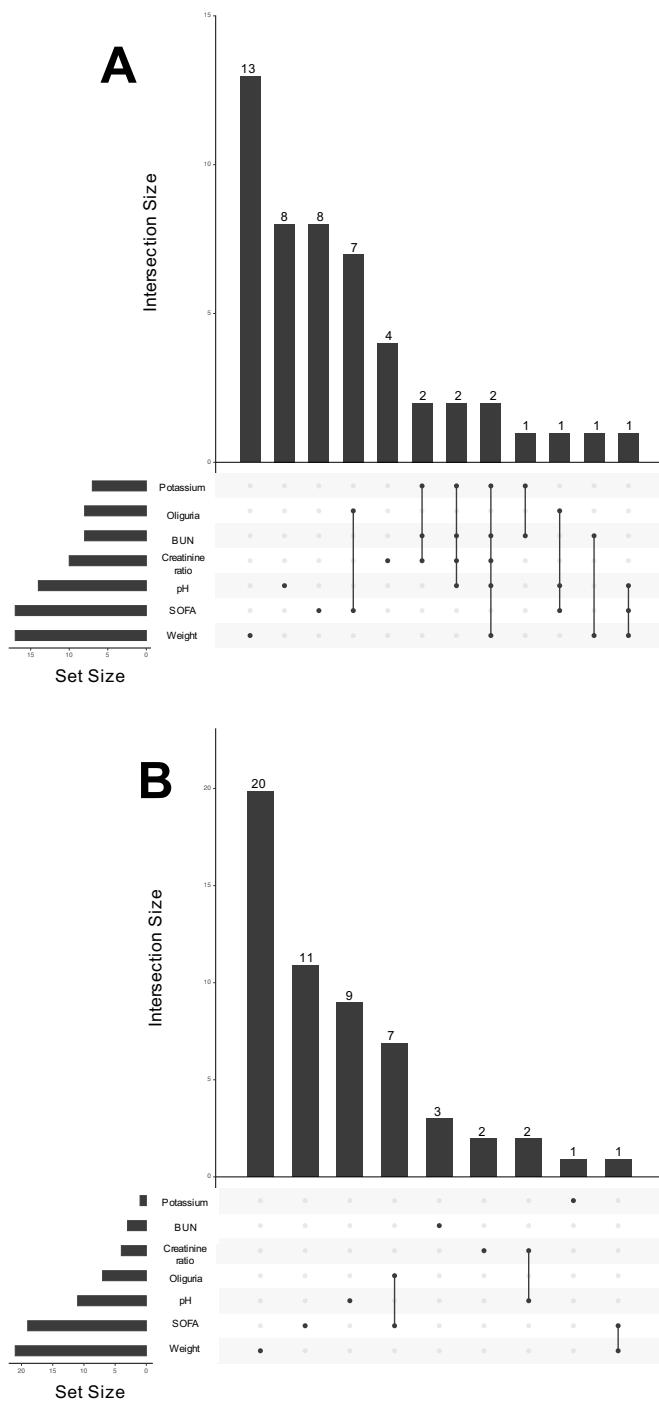


Figure S1. Missing data in the delayed (Panel A) and early (Panel B) strategy arms.

The tables show all missing data patterns observed. Vertical bars show the number of patients corresponding to the missing data pattern underneath. Horizontal bars show the number of patients with missing data corresponding to the variable aside. 500 patients (91%) had complete data with no missing candidate predictors in the delayed arms. 501 patients (90%) had complete data with no missing candidate predictors in the early arms. BUN=Blood Urea Nitrogen.

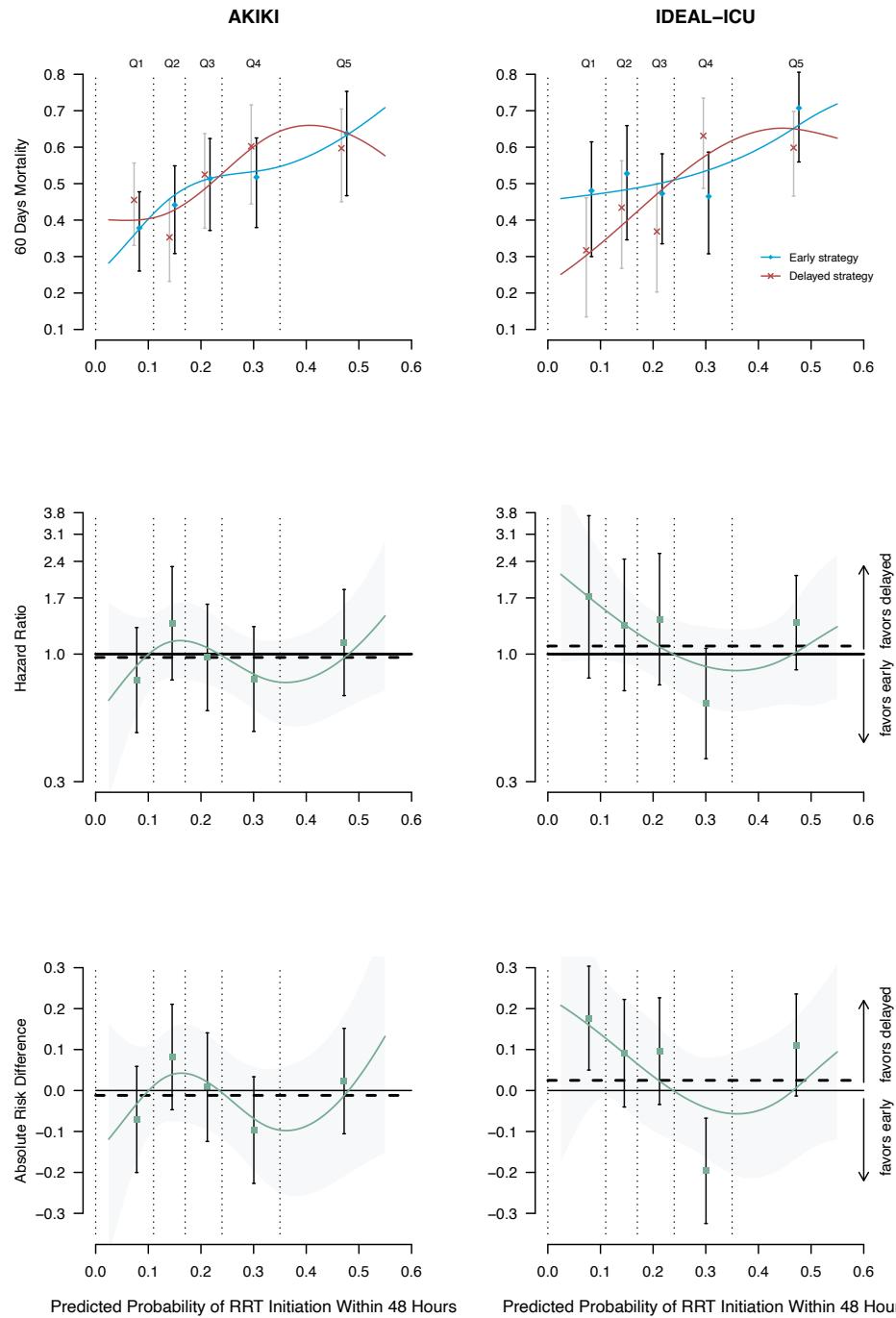


Figure S2. Heterogeneity of treatment effect (early vs delayed strategy) across different levels of risk of RRT initiation within 48 hours after allocation a delayed strategy in the AKIKI and IDEAL-ICU samples.

This figure presents the heterogenous of treatment effects of an early vs a delayed strategy of RRT initiation as a function of the baseline risk of RRT initiation within 48 hours after a delayed strategy. The dashed lines indicate the average treatment effects in the corresponding sample. Q1 = first fifth of risk (lowest), Q2 = second fifth of risk, Q3 = third fifth of risk, Q4 = fourth fifth of risk, Q5 = last fifth of risk (highest).

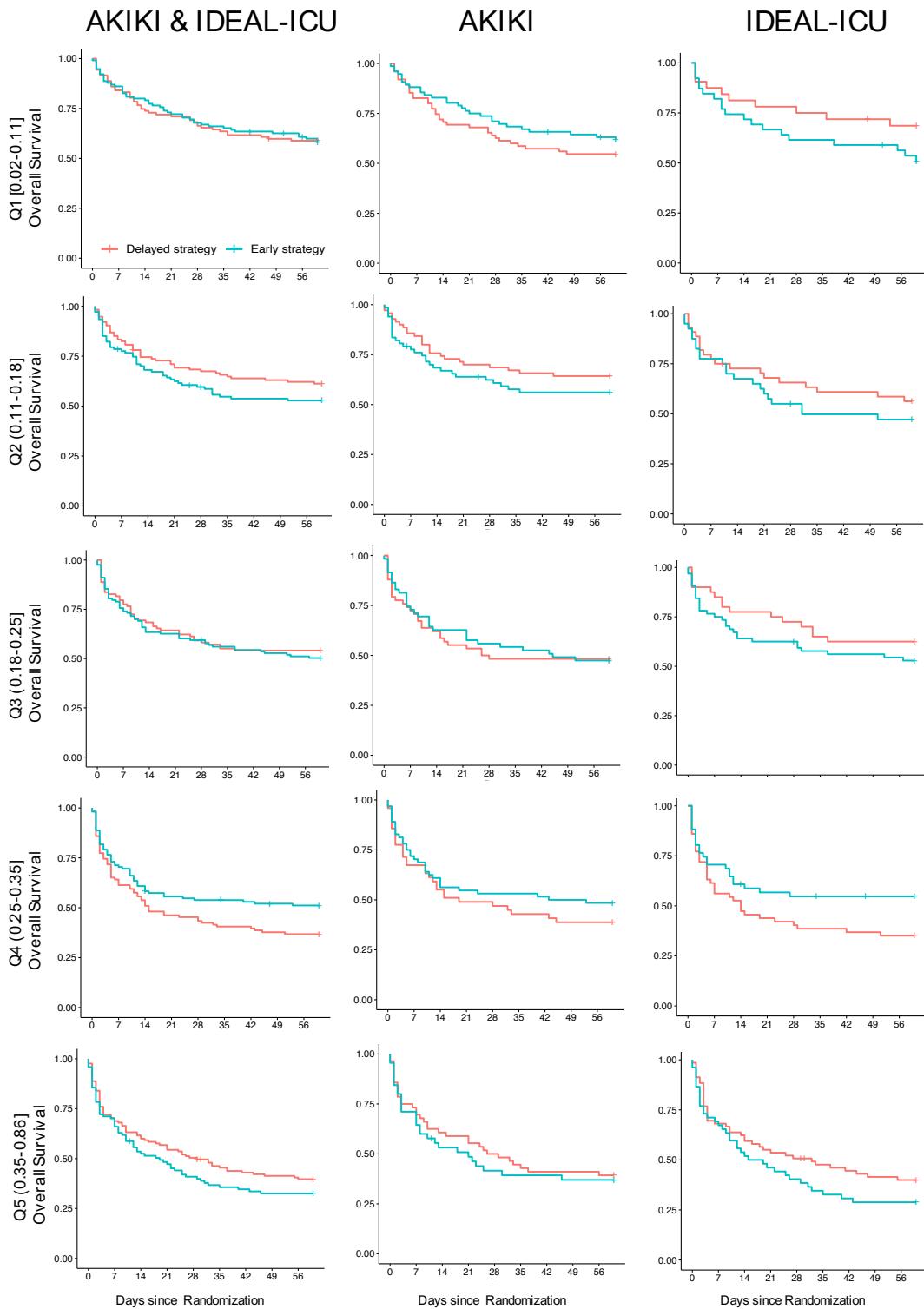


Figure S3. Kaplan-Meier estimates of survival at 60 days in each fifth of risk for the AKIKI, IDEAL-ICU and pooled samples.

Q1 = first fifth of risk (lowest), Q2 = second fifth of risk, Q3 = third fifth of risk, Q4 = fourth fifth of risk, Q5 = last fifth of risk (highest). Kaplan-Meier curves on the left-hand side are identical to those depicted in Figure 3 of the main text.

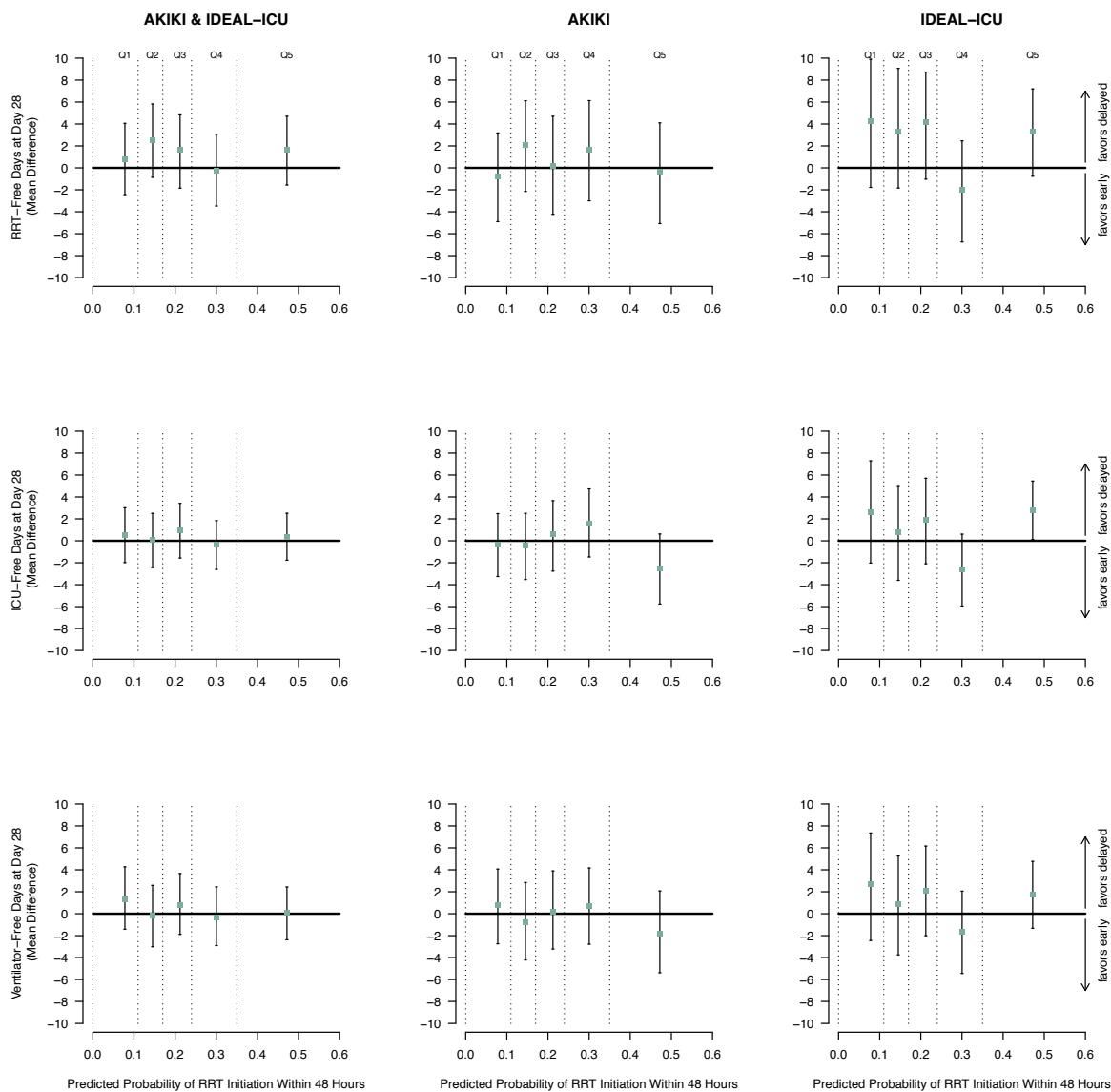


Figure S4. Results on secondary outcomes on the mean difference scale in each fifth of risk for the AKIKI, IDEAL-ICU and pooled samples.

Q1 = first fifth of risk (lowest), Q2 = second fifth of risk, Q3 = third fifth of risk, Q4 = fourth fifth of risk, Q5 = last fifth of risk (highest).

Chapter 5

A comprehensive framework for the evaluation of individual treatment rules from observational data

Harmony shall always come before perfection.

Buddhist maxim

La musique associée à ce chapitre est la symphonie n° 11, premiers mouvements de Philip Glass.

RESEARCH ARTICLE

A comprehensive framework for the evaluation of individual treatment rules from observational data

François Grolleau^{1,2,3} | François Petit^{1,2} | Raphaël Porcher^{1,2,3}

¹Université Paris Cité, Paris, France

²Centre de Recherche Épidémiologie et Statistiques (CRESS-UMR1153), INSERM, INRAE, Paris, France

³Centre d'Épidémiologie Clinique, Assistance Publique-Hôpitaux de Paris, Hôtel-Dieu, Paris, France

Correspondence

François Grolleau, Centre d'Épidémiologie Clinique, Hôpital Hôtel-Dieu, 1 Place du Parvis Notre-Dame Paris, 75004, Paris, France. Email: francois.grolleau@aphp.fr

Abstract

Individualized treatment rules (ITRs) are deterministic decision rules that recommend treatments to individuals based on their characteristics. Though ubiquitous in medicine, ITRs are hardly ever evaluated in randomized controlled trials. To evaluate ITRs from observational data, we introduce a new probabilistic model and distinguish two situations: i) the situation of a newly developed ITR, where data are from a population where no patient implements the ITR, and ii) the situation of a partially implemented ITR, where data are from a population where the ITR is implemented in some unidentified patients. In the former situation, we propose a procedure to explore the impact of an ITR under various implementation schemes. In the latter situation, on top of the fundamental problem of causal inference, we need to handle an additional latent variable denoting implementation. To evaluate ITRs in this situation, we propose an estimation procedure that relies on an expectation-maximization algorithm. In Monte Carlo simulations our estimators appear unbiased with confidence intervals achieving nominal coverage. We illustrate our approach on the MIMIC-III database, focusing on ITRs for dialysis initiation in patients with acute kidney injury.

KEYWORDS:

Personalized medicine; Causal inference; Mixture of experts; Expectation-maximization algorithm.

1 | INTRODUCTION

Individualized treatment rules (ITRs) are decision rules that recommend treatments to individuals based on their observed characteristics to maximize favorable outcomes on average. ITRs are widespread in medicine. In fact, most guidelines as well as the recently released computerized clinical decision support tools can be viewed as ITRs.^{1,2} Notable examples include decision tools for revascularization strategies in patients with coronary artery disease³ and for the personalization of blood pressure targets in hypertensive patients⁴. For evaluating the impact of an ITR, the gold standard would be to conduct a randomized controlled trial (RCT) comparing the implementation of that ITR to usual care. Yet, there are practical challenges to conducting

The authors gratefully acknowledge the Agence Nationale de la Recherche who partially funded this work under grant agreement no. ANR-18-CE36-0010-01. François Petit was supported by the IdEx Université Paris Cité, ANR-18-IDEX-0001. Raphaël Porcher acknowledges the support of the French Agence Nationale de la Recherche as part of the “Investissements d’avenir” program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute).

such RCTs.⁵ As ITRs often recommend treatments similar to usual care, the expected population-level effect is likely small and necessitating very large sample sizes. Moreover, health agency oversight is less stringent for the implementation of ITRs than for drug compounds and so, both the incentives and funding opportunities for conducting RCTs of ITRs are scarce. In practice, these RCTs remain rare. As a result, many ITRs are being implemented despite the lack of evidence supporting their benefit. In this paper, we develop a framework to evaluate from observational data the impact of ITRs.

To fit with most ITRs available in medicine (e.g., computerized clinical decision support tools, or guidelines), we view ITRs as deterministic maps recommending one of two treatment options. To make inference accounting for real-life prescription of treatment by physicians, we consider that deterministic ITRs are stochastically implemented with a probability of implementation depending on patient characteristics. Critically, we then distinguish two situations: i) the ITR was just released and treatment prescription was never based on it in the population, or ii) the ITR was available and, for some patients in the population, treatment prescription was based on it. We term these two situations the *new ITR* and the *partially implemented ITR* situations, respectively. In the former situation, we propose to numerically explore the benefit an ITR may have under different implementation schemes. In the latter situation, inference is more challenging as we are typically given observational data where we do not know which patients had implemented the ITR. That is, on top of the fundamental problem of causal inference, we need to handle an additional latent variable denoting implementation. To address this situation, we develop a new probability model and rely on a mixture of experts fitted via an EM algorithm for inference.

ITR estimation and evaluation has been considered in the literature of both statistics^{6,7,8} and machine learning.^{9,10} Works most related to ours include the evaluation of stochastic rules,¹¹ biomarker performance,¹² and ITR value accounting for the number of treated units.¹³ We note that encouragement designs¹⁴ and instrumental variable^{15,16} methods tackle problems related to, but subtly different from the question of estimating the effect of an ITR. To our knowledge, no work has focused on data originating from a *partially implemented ITR situation*, nor pursued to develop a comprehensive framework for the evaluation of ITRs from observational data.

This article is organized as follows. In the evaluation metrics section, we introduce our causal model as well as our three estimands of interest: the Average Rule Effect (ARE), the Average Implementation Effect (AIE), and the Maximal Implementation Gain (MIG). In the inference section, we provide a method to estimate the ARE, AIE, and MIG and compute their standard error in both the *new ITR* and *partially implemented ITR* situations. In the simulation section, we study the properties of our estimators in the more challenging *partially implemented ITR situation*. Finally, in the application section, we illustrate our approach on the MIMIC-III database, focusing on ITRs for dialysis initiation in patients with acute kidney injury. We evaluate two ITRs corresponding to the *new ITR* and the *partially implemented ITR* situations. The computer code for simulation studies and data applications is available at <https://github.com/fcgrolleau/ITReval>.

2 | SETUP AND EVALUATION METRICS

Following Neyman-Rubin causal model, we consider that a patient with observed outcome Y has two potential outcomes $Y^{a=0}$ and $Y^{a=1}$ representing the outcome s/he would achieve if, possibly contrary to fact, s/he had received treatment option $A = 0$ or $A = 1$ respectively.^{17,18} Without loss of generality, we consider $A = 1$ indicates that a patient received a specific treatment, and $A = 0$ indicates s/he received a control. Additionally, we consider for each patient, a vector of pre-treatment covariates X with values in \mathcal{X} .

We assume that we are given an ITR that is, a deterministic map $r : \mathcal{X} \mapsto \{0; 1\}$ which assigns a treatment option to each patient with covariates x . We model the implementation of the rule by the binary random variable S where $S = 1$ indicates that a patient's physician consulted the ITR and followed its recommendation to prescribe treatment (e.g., the physician accesses

the ITR and takes its recommendation seriously). For the rest of this paper, we term $S = 1$ as implementing the ITR. On the contrary, $S = 0$ indicates that a patient's physician either did not consult the ITR or consulted it but prescribed a treatment opposite to what the ITR recommended (e.g., the physician has no access to the ITR or s/he does not take its recommendation seriously). The formal definition of $S = 1$ and $S = 0$ is given in Table 1. Note that when $S = 0$ the prescribed treatment may still match the ITR's recommended treatment—i.e., the physician did not consult the ITR but, based on other grounds, s/he prescribed a treatment identical to the ITR's recommendation. We define the stochastic implementation function as the conditional distribution $\rho(x) = \mathbb{E}[S|X = x]$.^a

Table 1 The precise definition of $S = 1$ (ITR is implemented) and $S = 0$ (ITR is not implemented).

		A physician consults the ITR	
		No	Yes
The prescription matches the ITR's recommendation i.e., $A = r(X)$	No	$S=0$	$S=0$
	Yes	$S=0$	$S=1$

We define the propensity score π as the conditional distribution $\pi(x) = \mathbb{E}[A|X = x]$ and the treatment-specific prognostic functions μ_0, μ_1 as the functions satisfying $\mu_1(x) = \mathbb{E}[Y^{a=1}|X = x]$ and $\mu_0(x) = \mathbb{E}[Y^{a=0}|X = x]$. We denote τ the conditional average treatment effect (CATE) function i.e., $\tau(x) = \mathbb{E}[Y^{a=1} - Y^{a=0}|X = x] = \mu_1(x) - \mu_0(x)$.

2.1 | A probability model for the data generating mechanism

Our goal in this subsection is to introduce a new causal model that allows us to determine the causal effect of implementing versus not implementing an ITR. We introduce $A^{s=1}$, the potential treatment that would be given to a patient if her/his physician implemented the ITR, i.e., $A^{s=1} = r(X)$, and $A^{s=0}$ the potential treatment s/he would be given if her/his physician did not implement the ITR. Similarly, we define $Y^{s=1}$ and $Y^{s=0}$, patient's potential outcomes when physicians do or do not implement the ITR, respectively. In addition, we define the variables with superscript $(-)^*$ the counterfactual variables that are observable in the situation where some physicians sometimes implement the ITR. The variables S^*, A^* , and Y^* respectively indicate the implementation status, the prescribed treatment, and the outcome in this situation where the ITR is partially implemented. We consider that we can easily identify which of the following two situations we are dealing with:

- A.** A situation where physicians never implement the ITR because it was not available. In this situation, the variables with superscript $(-)^{s=0}$ are observed and we have $S = 0, A = A^{s=0}, Y = Y^{s=0}$. From this point onward, we call this situation the *new ITR situation*.
- B.** A situation where some physicians sometimes implement the ITR to prescribe treatment. In this situation, the variables with superscript $(-)^*$ are all observable (though in practice S^* , the implementation status, is often not collected) and we have $S = S^*, A = A^*, Y = Y^*$. For the remainder of this paper, we refer to this situation as the *partially implemented ITR situation*.

^aNote that we consider here the stochastic implementation of a deterministic rule r through ρ . This is different from defining a function $\mathcal{X} \rightarrow [0; 1]$ which would assign to each value x a probability to allocate treatment $A = 1$. This would correspond to what we call a stochastic rule — which r is not.

Importantly, we consider that all counterfactual variables exist in both situations. However, we consider that these two situations are mutually exclusive. In particular, we consider that we can straightforwardly identify which of these two situations we are dealing with. That is, formally, we consider we can always observe a binary random variable U where $U = 1$ indicates we are in a *partially implemented ITR situation*, and $U = 0$ indicates we are in a *new ITR situation*. This variable U allows us to precisely formulate the link between observed random variables and potential outcomes as

$$Y = UY^* + (1 - U)Y^{s=0},$$

$$S = US^*,$$

$$A = UA^* + (1 - U)A^{s=0}.$$

In this work, we do not consider counterfactuals with respect to U . Rather, to evaluate the effect of an ITR, we focus on the variables $Y^{s=0}$, $Y^{s=1}$, and Y^* . For clarity, in the remainder of this paper, we avoid making reference to U . To identify causal effects, we rely on the subsequent assumptions of exclusion restriction, exchangeability, and overlap.

Assumption 1 (exclusion restriction). The effect of the ITR on the outcome Y is only mediated through the treatment, that is,

$$Y^{s=1} = A^{s=1}Y^{a=1} + (1 - A^{s=1})Y^{a=0}, \quad (1)$$

$$Y^{s=0} = A^{s=0}Y^{a=1} + (1 - A^{s=0})Y^{a=0}, \quad (2)$$

$$A^* = S^*A^{s=1} + (1 - S^*)A^{s=0}, \quad (3)$$

$$Y^* = A^*Y^{a=1} + (1 - A^*)Y^{a=0}. \quad (4)$$

Assumption 2 (exchangeability). All confounders and variables causing implementation are measured, that is,

$$\{Y^{a=1}, Y^{a=0}\} \perp\!\!\!\perp A^{s=0}|X, \quad (5)$$

$$\{Y^{a=1}, Y^{a=0}\} \perp\!\!\!\perp A^*|X, \quad (6)$$

$$A^{s=0} \perp\!\!\!\perp S^*|X. \quad (7)$$

In the *new ITR situation*:

- equation (5) has the same interpretation as the classic no unmeasured confounder assumption,
- the interpretation of equation (6) is that we already record all the confounding variables that will be in effect under the future implementation of the ITR,^b
- the interpretation of equation (7) is that we have measured the variables that may cause future implementation of the ITR.

In the *partially implemented ITR situation*:

- equation (6) has the same interpretation as the classic no unmeasured confounder assumption,
- the interpretation of equation (5) is that, for the patients whose physician did not implement the ITR, all the confounding variables are measured,
- the interpretation of equation (7) is that we have measured the variables that cause the current implementation of the ITR.

Assumption 3 (overlap). Within all realistic levels of covariates, the patients could receive either treatment—including in the absence of ITR implementation. That is, denoting $\pi^{s=0}$ the propensity score in the absence of ITR implementation, i.e., $\pi^{s=0}(x) = \mathbb{E}[A^{s=0}|X = x]$,

$$\forall x \in \mathcal{X}, \quad 0 < \pi(x) < 1, \quad \text{and} \quad 0 < \pi^{s=0}(x) < 1.$$

^bAssuming that we measured all the prognostic variables and that these will not have changed in the future is enough to satisfy this assumption.

Observing the overlap assumption, we see that as the propensity score functions π and $\pi^{s=0}$ can never be deterministic rules, they are thus stochastic rules. We define two additional stochastic rules: the propensity score under stochastic implementation π^* that is $\pi^*(x) = \mathbb{E}[A^*|X = x]$, and the stochastic implementation function under implementation ρ^* i.e., $\rho^*(x) = \mathbb{E}[S^*|X = x]$.

Summarizing the exclusion restriction equations (1), (2), (3), (4) and the exchangeability equations (5), (6), (7), the data generating mechanism in the *new ITR* and the *partially implemented ITR* situations can be represented by the probabilistic graphical models in Figure 1A and 1B, respectively.

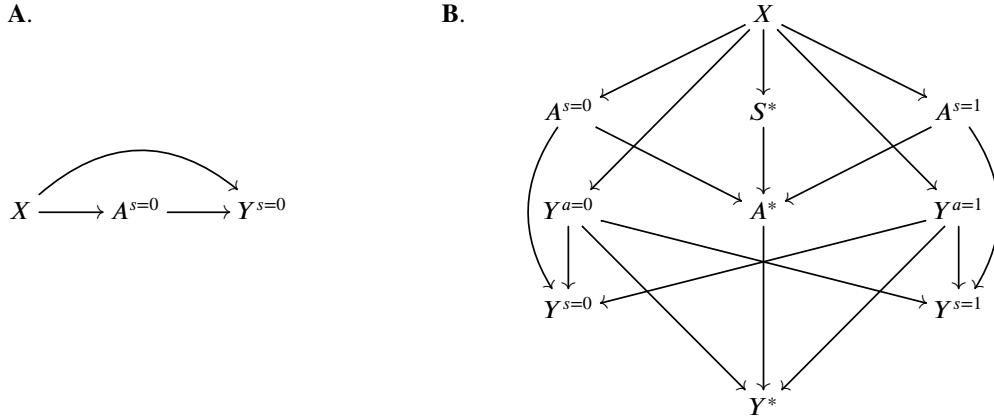


Figure 1 The probabilistic graphical model associated with the data generating mechanism in the *new ITR situation* (Panel A) and the *partially implemented ITR situation* (Panel B).

In our setup, it will prove convenient to define q_1 , the prognostic function under ITR implementation, as $q_1(x) = \mathbb{E}[Y^{s=1}|X = x]$ and q_0 , the prognostic function in the absence of ITR implementation, as $q_0(x) = \mathbb{E}[Y^{s=0}|X = x]$. Conditioning equation (1) with respect to X leads to $q_1(x) = r(x)\mu_1(x) + \{1 - r(x)\}\mu_0(x)$.

2.2 | Estimands of interest

We now introduce three estimands. First, the *Average Rule Effect* (ARE) of an ITR r :

$$\Delta(r) = \mathbb{E}[Y^{s=1} - Y^{s=0}].$$

This represents the population-level effect of the ITR on outcome Y in a randomized trial comparing a group where patients are systematically given the treatment recommended by ITR to usual care in the absence of ITR implementation.

Second, we define the *Average Implementation Effect* (AIE) of r as

$$\Lambda(r, \rho^*) = \mathbb{E}[Y^* - Y^{s=0}].$$

This represents the population-level effect of the ITR on outcome Y in a randomized trial comparing a group where physicians are provided with the ITR's treatment recommendation to usual care under no implementation. We may thus consider that in the experimental group the treatment A is prescribed according to a stochastic implementation ρ^* of r . Note that the ARE can be considered as a special case of the AIE where the stochastic implementation is perfect ($\rho^* \equiv 1$). We however single it out because this representation allows to assess whether the ITR has a potential population-level benefit, or if it is instead poorly designed.

Last, we define the *Maximal Implementation Gain* (MIG) of r :

$$\Gamma(r, \rho^*) = \mathbb{E}[Y^{s=1} - Y^*].$$

This represents the difference in average outcome between a full implementation of the ITR and the current or future partial implementation of the rule. From these definitions, it follows that

$$\Delta(r) = \Lambda(r, \rho^*) + \Gamma(r, \rho^*).$$

3 | INFERENCE

We assume we are given a single random sample of n independent and identically distributed (i.i.d.) units $(X_i^T, A_i, Y_i)_{1 \leq i \leq n}$ from a target population. As in the previous section, we distinguish between data originating from *new* and *partially implemented* ITR situations:

- A. In data originating from the *new ITR situation*, we have $S_i = 0$, $A_i = A_i^{s=0}$, and $Y_i = Y_i^{s=0}$. Note that the second equality implies $\pi = \pi^{s=0}$. In this situation, we drop all $(-)^{s=0}$ superscripts and use π rather than $\pi^{s=0}$ for clarity. Clearly, estimation of the AIE and MIG is not possible from data alone in this situation. Nonetheless, in the next section, we propose to explore their behaviour by hypothesizing implementation schemes.
- B. In data originating from the *partially implemented ITR situation*, we have $S_i = S_i^*$, $A_i = A_i^*$, and $Y_i = Y_i^*$. Note that the first two equalities imply $\rho = \rho^*$ and $\pi = \pi^*$. In this situation, we thus drop all $(-)^*$ superscripts for clarity. Because we expect that S_i will not have been collected in this situation, we treat it as a latent variable.

3.1 | New individualized treatment rule situation

3.1.1 | Average rule effect

Using exclusion restriction (1), exchangeability (2), and positivity (3), in the *new ITR situation*, we have

$$\begin{aligned}\Delta(r) &= \mathbb{E}[q_1(X) - Y] \\ &= \mathbb{E}\left[\left\{r(X)\frac{A}{\pi(X)} + \{1 - r(X)\}\frac{1 - A}{1 - \pi(X)} - 1\right\}Y\right].\end{aligned}$$

These equations suggest the following two estimators for $\Delta(r)$

$$\hat{\Delta}_Q(r) = n^{-1} \sum_{i=1}^n r(X_i)\hat{\mu}_1(X_i) + \{1 - r(X_i)\}\hat{\mu}_0(X_i) - Y_i, \quad (8)$$

$$\hat{\Delta}_{IPW}(r) = n^{-1} \sum_{i=1}^n \left[r(X_i) \frac{A_i}{\hat{\pi}(X_i)} + \{1 - r(X_i)\} \frac{1 - A_i}{1 - \hat{\pi}(X_i)} - 1 \right] Y_i, \quad (9)$$

where as for all estimators proposed hereafter, $\mu_0(\cdot)$, $\mu_1(\cdot)$, and $\pi(\cdot)$ can be estimated via any supervised learning method from observations $(X_i^\mu, Y_i)_{i: A_i=0}$, $(X_i^\mu, Y_i)_{i: A_i=1}$, and $(X_i^\pi, A_i)_{1 \leq i \leq n}$ respectively.^c An augmented counterpart of these estimators can be derived from Zhang et al.¹⁹:

$$\hat{\Delta}_{AIPW}(r) = n^{-1} \sum_{i=1}^n \left[\frac{C_i^r Y_i}{\hat{\pi}(X_i)C_i^r + \{1 - \hat{\pi}(X_i)\}(1 - C_i^r)} - \frac{C_i^r - [\hat{\pi}(X_i)C_i^r + \{1 - \hat{\pi}(X_i)\}(1 - C_i^r)]}{\hat{\pi}(X_i)C_i^r + \{1 - \hat{\pi}(X_i)\}(1 - C_i^r)} \hat{q}_1(X_i) - Y_i \right] \quad (10)$$

^cHere, X_i^μ and X_i^π denote two subsets of the relevant variables contained in X_i .

where we set $C_i^r = \mathbb{1}\{r(X_i) = A_i\}$ and $\hat{q}_1(X_i) = r(X_i)\hat{\mu}_1(X_i) + \{1 - r(X_i)\}\hat{\mu}_0(X_i)$ for clarity. We refer the reader to Tsiatis et al.²⁰ section 3.3.3 p. 61 for an extensive study of this specific case and the derivation of approximate large sample distribution. Using the CATE, the ARE can also be reformulated as

$$\Delta(r) = \mathbb{E} [\{r(X) - \pi(X)\}\tau(X)] \quad (11)$$

(a proof is given in Appendix A). This leads to the following estimator

$$\hat{\Delta}_{CATE}(r) = n^{-1} \sum_{i=1}^n \{r(X_i) - \hat{\pi}(X_i)\}\hat{\tau}(X_i).$$

Though the latter estimator requires to estimate the CATE τ , and hence may be less practical than estimators (8), (9), or (10), the equation (11) makes explicit the respective contribution of τ , r and π to the ARE.

3.1.2 | AIE, MIG under the modeling of the stochastic implementation functions

When the ITR is new and has never been deployed, the way in which it will be implemented is unpredictable. Hence, the AIE and the MIG cannot be estimated from data alone. However, it can be interesting to study numerically how the AIE and MIG would vary under different stochastic implementation schemes as this can provide information about the appropriateness of future ITR deployment. In the *new ITR situation*, it is possible to show that

$$\Lambda(r, \rho^*) = \mathbb{E} [\{\pi^*(X) - \pi(X)\}\tau(X)] \quad (12)$$

$$= \mathbb{E} [\rho^*(X)\{r(X) - \pi(X)\}\tau(X)], \quad (13)$$

and

$$\begin{aligned} \Gamma(r, \rho^*) &= \mathbb{E} [\{r(X) - \pi^*(X)\}\tau(X)] \\ &= \mathbb{E} [\{1 - \rho^*(X)\}\{r(X) - \pi(X)\}\tau(X)] \end{aligned}$$

(a proof is given in Appendix B). Hence, given an estimate $\hat{\tau}$ of the CATE function see Jacob²¹ for a review of the available estimation methods, an estimate $\hat{\pi}$ of the propensity score and a numerical model of ρ^* for a future stochastic implementation function, estimates of the AIE and the MIG are computable via

$$\hat{\Lambda}_{CATE}(r, \rho^*) = n^{-1} \sum_{i=1}^n \rho^*(X_i)\{r(X_i) - \hat{\pi}(X_i)\}\hat{\tau}(X_i),$$

and

$$\hat{\Gamma}_{CATE}(r, \rho^*) = n^{-1} \sum_{i=1}^n \{1 - \rho^*(X_i)\}\{r(X_i) - \hat{\pi}(X_i)\}\hat{\tau}(X_i).$$

Below, we propose, three schemes that model the form the stochastic implementation function may take in future deployment of the ITR:

- The random implementation scheme, where we model $\rho^*(\cdot)$ as

$$\rho_{rd,\alpha}^*(x) = \alpha$$

with $\alpha \in [0; 1]$ a parameter modelling the random implementation such that, uniformly for all patients, higher values of α are associated with higher probabilities of following the rule. This model of ρ^* describes a situation where patients are treated according to an implementation of the rule at random with probability α regardless of their characteristics.

- The cognitive bias scheme, where we model $\rho^*(\cdot)$ as

$$\rho_{cb,\alpha}^*(x) = \{1 - |r(x) - \pi(x)|\}^{\frac{1}{2} \log \frac{\alpha+1}{1-\alpha}}$$

with $\alpha \in [0; 1]$ a cognitive bias parameter such that higher values of α are associated with lower probabilities of following the rule for a given gap between the recommendation from the ITR and usual care under no implementation. This implementation scheme describes a situation where physicians follow the ITR recommendation more often when recommendations are similar to current practices and this trend to resist change increases as α increases.

- The confidence level scheme, where we assume that the ITR was constructed from estimated CATEs, $\tilde{\tau}(x)$, as in for instance $r(x) = \mathbb{1}[\tilde{\tau}(x) < 0]$, when $Y = 1$ denotes mortality. For this scheme to be actionable, $\tilde{\tau}(x)$ and their standard errors $se_{\tilde{\tau}(x)}$ must be provided along the ITR they helped build. Under such conditions, we model $\rho^*(\cdot)$ as

$$\rho_{cl,\alpha}^*(x) = \mathbb{1}\{[\tilde{\tau}(x) - q_{1-\alpha/2}se_{\tilde{\tau}(x)}]\{\tilde{\tau}(x) + q_{1-\alpha/2}se_{\tilde{\tau}(x)}\} > 0\}$$

for CATEs provided on an absolute scale (i.e., individual absolute risk difference) with $\alpha \in [0; 1]$ a type I error parameter such that smaller values of α lead to wider confidence intervals for $\tilde{\tau}(x)$. This scheme describes a situation where physicians follow the ITR recommendation only when there is evidence that $\tau(x) \neq 0$ at significance level α .

3.1.3 | Illustrative examples

In this section, we aim to provide a sense of what our method is trying to achieve when applied in the *new ITR situation*. For that purpose, in this subsection, we provide a toy model. Observing Formula (12), we see that the AIE of a new ITR r gets far off from zero as the difference between current treatment allocation and future treatment allocation under a stochastic implementation of r increases. More precisely, observing Formula (13), we note that the AIE of a new ITR r gets far off from zero as patients with common levels of covariates x have i) a high probability $\rho^*(x)$ of implementing the rule, and/or ii) a difference $r(x) - \pi(x)$ between recommendation from the ITR and usual care under no implementation far off from zero, and/or iii) large CATEs $\tau(x)$.

For illustration purposes, we imagined a disease for which only one patient characteristic, the age x , is relevant to treatment decision-making. In a population of patients with mean age 50 (standard deviation 15), we wish to evaluate the effectiveness of an ITR r with respect to the occurrence of an unfavorable binary outcome (i.e., 10-year mortality). In our two examples, ground truth is such that the treatment is beneficial for patients aged 40 to 60, detrimental for patients aged 60 to 80, and has almost no effect outside these ranges. For the sake of simplicity, we suppose that in both examples r is $r(X) = \mathbb{1}[\tau(X) < 0]$ that is, the rule is optimal (Figure 2 Panels A and B).

In our first example (Figure 2A), the usual care under no implementation is such that younger patients are treated more often while in our second example (Figure 2B), older patients are treated more often. In the random implementation schemes (Figure 2 Panels C and D), physicians follow the ITR's recommendation at random with probability 1/3 (red lines) or 2/3 (green lines). In the cognitive bias schemes (Figure 2 Panels E and F), physicians follow the ITR's recommendation more often when the ITR's recommendation tracks the usual care under no implementation. Cognitive bias parameter is 2/3 (red lines) or 1/3 (green lines), and higher parameter values are associated with lower probabilities of complying with the ITR. In the confidence level schemes (Figure 2 Panels G and H), physicians follow the ITR's recommendation only when confidence intervals for the predicted CATEs do not cross zero. Type I error parameters for the confidence intervals are 0.05 (red lines) or 0.45 (green lines) with higher values associated with tighter confidence intervals and therefore higher probabilities of implementing the ITR.

Despite the fact that both examples relied on the implementation of an identical ITR based on the true CATE function, the population-level benefit of this ITR is different between examples for all schemes. The ARE of the deterministic rule was -0.16 in the population from example 1 and -0.24 in the population from example 2, indicating an 8% greater benefit of implementing the ITR in population 2 than in population 1 if physicians always followed the ITR's recommendation. Similarly, in the stochastic implementation schemes, the population-level benefit of the ITR differ in the two example populations. In the random implementation scheme, AIEs are -0.11 and -0.05 in population 1 (Figure 2C) versus -0.16 and -0.08 in population

2 (Figure 2D) for random implementation parameters 2/3 and 1/3 respectively. We find similar differences in AIEs between population 1 and population 2 in the cognitive bias scheme (Figure 2 Panels E and F) and confidence level scheme (Figure 2 Panels G and H).

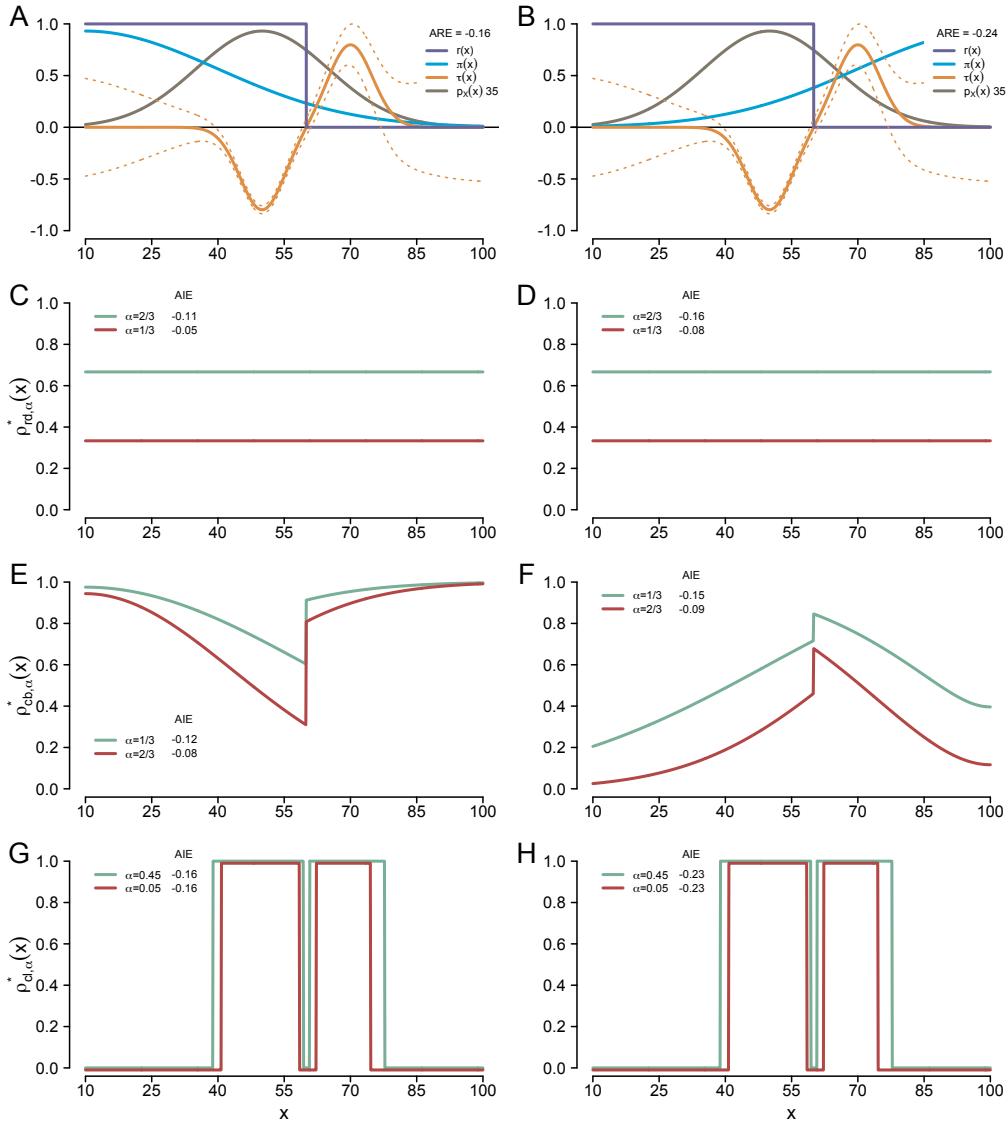


Figure 2 Panel A displays our first illustrative example where under no implementation usual care is such that younger patients are treated more often. Panel B displays our second illustrative example where under no implementation usual care is such that older patients are treated more often. The random implementation scheme is given panels C and D for both examples, respectively, the cognitive bias scheme on panels E and F, and the confidence level scheme panels G and H. The dotted lines correspond to the CATEs plus/minus its standard errors. AIEs are reported for each of implementation schemes and lower values indicate greater benefit from ITR implementation. We denote p_X the probability density of X (we re-scale $p_X(x)$ by a factor 35 for illustration purposes).

3.2 | Partially implemented individualized treatment rule situation

In this subsection, our aim is to estimate the ARE, AIE and MIG with data sampled from a population where the ITR r is partially implemented. Two cases have to be distinguished depending on whether the variable S is collected. If S is collected, estimating the ARE, AIE and MIG can be achieved by using a suitable adaptation of the IPW/AIPW estimators for the average treatment effect.²² However, in practice, we expect that the variable S will not have been collected. Hence, we focus our attention on the case where S needs to be regarded as a latent variable. Recall that in this subsection, we are dealing with a partially implemented ITR where the observed treatment follows either from the ITR being implemented or from the physicians disregarding the ITR to make treatment decisions. Inference in the partially implemented ITR situation relies on assumptions (1-3). Because in this situation estimation of the MIG is more straightforward than estimation of the ARE and AIE, we distinguish between the two cases.

3.2.1 | Maximal implementation gain

We start by studying the MIG, as neither S nor ρ play a role for this estimand in the *partially implemented ITR situation*. In fact in this situation, using exclusion restriction (1), exchangeability (2), and positivity (3), we have

$$\begin{aligned}\Gamma(r, \rho) &= \mathbb{E} [q_1(X) - Y] \\ &= \mathbb{E} \left[\left\{ r(X) \frac{A}{\pi(X)} + \{1 - r(X)\} \frac{1 - A}{1 - \pi(X)} - 1 \right\} Y \right].\end{aligned}$$

This suggests the estimators

$$\hat{\Gamma}_Q(r, \rho) = n^{-1} \sum_{i=1}^n r(X_i) \hat{\mu}_1(X_i) + \{1 - r(X_i)\} \hat{\mu}_0(X_i) - Y_i,$$

and

$$\hat{\Gamma}_{IPW}(r, \rho) = n^{-1} \sum_{i=1}^n \left[r(X_i) \frac{A_i}{\hat{\pi}(X_i)} + \{1 - r(X_i)\} \frac{1 - A_i}{1 - \hat{\pi}(X_i)} - 1 \right] Y_i.$$

The derivation is similar to that of the ARE in the *new ITR situation* (equations 8 and 9). We refer the reader to section 3.1.1, equation (10) for an augmented version of this estimator.

3.2.2 | Average rule effect and average implementation effect

Note that the MIG estimand is distinct from the ARE and AIE in that it does not involve the expectation term $\mathbb{E}(Y^{s=0})$. In contrast, the ARE and AIE depend on the pairs of expectations $\mathbb{E}(Y^{s=1})$, $\mathbb{E}(Y^{s=0})$ and $\mathbb{E}(Y)$, $\mathbb{E}(Y^{s=0})$ respectively. The quantity $\mathbb{E}(Y)$ is straightforward to estimate. The expectation $\mathbb{E}(Y^{s=1})$ can be estimated by various means, for instance by taking the expectation of $\hat{q}_1(X)$ as in section 3.2.1.

Estimation of $\mathbb{E}(Y^{s=0})$ is more challenging than that of $\mathbb{E}(Y^{s=1})$ because, substitution of $Y^{s=0}$ by its definition in (2) involves the potential outcome $A^{s=0}$ which is not directly identifiable from equation (3) as S is a latent variable. Our approach to estimate $\mathbb{E}(Y^{s=0})$ relies on the following result.

Lemma 1

In the partially implemented ITR situation, the following relations holds

$$(i) \quad \pi(x) = \rho(x)r(x) + \{1 - \rho(x)\}\pi^{s=0}(x), \tag{14}$$

$$(ii) \quad q_0(x) = \pi^{s=0}(x)\mu_1(x) + \{1 - \pi^{s=0}(x)\}\mu_0(x). \tag{15}$$

A proof of the lemma is given in Appendix C. Lemma 1 suggests that the functions π and q_0 may be represented by a particular type of mixture model called the mixture of experts model.²³ This model lays out a mixture of regression models (experts) where

the proportions of the mixture (gating network) depend on the covariates. Observing equation (14), we see that π can be viewed as a mixture of the known expert r and the unknown expert $\pi^{s=0}$, while the proportions of the mixture are given by the unknown gating network ρ . Several authors, including Teicher;²⁴ Jiang and Tanner;²⁵ Allman et al.²⁶ have studied the identifiability of mixture models. However, to the best of our knowledge, their results are not directly applicable to the specific type of mixture we are considering in equation (14). In Appendix D, we prove, under mild technical conditions on the treatment rule r , that if the covariates are continuous then the mixtures of experts considered in this paper are identifiable. Equation (15) suggest to rewrite $\Delta(r)$ and $\Lambda(r, \rho)$ as

$$\begin{aligned}\Delta(r) &= \mathbb{E} [q_1(X) - q_0(X)] \\ &= \mathbb{E} [\{r(X) - \pi^{s=0}(X)\}\tau(X)]\end{aligned}$$

and

$$\begin{aligned}\Lambda(r, \rho) &= \mathbb{E} [Y - q_0(X)] \\ &= \mathbb{E} [Y - \mu_1(X)\pi^{s=0}(X) - \mu_0(X)\{1 - \pi^{s=0}(X)\}].\end{aligned}$$

Because $\pi^{s=0}$ is unknown, we propose to estimate it via the procedure detailed in Algorithm 1. This procedure details an EM algorithm, based on the fitting algorithm of Xu and Jordan²⁷ and Jordan and Jacobs²³ where we posit parametric models for $\pi^{s=0}(\cdot)$ and $\rho(\cdot)$. In Appendix E, we detail how this algorithm was derived and how it could readily be extended to the case of non-parametric models. Figure 3 depicts the graphical model for the approach to estimating $\pi^{s=0}(\cdot)$. Estimating $\mu_0(\cdot)$, $\mu_1(\cdot)$, and $\tau(\cdot)$ as in section 3.1, mixture of experts estimators for the ARE and AIE are then given by

$$\hat{\Delta}_{ME}(r) = n^{-1} \sum_{i=1}^n \{r(X_i) - \hat{\pi}^{s=0}(X_i)\}\hat{\tau}(X_i),$$

and

$$\hat{\Lambda}_{ME}(r, \rho) = n^{-1} \sum_{i=1}^n Y_i - \hat{\mu}_1(X_i)\hat{\pi}^{s=0}(X_i) - \hat{\mu}_0(X_i)\{1 - \hat{\pi}^{s=0}(X_i)\}.$$

Assuming that all relevant parametric models are correctly specified, we note that each estimator $\hat{\Delta}_{ME}(r)$ and $\hat{\Lambda}_{ME}(r, \rho)$ jointly solve a set of “stacked” estimating equations. In particular, the parameters of $\pi^{s=0}(\cdot)$ jointly solve the derivative of the log-likelihood given in equation (E2) from the Appendix. The $\hat{\Delta}_{ME}(r)$ and $\hat{\Lambda}_{ME}(r, \rho)$ estimators can thus be viewed as M-estimators and it follows that under correct model specification, they are \sqrt{n} -consistent and asymptotically normal with asymptotic variances calculable via the empirical sandwich estimator. For a clear review of M-estimation and the stacked estimating equation method, we refer the reader to Stefanski and Boos.²⁸ For simplicity, in the remainder of this paper, we propose to estimate the variance of the estimators $\hat{\Delta}_{ME}(r)$ and $\hat{\Lambda}_{ME}(r, \rho)$ via the bootstrap. We assess the validity of this strategy in Monte Carlo simulations.

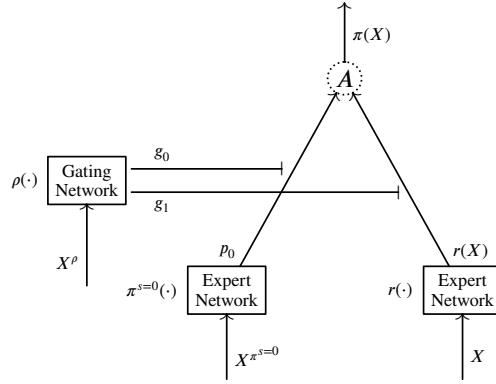


Figure 3 The graphical representation of the mixture of experts fitted by Algorithm 1. Note that we consider r as a known deterministic expert network, while both the expert network $\pi^{s=0}$ and the gating network ρ are unknown stochastic rules.

Algorithm 1 The EM procedure for estimating $\pi^{s=0}(\cdot)$ in the partially implemented ITR situation.

Input: The ITR $r : \mathcal{X} \mapsto \{0; 1\}$, and data $(X_i^T, A_i)_{1 \leq i \leq n}$ where $X_i^{\pi^{s=0}}$ and X_i^ρ are two relevant subsets of the variables contained in X_i .

Initialize the prior probabilities associated with the nodes of the tree as $g_{0,i} \leftarrow 0.5$, and $g_{1,i} \leftarrow 0.5$.

Initialize the parameters ζ of the expert $\pi^{s=0}(\cdot)$ at random e.g., $\zeta \sim \mathcal{N}(0, D)$ with D a diagonal matrix.

Compute individual contributions to r 's likelihood as $P_{1,i} \leftarrow r(X_i)^{A_i} \{1 - r(X_i)\}^{1-A_i}$.

Compute individual predictions from the initiated expert network $\pi^{s=0}(\cdot)$ as $p_{0,i} \leftarrow \text{expit}(\zeta^T X_i^{\pi^{s=0}})$.

Iterate until convergence on the parameters ζ :

Compute individual contributions to $\pi^{s=0}$'s likelihood as $P_{0,i} \leftarrow p_{0,i}^{A_i} (1 - p_{0,i})^{1-A_i}$.

Compute the posterior probabilities associated with the nodes of the tree as

$$h_{0,i} \leftarrow \frac{g_{0,i} P_{0,i}}{g_{0,i} P_{0,i} + g_{1,i} P_{1,i}} \quad \text{and} \quad h_{1,i} \leftarrow \frac{g_{1,i} P_{1,i}}{g_{0,i} P_{0,i} + g_{1,i} P_{1,i}}.$$

▷ E-step

For the gating network $\rho(\cdot)$ estimate parameters γ by solving the IRLS problem

$$\gamma \leftarrow \arg \max_{\gamma} \sum_{i=1}^n h_{1,i} \ln \{ \text{expit}(\gamma^T X_i^\rho) \} + (1 - h_{1,i}) \ln \{ 1 - \text{expit}(\gamma^T X_i^\rho) \}$$

▷ M-step

For the expert network $\pi^{s=0}(\cdot)$ estimate parameters ζ by solving the IRLS problem

$$\zeta \leftarrow \arg \max_{\zeta} \sum_{i=1}^n h_{0,i} \left[A_i \ln \{ \text{expit}(\zeta^T X_i^{\pi^{s=0}}) \} + (1 - A_i) \ln \{ 1 - \text{expit}(\zeta^T X_i^{\pi^{s=0}}) \} \right]$$

Update the prior probabilities associated with the nodes of the tree as

$$g_{1,i} \leftarrow \text{expit}(\gamma^T X_i^\rho) \quad \text{and} \quad g_{0,i} \leftarrow 1 - g_{1,i}.$$

Update the predictions from the expert network $\pi^{s=0}(\cdot)$ as $p_{0,i} \leftarrow \text{expit}(\zeta^T X_i^{\pi^{s=0}})$.

Return: $\hat{\pi}^{s=0}(x) = \text{expit}(\zeta^T x)$.

4 | SIMULATIONS

4.1 | Setup

In this section, we study the properties of the MIG, ARE and AIE estimators in the partially implemented ITR situation. To this end, we simulate data analysis in a setting where an ITR was partially implemented. We generate synthetic datasets comprising six Bernoulli, log-normally and normally distributed, correlated, covariates $X = (X_1, X_2, \dots, X_6)$ as follows.

Step 1 We randomly generate correlated intermediate covariates X'_1, X'_2, \dots, X'_6 from a multivariate gaussian distribution

$$(X'_1, X'_2, \dots, X'_6)^T \sim \mathcal{N}(0, \Sigma).$$

To generate Σ , we chose 6 eigenvalues $(\lambda_1, \lambda_2, \dots, \lambda_6) = (1, 1.2, 1.4, 1.6, 1.8, 2)$, and sample a random orthogonal matrix O of size 6×6 . The covariance matrix Σ is obtained via

$$\Sigma = O \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_6 \end{bmatrix} O^T.$$

Step 2 To allow for the Bernoulli or log-normal distribution of covariates, we generate X_1, X_2, \dots, X_6 as follows $(X_1, X_2) = (\mathbb{1}\{X'_1 < 0\}, \mathbb{1}\{X'_2 < 0\})$, $(X_3, X_4, X_5) = (\exp(X'_3), \exp(X'_4), \exp(X'_5))$, $X_6 = X'_6$. We add $X_0 \equiv 1$ to allow for intercepts.

Step 3 We generate data from the covariates in this manner:

$$\begin{aligned} S|X = x &\sim \text{Bernouilli}(\text{expit}(\gamma^T x)), & Y^{a=0}|X = x &\sim \text{Bernouilli}(\text{expit}(\alpha^T x)), \\ r(X) = \mathbb{1}\{\delta^T x < 0\}, & & Y^{a=1}|X = x &\sim \text{Bernouilli}(\text{expit}(\beta^T x)), \\ A^{s=0}|X = x &\sim \text{Bernouilli}(\text{expit}(\zeta^T x)), & Y^{s=0} &= A^{s=0}Y^{a=1} + (1 - A^{s=0})Y^{a=0}, \\ A^{s=1} &= r(X), & Y^{s=1} &= A^{s=1}Y^{a=1} + (1 - A^{s=1})Y^{a=0}, \\ A &= SA^{s=1} + (1 - S)A^{s=0}, & Y &= AY^{a=1} + (1 - A)Y^{a=0} \end{aligned}$$

with

$$\begin{aligned} \gamma &= (0, 0, 0, 0, 0, 0, 1)^T, & \delta &= (0.05, -0.5, 0.5, -0.5, 0.5, 0, 0)^T, \\ \alpha &= (0, -0.3, -0.05, 0.5, -0.15, -0.2, 0)^T, & \beta &= (0, -0.2, 0.05, 0.3, -0.1, -0.1, 0)^T \end{aligned}$$

and we vary ζ .

In scenario A, we set $\zeta = \delta$ which corresponds to a situation where treatment allocation in the absence of the ITR is different from the ITR. In scenario B, we set $\zeta = (0, 0, 0, 0, 0, 0)$ which corresponds to a situation where treatment allocation in the absence of the ITR is random with probability 0.5. In scenario C, we set $\zeta = -\delta$ which corresponds to a situation where treatment allocation in the absence of the ITR resembles the ITR. In each scenario we generate a target population of two million individuals from which we approximate ground truth for our estimands and drew random samples. We vary the sample size: $n = 200, 800, 2000$. The potential outcomes as well as the variable S are regarded as unobserved variables. Models for μ_0 and μ_1 are correctly specified with $X^\mu = (X_2, X_3, X_4, X_5, X_6)$ as explanatory variables. We fit the mixture of expert in equation (14) with Algorithm 1, specifying the gating network ρ with $X^\rho = X_6$ and the expert network $\pi^{s=0}$ with variables

$X^{\pi^{s0}} = (X_1, X_2, X_3, X_4, X_5)$. For each scenario/sample size combination, we implement 1000 simulation iterations and 999 bootstrap replications to generate confidence intervals.

4.2 | Results

The results of our simulations are reported in Table 2 and Figure 4. The MIG estimator $\hat{\Gamma}_Q(r, \rho)$, which does not rely on an EM procedure, exhibits its theoretical properties of unbiasness and consistency. The ARE estimator $\hat{\Delta}_{ME}(r)$ and the AIE estimator $\hat{\Lambda}_{ME}(r, \rho)$ which both rely on the EM procedure also appear unbiased and consistent. Their standard error is comparable to that of the MIG estimator $\hat{\Gamma}_Q(r, \rho)$. Ninety five percent bootstrap confidence intervals achieve close to nominal coverage for all three estimators.

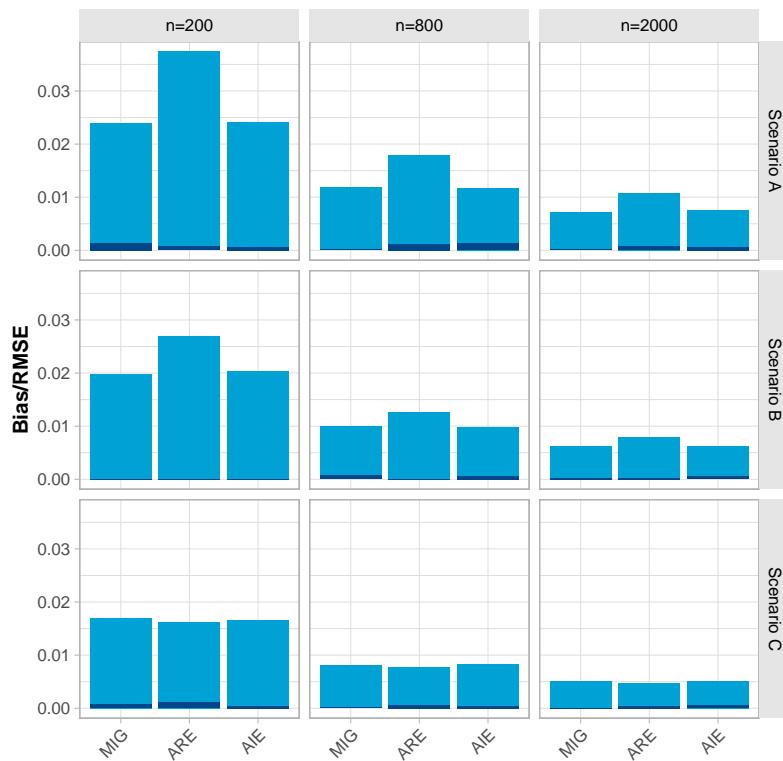


Figure 4 Absolute bias and Root Mean Square Error (RMSE) for the MIG, ARE and AIE estimators across 1000 simulation iterations in nine scenario/sample size combinations. Absolute bias is the darker portion of each bar ; RMSE corresponds to the total bar size. The MIG, ARE and AIE estimators are from $\hat{\Gamma}_Q(r, \rho)$, $\hat{\Delta}_{ME}(r)$, $\hat{\Lambda}_{ME}(r, \rho)$ respectively.

5 | APPLICATIONS ON THE MIMIC-III DATABASE

The Multi-Parameter Intelligent Monitoring in Intensive Care III (MIMIC-III) database is a publicly available electronic health record that contain data from 53,423 patients hospitalized in intensive care at Beth Israel Deaconess Medical Center from 2001 to 2012.²⁹ From these, we include the 3,748 intensive care unit adult patients with severe acute kidney injury who had received

either invasive mechanical ventilation or vasopressor infusion. We report the full inclusion/exclusion criteria in the Appendix F and the inclusion flow-diagram in the Appendix G. The patients we include are eligible for recommendation from both ITRs described below. For the sake of focusing on the estimators in our methodology, we handle patients with missing data by conducting a single imputation using chained equations.³⁰

In this section, we consider two example ITRs. In the first example, we evaluate a new ITR for dialysis initiation.^d This last ITR was not available at the time of data collection and decision to initiate dialysis never followed from its implementation. In the second example, we evaluate an ITR that was partially implemented at the time of data collection. Specifically, we evaluate the impact of an ITR that recommends initiating dialysis in the most severe patients based on the Sequential Organ Failure Assessment (SOFA) score.³¹

Table 2 Simulation results for the MIG, AIE and ARE estimators under all nine scenario/sample size combinations. The MIG, ARE and AIE estimators are from $\hat{\Gamma}_Q(r, \rho)$, $\hat{\Delta}_{ME}(r)$, $\hat{\Lambda}_{ME}(r, \rho)$ respectively. SE: standard error; RMSE: root mean squared error; CI: 95% confidence interval. Coverage probabilities are for 95% confidence intervals.

<i>n</i>	Scenario A			Scenario B			Scenario C		
	MIG	ARE	AIE	MIG	ARE	AIE	MIG	ARE	AIE
True value	-0.013	-0.026	-0.013	-0.008	-0.016	-0.008	-0.004	-0.007	-0.003
Bias									
200	-0.001	-0.001	0.001	0.000	0.000	0.000	-0.001	-0.001	0.000
800	0.000	0.001	0.001	-0.001	0.000	0.001	0.000	-0.001	0.000
2000	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	-0.001
Empirical SE									
200	0.024	0.037	0.024	0.020	0.027	0.020	0.017	0.016	0.017
800	0.012	0.018	0.012	0.010	0.013	0.010	0.008	0.008	0.008
2000	0.007	0.011	0.008	0.006	0.008	0.006	0.005	0.005	0.005
RMSE									
200	0.024	0.037	0.024	0.020	0.027	0.020	0.017	0.016	0.017
800	0.012	0.018	0.012	0.010	0.013	0.010	0.008	0.008	0.008
2000	0.007	0.011	0.008	0.006	0.008	0.006	0.005	0.005	0.005
Coverage									
200	0.964	0.954	0.937	0.953	0.976	0.941	0.936	0.979	0.944
800	0.945	0.953	0.949	0.945	0.956	0.947	0.953	0.957	0.949
2000	0.959	0.951	0.934	0.948	0.946	0.950	0.962	0.946	0.947
CI width									
200	0.099	0.153	0.096	0.080	0.113	0.082	0.067	0.073	0.067
800	0.047	0.071	0.046	0.039	0.050	0.039	0.033	0.031	0.033
2000	0.029	0.044	0.029	0.024	0.031	0.024	0.021	0.019	0.021

^dIn this section, we use the term “dialysis” loosely to refer to all kidney support therapies suitable for acute kidney injury patients i.e., including but not limited to intermittent hemodialysis and continuous hemofiltration.

5.1 | New ITR: dialysis initiation based on a combination of six biomarkers

Grolleau et al.³² have recently developed a new ITR for dialysis initiation in the intensive care unit using data from two RCTs. Briefly, this new ITR recommends initiating dialysis within 24 hours only in specific patients based on a combination of six biomarkers (SOFA score, pH, potassium, blood urea nitrogen, weight and, the prescription of immunosuppressive drug). Following the methodology detailed in section 3.1, we evaluate the impact of this new ITR on 60-day mortality. We estimate the ARE using a double robust estimator as detailed in section 3.1.1. We explore the impact of various degrees of implementation under either cognitive bias or confidence level schemes. The estimated values of AIE are given in Figure 5. Estimation of the ARE shows a trend for benefit from the implementation of the new ITR ($\hat{\Delta}_{AIPW}(r) = -0.02$; 95% confidence interval [-0.06 to 0.01]). Note that the ARE estimate is not equal to estimation of the AIE under full implementation as, contrary to the ARE case, for the AIE we use a CATE model. The variables included in each model are reported in Appendix H.

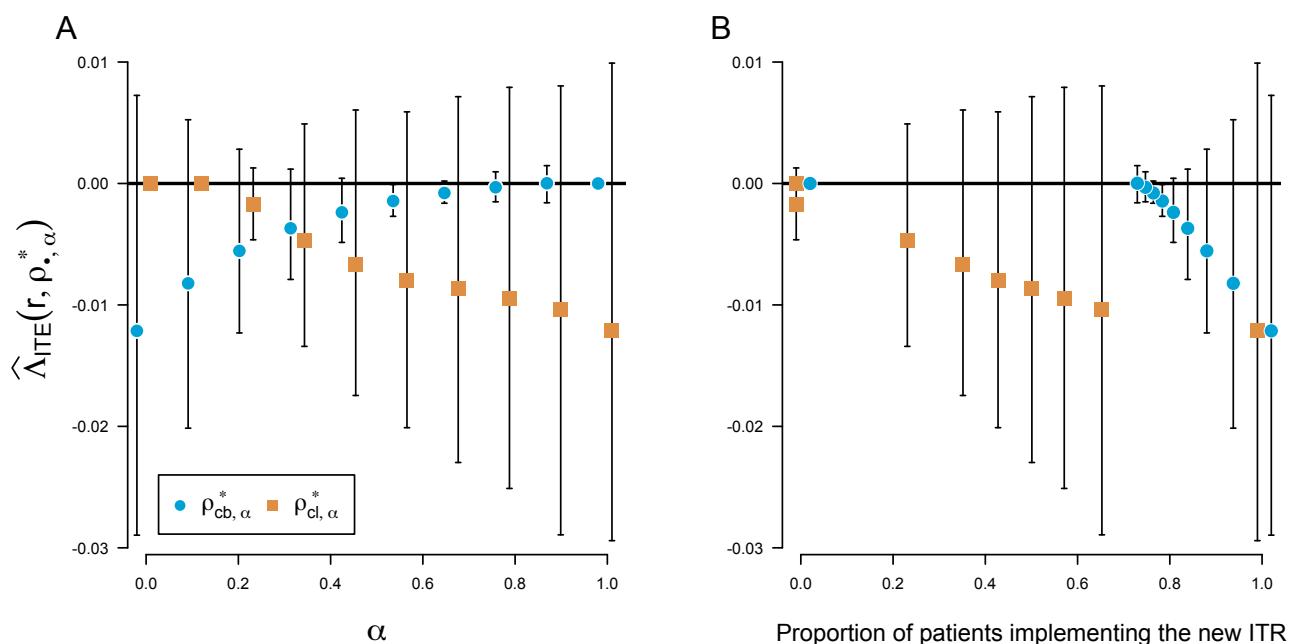


Figure 5 Evaluation of the impact of a new ITR (i.e., dialysis initiation within 24 hours only in specific patients based on a combination of six biomarkers) on 60-day mortality using the MIMIC-III observational database. Ninety-five percent confidence intervals are from the bootstrap. Blue diamonds are for the cognitive bias scheme; orange diamonds are for the confidence level scheme. Panels A depict the AIE for different values of implementation parameter α , Panels B depict the AIE as a function of the proportion of (future) patients implementing the new ITR: $n^{-1} \sum_{i=1}^n \rho_{\cdot, \alpha}^*(X_i)$. More negative values of the AIE indicate greater benefit from ITR implementation. Ninety-five percent confidence intervals are from the bootstrap.

5.2 | Partially implemented ITR: dialysis initiation based on SOFA scores

We evaluate the impact on 60-day mortality of a partially implemented, yet never evaluated, ITR that recommended initiating dialysis within 24 hours only in the patients with a Sequential Organ Failure Assessment (SOFA) score greater than 11. Following the methodology of section 3.2, we posit models for the treatment-specific prognosis functions, the propensity score in the

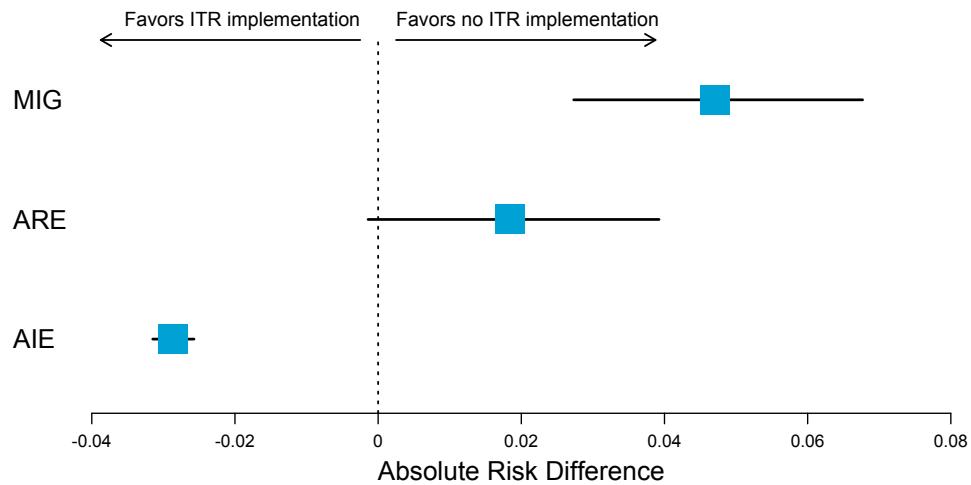


Figure 6 Evaluation of the impact of a partially implemented ITR (i.e., dialysis initiation within 24 hours only in the patients with a SOFA score greater than 11) on 60-day mortality using the MIMIC-III observational database. Ninety-five percent confidence intervals are from the bootstrap. MIG=Maximal Implementation Gain. ARE=Average Rule Effect. AIE=Average Implementation Effect.

absence of ITR implementation (expert network) and the stochastic implementation function (gating network). Specification of propensity score in the absence of ITR implementation include the variables thought to have caused treatment initiation while specification of the stochastic implementation function include all variables thought to be associated with ITR implementation. The variables included in each model are reported in I. The estimates of the MIG, ARE and AIE are given in Figure 6. Estimation of the MIG shows evidence of harm from further implementing the ITR ($\hat{\Delta}_Q(r, \rho) = 4.7\%$; 95% confidence interval [2.7% to 6.8%]). Similarly, estimation of the ARE shows a trend for harm in implementing the ITR in all patients versus in no one ($\hat{\Delta}_Q(r) = 1.8\%$; 95% confidence interval [-0.1% to 3.9%]) indicating that the ITR may be poorly designed. However, estimation of the AIE shows that the withdrawal of the ITR would on average yield outcomes worse than in the current situation ($\hat{\Delta}_Q(r, \rho) = -2.9\%$; 95% confidence interval [-3.1% to -2.6%]). This suggest that even though the ITR may be poorly designed, physicians identify correctly the patients who benefit from ITR implementation. In sum, these results indicate that neither full nor null implementation of the ITR would improve patient outcomes (at the population level). Rather, either one of these changes in ITR implementation, our analysis suggest, would worsen patient outcomes (at the population level). From a policy-maker standpoint, the best thing to do under such conditions may be to develop and subsequently evaluate a new ITR.

6 | DISCUSSION

Our goal was to construct an ecosystem for the evaluation of ITRs that will ultimately benefit patients. We believe that the probability model and inferential approach we introduced in this paper provide actionable tools to move this agenda forward. Below, we discuss some limitations of our approach.

In the *new ITR situation*, the exploration of the AIE relies on assuming future implementation schemes. Though sensible, the three implementation schemes we propose are subjective. Other realistic implementation schemes can be assumed and readily implemented in our methodology.

In the *partially implemented ITR situation*, inference relies on assuming a new probabilistic model. This model is largely inspired by the Neyman-Rubin causal model. As in the original model, our model requires assuming that the effect of an ITR is mediated only by the treatment prescribed by physicians to their patients. This exclusion restriction assumption may not hold in some specific settings. In cases where the decision to implement the ITR is taken by the patients—not the physicians—it is possible that merely seeing an ITR's recommendation affects outcomes. For instance, if an ITR recommends a patient treatment A, this patient may choose treatment B and compensate for not implementing the ITR by taking another effective treatment say C. This indirect effect of the ITR through treatment C would not be accounted for in our framework. With respect to the exchangeability assumption, our methodology relies on expert knowledge of the variables causing ITR implementation. This could include information about patients' physicians. For some research questions, the relevant variables may not be available thereby, limiting the usefulness of our approach. However, we do not believe that any statistical method can provide helpful workarounds under such conditions. Finally, as in the usual average treatment effect, one may be tempted to estimate the prognostic function q_0 directly, rather than estimating the propensity score π from a mixture model. As $Y_i^{s=0}$ are not observed, this would require to posit a hierarchical mixture of experts model. Though compelling at first glance, this approach may be impractical as hierarchical mixture of experts were shown to have likelihoods with arbitrary bad local maxima yielding EM algorithms sensitive to initialization conditions.³³ In contrast, for the simpler mixture considered in this paper, we were able to prove identifiability under mild technical conditions. Our simulations suggest that our mixture model's likelihood is well-behaved on finite samples so that an EM algorithm with a random initialization is successful with high probability in estimating the true value of the parameters. In fact, we expect that, for the specific mixture considered in this paper, under some conditions on $(X_i, A_i)_{1 \leq i \leq n}$ and r , the model's likelihood is a concave function of the parameters. As it turns out to be mathematically challenging to give a satisfying statement of these conditions, we plan to dedicate future theoretical work to this question.

Other future directions for our work include extensive simulations with misspecified networks and non-parametric networks as well as the extension of the proposed framework to non-binary treatments and dynamic ITRs.

ACKNOWLEDGMENTS

We thank Dr. Viet Thi Tran and Prof. Stéphane Gaudry for their insightful comments on epidemiological and critical care applications.

References

- Shalit U. Can we learn individual-level treatment policies from clinical data?. *Biostatistics* 2020; 21(2): 359–362.
- Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ digital medicine* 2020; 3: 17.
- Takahashi K, Serruys PW, Fuster V, et al. Redevelopment and validation of the SYNTAX score II to individualise decision making between percutaneous and surgical revascularisation in patients with complex coronary artery disease: secondary analysis of the multicentre randomised controlled SYNTAXES trial with external cohort validation. *The Lancet* 2020; 396(10260): 1399–1412.
- Basu S, Sussman JB, Rigdon J, Steimle L, Denton BT, Hayward RA. Benefit and harm of intensive blood pressure treatment: Derivation and validation of risk models using data from the SPRINT and ACCORD trials. *PLoS Medicine* 2017; 14(10):

- e1002410.
5. Tannock IF, Hickman JA. Limits to Personalized Cancer Medicine. *The New England Journal of Medicine* 2016; 375(13): 1289–1294.
 6. Qian M, Murphy SA. Performance guarantees for individualized treatment rules. *The Annals of Statistics* 2011; 39(2): 1180–1210.
 7. Zhao Y, Zeng D, Rush AJ, Kosorok MR. Estimating Individualized Treatment Rules Using Outcome Weighted Learning. *Journal of the American Statistical Association* 2012; 107(449): 1106–1118.
 8. Luedtke AR, Laan v. dMJ. Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. *The Annals of Statistics* 2016; 44(2): 713–742.
 9. Kallus N. Balanced policy evaluation and learning. *Proceedings of the 32nd International Conference on Neural Information Processing Systems* 2018: 8909–8920.
 10. Thomas PS, Brunskill E. Data-efficient off-policy policy evaluation for reinforcement learning. *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48* 2016: 2139–2148.
 11. Díaz I, Laan v. dMJ. Assessing the causal effect of policies: an example using stochastic interventions. *The International Journal of Biostatistics* 2013; 9(2): 161–174.
 12. Janes H, Brown MD, Pepe MS, Huang Y. An Approach to Evaluating and Comparing Biomarkers for Patient Treatment Selection. *The International Journal of Biostatistics* 2014; 10(1): 99–121.
 13. Imai K, Li ML. Experimental Evaluation of Individualized Treatment Rules. *Journal of the American Statistical Association* 2021; 0(0): 1–15. DOI: 10.1080/01621459.2021.1923511.
 14. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Statistics in Medicine* 2014; 33(13): 2297–2340.
 15. Angrist JD, Imbens GW. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association* 1995; 90(430): 431–442.
 16. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics* 2002; 58(1): 21–29.
 17. Neyman J. On the application of probability theory to agricultural experiments. Essay on Principles. Section 9 (translation published in 1990). *Statistical Science* 1923; 5: 472–480.
 18. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 1974; 66(5): 688–701.
 19. Zhang B, Tsiatis AA, Laber EB, Davidian M. A Robust Method for Estimating Optimal Treatment Regimes. *Biometrics* 2012; 68(4): 1010–1018.
 20. Tsiatis AA, Davidian M, Holloway ST, Laber EB. *Dynamic Treatment Regimes: Statistical Methods for Precision Medicine*. CRC Press . 2019.
 21. Jacob D. CATE meets ML. *Digital Finance* 2021; 3(2): 99–148.

22. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine* 2004; 23(19): 2937–2960.
23. Jordan MI, Jacobs RA. Hierarchical Mixtures of Experts and the EM Algorithm. *Neural Computation* 1994; 6(2): 181–214.
24. Teicher H. Identifiability of Finite Mixtures. *The Annals of Mathematical Statistics* 1963; 34(4): 1265 – 1269.
25. Jiang W, Tanner M. On the identifiability of mixtures-of-experts. *Neural Networks* 1999; 12(9): 1253-1258.
26. Allman ES, Matias C, Rhodes JA. Identifiability of parameters in latent structure models with many observed variables. *The Annals of Statistics* 2009; 37(6A): 3099 – 3132.
27. Xu L, Jordan MI. EM Learning on A Generalized Finite Mixture for Combining Multiple Classifiers. *World Congress on Neural Networks* 1993; 4: 227–230.
28. Stefanski LA, Boos DD. The Calculus of M-Estimation. *American Statistician* 2002; 56(1): 29–38.
29. Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Scientific Data* 2016; 3(1).
30. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine* 2011; 30(4): 377–399.
31. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Medicine* 1996; 22(7): 707–710.
32. Grolleau F, Porcher R, Barbar S, et al. Personalization of renal replacement therapy initiation: a secondary analysis of the AKIKI and IDEAL-ICU trials. *Critical Care* 2022; 26(1): 64.
33. Jin C, Zhang Y, Balakrishnan S, Wainwright MJ, Jordan MI. Local maxima in the likelihood of gaussian mixture models: Structural results and algorithmic consequences. *Advances in Neural Information Processing Systems* 2016; 29.

How to cite this article: Grolleau F., Petit F., Porcher R. (2023), A comprehensive framework for the evaluation of individual treatment rules from observational data, *arXiv:2207.06275*.

Supplementary Materials for “A comprehensive
framework for the evaluation of individual
treatment rules from observational data”

François Grolleau, François Petit, and Raphaël Porcher

A Proof for the ARE formula in the new ITR situation

$$\begin{aligned}
\Delta(r) &= \mathbb{E}[Y^{s=1} - Y^{s=0}] \\
&= \mathbb{E}[r(X)Y^{a=1} + \{1 - r(X)\}Y^{a=0} - AY^{a=1} - (1 - A)Y^{a=0}] \\
&= \mathbb{E}[\{r(X) - A\}\{Y^{a=1} - Y^{a=0}\}] \\
&= \mathbb{E}\left[\mathbb{E}[\{r(X) - A\}\{Y^{a=1} - Y^{a=0}\}|X]\right] \\
&= \mathbb{E}[\{r(X) - \pi(X)\}\tau(X)]
\end{aligned} \tag{A1}$$

where equality in (A1) uses equation (5) from assumption 2 and the fact that in the new ITR situation $A = A^{s=0}$.

B Proof for the AIE/MIG formulas in the new ITR situation

$$\begin{aligned}
\Lambda(r, \rho^*) &= \mathbb{E}[Y^* - Y^{s=0}] \\
&= \mathbb{E}[A^*Y^{a=1} + (1 - A^*)Y^{a=0} - A^{s=0}Y^{a=1} - (1 - A^{s=0})Y^{a=0}] \\
&= \mathbb{E}[\{A^* - A\}\{Y^{a=1} - Y^{a=0}\}] \\
&= \mathbb{E}\left[\mathbb{E}[\{A^* - A\}\{Y^{a=1} - Y^{a=0}\}|X]\right] \\
&= \mathbb{E}[\{\pi^*(X) - \pi(X)\}\tau(X)] \tag{B1} \\
&= \mathbb{E}[\rho^*(X)\{r(X) - \pi(X)\}\tau(X)] \tag{B2}
\end{aligned}$$

where equality in (B1) uses equations (5) and (6) from assumption 2 and the fact that in the new ITR situation $A = A^{s=0}$. The equality in (B2) follows from

$$\begin{aligned}
\pi^*(X) &= \mathbb{E}[A^*|X] \\
&= \mathbb{E}[S^*A^{s=1} + (1 - S^*)A^{s=0}|X] \\
&= \rho^*(X)r(X) + \{1 - \rho^*(X)\}\pi(X)
\end{aligned}$$

where the last equality uses the fact that $A^{s=1} = r(X)$ and equation (7) from assumption 2. The proof for the $\Gamma(r, \rho^*)$ formula follows a similar argument.

C Proof of lemma 1 in the partially implemented ITR situation

$$\begin{aligned}
(i) \quad \pi(X) &= \mathbb{E}[A|X] \\
&= \mathbb{E}[Sr(X) + (1 - S)A^{s=0}|X] \\
&= r(X)\mathbb{E}[S|X] + \mathbb{E}[(1 - S)A^{s=0}|X] \\
&= r(X)\rho(X) + \mathbb{E}[1 - S|X]\mathbb{E}[A^{s=0}|X] \\
&= r(X)\rho(X) + \{1 - \rho(X)\}\pi^{s=0}(X)
\end{aligned} \tag{C1}$$

Equality in (C1) relies on equation (7) from assumption 2 and the fact that in the partially implemented ITR situation $S = S^*$.

$$\begin{aligned}
(ii) \quad q_0(X) &= \mathbb{E}[Y^{s=0}|X] \\
&= \mathbb{E}[A^{s=0}Y^{a=1} + (1 - A^{s=0})Y^{a=0}|X] \\
&= \mathbb{E}[A^{s=0}Y^{a=1}|X] + \mathbb{E}[(1 - A^{s=0})Y^{a=0}|X] \\
&= \mathbb{E}[A^{s=0}|X]\mathbb{E}[Y^{a=1}|X] + \mathbb{E}[(1 - A^{s=0})|X]\mathbb{E}[Y^{a=0}|X] \\
&= \pi^{s=0}(X)\mu_1(X) + \{1 - \pi^{s=0}(X)\}\mu_0(X)
\end{aligned} \tag{C2}$$

Equality in (C2) relies on equation (5) from assumption 2.

D Identifiability result

In this appendix, we prove an identifiability result for the model given by equation (3.14)

$$\pi(x, \theta) = \rho(x, \alpha, a)r(x) + \{1 - \rho(x, \alpha)\}\pi^{s=0}(x, \beta, b)$$

where $\theta = (\alpha, a, \beta, b) \in \mathbb{R}^{2d+2}$.

Theorem. *Assume that \mathcal{X} is an open subset of \mathbb{R}^d , that $r: \mathcal{X} \rightarrow \{0; 1\}$ is such that the sets $\{r = 0\}$ and $\{r = 1\}$ have non-empty interior and that the applications $\rho(\cdot, \alpha, a): \mathcal{X} \rightarrow \mathbb{R}$ and $\pi^{s=0}(\cdot, \beta, b): \mathcal{X} \rightarrow \mathbb{R}$ are of the form*

$$\begin{aligned}\rho(x, \alpha, a) &= \text{expit}(\alpha^T x + a), \\ \pi^{s=0}(x, \beta, b) &= \text{expit}(\beta^T x + b).\end{aligned}$$

Then the model $\pi(x, \theta)$ is identifiable.

Proof. Let θ and θ' in \mathbb{R}^{d+2} such that $\pi(x, \theta) = \pi(x, \theta')$. Consider the functions

$$\begin{aligned}f(x) &: \mathbb{R}^d \rightarrow \mathbb{R}, x \mapsto \{(1 - \rho(x, \alpha, a))\pi^{s=0}(x, \beta, b)\}^{-1} \\ g(x) &: \mathbb{R}^d \rightarrow \mathbb{R}, x \mapsto \{(1 - \rho(x, \alpha', a'))\pi^{s=0}(x, \beta', b')\}^{-1}\end{aligned}$$

Since \exp is an entire function, f and g can be written as power i.e. there exist sequences of real numbers c_ν and c'_ν such that for all $x \in \mathbb{R}^d$,

$$f(x) = \sum_{\nu \in \mathbb{N}^d} c_\nu x^\nu \quad \text{and} \quad g(x) = \sum_{\nu \in \mathbb{N}^d} c'_\nu x^\nu \quad (\text{D3})$$

where $\nu = (\nu_1, \nu_2, \dots, \nu_d)^T \in \mathbb{N}^d$, $x = (x_1, \dots, x_d)^T \in \mathbb{R}^d$, and x^ν is a shorthand for $x_1^{\nu_1} x_2^{\nu_2} \dots x_d^{\nu_d}$. Moreover, we have the following identities

$$\begin{aligned}\forall x \in \{r = 0\}, (1 - \rho(x, \alpha, a))\pi^{s=0}(x, \beta, b) \\ = (1 - \rho(x, \alpha', a'))\pi^{s=0}(x, \beta', b'),\end{aligned} \quad (\text{D4})$$

$$\begin{aligned}\forall x \in \{r = 1\}, \rho(x, \alpha, a) + \{1 - \rho(x, \alpha)\}\pi^{s=0}(x, \beta, b) \\ = \rho(x, \alpha', a') + \{1 - \rho(x, \alpha')\}\pi^{s=0}(x, \beta', b').\end{aligned} \quad (\text{D5})$$

Equation (D4) implies that on $\{r = 0\}$, $f \equiv g$. Since, f and g are power series that coincide on $\{r = 0\}$ which has a non-empty interior, the principle of analytic continuation implies that $f \equiv g$ on \mathbb{R}^d . Hence, Equation (D4) holds on

\mathbb{R}^d and in particular on $\{r = 1\}$. Subtracting Equation (D4) from Equation (D5), we obtain that

$$\forall x \in \{r = 1\}, \rho(x, \alpha, a) = \rho(x, \alpha', a')$$

Since $\text{expit} : \mathbb{R} \rightarrow \mathbb{R}$ is a bijection, we have that for every $x \in \{r = 1\}$, $\alpha^T x + a = \alpha'^T x + a'$. Since $\{r = 1\}$ has a non-empty interior, we get $(\alpha, a) = (\alpha', a')$. This last equality together with Equation (D4) implies that for every $x \in \mathbb{R}^d$ $\pi^{s=0}(x, \beta, b) = \pi^{s=0}(x, \beta', b')$. It follows that $(\beta, b) = (\beta', b')$. Hence the model $\pi(x, \theta)$ is identifiable. \square

E Derivation of an EM algorithm for the mixture of experts presented in section 3.2.2

Here, we follow the structure of the exposition of Jordan and Jacobs (1994). For simplicity, we start assuming that the true models for the expert $\pi^{s=0}$ and the gating network ρ are parametric models. Under this assumption, we characterize the probability of generating A_i from X_i separately for the case where $S_i = 1$ and the case where $S_i = 0$. When $S_i = 1$, this probability is either 0 or 1 that is,

$$\begin{aligned} P_1(A_i|X_i) &= \mathbb{1}\{A_i = r(X_i)\} \\ &= r(X_i)A_i + \{1 - r(X_i)\}(1 - A_i) \\ &= r(X_i)^{A_i}\{1 - r(X_i)\}^{1-A_i}. \end{aligned}$$

When $S_i = 0$, the probability is

$$P_0(A_i|X_i, \zeta^0) = \pi^{s=0}(X_i; \zeta^0)^{A_i} (1 - \pi^{s=0}(X_i; \zeta^0))^{1-A_i}$$

where ζ^0 denotes the true value of the parameter for the expert $\pi^{s=0}$. The total probability of generating A_i from X_i is thus

$$P(A_i|X_i, \theta^0) = \rho(X_i; \gamma^0)P_1(A_i|X_i) + \{1 - \rho(X_i; \gamma^0)\}P_0(A_i|X_i, \zeta^0) \quad (\text{E1})$$

where we denote γ^0 the true value of the parameter for the gating network ρ and we write $\theta^0 = (\gamma^0, \zeta^0)^T$. We utilize (E1) without the superscripts to refer to the probability model defined by the mixture, irrespective of any reference to a “true” model. Given a data set $\mathcal{D} = (X_i^T, A_i)_{1 \leq i \leq n}$, the log-likelihood is obtained by taking the log of the product of n densities of the form of (E1), which yields the following log-likelihood:

$$l(\theta; \mathcal{D}) = \sum_{i=1}^n \ln \left\{ \rho(X_i; \gamma)P_1(A_i|X_i) + \{1 - \rho(X_i; \gamma)\}P_0(A_i|X_i, \zeta) \right\}. \quad (\text{E2})$$

We refer to (E2) as the “incomplete-data log-likelihood” as it involves no reference to the latent random variable S_i . If S_i were known, then the maximum likelihood problem would decouple into a separate set of two regression problems: one for the gating network ρ and another for the unknown expert $\pi^{s=0}$. These problems would be solved independently of each other, yielding a rapid one-pass learning algorithm. Of course, S_i are not known, but we can specify a probability model that links them to the observable data. To this end, we posit

the following probability model written in terms of the S_i :

$$P(A_i, S_i | X_i, \theta) = \left\{ \rho(X_i; \gamma) P_1(A_i | X_i) \right\}^{S_i} \left[\{1 - \rho(X_i; \gamma)\} P_0(A_i | X_i, \zeta) \right]^{1-S_i}.$$

Taking the logarithm of this probability model yields the following “complete-data log-likelihood”:

$$l_c(\theta; \mathcal{C}) = \sum_{i=1}^n S_i \ln \{\rho(X_i; \gamma) P_1(A_i | X_i)\} + (1 - S_i) \ln \{(1 - \rho(X_i; \gamma)) P_0(A_i | X_i, \zeta)\} \quad (\text{E3})$$

where $\mathcal{C} = (X_i^T, A_i, S_i)_{1 \leq i \leq n}$ denotes a “complete data” set. Note the relationship of the complete-data log-likelihood in equation (E3) to the incomplete-data log-likelihood in equation (E2). The use of the latent variables S_i allowed the logarithm to be broken into a sum of logarithms, substantially simplifying the maximization problem. However, notice that the complete-data likelihood is a random variable, because the variables S_i are in fact unknown. Taking expectation with respect to the value of the parameters at the p -th iteration denoted $\theta^{(p)}$, we can conduct the E step of the EM algorithm to obtain the deterministic (i.e., free of the random variables S_i) function Q :

$$\begin{aligned} Q(\theta, \theta^{(p)}) &\stackrel{\text{def}}{=} \mathbb{E}[l_c(\theta; \mathcal{C}) | \mathcal{D}] \\ &= \sum_{i=1}^n h_{1,i} \ln \rho(X_i; \gamma) + (1 - h_{1,i}) \ln (1 - \rho(X_i; \gamma)) \\ &\quad + \sum_{i=1}^n (1 - h_{1,i}) \left[A_i \ln \{\pi^{s=0}(X_i; \zeta)\} + (1 - A_i) \ln \{1 - \pi^{s=0}(X_i; \zeta)\} \right] \\ &\quad + \sum_{i=1}^n h_{1,i} \ln P_1(A_i | X_i) \end{aligned} \quad (\text{E4})$$

where we denote $h_{1,i} = \frac{P_1(A_i | X_i) \rho(X_i; \gamma^{(p)})}{\rho(X_i; \gamma^{(p)}) P_1(A_i | X_i) + (1 - \rho(X_i; \gamma^{(p)})) P_0(A_i | X_i, \zeta^{(p)})}$ and we have used the fact that

$$\begin{aligned} \mathbb{E}[S_i | \mathcal{D}] &= \mathbb{P}(S_i = 1 | A_i, X_i, \theta^{(p)}) \\ &= \frac{\mathbb{P}(A_i | S_i = 1, X_i, \theta^{(p)}) \mathbb{P}(S_i = 1 | X_i, \theta^{(p)})}{\mathbb{P}(A_i | X_i, \theta^{(p)})} \\ &= \frac{\mathbb{P}(A_i | S_i = 1, X_i, \theta^{(p)}) \mathbb{P}(S_i = 1 | X_i, \theta^{(p)})}{\mathbb{P}(S_i = 1 | X_i, \theta^{(p)}) \mathbb{P}(A_i | S_i = 1, X_i, \theta^{(p)}) + \mathbb{P}(S_i = 0 | X_i, \theta^{(p)}) \mathbb{P}(A_i | S_i = 0, X_i, \theta^{(p)})} \\ &= \frac{P_1(A_i | X_i) \rho(X_i; \gamma^{(p)})}{\rho(X_i; \gamma^{(p)}) P_1(A_i | X_i) + (1 - \rho(X_i; \gamma^{(p)})) P_0(A_i | X_i, \zeta^{(p)})} \\ &= h_{1,i}. \end{aligned}$$

The M step requires maximizing $Q(\theta, \theta^{(p)})$ with respect to the gating network parameters γ and the expert parameters ζ . Examining equation (E4), we see that the parameters γ and ζ influence the Q function only through the first and second summation terms. Thus the M step reduces to the following separate maximization problems:

$$\gamma^{(p+1)} = \arg \max_{\gamma} \sum_{i=1}^n h_{1,i} \ln \rho(X_i; \gamma) + (1 - h_{1,i}) \ln (1 - \rho(X_i; \gamma)) \quad (\text{E5})$$

$$\zeta^{(p+1)} = \arg \max_{\zeta} \sum_{i=1}^n (1 - h_{1,i}) \left[A_i \ln \{\pi^{s=0}(X_i; \zeta)\} + (1 - A_i) \ln \{1 - \pi^{s=0}(X_i; \zeta)\} \right] \quad (\text{E6})$$

Each of these maximization problems is itself a maximum likelihood problem. Equation (E5) is a maximum likelihood problem with observations $(X_i^T, h_{1,i})_{1 \leq i \leq n}$. Equation (E6) is a weighted maximum likelihood problem with observations $(X_i^T, A_i)_{1 \leq i \leq n}$ and observation weights $1 - h_{1,i}$. Conducting the M-step maximizes the value of Q , the expectation of the complete-data log-likelihood. (Dempster *and others*, 1977) proved than an increase in Q implies an increase in the incomplete-data log-likelihood and so we have

$$l(\theta^{(p+1)}; \mathcal{D}) \geq l(\theta^{(p)}; \mathcal{D}).$$

For clarity and ease of implementation, in Algorithm 1, we assume the Bernoulli distribution for both the expert and gating networks models and thus write $\pi^{s=0}(x; \zeta) = \text{expit}(\zeta^T x)$ and $\rho(x; \gamma) = \text{expit}(\gamma^T x)$. Note that viewing equations (E5) and (E6) as supervised learning problems with relevant observations and weights, nonparametric models such as boosting, random forests, or deep neural networks may also be used for estimating $\rho(\cdot)$ and $\pi^{s=0}(\cdot)$. That option is available, although we lose the EM proof of convergence (cf. Jordan and Xu (1995)). In Algorithm 1, we simplify notations as follows: $P_{1,i}$ and $P_{0,i}$ stand for $P_1(A_i|X_i)$ and $P_0(A_i|X_i, \zeta^{(p)})$ respectively; $g_{1,i}$ and $g_{0,i}$ stand for $\rho(X_i; \gamma^{(p)})$ and $1 - \rho(X_i; \gamma^{(p)})$ respectively; and we note $1 - h_{1,i}$ as $h_{0,i}$.

F Inclusion and exclusion criteria from the MIMIC-III analysis

Inclusion criteria were (all needed be fulfilled):

- Admission to an intensive care unit
- Age greater than 18 years on the day of intensive care admission
- Evidence of severe acute kidney injury (stage 3 in the Kidney Disease Improving Global Outcomes classification) during the stay in intensive care
- Initiation of either mechanical ventilation or intravenous vasopressors during the stay in intensive care, prior to severe acute kidney injury

Exclusion criteria were (all needed be absent):

- End-stage renal kidney disease at intensive care admission
- Renal replacement therapy initiated prior to severe acute kidney injury
- Patients included in the study for an earlier episode of severe acute kidney injury in intensive care
- Patients expected to die within three days

G Patient inclusion diagram from the MIMIC-III database

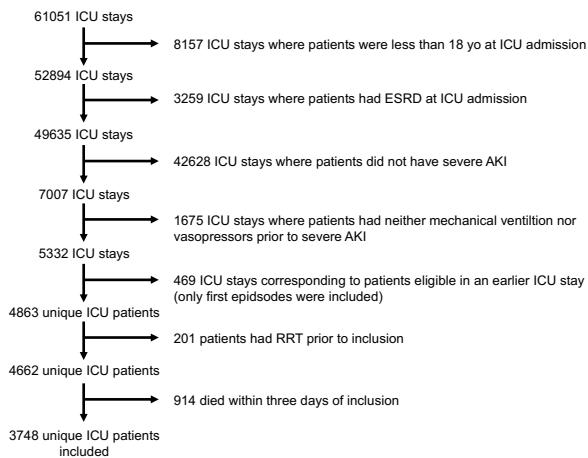


Figure G1: Patient inclusion diagram from the MIMIC-III database.
 ICU=Intensive Care Unit. ESRD=End-Stage Kidney Renal Disease.
 AKI=Acute Kidney Injury. RRT=Renal Replacement Therapy.

H Model specification in the MIMIC-III analysis for the new ITR: dialysis initiation based on a combination of six biomarkers

For this analysis, we specify the models as follows.

- Propensity score model

$$\hat{\pi}(X^\pi) = \mathbb{P}(\text{Dialysis_initiation_within_24h} | \text{Patient_characteristics}) \\ \sim \text{Age} + \text{Weight} + \text{BUN} + \text{pH} + \text{Potassium} + \text{SOFA} + \text{Immunosuppressive_drug}$$

- Prognosis model

$$\hat{\mathbb{E}}[Y|X, A] = \mathbb{P}(\text{Death_at_day_60} | \text{Patient_characteristics}, \text{Dialysis_initiation_within_24h}) \\ \sim \text{Age} + \text{Weight} + \text{BUN} + \text{pH} + \text{Potassium} + \text{SOFA} \\ + \text{Dialysis_initiation_within_24h} \times [\text{Age} + \text{Weight} + \text{BUN} + \text{pH} + \text{Potassium} + \\ \text{SOFA}]$$

- ITE model

$$\hat{\tau}(X) = \hat{\mathbb{E}}[Y|X, A = 1] - \hat{\mathbb{E}}[Y|X, A = 0]$$

- Prognostic function under the ITR

$$\hat{q}_1(X) = r(X)\hat{\mathbb{E}}[Y|X, A = 1] + \{1 - r(X)\}\hat{\mathbb{E}}[Y|X, A = 0]$$

BUN=Blood Urea Nitrogen. SOFA=Sequential Organ Failure Assessment.

ITE=Individualized Treatment Effect. ITR=Individualized Treatment Rule.

I Model specification in the MIMIC-III analysis for a partially implemented ITR: dialysis initiation based on SOFA scores

For this analysis, we specify the models as follows.

- Treatment-specific prognosis functions

$$\hat{\mu}_0(X^\mu) = \mathbb{P}(\text{Death_at_day_60} | \text{Patient_characteristics}) \\ \sim \text{Age} + \text{SOFA}$$

$$\hat{\mu}_1(X^\mu) = \mathbb{P}(\text{Death_at_day_60} | \text{Patient_characteristics}) \\ \sim \text{Age} + \text{SOFA}$$

- Propensity score in the absence of ITR implementation (expert network)

$$\hat{\pi}^{s=0}(X^{\pi^{s=0}}) = \mathbb{P}(\text{Dialysis_initiation_within_24h} | \text{Patient_characteristics}, S = 0) \\ \sim \text{Age} + \text{Weight} + \text{BUN} + \text{pH} + \text{Potassium} + \text{SOFA}$$

- Stochastic implementation function (gating network)

$$\hat{\rho}(X^\rho) \sim \text{Age} + \text{BUN} + \text{pH} + \text{Potassium}$$

BUN=Blood Urea Nitrogen. SOFA=Sequential Organ Failure Assessment.

References

- DEMPSTER, ARTHUR P, LAIRD, NAN M AND RUBIN, DONALD B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)* **39**(1), 1–22.
- JORDAN, MICHAEL I. AND JACOBS, ROBERT A. (1994, March). Hierarchical Mixtures of Experts and the EM Algorithm. *Neural Computation* **6**(2), 181–214.
- JORDAN, MICHAEL I AND XU, LEI. (1995). Convergence results for the em approach to mixtures of experts architectures. *Neural networks* **8**(9), 1409–1431.

Chapter 6

Personalizing renal replacement therapy initiation in the intensive care unit: a statistical reinforcement learning-based strategy with external validation on the AKIKI randomized controlled trials

We shall find it convenient only to speak of things existing when they are in time, that is to say, when we can point to some time at which they exist.

[...]

The world of existence is fleeting, vague, without sharp boundaries, without any clear plan or arrangement, but it contains all thoughts and feelings, all the data of sense, and all physical objects, everything that can do either good or harm, everything that makes any difference to the value of life and the world.

Bertrand Russel, *The problems of philosophy*

La musique associée à ce chapitre est ‘On the Nature of Daylight’ de Max Richter.

Personalizing renal replacement therapy initiation in the intensive care unit: a statistical reinforcement learning-based strategy with external validation on the AKIKI randomized controlled trials

François GROLLEAU,¹ François PETIT*,² Stéphane GAUDRY*,³ Élise DIARD,⁴ Jean-Pierre QUENOT,⁵ Didier DREYFUSS,⁶ Viet-Thi TRAN,⁷ Raphaël PORCHER.⁸

* These authors contributed equally as second authors.

Manuscript words count: 3497

Source of Funding and Conflicts of Interest:

François Petit was supported by a “Chaire d’excellence” (excellence fellowship) from the IdEx Université Paris Cité, ANR-18-IDEX-0001. Raphaël Porcher acknowledges the support of the French Agence Nationale de la Recherche as part of the “Investissements d’avenir” program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute). The authors have disclosed that they do not have any conflicts of interest.

Corresponding author:

François Grolleau

Hôtel-Dieu Hospital, 1 place du Parvis Notre-Dame 75004 Paris, France.

Email: francois.grolleau@aphp.fr

Tel: +33 1 42 34 89 87

¹ M.D., M.P.H., Ph.D. Candidate, Assistant Professor, Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), and Centre d’Epidémiologie Clinique, AP-HP, Hôpital Hôtel Dieu, F-75004 Paris, France.

² Ph.D., Junior Professor, Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), F-75004 Paris, France.

³ M.D., Ph.D., Professor, AP-HP, Hôpital Avicenne, Service de Réanimation Médico-Chirurgicale, Université Paris 13, Bobigny, Health Care Simulation Center, UFR SMBH, Sorbonne Paris Cité and INSERM UMR S1155 “Common and Rare Kidney Diseases: from Molecular Events to Precision Medicine,” Paris, France.

⁴ M.S., Information technology developer, Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), and Centre d’Epidémiologie Clinique, AP-HP, Hôpital Hôtel Dieu, F-75004 Paris, France.

⁵ M.D., Ph.D., Professor, Department of Intensive Care, François Mitterrand University Hospital, Lipness Team, INSERM Research Center, LNC-UMR1231 and LabEx LipSTIC, and INSERM CIC 1432, Clinical Epidemiology, University of Burgundy, Dijon, France.

⁶ M.D., Professor Emeritus, Université Paris Cité, Service de Médecine Intensive-Réanimation, Hôpital Louis Mourier, AP-HP and INSERM UMR S1155 “Common and Rare Kidney Diseases: from Molecular Events to Precision Medicine,” Sorbonne Université, Paris, France.

⁷ M.D., Ph.D., Associate Professor, Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), and Centre d’Epidémiologie Clinique, AP-HP, Hôpital Hôtel Dieu, F-75004 Paris, France.

⁸ Ph.D., Professor, Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), and Centre d’Epidémiologie Clinique, AP-HP, Hôpital Hôtel Dieu, F-75004 Paris, France.

Abstract

Background: Trials sequentially randomizing patients each day have never been conducted for renal replacement therapy (RRT) initiation. We used clinical data from routine care and trials to learn and validate optimal dynamic strategies for RRT initiation in the intensive care unit (ICU).

Methods: We included participants from the MIMIC-III database for development, and AKIKI and AKIKI2 (two randomized controlled trials on RRT timing) for validation. Participants were eligible if they were adult ICU patients with severe acute kidney injury, receiving invasive mechanical ventilation, catecholamine infusion, or both. We used a doubly-robust estimator to learn when to start RRT after the occurrence of severe acute kidney injury given a patient's evolving characteristics—for three days in a row. The ‘crude strategy’ aimed to maximize hospital-free days at day 60 (HFD60). The ‘stringent strategy’ recommended initiating RRT only when there was evidence at the 0.05 threshold that a patient would benefit from initiation. For external validation, we evaluated the causal effects of implementing our learned strategies *versus* following current best practices on HFD60.

Results: We included 3 748 patients in the development set (median age 69y [IQR 57–79], median SOFA score 9 [IQR 6–12], 1 695 [45.2%] female), and 1 068 in the validation set (median age 67y [IQR 58–75], median SOFA score 11 [IQR 9–13], 344 [32.2%] female). Through external validation, we found that compared to current best practices, the crude and stringent strategies improved average HFD60 by 13.7 [95% CI -5.3–35.7], and 14.9 [95% CI -3.2–39.2] days respectively. Contrasted to current best practices where 38% of patients initiated RRT within three days, with the stringent strategy, we estimated that only 14% of patients would.

Conclusion: We developed a practical and interpretable dynamic decision support system for RRT initiation in the ICU. Its implementation could improve the average number of days that ICU patients spend alive and outside the hospital.

Key words: acute kidney injury, renal replacement therapy, personalized medicine, causal inference, reinforcement learning.

Introduction

In intensive care units (ICU), acute kidney injury (AKI) affects about one in two patients, and its onset is associated with high mortality and long-term sequelae (1). Renal replacement therapy (RRT) is an invasive but potentially life-saving treatment for AKI (2). Because AKI is a heterogeneous and rapidly evolving syndrome (3), controversies on the timing and selection of patients for initiating RRT have long prevailed (4). In the last decade, multicenter randomized trials, compared early versus delayed RRT initiation strategies, but the analyzes (5–7) and meta-analyzes (8, 9) of these trials failed to show significant differences in patient-important outcomes at the population level. As such negative trial findings are widespread in critical care, identifying individualized treatment effects has been judged a research priority (10).

Physicians' attempts to deliver timely interventions tailored to patients' characteristics have a long history (11). While in some diseases, biological insight proved decisive in moving precision medicine forward (12), AKI—due to its heterogeneous syndromic nature—is less amenable to this approach. Recently, authors proposed algorithms for RRT initiation in the ICU (13, 14), but the need for validated data-driven decision support tools remains (15). Previously, we developed a decision support tool based on clinical trial data and considered the static case where the decision to initiate RRT is only pondered at AKI onset (16). Yet, for such decision tools to be actionable and consistent with practice, they must go beyond the static case and account for the fundamentally dynamic nature of AKI. In fact, when a decision support tool recommends not initiating RRT for a given patient on a given day, it ought to re-evaluate its recommendation on the next day considering the evolution of the patient's characteristics.

To learn an optimal RRT initiation strategy under this setting, the ideal method would be to conduct a Sequential Multiple Assignment Randomized Trial (SMART) where AKI patients are sequentially randomized each day to either initiate treatment or not (17). Due to cost, time, and practical constraints, SMART trials have never been conducted in the ICU. However, recent developments in statistics and computer science provided robust methods to learn and evaluate optimal treatment initiation strategies from observational data (18, 19). To our knowledge, only a single monocenter study has analyzed clinical data in an attempt to develop a dynamic decision support system for RRT initiation (20).

In this paper, we used reinforcement learning methods on data from electronic health records to estimate optimal dynamic strategies for RRT initiation in ICU patients with AKI. Then, in an external validation step, we used data from two large multicenter randomized trials on RRT timing to estimate the benefit of implementing these strategies.

Methods

Sources of data

The development sample included participants from the Multi-Parameter Intelligent Monitoring in Intensive Care III (MIMIC-III) database. MIMIC-III is a project maintained by the Laboratory for Computational Physiology at the Massachusetts Institute of Technology which contains routinely collected data from 61,051 distinct ICU admissions of adult patients admitted between 2001 and 2012 (21). For reproducibility, we used the database's official code repository to extract all relevant variables (22). As out-of-hospital mortality was not available in the latest version of the MIMIC project, we used MIMIC-III version 1.4.

The validation sample included participants from the AKIKI and AKIKI2 trials, two multicenter RCTs conducted in France (5, 23). The AKIKI trial was conducted at 31 ICUs from Sept 2013 through Jan 2016 and recruited 619 patients with stage 3 KDIGO-AKI who required mechanical ventilation, catecholamine infusion, or both. Included patients were 1:1 randomized to either an early RRT initiation strategy or to a standard-delayed initiation strategy. The AKIKI2 trial was embedded in a cohort recruiting at 39 ICUs from May 2018 through Oct 2019. In AKIKI2, eligibility criteria for the cohort were identical to the eligibility criteria from the original AKIKI trial. Of the 767 patients included in the cohort, 278 met one or more randomization criteria (oliguria for more than 72h or blood urea nitrogen concentration greater than 112 mg/dL) and were 1:1 allocated to either a standard-delayed RRT initiation strategy or to a more-delayed strategy.

Population

Eligible patients were adults (18 years of age or older) hospitalized in the ICU with stage 3 KDIGO AKI who were receiving (or had received for this episode) invasive mechanical ventilation, catecholamine infusion, or both. Staging in the KDIGO classification was based on serum creatinine and/or urine output with higher stages indicating greater severity (24). As the latest clinical guidelines recommend a standard-delayed strategy of RRT initiation (25), we chose this strategy as the reference “best practice” upon which to improve. Precisely, our target population was made of individuals whose physicians implemented a standard-delayed strategy. In both AKIKI trials, the standard-delayed strategy suggested initiating RRT if one of the following criteria occurred: severe hyperkaliemia and/or metabolic acidosis, pulmonary edema resistant to diuretics, oliguria for more than 72 hours, blood urea nitrogen level higher

than 112 mg per deciliter. In the current study, we used the same exclusion criteria as in the AKIKI trials i.e., moribund state (patient likely to die within 72h), end-stage kidney disease (i.e., patient with creatinine clearance < 15ml/ml), patients having received RRT before inclusion, and patients already included at a previous date.

Setup and timepoints for learning dynamic RRT initiation strategies

From a clinical standpoint, the decision to start RRT is considered difficult in the first days following severe AKI. After three days, this decision often becomes straightforward, as most patients have either recovered or deteriorated. We focused on developing a when-to-treat strategy for RRT initiation in the first 72 hours following the onset of severe AKI (i.e., stage 3 KDIGO-AKI). Specifically, we learned a strategy that—for three days in a row after severe AKI onset—assessed the need to start RRT given a patient's evolving characteristics. Our strategy was non-stationary i.e., the decision rules for RRT initiation could differ depending on the day. We considered three decision timepoints at 0, 24, and 48 hours after severe AKI onset (**Figure 1**). The strategy was developed so that, at each timepoint, it used clinical and biological information gathered prior to this timepoint as inputs and outputted a recommendation to either initiate RRT within 24 hours or not. We considered that once RRT had been recommended (or initiated in contradiction with the strategy's recommendation), the strategy would persist in recommending RRT for all subsequent decision timepoints. This so-called regularity in the strategy's behavior indicates that we did not consider when to stop RRT in the three days following severe AKI onset.

Primary outcome

The primary outcome was hospital-free days at day 60 (HFD60). This outcome was chosen because i) it was a good compromise between patient-centeredness and pragmatism (26); and ii) it reduced the risk that the learned strategy had unexpected side effects—a well-known issue in reinforcement learning (27). For instance, using short-term mortality as a primary outcome, the model may learn a strategy that maximizes survival at the cost of keeping patients alive in the ICU as long as possible.

Learning an optimal strategy

To learn an optimal strategy, we used a doubly robust estimator with weighted least squares (dWOLS) (28). This method relies on estimating blip functions for each decision timepoint. Given the evolving characteristics of an individual up to timepoint t , the blip function predicts the effect of initiating RRT at t versus not initiating it, under the assumption that optimal treatment decisions are taken from timepoint $t + 1$ onwards. We derived two strategies from the estimated blip functions. We termed “crude” the strategy that recommended RRT initiation to the patients with positive values of blips, and “stringent” the strategy that recommended initiating RRT only when there was evidence at the 0.05 significance level that a patient would benefit from RRT initiation (i.e., positive lower bound for the blip’s 95% confidence interval). As stated before, once RRT initiation was recommended, both strategies persisted in recommending RRT regardless of the blips at subsequent timepoints. Patients who died within three days of AKI were excluded from the development sample, considering no relevant information would be learned from these patients’ data. Indeed, it seemed unlikely that patients who died within three days of severe AKI could have been discharged from the hospital under a different RRT initiation strategy: we expected their outcome to be the same under all strategies (HFD60 is zero for all patients who die in the hospital). However, these patients were not excluded from the validation sample, to avoid time-dependent selection bias. More details on dWOLS estimation and inference are given in the appendix (pp 3-4).

External validation

To match our target “best practice” population, we included all patients from AKIKI and AKIKI2 who had received a standard-delayed strategy. As the AKIKI2 patients randomized to a more-delayed strategy were compatible with a standard-delayed strategy until they met a randomization criterion, we excluded these patients but duplicated the patients randomized to the standard-strategy arm according to the cloning and censoring principle used for emulating target trials from observational data (29). To estimate hospital mortality and the proportion of patients who would initiate RRT within three days under a strategy, we used importance sampling for policy evaluation in reinforcement learning (30).

To evaluate the effect of new strategies on HFD60, we considered current best practices (i.e., the standard-delayed strategy from the AKIKI trials) as a common control and compared it to the following three strategies: i) the crude strategy, ii) the stringent strategy, and iii) a strategy that recommends initiating RRT in all patients within 24 hours after severe AKI onset.

We estimated the causal effect of implementing each of these strategies compared to current best practices using the cross-fitted advantage doubly robust estimator for strategy evaluation with terminal states (18). This estimator allows estimating the mean difference in the outcome that would have been observed under any given strategy and the outcome that would have been observed under a reference strategy. We provide more details on importance sampling for policy evaluation and the advantage doubly robust estimator in the appendix (pp 4-5).

Ethical approval and research transparency

The MIMIC-III analysis received approval from the Institutional Review Boards of the Massachusetts Institute of Technology and Beth Israel Deaconess Medical Center (BIDMC). The AKIKI and AKIKI2 trials received approval from competent French legal authority (Comité de Protection des Personnes d'Ile de France VI, ID RCB 2013-A00765-40, NCT01932190 for AKIKI and ID RCB 2017-A02382-51, NCT03396757 for AKIKI 2). The funding sources were not involved in the study design; collection, analysis, and interpretation of data; writing of the manuscript; or the process of submission for publication. Two authors (FG, RP) had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the analysis. Analyzes were conducted using R version 4.2.1 for strategy learning as well as plotting, and Python 3.8.8 for strategy evaluation. The code used in this study is available at <https://github.com/fcgrolleau/dynamic-rrt>.

Results

Learning optimal dynamic strategies for RRT initiation

1. Patients

From 2001 through 2012, a total of 3 748 ICU patients with AKI recruited at a tertiary teaching hospital (BIDMC — Harvard Medical School) met eligibility criteria and were included in the development set (**Figure S1, Panel A**). Almost half of individuals were females (n=1 695; 45.2%). At enrollment, patients had a mean SOFA score of 9 (interquartile range [IQR], 6–12). All patients had severe AKI (i.e., stage 3 KDIGO-AKI) which diagnosis was most often based on urine output (n=3 328; 88.8%). At enrollment, the median serum creatinine and urine output were 1.40 mg/dL (IQR, 0.90–2.40) and 0.28 ml/kg/h (IQR, 0.22–0.29) respectively. During the follow-up, 400 (10.7%) patients initiated RRT within three days of severe AKI, and 892 (23.8%) died during hospitalization. The mean and median HFD60 were 33.6 and 42.9 days, (IQR, 0.9–51.7) respectively. Additional baseline and evolving characteristics for these patients are given in **Table 1**.

2. Learned strategies

For patients with severe AKI who have never initiated RRT at a given decision timepoint, decision rules whether to initiate RRT in the next 24 hours were derived from the models we estimated at each timepoint (blip functions). Estimated parameters are given with didactical instructions for calculations in **Table 2** (their covariances are given in **Table S1**). In **Figure 2A** we display the recommendations from the two learned strategies (i.e., the crude and stringent strategies) along with the uncertainty in the recommendation for each patient in the development set. We present in **Table 3**, three illustrative examples where the learned strategies were applied for individualizing the decision to initiate RRT within 72 hours of severe AKI. The apparent effect (i.e., in the development set) of implementing the crude strategy versus implementing the MIMIC-III RRT initiation strategy was a 6.6 days improvement in mean HFD60.

External validation

1. Patients

From 2013 through 2019, a total of 931 unique ICU patients with AKI from the AKIKI and AKIKI2 trials met eligibility criteria and were included in the validation set. After cloning and censoring, these corresponded to a sample of 1 068 individuals from a population who have received current best practices (i.e., a standard-delayed strategy, *see Figure S1, Panel B*). About a third of individuals ($n=344$; 32.2%) from the validation set were females. For most patients, severe AKI was associated with septic shock and the mean SOFA score was 11 (IQR, 9–13). A drop in urine output triggered stage 3 KDIGO-AKI diagnosis in 401 patients (37.5%). At enrollment, the median serum creatinine and urine output were 3.39 mg/dL (IQR, 2.57–4.33) and 0.12 ml/kg/h (IQR, 0.04–0.34) respectively. During follow-up, 482 (45.1%) died during hospitalization, while 405 (38%) and 99 (20.5%) respectively initiated RRT or died within three days of severe AKI. The mean HFD60 was 14.2 days (median 0, IQR, 0–30.3).

2. External validation of the learned strategies

In the external validation population, we estimated that, under the crude strategy, 41% of patients would die during hospitalization and 53% would initiate RRT within three days. Under the stringent strategy, we estimated that 38% of patients would die during hospitalization and 14% of patients would initiate RRT within three days. Recommendations from the learned strategies along with the uncertainty in individual-patient recommendations are given for all patients in the validation set in **Figure 2B**. The discrepancies between current best practices and the recommendations from the learned strategies are shown in **Figure 3**. We found that compared to current best practices (i.e., the standard-delayed strategy from the AKIKI trials), the crude and stringent strategies yielded a 13.7 days and 14.9 days improvement in mean HFD60 respectively (**Figure 4**).

Discussion

Summary of findings

In this study, we used electronic health record data to learn dynamic RRT initiation strategies for ICU patients with severe AKI. Then, using data from two large RCTs of RRT timing we conducted external validation: compared to current best practices (i.e., a standard-delayed strategy), we found that the crude strategy may improve HFD60 by 13.7 days on average. Note that even though the crude strategy may recommend RRT initiation sooner than the standard-delayed strategy, it is not an early strategy. Consistent with previous trials (5–7), we showed that a strategy that recommends RRT initiation in all patients within 24 hours of severe AKI may yield outcomes similar to or worse than that of a standard-delayed strategy. In contrast to early strategies, the crude strategy identified that only 53% of patients required RRT initiation in the three days following severe AKI. Of note, in the STARRT-AKI and AKIKI arms corresponding to current best practices (i.e., the arms termed standard and delayed respectively), rates of RRT initiations a week after severe AKI were 59% and 55%. We believe that the benefit of the crude strategy stems from its ability to identify earlier the patients who will ultimately require RRT. That said, we found that the stringent strategy may also improve patients' HFD60 all the while reducing RRT prescriptions in the three days following severe AKI. This suggests that the individual-patient confidence intervals given by the crude strategy provide important information for deciding the initiation of RRT. Entailing less frequent usage of RRT, the stringent strategy could have the benefit of not only improving patient-important outcomes but also saving health resources.

Our methodology aimed at developing interpretable linear decision rules together with confidence intervals to guide clinicians at the bedside. For greater transparency and interpretability, we released a user-friendly online implementation of our learned strategies at <http://dynamic-rrt.eu>. Using the time-varying characteristics of a patient as input, clinicians can with this web application obtain individual-patient recommendations from the crude strategy along its 95% confidence intervals. With respect to interpretability, we noticed that on the first day, the crude strategy recommended RRT initiation more often in older patients with higher values of serum creatinine and serum potassium. On the second day, it seemed inclined to recommend RRT initiation in patients with stable arterial pH having a critical combination of low urine output and high blood urea nitrogen levels. Only on the third day did the learned strategies appear more aggressive recommending RRT initiation in most patients who had not

recovered kidney function (i.e., patients with persisting low urine output and high blood urea nitrogen levels).

In this work, we chose to use HFD60 rather than mortality as the primary outcome. Mortality at a given timepoint conveys limited statistical information, as it contains only two possible values. In recent years, there has been an increased focus on patient-centered non-mortality outcomes such as event-free day endpoints in ICU research (26, 31). In a dynamic reinforcement learning setting, there is however one more reason not to use survival as the outcome to optimize. Using survival as a distal reward signal may push the system to find a strategy that maximizes survival at the cost of unnecessary invasive procedures. Practically, the model could use its many degrees of freedom to learn a strategy that increases 60-day survival but decreases hospital and ICU discharge. On the contrary, optimizing over HFD60 is unlikely to yield longer hospital or ICU stays (32).

Strength and limitations

To our knowledge, this study is the first to provide a validated dynamic decision support system for RRT initiation in the ICU. We believe the implications of our work are not only clinical but also methodological as the approach we used can be adapted for the timely initiation of a wide variety of treatments in medicine. However, our study has several limitations. First, we considered only regular strategies, i.e., we did not allow for strategies to recommend stopping RRT before the third day if it had been initiated earlier. Disregarding the opportunities to stop treatment had a strong statistical advantage as it decreased the opportunities for a mismatch between prescribed and recommended treatments, thereby reducing variance in strategy learning and strategy evaluation. From a clinical standpoint, finding an optimal stopping strategy would rather be a distinct question that is more relevant after the third day. Second, we acknowledge that the effect size from implementing our learned strategies, though clinically relevant, was not statistically significant at the conventional 0.05 threshold. In reinforcement learning, learned strategies have long been tested on their training data, and inference for strategy evaluation is still rarely provided as reaching statistical significance often requires huge sample sizes (33). In this study, we performed external validation and estimated confidence intervals of the strategies' benefits. This transparent approach indicates that developing more robust strategies may require training and testing on larger databases, perhaps coupling multiple electronic health records. Third, we concede that given infinitely large sample sizes, methods that leverage computation rather than expert knowledge (e.g., methods such as deep Q networks) may ultimately be more effective. Nevertheless, we believe that as

even large electronic health records yield small effective sample sizes, encoding expert knowledge in the feature engineering process remains essential. Compared to black-box algorithms, we trust this human-centric approach is more likely to convince clinicians as it offers a window for interpretability.

Implication for future research

As is true of traditional drugs, new individualized strategies will require proper testing in clinical settings before they can be deployed (34). This could be done for instance in a cluster randomized controlled trial comparing physicians alone to physicians assisted by the clinical decision support system. Alternatively, new trial designs could help to improve the learned strategy while it is being prospectively evaluated (35). Finally, if kidney damage markers (e.g., C-C motif chemokine ligand 14) demonstrate their clinical utility (36), new strategies leveraging this information may be developed. In the long run, these developments may help bridge the gap between biological knowledge and actionable data-driven approaches. We believe that fostering collaborations of clinical experts, methodologists, and mathematicians all genuinely interested in AKI and reinforcement learning is key. This, we hope, will continue to move personalized medicine forward for the benefit of intensive care patients.

In conclusion, we developed a dynamic RRT initiation strategy and confirmed via external validation that its implementation could increase the average number of days that ICU patients spend alive and outside the hospital. This interpretable strategy relies on routinely collected data and provides confidence intervals to guide decision-making at the bedside. It will require prospective testing and refinements before it can be broadly deployed in practice.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/XX>.

Supplementary Methods: **Appendix A** Setup notations. **Appendix B** Summary of notations introduced in the appendix. **Appendix C** Doubly robust dynamic treatment regimen via weighted least squares. **Appendix D** Variable selection. **Appendix E** Missing data management. **Appendix F** Importance sampling for policy evaluation. **Appendix G** Advantage doubly robust estimator.

Supplementary Results: **Table S1.** Variance-covariance matrices of blip parameter estimates for the learned strategy based on multiple imputation analysis of one hundred data sets. **Figure S1.** Flow diagrams for the development set (A) and validation set (B). **Figure S2.** Comparison of recommendations from the original (A) or stringent (B) learned strategy and the RRT prescriptions received in the development set. **Figure S3.** Missing data patterns in the development set (A) and validation set (B).

ABBREVIATIONS

AKI: Acute Kidney Injury

BIDMC: Beth Israel Deaconess Medical Center

CI: Confidence Interval

HFD60: Hospital-Free Days at day 60

ICU: Intensive Care Unit

IQR: Interquartile Range

KDIGO: Kidney Disease: Improving Global Outcomes

RCT: Randomized Controlled Trial

RRT: Renal Replacement Therapy

SMART: Sequential Multiple Assignment Randomized Trial

SOFA: Sequential Organ Failure Assessment

DECLARATIONS

Ethics approval and consent to participate

The MIMIC-III analysis received approval from the Institutional Review Boards of the Massachusetts Institute of Technology and Beth Israel Deaconess Medical Center (BIDMC). The AKIKI and AKIKI2 trials received approval from competent French legal authority (Comité de Protection des Personnes d'Ile de France VI, ID RCB 2013-A00765-40, NCT01932190 for AKIKI and ID RCB 2017-A02382-51, NCT03396757 for AKIKI 2) and consent of patient or relatives was obtained before inclusion.

Consent for publication

All authors have consented to the publication of the present manuscript, should the article be accepted by the Editor-in-chief upon completion of the refereeing process.

Availability of data and materials

The MIMIC-III data is publicly available at <https://mimic.mit.edu>. Anonymous participant data from the AKIKI trials is available under specific conditions. Proposals will be reviewed and approved by the sponsor, scientific committee, and staff on the basis of scientific merit and absence of competing interests. Once the proposal has been approved, data can be transferred through a secure online platform after the signing of a data access agreement and a confidentiality agreement.

Competing interests

The authors have disclosed that they do not have any conflicts of interest.

Funding

FP was supported by a “Chaire d’excellence” (excellence fellowship) from the IdEx Université Paris Cité, ANR-18-IDEX-0001. RP acknowledges the support of the French Agence Nationale de la Recherche as part of the “Investissements d’avenir” program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute). The authors have disclosed that they do not have any conflicts of interest.

Authors' contributions

FG, RP, FP, and VTT conceived the study. FG wrote the codes and did the computational analysis with input from RP and FP. SG, JPQ, and DD provided data from the AKIKI trials. ED designed the Sankey diagrams and contributed to the user interface. FG drafted the manuscript with inputs from RP, VTT, FP, SG, DD, and JPQ. All the authors read the paper and suggested edits. RP supervised the project. FG and RP accessed and verified the data. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Acknowledgements

The authors thank Cynthia T. Chen (Westaf) for editing. We thank all patients included in the AKIKI trials as well as their surrogates. We express our gratitude to the medical and nursing teams that participated in these trials.

References

1. Hoste EAJ, Bagshaw SM, Bellomo R, Cely CM, Colman R, Cruz DN, *et al.* Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study. *Intensive Care Med* 2015;41:1411–1423.
2. Gaudry S, Palevsky PM, Dreyfuss D. Extracorporeal Kidney-Replacement Therapy for Acute Kidney Injury. *N Engl J Med* 2022;386:964–975.
3. Ronco C, Bellomo R, Kellum JA. Acute kidney injury. *Lancet* 2019;394:1949–1964.
4. Ostermann M, Bellomo R, Burdmann EA, Doi K, Endre ZH, Goldstein SL, *et al.* Controversies in acute kidney injury: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Conference. *Kidney Int* 2020;98:294–309.
5. Gaudry S, Hajage D, Schortgen F, Martin-Lefevre L, Pons B, Boulet E, *et al.* Initiation Strategies for Renal-Replacement Therapy in the Intensive Care Unit. *N Engl J Med* 2016;375:122–133.
6. Barbar SD, Clere-Jehl R, Bourredjem A, Hernu R, Montini F, Bruyère R, *et al.* Timing of Renal-Replacement Therapy in Patients with Acute Kidney Injury and Sepsis. *N Engl J Med* 2018;379:1431–1442.
7. STARRT-AKI Investigators, Canadian Critical Care Trials Group, Australian and New Zealand Intensive Care Society Clinical Trials Group, United Kingdom Critical Care Research Group, Canadian Nephrology Trials Network, Irish Critical Care Trials Group, *et al.* Timing of Initiation of Renal-Replacement Therapy in Acute Kidney Injury. *N Engl J Med* 2020;383:240–251.
8. Fayad AII, Buamscha DG, Ciapponi A. Timing of renal replacement therapy initiation for acute kidney injury. *Cochrane Database Syst Rev* 2018;12:CD010612.
9. Gaudry S, Hajage D, Benichou N, Chaïbi K, Barbar S, Zarbock A, *et al.* Delayed versus early initiation of renal replacement therapy for severe acute kidney injury: a systematic review and individual patient data meta-analysis of randomised clinical trials. *Lancet* 2020;395:1506–1515.
10. Semler MW, Bernard GR, Aaron SD, Angus DC, Biros MH, Brower RG, *et al.* Identifying Clinical Research Priorities in Adult Pulmonary and Critical Care. NHLBI Working Group Report. *Am J Respir Crit Care Med* 2020;202:511–523.
11. Phillips CJ. Precision Medicine and Its Imprecise History. *Harvard Data Science Review* 2020;2:..

12. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE, Davidson NE, *et al.* Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005;353:1673–1684.
13. Gaudry S, Quenot J-P, Hertig A, Barbar SD, Hajage D, Ricard J-D, *et al.* Timing of Renal Replacement Therapy for Severe Acute Kidney Injury in Critically Ill Patients. *Am J Respir Crit Care Med* 2019;199:1066–1075.
14. Bagshaw SM, Hoste EA, Wald R. When should we start renal-replacement therapy in critically ill patients with acute kidney injury: do we finally have the answer? *Critical Care* 2021;25:179.
15. Schaub JA, Heung M. Precision Medicine in Acute Kidney Injury: A Promising Future? *Am J Respir Crit Care Med* 2019;199:814–816.
16. Grolleau F, Porcher R, Barbar S, Hajage D, Bourredjem A, Quenot J-P, *et al.* Personalization of renal replacement therapy initiation: a secondary analysis of the AKIKI and IDEAL-ICU trials. *Critical Care* 2022;26:64.
17. Almirall D, Nahum-Shani I, Sherwood NE, Murphy SA. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. *Transl Behav Med* 2014;4:260–274.
18. Nie X, Brunsell E, Wager S. Learning when-to-treat policies. *Journal of the American Statistical Association* 2021;116:392–409.
19. Tsiatis AA, Davidian M, Holloway ST, Laber EB. *Dynamic Treatment Regimes: Statistical Methods for Precision Medicine*. CRC Press; 2019.
20. Morzywołek P, Steen J, Vansteelandt S, Decruyenaere J, Sterckx S, Van Biesen W. Timing of dialysis in acute kidney injury using routinely collected data and dynamic treatment regimes. *Crit Care* 2022;26:365.
21. Johnson AEW, Pollard TJ, Shen L, Lehman L-WH, Feng M, Ghassemi M, *et al.* MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035.
22. Johnson AEW, Stone DJ, Celi LA, Pollard TJ. The MIMIC code repository: enabling reproducibility in critical care research. *J Am Med Inform Assoc* 2018;25:32–39.
23. Gaudry S, Hajage D, Martin-Lefevre L, Lebbah S, Louis G, Moschietto S, *et al.* Comparison of two delayed strategies for renal replacement therapy initiation for severe acute kidney injury (AKIKI 2): a multicentre, open-label, randomised, controlled trial. *The Lancet* 2021;397:1293–1300.
24. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO clinical practice guideline for acute kidney injury. *Kidney Int Suppl* 2012;2:1–138.

25. Evans L, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, *et al.* Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock 2021. *Crit Care Med* 2021;49:e1063–e1143.
26. Auriemma CL, Taylor SP, Harhay MO, Courtright KR, Halpern SD. Hospital-free days: a pragmatic and patient-centered outcome for trials among critically and seriously ill patients. *Am J Respir Crit Care Med* 2021;204:902–909.
27. Sutton RS, Barto AG. 17.4 Designing reward signals. *Reinforcement learning: An introduction* MIT press; 2018.
28. Wallace MP, Moodie EEM. Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics* 2015;71:636–644.
29. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol* 2016;183:758–764.
30. Precup D. Eligibility traces for off-policy policy evaluation. *Computer Science Department Faculty Publication Series* 2000; p. 80.
31. Harhay MO, Casey JD, Clement M, Collins SP, Gayat É, Gong MN, *et al.* Contemporary strategies to improve clinical trial design for critical care research: insights from the First Critical Care Clinical Trialists Workshop. *Intensive Care Med* 2020;46:930–942.
32. Hadfield-Menell D, Russell SJ, Abbeel P, Dragan A. Cooperative inverse reinforcement learning. *Advances in neural information processing systems* 2016;29:..
33. Gottesman O, Johansson F, Komorowski M, Faisal A, Sontag D, Doshi-Velez F, *et al.* Guidelines for reinforcement learning in healthcare. *Nat Med* 2019;25:16–18.
34. Komorowski M. Clinical management of sepsis can be improved by artificial intelligence: yes. *Intensive Care Med* 2020;46:375–377.
35. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, *et al.* Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health Psychol* 2015;34S:1220–1228.
36. Ostermann M, Zarbock A, Goldstein S, Kashani K, Macedo E, Murugan R, *et al.* Recommendations on Acute Kidney Injury Biomarkers From the Acute Disease Quality Initiative Consensus Conference: A Consensus Statement. *JAMA Netw Open* 2020;3:e2019209.

Tables

Table 1 Baseline and evolving characteristics of the patients from the development set (MIMIC-III) and the validation set (AKIKI trials).

	MIMIC-III (n=3 748)	AKIKI trials (n=1 068)
Baseline characteristics*		
Age (year)	69 [57–79]	67 [58–75]
Female gender	1 695 (45.2)	344 (32.2)
Weight (kg)	89 [73–107]	81 [69–95]
Non-corticosteroid immunosuppressive drug	62 (1.7)	53 (5.0)
SOFA score (0 to 24)	9 [6–12]	11 [9–13]
Serum creatinine (mg/dL)	1.40 [0.90–2.40]	3.39 [2.57–4.33]
Blood urea nitrogen (mg/dL)	29 [19–47]	56 [39–78]
Serum potassium (mmol/L)	4.2 [3.9–4.7]	4.4 [3.9–5.0]
Arterial blood pH	7.38 [7.33–7.42]	7.31 [7.24–7.37]
Urine output (ml/kg/h)	0.28 [0.22–0.29]	0.12 [0.04–0.34]
Characteristics at H24†		
Blood urea nitrogen (mg/dL)	34 [21–53]	64 [48–90]
Serum potassium (mmol/L)	4.1 [3.8–4.5]	4.4 [3.9–5.0]
Arterial blood pH	7.38 [7.33–7.42]	7.31 [7.25–7.38]
Urine output (ml/kg/h)	0.38 [0.24–0.64]	0.28 [0.08–0.70]
Characteristics at H48‡		
Blood urea nitrogen (mg/dL)	35 [21–56]	67 [48–92]
Serum potassium (mmol/L)	4.1 [3.8–4.4]	4.3 [3.8–4.9]
Arterial blood pH	7.39 [7.34–7.43]	7.34 [7.27–7.40]
Urine output (ml/kg/h)	0.58 [0.31–0.98]	0.41 [0.09–0.88]

Data are n (%) or median [IQR]. IQR=Interquartile range. SOFA score=Sequential Organ Failure Assessment score. To convert the values for creatinine to micrograms per liter, multiply by 88.4. To convert values for blood urea nitrogen to millimoles per liter, multiply by 0.357.

*Characteristics measured just before the first decision timepoint. †Characteristics measured just before the second decision timepoint.

‡Characteristics measured just before the third decision timepoint.

Table 2 Blip parameter estimates from the learned strategies. Estimations based on the multiple imputation analysis of one hundred data sets.

Tailoring covariate	$\hat{\psi}$	(95% CI)
First decision ^a		
Intercept ₁	-39.589	(-63.885 to -15.294)
Age _{t=1} (years)	0.245	(0.035 to 0.454)
Creatinine _{t=1} (mg/dL)	1.349	(-0.317 to 3.015)
Potassium _{t=1} (mmol/L)	3.409	(-0.547 to 7.364)
Second decision ^b		
Intercept ₂	-7.747	(-23.343 to 7.849)
SOFA score _{t=2}	0.514	(-0.372 to 1.400)
Blood urea nitrogen _{t=2} (mg/dL)	0.095	(-0.033 to 0.223)
pH _{t=1} - pH _{t=2}	-63.874	(-118.998 to -8.750)
Urine output _{t=1} + Urine output _{t=2} (ml/kg/h)	-7.734	(-15.303 to -0.165)
Third decision ^c		
Intercept ₃	5.397	(-14.443 to 25.237)
Urine output _{t=3} (ml/kg/h)	-19.316	(-34.365 to -4.268)
Blood urea nitrogen _{t=3} /Blood urea nitrogen _{t=1}	1.922	(-10.974 to 14.818)

The crude strategy includes the following three decision rules that we derived from the blip parameter estimates $\hat{\psi}$. Decision rules are applicable to patients with severe AKI who have never initiated RRT at a given decision timepoint and whose previous recommendations from the crude strategy were never to initiate RRT (else, the crude strategy persist in its choice to initiate RRT). At the first decision timepoint (beginning of day 1), RRT should be initiated within 24 hours if the linear combination $-39.589 + 0.245 \times \text{age}_{t=1} (\text{years}) + 1.349 \times \text{creatinine}_{t=1} (\text{mg/dL}) + 3.409 \times \text{potassium}_{t=1} (\text{mmol/L})$ is positive. At the second decision timepoint (beginning of day 2), RRT should be initiated within 24 hours if $-7.747 + 0.514 \times \text{SOFA}_{t=2} + 0.095 \times \text{blood urea nitrogen}_{t=2} (\text{mg/dL}) - 63.874 \times |\text{pH}_{t=1} - \text{pH}_{t=2}| - 7.734 \times [\text{urine output}_{t=1} + \text{urine output}_{t=2}]$ is positive. At the third decision timepoint (beginning of day 3), RRT should be initiated within 24 hours if $5.397 - 19.316 \times \text{urine output}_{t=3} (\text{ml/kg/h}) + 1.922 \times [\text{blood urea nitrogen}_{t=3} / \text{blood urea nitrogen}_{t=1}]$ is positive. The $(-)$ _{t=1}, $(-)$ _{t=2}, $(-)$ _{t=3} subscripts refer to values measured just before the first, second, and third decision time point respectively (i.e., at the time of stage 3 KDIGO-AKI onset, stage 3 KDIGO-AKI + 24 hours, and stage 3 KDIGO-AKI + 48 hours respectively). AKI=Acute Kidney Injury. KDIGO=Kidney Disease Improving Global Outcomes. SOFA score=Sequential Organ Failure Assessment score.

^a in the development set n=3748, in the validation set n=1068.

^b in the development set n=3570, in the validation set n=869.

^c in the development set n=3431, in the validation set n=718.

Table 3 Use of the learned strategies for individualized decision-making in three illustrative examples. In patient one (a man aged 58 years), disease severity (as described by SOFA score and arterial blood pH) and kidney function (as described by blood urea nitrogen, serum creatinine, and urine output) remain stable, and the crude strategy suggests against initiating RRT in the 72 hours following stage 3 KDIGO-AKI onset. In patient two (a woman aged 60 years), disease severity lessens over time, but kidney function deteriorates, and the crude strategy suggests initiating RRT on the third day following stage 3 KDIGO-AKI. In patient three (a woman aged 65 years), disease severity is stabilized 24 hours after stage 3 KDIGO-AKI onset, but kidney function has become critical, and the crude strategy suggests initiating RRT on the second day. Note that once a learned strategy recommends initiating RRT it persists in its recommendation until the third day regardless of patients' subsequent characteristics. AKI=Acute Kidney Injury. KDIGO=Kidney Disease Improving Global Outcomes. RRT=Renal Replacement Therapy. SOFA score=Sequential Organ Failure Assessment score.

	Patient one			Patient two			Patient three		
	First decision timepoint	Second decision timepoint	Third decision timepoint	First decision timepoint	Second decision timepoint	Third decision timepoint	First decision timepoint	Second decision timepoint	Third decision timepoint
Stationary characteristics									
Age (years)	58	58	58	60	60	60	65	65	65
Time-evolving characteristics									
SOFA score	10	10	11	16	15	13	12	12	—
Serum creatinine (mg/dL)	3.6	3.6	3.7	2.1	2.9	3.9	2.2	3.2	—
Blood urea nitrogen (mg/dL)	40	42	47	30	39	53	73	90	—
Serum potassium (mmol/L)	4.8	5.3	4.8	4.2	4.2	4.0	4.6	5.3	—
Arterial blood pH (mmol/L)	7.22	7.25	7.29	7.16	7.20	7.41	7.31	7.31	—
Urine output (ml/kg/min)	0.43	0.37	0.45	0.15	0.10	0.08	0.03	0.01	—
Learned strategy									
Blip (95% CI)	-4.2 (-8.0 to -0.4)	-6.7 (-12.7 to -0.7)	-1.0 (-7.2 to 5.1)	-7.8 (-12.4 to -3.1)	-0.8 (-7.0 to 5.4)	7.2 (0.2 to 14.3)	-5.0 (-9.5 to -0.6)	6.7 (-1.1 to 14.4)	—
Crude strategy's recommendation	Do not initiate	Do not initiate	Do not initiate	Do not initiate	Do not initiate	Initiate*	Do not initiate	Initiate†	Continue

*The stringent strategy would also recommend initiating RRT since the confidence interval shows evidence that patient two will benefit from RRT initiation (i.e., the confidence interval's lower bound is positive).

†Contrary to the crude strategy, the stringent strategy would not recommend initiating RRT since the confidence interval does not show evidence that patient three will benefit from RRT initiation

Figures

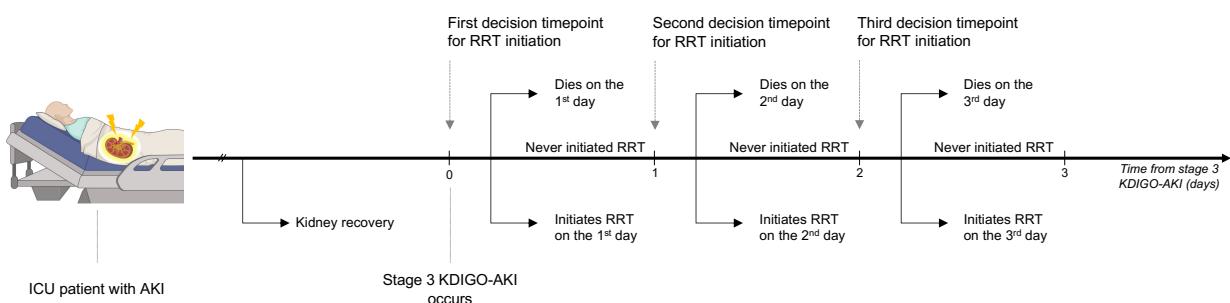


Figure 1 Possible trajectories of a single patient with acute kidney injury in our learning setup.

The first decision timepoint is defined as the time when stage 3 KDIGO-AKI occurs. In our setup, for a patient with stage 3 KDIGO-AKI, the decision rule to initiate RRT mimics that of clinicians i.e., decisions are re-evaluated every day—for three days in a row—given patients' evolving characteristics. Note that at a given decision timepoint a decision needs to be made only if a patient has neither initiated RRT nor died earlier. AKI=Acute Kidney Injury. ICU=Intensive Care Unit. KDIGO=Kidney Disease Improving Global Outcomes. RRT=Renal Replacement Therapy.

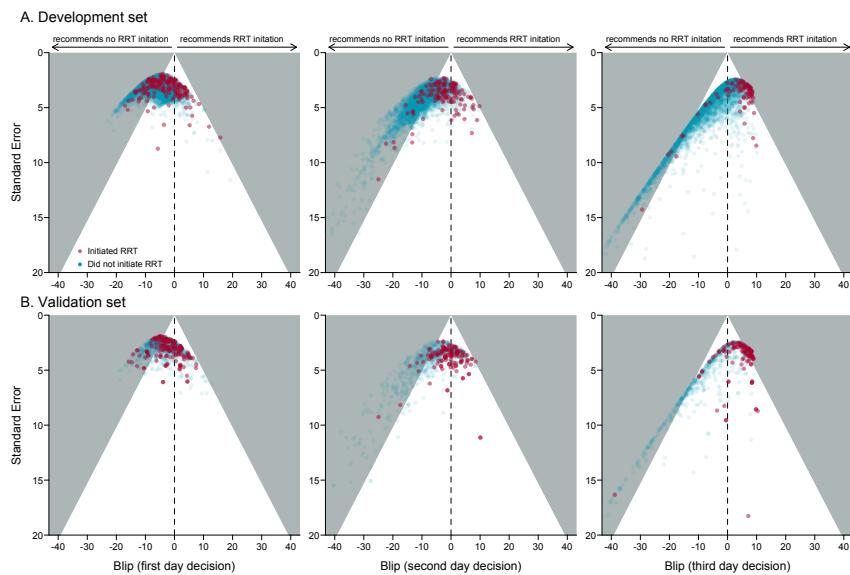


Figure 2 Recommendations from the learned strategies for patients in the development set (Panel A) and in the validation set (Panel B).

Each dot corresponds to a patient for whom a decision whether to initiate RRT needed to be made at the first (left-hand panels), second (middle panels), or third (right-hand panels) decision timepoint. Dot colors depict the RRT prescription observed for these patients. On the x-axis, predicted blips indicate on a HFD60 scale the magnitude of individual-patient harm (negative blips) or benefit (positive blips) from initiating RRT at a particular timepoint. Vertical dashed lines indicate no effect. Uncertainty in the individual-patient blips is represented on y-axis. Dots falling in gray-shaded areas represent patients for whom there is evidence of either harm (left-hand areas), or benefit (right-hand areas) from RRT initiation at the 0.05 alpha level. The crude strategy would recommend initiating RRT at a given timepoint if a patient's dot fell on the right-hand side of the dashed line. On the other hand, the stringent strategy would recommend initiating RRT at a given timepoint only if a patient's dot fell in the right-hand gray-shaded area.

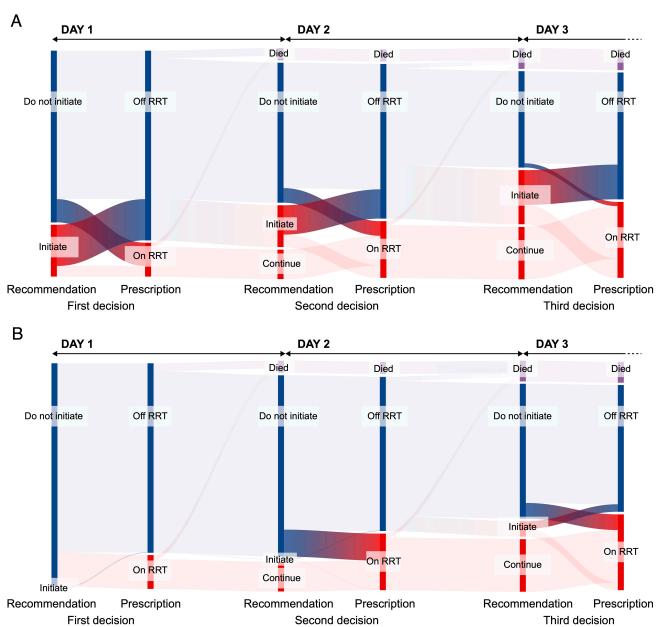


Figure 3 Comparison of recommendations from the crude (Panel A) or stringent (Panel B) strategy and the prescriptions received in the validation set.

Prescriptions received are denoted ‘On RRT’ or ‘Off RRT.’ The bar heights represent the proportions of patients in each category. At each decision timepoint, recommendation and prescription of RRT appear in red while the absence of recommendation or prescription of RRT is shown in blue. Discrepancies between recommendations and prescriptions are shown in brighter colors. Note that when patients initiated RRT (sometimes in contradiction with the strategy’s recommendation) the strategy never recommends stopping it afterward.

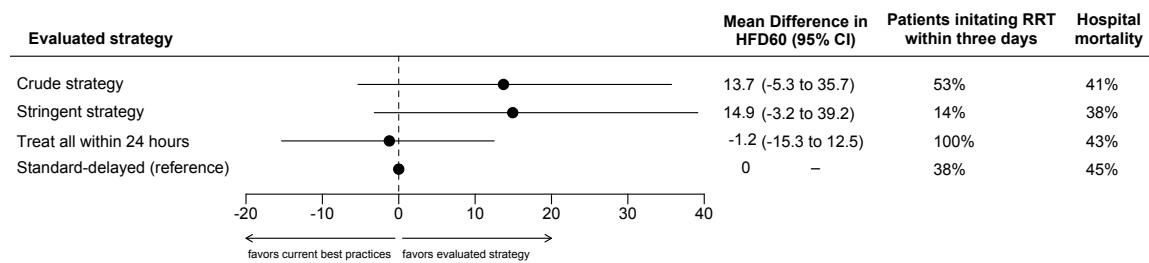


Figure 4 External validation of the learned strategies' benefit as compared to current best practices (i.e., a standard-delayed strategy).
The mean difference in HFD60 represent the causal effects of implementing a strategy compared to current best practices. The “crude strategy” refers to the strategy derived from the blip parameter estimates given in Table 2. The “stringent strategy” refers to a strategy that recommends initiating RRT only when there is evidence at the 0.05 threshold that a patient will benefit from RRT initiation. The “treat all within 24 hours strategy” designates a strategy to initiate RRT in all patients within 24 hours regardless of emergency criteria. HFD60=Hospital-Free Days at day 60. CI=Confidence Interval. RRT=Renal Replacement Therapy.

Personalizing renal replacement therapy initiation in the intensive care unit: a statistical reinforcement learning-based strategy with external validation on the AKIKI randomized controlled trials

François GROLLEAU,¹ François PETIT*,² Stéphane GAUDRY*,³ Élise DIARD,⁴ Jean-Pierre QUENOT,⁵ Didier DREYFUSS,⁶ Viet-Thi TRAN,⁷ Raphaël PORCHER.⁸

* These authors contributed equally as second authors.

ADITIONAL FILE

Table of Contents

Supplementary Methods	3
Setup notations.....	3
Summary of notations introduced in the appendix	3
Doubly robust dynamic treatment regimen via weighted least squares.....	4
Variable selection.....	4
Missing data management.....	5
Importance sampling for policy evaluation	5
Advantage doubly robust estimator	6
Supplementary Results.....	7
Table S1. Variance-covariance matrices of blip parameter estimates for the learned strategy based on multiple imputation analysis of one hundred data sets.	7
Figure S1. Flow diagrams for the development set (A) and validation set (B).	8
Figure S2. Comparison of recommendations from the original (A) or stringent (B) learned strategy and the RRT prescriptions received in the development set.	9
Figure S3. Missing data patterns in the development set (A) and validation set (B).	10
References.....	11

Supplementary Methods

Setup notations

- t : decision timepoint, $t \in \{1,2,3\}$.
- A_t : treatment observed at time t .
- H_t : history of variables collected up to time t including the treatment received before time t but excluding the treatment received at time t .
- H_t^f : subset of H_t relevant for prognosis.
- H_t^γ : subset of H_t relevant for the effect of treatment initiation at time t .
- Y : outcome of interest.
- $Y^{\bar{a}_t, \underline{a}_{t+1}^{opt}}$: potential outcome corresponding to the outcome that would have been observed if treatments a_1, \dots, a_t had been delivered at decision timepoints $1, \dots, t$ and (possibly contrary to fact) all subsequent treatment decisions had been optimal. We use over and underline notations to indicate the past and future treatments respectively i.e., $\bar{a}_t = (a_1, \dots, a_t)$ and $\underline{a}_t = (a_t, \dots, a_3)$.

Summary of notations introduced in the appendix

- $e_t(H_t) = \mathbb{E}[A_t|H_t]$: propensity score at time t .
- $f_t(h_t) = \mathbb{E}\left[Y^{\bar{a}_{t-1}, 0, \underline{a}_{t+1}^{opt}}|H_t = h_t\right]$: treatment-free function at time t .
- $\gamma_t(a_t, h_t) = \mathbb{E}\left[Y^{\bar{a}_{t-1}, a_t, \underline{a}_{t+1}^{opt}} - Y^{\bar{a}_{t-1}, 0, \underline{a}_{t+1}^{opt}}|H_t = h_t\right]$: blip function at time t .
- $\tilde{Y}_t = \widehat{\mathbb{E}}[Y^{\bar{a}_t, \underline{a}_{t+1}^{opt}}|H_t, A_t]$: pseudo-outcomes at time t .
- $\tilde{w}_t(H_t) = |A_t - \hat{e}_t(H_t)|$: overlap weights at time t .
- $\tau = (H_3, A_3, Y)$: the observable full trajectory of a patients.
- \mathcal{T} : the space of trajectories.
- \mathcal{H}_t : the space of histories of variables collected up to time t .
- $\pi = (\pi_1, \pi_2, \pi_3)$ with $\pi_t: \mathcal{H}_t \rightarrow \{0,1\}$ for $t = 1,2,3$: the non-stationary deterministic policy* π .
- $H_t^{(i)} = \Phi$: indicates that at time t , patient i is in a terminal state (i.e., death).

* Note that throughout the paper, we use the term strategy rather than policy. For the remainder of this appendix, these can be taken to be synonymous.

Doubly robust dynamic treatment regimen via weighted least squares

The procedure (termed dWOLS) was formally introduced and described in detail by Wallace and Moodie.¹ Succinctly, the method requires that for each decision timepoint $t = 1, 2, 3$, we posit models for the propensity scores $e_t(H_t) = \mathbb{E}[A_t|H_t]$ as well as the treatment-free $f_t(\cdot)$, and blip $\gamma_t(\cdot)$ functions. Treatment-free and blip functions are defined as $f_t(h_t) = \mathbb{E}\left[Y^{\bar{a}_{t-1}, 0, \underline{a}_{t+1}^{opt}}|H_t = h_t\right]$, and $\gamma_t(a_t, h_t) = \mathbb{E}\left[Y^{\bar{a}_{t-1}, a_t, \underline{a}_{t+1}^{opt}} - Y^{\bar{a}_{t-1}, 0, \underline{a}_{t+1}^{opt}}|H_t = h_t\right]$ so that $f_t(h_t) + \gamma_t(a_t, h_t) = \mathbb{E}\left[Y^{\bar{a}_t, \underline{a}_{t+1}^{opt}}|H_t = h_t, A_t = a_t\right]$.[†] The estimation of $f_t(\cdot)$ and $\gamma_t(\cdot)$ starts at $t = 3$ by regressing $Y^{\bar{a}_3, \underline{a}_3^{opt}} = Y$ onto $(H_3^f, A_3 H_3^\gamma)$ via weighted least squares with weights $\tilde{w}_3(H_3) = |A_3 - \hat{e}_3(H_3)|$. The procedure then follows a backward stepwise approach where we substitute all unobserved potential outcomes by pseudo-outcomes. Specifically, for $t = 2, 1$, we build pseudo-outcomes $\tilde{Y}_t = \hat{\mathbb{E}}[Y^{\bar{a}_t, \underline{a}_{t+1}^{opt}}|H_t, A_t]$ by taking naive outcomes Y and summing up subsequent regrets i.e., $\tilde{Y}_t = Y + \sum_{k=t+1}^3 \max(\hat{y}_k(1, H_t), 0 - \hat{y}_k(A_t, H_t))$. Pseudo-outcomes at time t represent the outcomes that would have been observed if treatment decisions had been optimal from time $t + 1$ onwards. We then regress \tilde{Y}_t onto $(H_t^f, A_t H_t^\gamma)$ via weighted least squares with weights $\tilde{w}_t(H_t) = |A_t - \hat{e}_t(H_t)|$. Using these overlap weights provide double robustness and enhance sample efficiency. Note that the dWOLS estimation procedure does not require making a Markovian assumption. It only requires assuming that for each decision timepoints either the variables causing renal replacement therapy (RRT) initiation, or the prognosis variables were measured. Because we considered that once initiated, RRT is not stopped in the three days following stage 3 KDIGO-AKI, analysis for each decision timepoint was limited to those participants who had not initiated RRT until this decision timepoint (as those who had initiated RRT had no treatment decision to make).

Variable selection

For each decision timepoint, the variables we considered for modeling the probability of RRT initiation within 24 hours were: blood urea nitrogen, serum potassium, arterial blood pH, and urine output. For each decision timepoint, the variables we considered for predicting hospital-free days at day 60 (HFD60) were: age, weight, gender, SOFA score, serum creatinine, blood urea nitrogen, serum potassium, arterial blood pH, and urine output. We considered the evolving values of the aforementioned variables prior to the decision timepoint of interest. The same variables were considered in development and validation analyzes.

[†] This last equality uses the sequential ignorability assumption i.e., $Y^{\bar{a}_{t-1}, a_t} \perp\!\!\!\perp A_t|H_t$ for all t .

Missing data management

Missing data were handled through multiple imputations by chained equations using outcomes as well as all aforementioned predictors in the imputation models. One hundred independent imputed data sets were generated and analyzed separately. Variance-covariance matrices of blip functions parameters were estimated using the bootstrap (999 iterations). Estimates were then pooled using Rubin's rules.

Importance sampling for policy evaluation

We used importance sampling for policy evaluation in reinforcement learning² to estimate hospital mortality as well as the proportion of patients who would initiate RRT within three days under a learned strategy. The approach is similar to inverse propensity weighting as used in the context of marginal structural modeling in epidemiology.³ Succinctly, denoting $\tau = (H_3, A_3, Y) \in \mathcal{T}$ the observable full trajectory of a patient and $R: \mathcal{T} \rightarrow \mathbb{R}$ any reward function of the trajectory, the expected reward under a different strategy, say the non-stationary deterministic strategy π i.e., $\pi = (\pi_1, \pi_2, \pi_3)$ with $\pi_t: \mathcal{H}_t \rightarrow \{0,1\}$ for $t = 1, 2, 3$, can be estimated by

$$\widehat{\mathbb{E}}_{\tau \sim \pi}[R(\tau)] = n^{-1} \sum_{i=1}^n R(\tau^{(i)}) \prod_{k=1}^3 \frac{\mathbb{I}[\pi_k(H_k^{(i)}) = A_k^{(i)}]}{\hat{e}_k(H_k^{(i)})^{A_k^{(i)}} \{1 - \hat{e}_k(H_k^{(i)})\}^{1-A_k^{(i)}}}. \quad (1)$$

To estimate hospital mortality, the reward function we used was $R(\tau^{(i)}) = 1$ if patient i died in the hospital and $R(\tau^{(i)}) = 0$ otherwise. To estimate the proportion of patients who would initiate RRT within three days under our learned strategies, we used the reward function $R(\tau^{(i)}) = 1 - \prod_{k=1}^3 \mathbb{I}[A_k^{(i)} = 0]$ which outputs one whenever patient i initiated RRT at any time along their observed trajectory. The estimator above straightforwardly handles the patients who died before day 3, provided we consider that the strategy π stops prescribing treatment once a patient has died i.e., $\pi_k(\Phi) = 0$, and that no RRT was prescribed to the patients who have died i.e., $e_k(\Phi) = 0$.[‡] The variables we considered for modeling the propensity scores are identical to those given in the previous section. To improve efficiency, we used the weighted version of the estimator above that is given in equation 3 from Precup et al.²

[‡] For the sake of clarity, we denoted $H_t^{(i)} = \Phi$ when patient i is in a terminal state (i.e., death) at time t .

Advantage doubly robust estimator

Although the estimator given in equation 1 accounts for the patients who died before day 3, it only uses trajectories that match the policy π exactly, which can make policy evaluation sample inefficient. To estimate the causal effect of implementing the original, stringent, or treat all strategies compared to current best practices, we used the cross-fitted advantage doubly robust (ADR) estimator with terminal state for strategy evaluation given in the Algorithm 2 from Nie et al.⁴ The ADR estimator allows the evaluation of when-to-treat-policies exploiting the subparts of trajectories that match the policy π . The original, stringent, and treat all strategies are all regular when-to-treat policies in the sense of Definition 1b from Nie et al. The ADR estimator is more data efficient but also more robust than the estimator given in equation (1). Briefly, this estimator relies on the decomposition into a sum of local advantages of the relative value of any given strategy in comparison to that of the never-treating policy $\mathbf{0}$, following Lemma 1 of Murphy.⁵ The ADR estimand is $\Delta(\pi, \mathbf{0}) = \mathbb{E}_{\tau \sim \pi}(Y) - \mathbb{E}_{\tau \sim \mathbf{0}}(Y)$ where, π denotes the strategy to be tested, and zero is the never-treating policy. The causal effects of implementing the original, stringent, or treat all strategies compared to a “best practices policy” denoted π_{bp} , are given by

$$\widehat{\Delta}(\pi, \pi_{bp}) = \widehat{\Delta}(\pi, \mathbf{0}) - \widehat{\Delta}(\pi_{bp}, \mathbf{0}).$$

As in the dWOLS procedure, the ADR estimator does not need any structural (e.g., Markovian) assumptions. As Nie et al.,⁴ we estimated all the nuisance components using cross-fitting to reduce the effect of own-observation bias.

For each decision timepoint, the variables we considered for modeling the probability of RRT initiation within 24 hours were: blood urea nitrogen, serum potassium, arterial blood pH, and urine output. For each decision timepoint the variables we considered for predicting hospital-free days at day 60 (HFD60) were: age, weight, gender, SOFA score, serum creatinine, blood urea nitrogen, serum potassium, arterial blood pH, and urine output.

Missing data were handled through multiple imputations by chained equations using outcomes as well as all aforementioned predictors in the imputation models. Twenty independent imputed data sets were generated and analyzed separately. The variances of the estimators were estimated using the bootstrap (999 iterations). Estimates were then pooled using Rubin’s rules.

Supplementary Results

Table S1. Variance-covariance matrices of blip parameter estimates for the learned strategy based on multiple imputation analysis of one hundred data sets.

Denoting H_t a patient's vector of covariates at decision timepoint t ; \widehat{M}_t the estimated variance-covariance matrix from decision timepoint t ; $\widehat{\psi}_t$ the blip parameter estimates from decision timepoint t , 95% confidence intervals for the individual blips can be calculated as

$$\widehat{\psi}_t^T H_t \pm 1.96 \times \sqrt{H_t^T \widehat{M}_t H_t}.$$

Decision point	Intercept	First variable	Second variable	Third variable	Fourth variable
First decision	Intercept _{t=1}	Age _{t=1}	Creatinine _{t=1}	Potassium _{t=1}	—
Intercept _{t=1}	153.66	-0.843	-0.920	-20.615	
Age _{t=1}	-0.843	0.011	-0.005	0.039	
Creatinine _{t=1}	-0.920	-0.005	0.722	-0.263	
Potassium _{t=1}	-20.615	0.039	-0.263	4.073	
Second decision	Intercept _{t=2}	SOFA score _{t=2}	Blood urea nitrogen _{t=2}	$ \text{pH}_{t=1} - \text{pH}_{t=2} $	Urine output _{t=1} + Urine output _{t=2}
Intercept _{t=2}	63.319	-2.711	-0.270	-74.776	-9.185
SOFA score _{t=2}	-2.711	0.204	0.001	1.934	0.174
Blood urea nitrogen _{t=2}	-0.270	0.001	0.004	0.077	-0.022
$ \text{pH}_{t=1} - \text{pH}_{t=2} $	-74.776	1.934	0.077	791.019	1.372
Urine output _{t=1} + Urine output _{t=2}	-9.185	0.174	-0.022	1.372	14.914
Third decision	Intercept _{t=3}	Urine output _{t=3}	Blood urea nitrogen _{t=3} / Blood urea nitrogen _{t=1}	—	
Intercept _{t=3}	102.467	-23.944	-63.053		
Urine output _{t=3}	-23.944	58.95	5.045		
Blood urea nitrogen _{t=3} / Blood urea nitrogen _{t=1}	-63.053	5.045	43.292		

Units are years for age; mg/dL for creatinine; mmol/L for potassium; mg/dL for blood urea nitrogen; ml/kg/h for urine output. The $(-)_{t=1}$, $(-)_{t=2}$, $(-)_{t=3}$ subscripts refer to values measured just before the first, second, and third decision time point respectively.

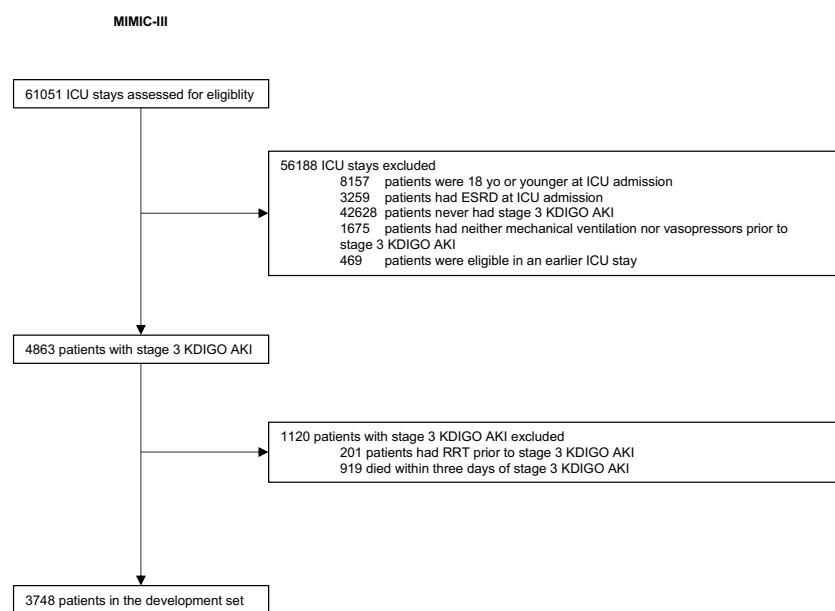
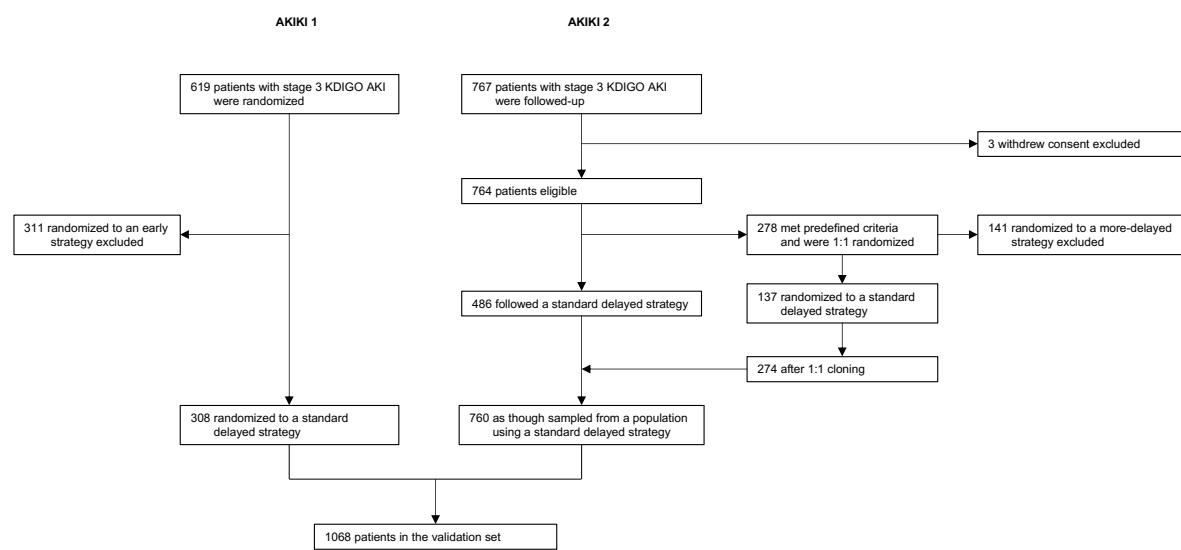
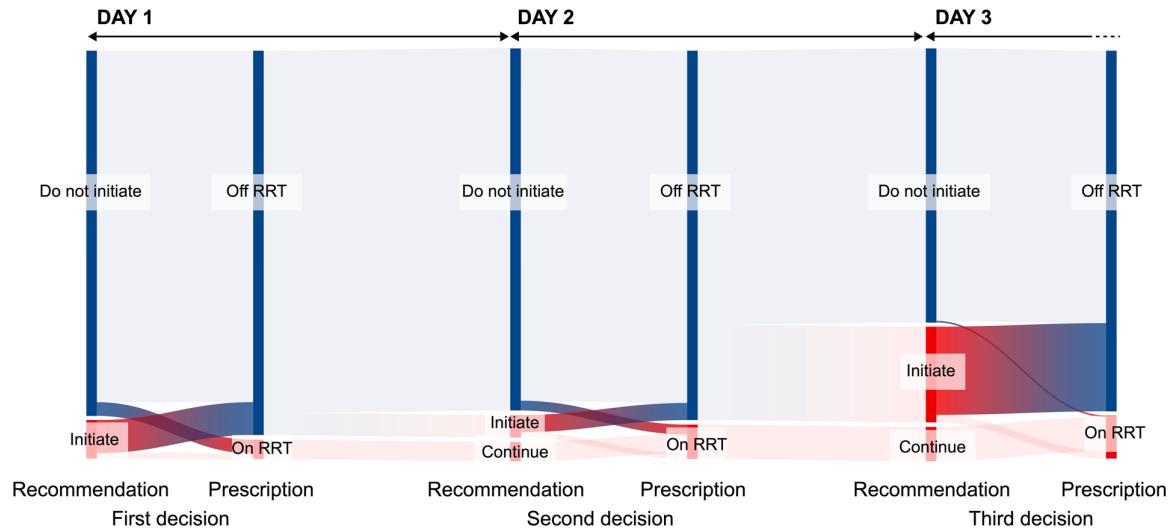
Figure S1. Flow diagrams for the development set (A) and validation set (B).**A****B**

Figure S2. Comparison of recommendations from the original (A) or stringent (B) learned strategy and the RRT prescriptions received in the development set.

A



B

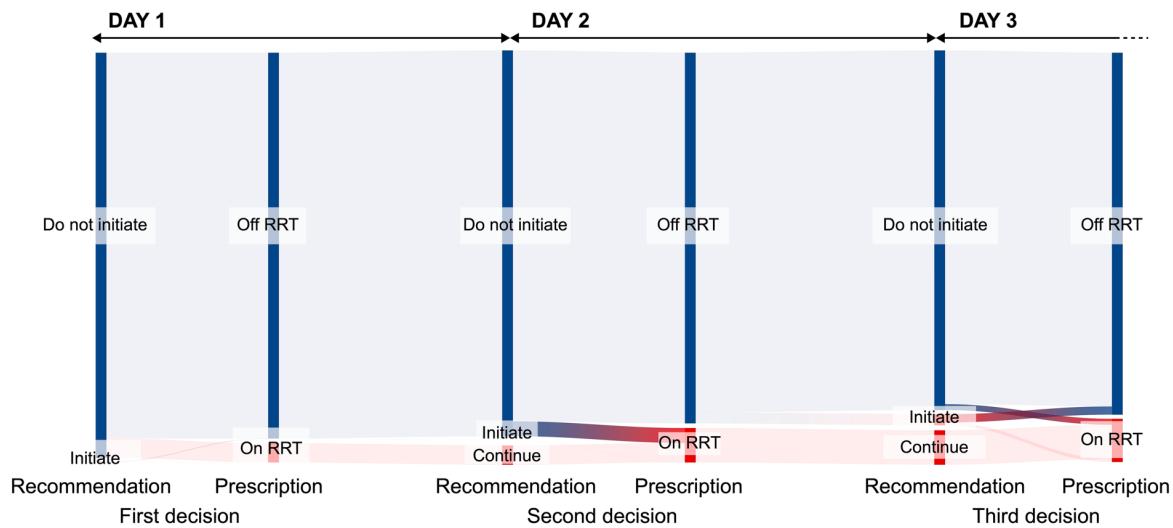
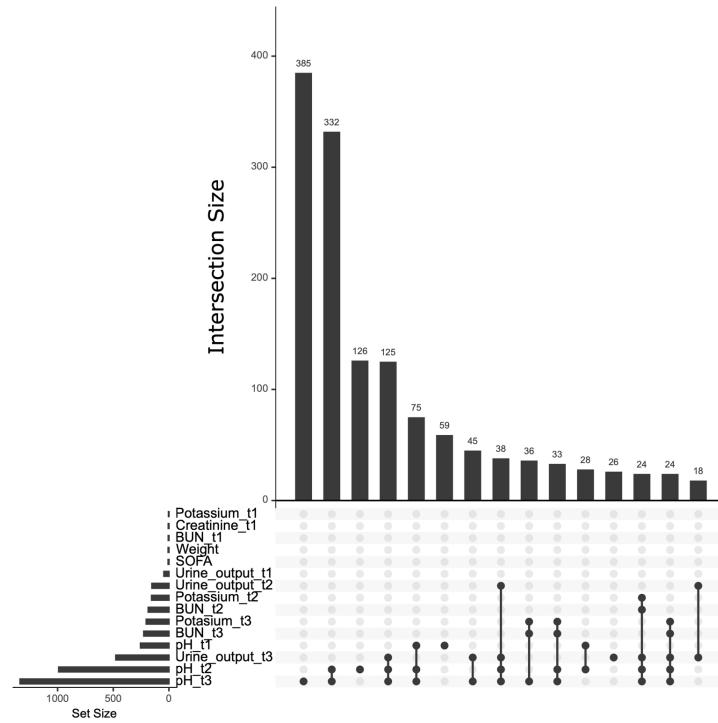
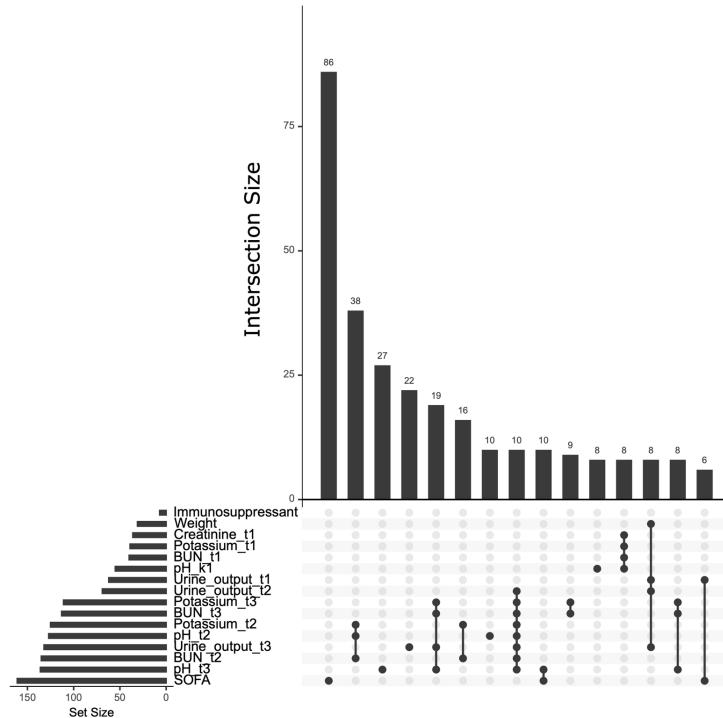


Figure S3. Missing data patterns in the development set (A) and validation set (B).**A****B**

References

- 1 Wallace MP, Moodie EEM. Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics* 2015; **71**: 636–44.
- 2 Precup D. Eligibility traces for off-policy policy evaluation. *Computer Science Department Faculty Publication Series* 2000; p. 80.
- 3 Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; **11**: 550–60.
- 4 Nie X, Brunskill E, Wager S. Learning when-to-treat policies. *Journal of the American Statistical Association* 2021; **116**: 392–409.
- 5 Murphy SA. A Generalization Error for Q-Learning. *J Mach Learn Res* 2005; **6**: 1073–97.

Chapter 7

Discussion et perspectives

I don't feel frightened by not knowing things, by being lost in a mysterious universe without any purpose, which is the way it really is, as far as I can tell.

Richard P. Feynman, *The pleasure of finding things out*

7.1 La causalité est-elle fondamentale ou simplement utile ?

Dans cette thèse abordant l’inférence causale nous n’avons pas encore évoquée l’adage classique “corrélation n’est pas causalité”. Dans cette section nous faisons un bref rappel de la notion de causalité dans le contexte médical. Une réinterprétation personnelle visant à mettre en lumière l’intérêt des méthodes d’inférence causale est donnée ensuite.

Déclarer à tort qu’une association est non causale peut avoir des conséquences dramatiques en termes de santé publique. Dans les années 1960, le célèbre statisticien R.A. Fischer, inventeur de l’essai randomisé moderne, niait l’existence de tout lien causal entre le tabagisme et le cancer du poumon [79]. La corrélation positive constatée entre tabagisme et cancer du poumon était selon lui expliquée par l’existence — désormais réfutée — de prédispositions génétiques causant simultanément les deux événements. En l’absence d’un cadre formel pour l’inférence causale, ni la brillance de R.A. Fischer, ni la collecte supplémentaire de données n’ont permis d’éclairer le débat. En 1965, ce sont les critères de Bradford-Hill [80] qui ont permis d’évaluer plus précisément le lien causal entre tabagisme et cancer du poumon. Ces critères peuvent être vus comme un pas important en direction de la théorie moderne de l’inférence causale culminant avec l’attribution d’un prix Turing en 2011 (Judea Pearl, informatique), d’un prix Nobel en 2021 (Guido Imbens, économie) et d’un prix Rousseeuw en 2022 (James Robins et ses collaborateurs, statistiques).

Considérer à tort qu’une association est causale peut aussi avoir des conséquences délétères pour la santé des patients. En réanimation par exemple, il existe une corrélation positive forte entre surcharge hydrique (importance des œdèmes) et mortalité hospitalière [81]. Devant cette association certains médecins préconisent l’administration précoce de diurétiques pour réduire les œdèmes et diminuer la mortalité. Ce raisonnement est incorrect et potentiellement dangereux. La sévérité initiale des patients de réanimation est une cause

commune évidente pour la surcharge hydrique et la mortalité hospitalière. L'utilisation trop précoce de diurétiques expose les patients à un risque iatrogène (insuffisance rénale aiguë, etc.) sans preuve de bénéfice en termes de survie [82].

Les interprétations causales (correctes ou incorrectes) sont en un certain sens inéluctables chez les humains car celles-ci sont indispensables pour la communication des idées. Naturellement, sans faire d'efforts, nous produisons des raisonnements causaux, déterministes et discrets de type : “si X avait fait Y, il se serait produit Z” [83]. Chez l'être humain, il est communément admis que ce mécanisme cognitif d'interprétation causale est rapidement acquis au cours du développement [84] ; il est par ailleurs probable que les enfants ne développant aucune faculté d'interprétation causale décèdent précocement. En ce sens, la causalité peut être vue comme un très important *prior* humain.

Cependant, le constat de ce fort *prior* humain causal n'implique pas que les phénomènes naturels réels suivent eux-mêmes des lois causales. À ce sujet, le point de vue des physiciens est clair : au niveau fondamental (infiniment petit), tout changement est *réversible* et le temps n'ayant pas de flèche, la causalité ne peut pas être correctement définie [85]. En revanche, à une plus grande échelle, l'entropie devient importante, son augmentation permet de caractériser la flèche du temps et la causalité peut être définie précisément (pour une courte vidéo d'introduction voir [86]). En résumé, du point de vue des physiciens la causalité est une propriété émergente et non fondamentale de la nature.

Même si cette compréhension est correcte,¹ lors d'une démarche scientifique, il ne semble pas raisonnable de renoncer complètement à la formulation d'interprétations causales. Comme le dit Judea Pearl : “*physicists talk cause-effect in the cafeteria*” [88]. Malheureusement, du fait de l'adage classique “corrélation n'est pas causalité”, il persiste en recherche clinique un tabou en ce qui concerne la formulation de toute interprétation causale à l'issue de l'analyse de données observationnelles [89]. En pratique, la théorie moderne de l'inférence causale est le lien manquant entre l'analyse de données observationnelles et la réponse rigoureuse à une question scientifique. A mon sens, cette théorie réussit le tour de force d'explicitement rassembler : i) les inéluctables *priors* causaux pour permettre une communication efficace des idées entre humains et ii) l'évaluation statistique de phénomènes naturels afin de quantifier l'incertitude de chaque prédiction causale. Il est intéressant de remarquer que la théorie de l'inférence causale n'est pas falsifiable² mais simplement mathématiquement cohérente. C'est avant tout l'adoption de plus en plus large de cette théorie par une communauté diverse (épidémiologie, statistiques, économétrie, informatique etc.) qui témoigne de son importance scientifique.

7.2 Au-delà de l'inférence causale, comment créer des systèmes de médecine personnalisée plus intelligents ?

Lors de l'utilisation de méthodes d'inférence causale, une difficulté fréquemment rencontrée est la spécification des variables causant l'allocation du traitement et des variables causant la survenue du résultat. Pour atteindre cet objectif, les *priors* causaux sont parfois représentés dans un graphe orienté acyclique après une procédure de Delphi. Ce graphe doit refléter au mieux la connaissance experte du domaine en lien avec la question de

¹En opposition aux théories évoquées ci-dessus, Stephen Wolfram et al. essaient de montrer que la causalité (sous forme de lois déterministes discrètes ; voir *multiway causal graph* [87]) est fondamentale. À ce jour, ces approches sont reconnues comme incomplètes.

²Aucune donnée (même hypothétique) ne semble pouvoir invalider la théorie de l'inférence causale.

recherche. En recherche clinique, le niveau de certitude est souvent peu élevé pour la spécification de ces variables, l’arbitrage de leur rôle (instrumental et/ou pronostique) et la description de la forme des dépendances entre variables. En ce qui concerne la médecine personnalisée (estimateurs de CATE ou de politiques optimales), une difficulté supplémentaire est liée à la nécessité de spécifier des variables influençant la réponse au traitement (*tailoring variables*). En pratique, le niveau de certitude est encore plus faible pour la spécification des *tailoring variables* (et la forme des dépendances entre ces variables) que pour la spécification des variables causales instrumentales et pronostiques.

Une manière de résoudre ce problème est d’inclure un large set de *tailoring variables* candidates dans un modèle non-paramétrique conçu pour la médecine personnalisée et permettant la sélection de ces variables par une procédure de régularisation. Dans le cas d’une décision de traitement non-séquentielle par exemple, la procédure R-learner permet d’utiliser facilement une méthode de régularisation pour sélectionner les *tailoring variables* et pénaliser les dépendances complexes entre variables (voir section 2.3.2). En revanche, la majorité des méthodes disponibles pour estimer des stratégies dynamiques (voir section 2.5) ne permettent pas de sélectionner facilement les variables qui devraient influencer les décisions de traitement (où de pénaliser les dépendances complexes entre variables). De nouvelles méthodes développées dans cette direction pourraient permettre de limiter le surapprentissage en ce qui concerne l’estimation de stratégies dynamiques optimales (voir par exemple Zhang et Zhang [90]).

Il est intéressant de remarquer que les méthodes récentes de médecine personnalisée — pour estimer la fonction CATE ou une stratégie dynamique optimale — s’appuient essentiellement sur des techniques d’apprentissage (e.g., R-learner, Q-learning). Avec ces méthodes, il est supposé que toute ou presque toute l’information pertinente concernant la forme de la fonction CATE ou d’une stratégie dynamique optimale est contenue dans les données. Les avantages et inconvénients de cette approche sont similaires à ceux des méthodes d’apprentissage supervisé : en utilisant des modèles de réponse au traitement suffisamment souples sur un échantillon d’ entraînement de taille et de dimension suffisamment grande, il est garanti que, pour la population dont est issu l’échantillon, la performance — en terme de calibration pour le bénéfice ou de regret — sera élevée. Cependant, pour ces méthodes de médecine personnalisée les vitesses de convergences sont lentes et ne sont étudiées qu’à horizon asymptotique. Il existe un champ de recherche actif visant à améliorer l’efficience des estimateurs CATE et des méthodes d’apprentissage par renforcement [91].

Devant le niveau de certitude faible pour la spécification a priori des *tailoring variables* et de la forme des dépendances entre ces variables, une piste de recherche complémentaire à l’apprentissage pourrait être d’ajouter de l’information en utilisant des *priors* humains plus fondamentaux que la connaissance causale experte (par exemple la topologie ou l’arithmétique élémentaire voir [92]). La sélection du meilleur modèle pourrait être réalisée en utilisant des algorithmes de recherche discrets sur un espace vaste de “programmes possibles” (voir *program synthesis* [93]). À ce jour, pour des problèmes non triviaux, les approches de type *program synthesis* sont intractables numériquement. Dans le futur, ces approches devraient prendre une place plus importante en combinaison avec des méthodes d’apprentissage différentiables conventionnelles comme *deep learning* ou *gradient boosting*. Ces approches mixtes apprentissage/recherche pour la médecine personnalisée pourraient permettre le réentraînement efficient des modèles lors d’un changements de population cible. Un enjeu important sera le développement de méthodes de médecine personnalisée robustes aux changements dans le temps (e.g., un modèle de médecine personnalisée

reste performant si un service hospitalier change ses pratiques) et robustes aux changements dans l'espace (e.g., un modèle de médecine personnalisée développé sur des données d'hôpitaux universitaires français reste performant s'il est utilisé dans des hôpitaux de pays en développement).

7.3 Impasses et perspectives pour l'IA appliquée à la santé

Le terme intelligence artificielle (IA) est souvent mal défini et la mesure de l'intelligence d'un système est souvent confondue avec l'évaluation de ses compétences ou performances (par exemple la calibration d'un modèle) [94]. Pour les applications médicales, cette confusion entre compétence et intelligence d'un système pourrait nuire au progrès. En santé, l'environnement changeant rapidement (e.g., évolution des pratiques, émergence de nouvelles maladies), le développement de systèmes compétents n'est pas suffisant (voir *distribution shift* [95]). L'objectif devrait être le développement de systèmes adaptables — c'est-à-dire restant performants, ou à défaut devenant rapidement performants, dans des situations nouvelles [96]. Pour insister sur cette notion, une redéfinition utile de l'intelligence d'un système pourrait être “sa vitesse d'acquisition de nouvelles compétences pertinentes — en corrigant pour l'information liée à l'expérience du système (données antérieures ‘mémorisées’) et aux *priors* du système (codés par un humain)”.³

A ce jour, l'accès à des données cliniques de grande taille et de haute dimension est le frein principal au développement de l'IA en médecine. Dans les prochaines années, les méthodes permettant d'exploiter de manière systématique les données imparfaites (manquantes, bias de classement, etc.) et multimodales (images, texte, résultats biologiques, etc.) pourraient jouer un rôle important dans les applications médicales (voir *data-centric AI* [98], *tokenization* [99] et *multimodal foundation models* [96]). En comparaison, les stratégies de modélisation pour la prédiction devraient jouer un rôle plus limité ; celles-ci seront probablement de plus en plus souvent automatisées dans des pipelines (voir *automated machine learning* [100]). Par ailleurs, une augmentation des tailles effectives d'échantillons sera en partie possible avec l'utilisation de méthodes d'IA génératives (voir *synthetic data generation* [101]). Ces techniques pourraient être intéressantes pour les applications sur données non structurées (e.g., images, texte, audio) — le bénéfice étant probablement plus faible pour des données cliniques tabulaires.⁴

Des méthodes similaires sont attendues, dans les prochaines années, pour permettre la production de comptes rendus crédibles d'imagerie scanner et IRM (voir *generalist medical AI* [104]). L'application de modèles préentraînés (voir *pre-trained models* [105]) sur des données de taille limitée provenant de pays en développement pourrait permettre de rapidement adapter les modèles pour ces populations. Cette approche pourrait être utile pour toutes les régions du monde où la disponibilité de radiologues experts est insuffisante au regard de la quantité d'images produites [106]. Concernant ces modèles produisant du langage naturel (tels que GPT 4), il existe un risque avéré d'anthropomorphisme (“[le modèle] a compris la question”, “[le modèle] m'a menti”, etc.). Ce phénomène peut avoir deux conséquences délétères qu'il faudra réguler et surveiller : i) le déploiement de modèles non validés produisant du langage naturel, ii) l'altération de la relation médecin-malade.

³Cette définition peut être formalisée mathématiquement voir [97].

⁴La raison souvent évoquée est que les données cliniques tabulaires étant majoritairement discrètes, elles ne reposent pas sur une variété lisse latente qui puisse être estimée (voir *manifold hypothesis* [102, 103]).

Le futur est incertain pour les méthodes d'apprentissage par renforcement visant à développer des systèmes informatiques d'aide aux décisions cliniques. Pour les prises de décisions séquentielles, ces méthodes de médecine personnalisée sont matures ; en théorie, elles permettent de répondre rigoureusement à de nombreuses questions cliniques pertinentes. Cependant, à l'issue de cette thèse, il me semble peu probable de voir prochainement, de multiples applications cliniques concrètes reposant exclusivement sur des méthodes d'apprentissage par renforcement.

La principale difficulté est liée au fait que les méthodes d'apprentissage par renforcement requièrent l'utilisation de données de grande taille — provenant idéalement d'essais randomisés de type SMART [107, 108]. En médecine, contrairement à d'autres domaines d'applications, il est souvent impossible de définir un objectif intermédiaire pertinent (*proxy* ou *intermediate reward*) évaluable avant la fin de la trajectoire d'un individu.⁵ Pour utiliser ces méthodes en santé, il est en général nécessaire de compenser le recours à des données observationnelles de petite taille effective par une quantité importante de connaissance médicale experte. Bien que satisfaisante à court terme, cette approche centrée sur la connaissance experte est difficile à reconduire d'un problème à l'autre ; et elle est contre productive à long terme [110]. Une seconde difficulté est liée aux doutes légitimes des médecins vis-à-vis des résultats de ces méthodes qui ont une explicabilité limitée. À cet égard, la production d'arbres décisionnels et d'intervalles de confiance individuels n'offre pas — contrairement aux systèmes produisant du texte en langage naturel — l'illusion d'une compréhension intime des mécanismes.

En santé, les méthodes d'apprentissage par renforcement pourraient être plus largement utilisées dans d'autres contextes. Par exemple, pour améliorer les systèmes produisant des résumés cliniques en langage naturel au lit du malade (voir *in-context learning* [111], *large language models* [112] et *bedside decision support* [96]) grâce aux retours positifs ou négatifs de professionnels de santé (voir *reinforcement learning with human feedback* [113]). À ce jour, les succès des applications exclusives de méthodes d'apprentissage par renforcement se sont produits dans les situations où i) il est possible d'obtenir à faible coût une quantité importante de données précises et ii) les prises de décisions sont jugées sans risque (typiquement la robotique et les jeux vidéos). En santé publique, de telles situations sont rencontrées dans un cadre particulier : la personnalisation de l'envoi de notifications selon des données collectées sur smartphones afin d'aider des individus à atteindre un objectif intermédiaire comme faire de l'exercice, arrêter de fumer, contrôler la pression artérielle etc. (voir *just-in-time adaptive interventions* [114]).

Dans un autre contexte, les entreprises comme Facebook/Meta ont démontré que les méthodes d'apprentissage par renforcement [115] permettaient de facilement modifier les comportements d'individus pour maximiser le temps⁶ que ceux-ci passent en-ligne sur leur plateforme. Avec cet objectif, de manière agnostique, les modèles découvrent et exploitent les faiblesses d'individus souvent enfants ou adolescents et ceux-ci développent une addiction à la plateforme. Cette addiction, effet indésirable non-anticipé, est bien sûr cause de souffrance mentale et parfois de suicide [116]. En tant qu'épidémiologiste, le parallèle avec l'industrie du tabac dans les années 1960 me semble évident : intérêts financiers, déni de causalité, addiction intense des individus, séquelles à long termes et mortalité évitable

⁵Dans cette configuration où la quantité de données est déjà limitée, les méthodes d'apprentissage par renforcement sont peu efficientes ; voir *sparse reward problem* [109].

⁶Le résultat/*reward* à maximiser n'a jamais été décrit précisément par les compagnies ; il est souvent qualifié par les adjectifs “participation” ou “attention”.

augmentée. Au vu de l'ubiquité des plateformes, il me semble déraisonnable de ne pas considérer cette question comme un problème de santé publique majeur à résoudre sans délai.

Si les méthodes d'apprentissage par renforcement peuvent avoir une place pour aider les individus à faire de l'exercice, arrêter de fumer etc., il me semble improbable que celles-ci soient aussi efficaces pour atteindre ces objectifs, qu'elles ne le sont pour atteindre des objectifs commerciaux en exploitant les faiblesses d'individus à risque d'addiction. Des résultats de personnalisation décevante dans le domaine de la prévention cardiovasculaire via smartphone vont dans ce sens [117]. Pour modifier efficacement les comportements sans s'exposer à un risque d'addiction, il me semble primordial de laisser aux individus la possibilité de régulièrement changer l'objectif qu'ils souhaitent atteindre avec la technologie. Contrairement aux modèles d'apprentissage par renforcement qui maximisent un objectif prédéfini unique – potentiellement au détriment de tout le reste – il est essentiel de se rappeler que les individus humains s'épanouissent en redéfinissant leurs objectifs (l'adaptation au rôle de parent est un exemple classique évident). À ce sujet, une question reste sans réponse : quelles sont les conditions qui permettraient à un modèle d'apprentissage par renforcement d'aider une personne à mieux se connaître elle-même ?

Conclusion

Dans cette thèse nous avons choisi un exemple clinique concret, l'initiation de l'épuration extra-rénale en réanimation, pour illustrer l'utilisation de techniques avancées d'inférence causale et de médecine personnalisée telles que les essais émulés (Chapitre 3), l'estimation de l'hétérogénéité de l'effet du traitement (Chapitre 4), et l'apprentissage de stratégies dynamiques optimales (Chapitre 6). Les importantes leçons apprises peuvent être résumées en trois points.

- Aujourd'hui, très souvent, les données disponibles — y compris celles issues de dossiers patient informatisés (*Electronic Health Records*) — sont de petite taille effective pour développement de règles de médecine personnalisée. Par conséquent pour développer des modèles de médecine personnalisée performants et robustes, le chercheur doit — à ce jour — s'appuyer sur une quantité importante de connaissance médicale experte en testant un ensemble restreint d'hypothèses prédéfinies (voir Chapitre 4 et [60]).
- Le risque de sur-apprentissage et d'échec de généralisation est très élevé pour les modèles de médecine personnalisée. Pour cette raison, il est indispensable d'évaluer précisément la validité interne et/ou externe de ces modèles (voir Chapitre 6 et [54]).
- Pour un modèle de médecine personnalisée, confirmer sa bonne performance par validation externe n'est pas suffisant pour recommander l'utilisation du modèle en routine. Paradoxalement, pour justifier l'implémentation d'un modèle de médecine personnalisée, le bénéfice de l'utilisation de celui-ci doit pouvoir être mesuré à l'échelle populationnelle (voir Chapitre 5 et [57]).

Pour continuer à faire des progrès dans le domaine de la médecine personnalisée plusieurs pistes de recherche sont possibles. Certaines options nous semblent cependant prioritaires. D'abord, il faudra réaliser le travail infrastructurel et de régulation nécessaire pour améliorer la disponibilité de textes, images et autres données cliniques non structurées, en veillant à ne pas prendre de risque déraisonnable en ce qui concerne la confidentialité. Du point de vue méthodologique, il faudra proposer de nouveaux modèles de médecine personnalisée, plus généraux, explicables, et produisant du texte en langage naturel. Enfin, la place de l'apprentissage par renforcement pour la médecine personnalisée devra être redéfinie. Ces méthodes ne devront pas être abandonnées ; au contraire, en utilisant les retours d'utilisateurs (*reinforcement learning with human feedback*), elles pourraient permettre le réglage fin des modèles plus généraux mentionnés précédemment.

Ces éléments nous semblent importants à la date d'écriture (juin 2023). Ceux-ci devront rapidement être réévalués afin que l'objectif de la recherche soit maintenu aligné avec les valeurs des patients hospitaliers, mais également de tous les êtres humains pouvant bénéficier de stratégies de prévention personnalisée.

Bibliographie

1. Kellum, J. A. *et al.* Acute kidney injury. *Nature reviews Disease primers* **7**, 52 (2021).
2. Kellum, J. A. *et al.* Kidney disease: improving global outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury. *Kidney international supplements* **2**, 1–138 (2012).
3. Gaudry, S., Palevsky, P. M. & Dreyfuss, D. Extracorporeal kidney-replacement therapy for acute kidney injury. *New England Journal of Medicine* **386**, 964–975 (2022).
4. Teitelbaum, I. Peritoneal dialysis. *New England Journal of Medicine* **385**, 1786–1795 (2021).
5. Rabindranath, K. S., Adams, J., MacLeod, A. M. & Muirhead, N. Intermittent versus continuous renal replacement therapy for acute renal failure in adults. *Cochrane Database of Systematic Reviews* (2007).
6. Quenot, J.-P. *et al.* Very high volume hemofiltration with the Cascade system in septic shock patients. *Intensive Care Medicine* **41**, 2111–2120 (2015).
7. Vijayan, A., Santos, R. B. D., Li, T., Goss, C. W. & Palevsky, P. M. Effect of frequent dialysis on renal recovery: results from the acute renal failure trial network study. *Kidney International Reports* **3**, 456–463 (2018).
8. Gaudry, S. *et al.* Initiation strategies for renal-replacement therapy in the intensive care unit. *New England Journal of Medicine* **375**, 122–133 (2016).
9. Barbar, S. D. *et al.* Timing of renal-replacement therapy in patients with acute kidney injury and sepsis. *New England Journal of Medicine* **379**, 1431–1442 (2018).
10. Investigators, S.-A. *et al.* Timing of initiation of renal-replacement therapy in acute kidney injury. *New England Journal of Medicine* (2020).
11. Gaudry, S. *et al.* Delayed versus early initiation of renal replacement therapy for severe acute kidney injury: a systematic review and individual patient data meta-analysis of randomised clinical trials. *The Lancet* **395**, 1506–1515 (2020).
12. Evans, L. *et al.* Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Medicine* **47**, 1181–1247 (2021).
13. Gaudry, S. *et al.* Comparison of two delayed strategies for renal replacement therapy initiation for severe acute kidney injury (AKIKI 2): a multicentre, open-label, randomised, controlled trial. *The Lancet* **397**, 1293–1300 (2021).
14. Laffey, J. G. & Kavanagh, B. P. Negative trials in critical care: why most research is probably wrong. *The Lancet Respiratory Medicine* **6**, 659–660 (2018).

15. Semler, M. W. *et al.* Identifying clinical research priorities in adult pulmonary and critical care. NHLBI Working Group report. *American Journal of Respiratory and Critical Care Medicine* **202**, 511–523 (2020).
16. Phillips, C. J. Precision medicine and its imprecise history. *Harvard Data Science Review* **2** (2020).
17. Romond, E. H. *et al.* Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *New England Journal of Medicine* **353**, 1673–1684 (2005).
18. Messersmith, W. A. & Ahnen, D. J. Targeting EGFR in colorectal cancer. *New England Journal of Medicine* **359**, 1834–1836 (2008).
19. Druker, B. J. *et al.* Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *New England Journal of Medicine* **355**, 2408–2417 (2006).
20. Ostermann, M. *et al.* Recommendations on acute kidney injury biomarkers from the acute disease quality initiative consensus conference: a consensus statement. *JAMA network open* **3**, e2019209–e2019209 (2020).
21. Ronco, C., Bellomo, R. & Kellum, J. A. Acute kidney injury. *The Lancet* **394**, 1949–1964 (2019).
22. Murad, M. H., Asi, N., Alsawas, M. & Alahdab, F. New evidence pyramid. *BMJ Evidence-Based Medicine* **21**, 125–127 (2016).
23. Kent, D. M. *et al.* The predictive approaches to treatment effect heterogeneity (PATH) statement. *Annals of Internal Medicine* **172**, 35–45 (2020).
24. Gaudry, S. *et al.* Timing of renal replacement therapy for severe acute kidney injury in critically ill patients. *American Journal of Respiratory and Critical Care Medicine* **199**, 1066–1075 (2019).
25. Bagshaw, S. M., Hoste, E. A. & Wald, R. When should we start renal-replacement therapy in critically ill patients with acute kidney injury: do we finally have the answer? *Critical Care* **25**, 1–4 (2021).
26. Almirall, D., Nahum-Shani, I., Sherwood, N. E. & Murphy, S. A. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. *Translational Behavioral Medicine* **4**, 260–274 (2014).
27. Neyman, J. On the application of probability theory to agricultural experiments. Essay on Principles. Section 9 (translation published in 1990). *Statistical Science* **5**. Publisher: Institute of Mathematical Statistics, 472–480 (Nov. 1923).
28. Rubin, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66**, 688 (1974).
29. Rubin, D. B. Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*, 34–58 (1978).
30. Horvitz, D. G. & Thompson, D. J. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**, 663–685 (1952).
31. Rubinstein, R. Y. & Kroese, D. P. *Simulation and the Monte Carlo method* Section 5.7.1 Importance sampling & weighted samples (John Wiley & Sons, 2016).

32. Lunceford, J. K. & Davidian, M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine* **23**, 2937–2960 (2004).
33. Robins, J. M., Rotnitzky, A. & Zhao, L. P. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association* **89**, 846–866 (1994).
34. Schuler, M. S. & Rose, S. Targeted maximum likelihood estimation for causal inference in observational studies. *American Journal of Epidemiology* **185**, 65–73 (2017).
35. Chernozhukov, V. *et al.* Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal* **21**, C1–C68 (2018).
36. Tsiatis, A. A., Davidian, M., Holloway, S. T. & Laber, E. B. *Dynamic treatment regimes: Statistical methods for precision medicine* (CRC press, 2019).
37. Boos, D. D., Stefanski, L. A., *et al.* *Essential statistical inference* (Springer, 2013).
38. Colnet, B., Josse, J., Varoquaux, G. & Scornet, E. Risk ratio, odds ratio, risk difference... Which causal measure is easier to generalize? *arXiv preprint arXiv:2303.16008* (2023).
39. Lin, D. On the Breslow estimator. *Lifetime Data Analysis* **13**, 471–480 (2007).
40. Fine, J. Comparing nonnested Cox models. *Biometrika* **89**, 635–648 (2002).
41. Nie, X. & Wager, S. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika* **108**, 299–319 (2021).
42. Kennedy, E. H. Towards optimal doubly robust estimation of heterogeneous causal effects. *arXiv preprint arXiv:2004.14497* (2020).
43. Küngel, S. R., Sekhon, J. S., Bickel, P. J. & Yu, B. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences* **116**, 4156–4165 (2019).
44. Zhao, Y., Zeng, D., Rush, A. J. & Kosorok, M. R. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association* **107**, 1106–1118 (2012).
45. Hill, J. L. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics* **20**, 217–240 (2011).
46. Blatt, D., Murphy, S. A. & Zhu, J. A-learning for approximate planning. *Ann Arbor* **1001**, 48109–2122 (2004).
47. Hansotia, B. & Rukstales, B. Incremental value modeling. *Journal of Interactive Marketing* **16**, 35–46 (2002).
48. Hahn, P. R., Murray, J. S. & Carvalho, C. M. Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects (with discussion). *Bayesian Analysis* **15**, 965–1056 (2020).
49. Athey, S., Tibshirani, J. & Wager, S. Generalized random forests (2019).
50. Powers, S. *et al.* Some methods for heterogeneous treatment effect estimation in high dimensions. *Statistics in Medicine* **37**, 1767–1787 (2018).
51. Robinson, P. M. Root-N-consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, 931–954 (1988).

52. Athey, S. & Wager, S. Policy learning with observational data. *Econometrica* **89**, 133–161 (2021).
53. Bouvier, F. *et al.* Do machine learning methods lead to similar individualized treatment rules? A comparison study on real data. *arXiv preprint arXiv:2308.03398* (2023).
54. Van Klaveren, D., Steyerberg, E. W., Serruys, P. W. & Kent, D. M. The proposed ‘concordance-statistic for benefit’ provided a useful metric when modeling heterogeneous treatment effects. *Journal of Clinical Epidemiology* **94**, 59–68 (2018).
55. Cook, N. R. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* **115**, 928–935 (2007).
56. Xia, Y., Gustafson, P. & Sadatsafavi, M. Methodological concerns about “concordance-statistic for benefit” as a measure of discrimination in predicting treatment benefit. *Diagnostic and Prognostic Research* **7**, 10 (2023).
57. Imai, K. & Li, M. L. Experimental evaluation of individualized treatment rules. *Journal of the American Statistical Association* **118**, 242–256 (2023).
58. Sadatsafavi, M., Mansournia, M. A. & Gustafson, P. A threshold-free summary index for quantifying the capacity of covariates to yield efficient treatment rules. *Statistics in Medicine* **39**, 1362–1373 (2020).
59. Janes, H., Brown, M. D., Huang, Y. & Pepe, M. S. An approach to evaluating and comparing biomarkers for patient treatment selection. *The International Journal of Biostatistics* **10**, 99–121 (2014).
60. Takahashi, K. *et al.* Redevelopment and validation of the SYNTAX score II to individualise decision making between percutaneous and surgical revascularisation in patients with complex coronary artery disease: secondary analysis of the multicentre randomised controlled SYNTAXES trial with external cohort validation. *The Lancet* **396**, 1399–1412 (2020).
61. Hastie, T., Tibshirani, R., Friedman, J. H. & Friedman, J. H. *The elements of statistical learning: data mining, inference, and prediction* Chapter 5, 151–156 (Springer, 2009).
62. Jacobs, R. A., Jordan, M. I., Nowlan, S. J. & Hinton, G. E. Adaptive mixtures of local experts. *Neural Computation* **3**, 79–87 (1991).
63. Jordan, M. I. & Jacobs, R. A. Hierarchical mixtures of experts and the EM algorithm. *Neural Computation* **6**, 181–214 (1994).
64. Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: series B (methodological)* **39**, 1–22 (1977).
65. Jiang, W. & Tanner, M. A. On the identifiability of mixtures-of-experts. *Neural Networks* **12**, 1253–1258 (1999).
66. Moodie, E. E., Dean, N. & Sun, Y. R. Q-learning: Flexible learning about useful utilities. *Statistics in Biosciences* **6**, 223–243 (2014).
67. Murphy, S. A. A generalization error for Q-learning. *Journal of Machine Learning Research* **6**, 1073–1097 (2005).

68. Chakraborty, B. & Moodie, E. E. Statistical methods for dynamic treatment regimes. *Springer-Verlag*. doi **10**, 978–1 (2013).
69. Robins, J. M., Hernan, M. A. & Brumback, B. Marginal structural models and causal inference in epidemiology. *Epidemiology*, 550–560 (2000).
70. Precup, D. Eligibility traces for off-policy policy evaluation. *Computer Science Department Faculty Publication Series*, 80 (2000).
71. Zhang, B., Tsiatis, A. A., Laber, E. B. & Davidian, M. Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika* **100**, 681–694 (2013).
72. Nie, X., Brunskill, E. & Wager, S. Learning when-to-treat policies. *Journal of the American Statistical Association* **116**, 392–409 (2021).
73. Wallace, M. P. & Moodie, E. E. Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics* **71**, 636–644 (2015).
74. Li, Z., Chen, J., Laber, E., Liu, F. & Baumgartner, R. Optimal treatment regimes: a review and empirical comparison. *International Statistical Review* (2023).
75. Simoneau, G., Moodie, E. E., Nijjar, J. S., Platt, R. W., Investigators, S. E. R. A. I. C., et al. Estimating optimal dynamic treatment regimes with survival outcomes. *Journal of the American Statistical Association* **115**, 1531–1539 (2020).
76. Cho, H., Holloway, S. T., Couper, D. J. & Kosorok, M. R. Multi-stage optimal dynamic treatment regimes for survival outcomes with dependent censoring. *Biometrika* **110**, 395–410 (2023).
77. Andersen, P. K. & Pohar Perme, M. Pseudo-observations in survival analysis. *Statistical Methods in Medical Research* **19**, 71–99 (2010).
78. Overgaard, M., Parner, E. T. & Pedersen, J. Pseudo-observations under covariate-dependent censoring. *Journal of Statistical Planning and Inference* **202**, 112–122 (2019).
79. Stolley, P. D. When genius errs: RA Fisher and the lung cancer controversy. *American Journal of Epidemiology* **133**, 416–425 (1991).
80. Hill, A. B. *The environment and disease: association or causation?* 1965.
81. Lee, J. et al. Association between fluid balance and survival in critically ill patients. *Journal of Internal Medicine* **277**, 468–477 (2015).
82. Krzych, Ł. J. & Czempik, P. F. Impact of furosemide on mortality and the requirement for renal replacement therapy in acute kidney injury: a systematic review and meta-analysis of randomised trials. *Annals of Intensive Care* **9**, 1–9 (2019).
83. Kahneman, D. *Thinking, fast and slow* Chapter 6, section: “seeing causes and intentions”, 74–78 (Penguin, 2011).
84. Leslie, A. M. & Keeble, S. Do six-month-old infants perceive causality? *Cognition* **25**, 265–288 (1987).
85. Susskind, L. & Friedman, A. *Quantum mechanics: the theoretical minimum* Chapter 4: Time and change, 93–94 (Basic Books, 2014).
86. Carroll, S. *Do Cause and Effect Really Exist?* Minute Physics. 2017. <https://www.youtube.com/watch?v=3AMCcYnAsdQ>.

87. Wolfram, S. A class of models with the potential to represent fundamental physics. *arXiv preprint arXiv:2004.08210* (2020).
88. Pearl, J. *The Art and Science of Cause and Effect?* UCLA Faculty Research Lectureship Program. 1996. <http://bayes.cs.ucla.edu/BOOK-2K/causality2-epilogue.pdf>.
89. Hernán, M. A. The C-word: scientific euphemisms do not improve causal inference from observational data. *American Journal of Public Health* **108**, 616–619 (2018).
90. Zhang, B. & Zhang, M. Variable selection for estimating the optimal treatment regimes in the presence of a large number of covariates. *The Annals of Applied Statistics* **12**, 2335–2358 (2018).
91. Botvinick, M. *et al.* Reinforcement learning, fast and slow. *Trends in Cognitive Sciences* **23**, 408–422 (2019).
92. Spelke, E. S. & Kinzler, K. D. Core knowledge. *Developmental Science* **10**, 89–96 (2007).
93. Gulwani, S., Polozov, O., Singh, R., *et al.* Program synthesis. *Foundations and Trends in Programming Languages* **4**, 1–119 (2017).
94. Wang, P. On defining artificial intelligence. *Journal of Artificial General Intelligence* **10**, 1–37 (2019).
95. Finlayson, S. G. *et al.* The clinician and dataset shift in artificial intelligence. *New England Journal of Medicine* **385**, 283–286 (2021).
96. Moor, M. *et al.* Foundation models for generalist medical artificial intelligence. *Nature* **616**, 259–265 (2023).
97. Chollet, F. On the measure of intelligence. *arXiv preprint arXiv:1911.01547* (2019).
98. Sambasivan, N. *et al.* “Everyone wants to do the model work, not the data work”: *Data Cascades in High-Stakes AI* in proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (2021), 1–15.
99. Reed, S. *et al.* A generalist agent. *arXiv preprint arXiv:2205.06175*. Voir section 2.1 Tokenization (2022).
100. Imrie, F., Cebere, B., McKinney, E. F. & van der Schaar, M. AutoPrognosis 2.0: Democratizing Diagnostic and Prognostic Modeling in Healthcare with Automated Machine Learning. *arXiv preprint arXiv:2210.12090* (2022).
101. Yoon, J., Drumright, L. N. & Van Der Schaar, M. Anonymization through data synthesis using generative adversarial networks (ads-gan). *IEEE journal of biomedical and health informatics* **24**, 2378–2388 (2020).
102. Goodfellow, I., Bengio, Y. & Courville, A. Manifold Learning (Section 5.11.3). *Deep learning*. MIT press (2016).
103. Chollet, F. Fundamentals of machine learning (Section 5.2.1). *Deep learning with Python*. Manning Publications Co (2021).
104. Rajpurkar, P. & Lungren, M. P. The Current and Future State of AI Interpretation of Medical Images. *New England Journal of Medicine* **388**, 1981–1990 (2023).
105. Han, X. *et al.* Pre-trained models: Past, present and future. *AI Open* **2**, 225–250 (2021).

106. Jha, S. & Topol, E. J. Upending the model of AI adoption. *The Lancet* **401**, 1920 (2023).
107. Yu, C., Liu, J., Nemati, S. & Yin, G. Reinforcement learning in healthcare: A survey. *ACM Computing Surveys (CSUR)* **55**, 1–36 (2021).
108. Weltz, J., Volfovsky, A. & Laber, E. B. Reinforcement learning methods in public health. *Clinical therapeutics* **44**, 139–154 (2022).
109. Sutton, R. S. & Barto, A. G. *Reinforcement learning: An introduction* 17.4 Designing Reward Signals, 469–472 (MIT press, 2018).
110. Sutton, R. S. The bitter lesson. *Incomplete Ideas* **13** (2019).
111. Lampinen, A. K. *et al.* Can language models learn from explanations in context? *arXiv preprint arXiv:2204.02329* (2022).
112. Singhal, K. *et al.* Large language models encode clinical knowledge. *Nature*, 1–9 (2023).
113. Ouyang, L. *et al.* Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems* **35**, 27730–27744 (2022).
114. Nahum-Shani, I. *et al.* Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine* **52**, 446–462 (2018).
115. Gauci, J. *et al.* Horizon: Facebook’s open source applied reinforcement learning platform. *arXiv preprint arXiv:1811.00260* (2018).
116. Twenge, J. M., Haidt, J., Joiner, T. E. & Campbell, W. K. Underestimating digital media harm. *Nature Human Behaviour* **4**, 346–348 (2020).
117. Ghosh, S. *et al.* Did we personalize? Assessing personalization by an online reinforcement learning algorithm using resampling. *arXiv preprint arXiv:2304.05365* (2023).

Annexe A : Un modèle causal général pour l'estimation de CACE sans hypothèses de monotonicité ni d'exclusion restriction

I have a friend, a writer, who makes the core of his life an act of imagination. Is it escape or is it liberation? You tell me, I don't know anything about these things.

Philip Glass, *A portrait of philip in twelve parts*

Ce chapitre ne traite pas de la médecine personnalisée. Il fait le lien entre différents outils théoriques utilisés dans cette thèse.

La musique associée à ce chapitre est ‘別人的- 電影《孤味》片尾曲’ de Vivian Hsu.

Un modèle causal général pour l'estimation de CACE sans hypothèses de monotonicité ni d'exclusion restriction

1 Introduction

Les essais randomisés à deux bras avec une probabilité d'allocation fixe sont courants en recherche clinique. Les récentes recommandations internationales ICH E9 insistent sur la nécessité de clarifier les estimands lors de l'analyse de ces essais [1]. Dans ce travail nous nous focalisons sur un estimand particulier, le *Complier Average Causal Effect* (CACE),¹ qui a été étudié à la fois dans la littérature de biostatistique [2, 3] et d'économétrie [4, 5].

Le rationnel du CACE, s'appuie sur le fait que souvent, dans un essai randomisé, certains patients ne prennent pas le traitement correspondant au bras dans lequel ils ont été alloués. En d'autres termes, l'observance (*compliance*) des patients est imparfaite. Le CACE vise à évaluer l'effet de : traitement intervention versus traitement contrôle — uniquement chez les patients qui prendront le traitement tel qu'alloué (*the compliers*).

Classiquement les procédures d'estimations pour CACE s'appuient sur les hypothèses de monotonicité [6] (*no defiers assumption* : absence de patients “contradicteurs” prenant systématiquement le traitement opposé à celui qui leur est alloué) et/ou d'exclusion restriction [7] (*exclusion restriction assumption* : la randomisation n'a pas d'effet en soi, seul le traitement pris à un effet). L'hypothèse de monotonicité, n'est pas facilement testable dans les données et l'hypothèse d'exclusion restriction peut être violée dans les essais en ouvert (voir “biais de performance”, décrit classiquement en épidémiologie [8]).

Dans ce travail nous proposons un modèle causal général permettant d'estimer CACE sans faire d'hypothèse de monotonicité, ni d'hypothèse d'exclusion restriction. Notre méthodologie s'appuie sur l'hypothèse d'ignorabilité principale (*principal ignorability*) [9]. Cette hypothèse s'interprète comme le fait que l'information décrivant le statut latent observant (*complier*), traité systématique (*always-taker*), non-traité systématique (*never-taker*), ou contradicteur (*defier*) d'un patient est contenue dans ses covariables de pré-randomisation. La procédure d'estimation que nous proposons s'appuie sur l'utilisation de modèles de mélange d'experts dont les paramètres sont estimés par des algorithmes EM [10].

À notre connaissance, aucun modèle causal général n'a été proposé pour permettre d'estimer CACE sans hypothèses de monotonicité ni d'exclusion restriction dans un cadre d'inférence fréquentiste. En s'appuyant sur un cadre Bayesien, Imbens et Rubin ont développé un cadre causal général pour l'estimation de CACE [11]. Cependant, ces auteurs n'étudient pas l'identificabilité du modèle

¹En économétrie cet estimand est parfois dénommé *Local Average Treatment Effect*.

qu'ils utilisent dans leur procédure d'estimation. À notre connaissance aucune implémentation ne reposant que sur l'hypothèse d'ignorabilité principale pour facilement estimer CACE.

Nos contributions dans ce travail sont les suivantes : i) nous développons un modèle causal général permettant de dériver un nouvel estimateur pour CACE s'appuyant uniquement sur l'hypothèse d'ignorabilité principale, ii) nous prouvons l'identifiabilité du modèle sur lequel repose notre procédure d'estimation, iii) nous proposons des estimateurs à priori de variance inférieure dans les situations où les hypothèses de monotonicité et/ou d'exclusion restriction sont vérifiées, iv) nous proposons une extension de notre méthodologie permettant l'utilisation de modèles non-paramétriques (e.g., boosting, lasso, neural networks, etc.).

L'article est structuré de la manière suivante. Dans la section 2, nous introduisons notre modèle causal ; dans la section 3, nous introduisons des notations utiles pour décrire notre procédure d'estimation ; dans la section 4, nous reformulons l'estimand d'intérêt, CACE, à la lumière de notre modèle causal ; dans la section 5 nous détaillons notre procédure d'estimation en l'absence d'hypothèses de monotonicité et d'exclusion restriction ; dans la section 6, nous détaillons l'estimation de la variance asymptotique de notre estimateur de CACE ; dans la section 7, nous proposons d'autres estimateurs de CACE quand les hypothèses de monotonicité et/ou d'exclusion restriction sont vérifiées ; dans la section 8, nous détaillons les perspectives envisagées pour ce travail. Le théorème d'identifiabilité pour le modèle de mélange sur lequel repose notre procédure d'estimation est énoncé en annexe.

2 Modèle probabiliste et hypothèses causales

En se rappelant du contexte d'un essai randomisé avec allocation Z on a

$$Z \sim Bernoulli(p = \alpha), \alpha \in]0; 1[. \quad (\text{Exogeneity assumption})$$

Nous formalisons maintenant les notions de *complier*, *always taker*, *never taker*, *defier*. Pour cela, nous définissons les traitements potentiels

$$T^{s=c} \stackrel{\text{def}}{=} Z, \quad T^{s=a} \stackrel{\text{def}}{=} 1, \quad T^{s=n} \stackrel{\text{def}}{=} 0, \quad \text{et} \quad T^{s=d} \stackrel{\text{def}}{=} 1 - Z$$

qui correspondent au traitement que recevrait un *complier*, un *always taker*, un *never taker*, et un *defier* respectivement. Le s indique la “strate” *compliers*, *always takers*, *never takers*, *defiers* d'un patient. La strate d'un patient n'étant en pratique jamais observée on considérera dans ce travail qu'il s'agit d'une variable latente. Introduisons X , les covariables prérandomisation :

$$X \sim f_X.$$

Introduisons $(C, A, N, D)^T$, le vecteur aléatoire d'encodage 1 parmi 4 décrivant la strate des patients. C'est à dire les vecteurs $(1, 0, 0, 0)^T$, $(0, 1, 0, 0)^T$, $(0, 0, 1, 0)^T$, $(0, 0, 0, 1)^T$ indiquent qu'un patient est un *complier*, *always taker*, *never taker*, *defier* respectivement. Notons $\rho_c(X) \stackrel{\text{def}}{=} \mathbb{E}(C|X)$, $\rho_a(X) \stackrel{\text{def}}{=} \mathbb{E}(A|X)$, $\rho_n(X) \stackrel{\text{def}}{=} \mathbb{E}(N|X)$, et $\rho_d(X) \stackrel{\text{def}}{=} \mathbb{E}(D|X)$ les probabilités conditionnelles qu'un patient soit un *complier*, *always taker*, *never taker*, *defier* respectivement. Dans ces conditions, on a

:

$$(C, A, N, D)^T | X \sim \text{Multinomial}\left(n = 1, k = 4, p = (\rho_c(X), \rho_a(X), \rho_n(X), \rho_d(X))^T\right).$$

Remarquons que nous ne faisons aucune hypothèse paramétrique sur le mécanisme génératif de $(C, A, N, D)^T$. Le traitement pris par un patient est

$$\begin{aligned} T &\stackrel{\text{def}}{=} CT^{s=c} + AT^{s=a} + NT^{s=n} + DT^{s=d} \\ &= C \times Z + A \times 1 + N \times 0 + D \times (1 - Z). \end{aligned}$$

Supposons également l'existence de résultats potentiels élémentaires (*elementary potential outcomes*) :

$$Y^{s=k, z=l, t=m} | X \sim f_{Y^{s=k, z=l, t=m} | X}$$

pour $k \in \{c, a, n, d\}$, $l \in \{0, 1\}$ et $m \in \{0, 1\}$. En formalisant le problème de cette manière, on remarque par exemple que le résultat potentiel sous traitement d'un *never taker* n'a pas de sens. En clair, il n'est pas nécessaire d'introduire les variables² $Y^{s=n, z=0, t=1}$ et $Y^{s=n, z=1, t=1}$. Le modèle ci-dessus implique cependant que

$$Y^{s=k, z=l, t=m} \perp\!\!\!\perp (C, A, N, D) | X \quad \forall k, l, m. \quad (\text{Principal ignorability})$$

On introduit les résultats potentiels $Y^{t=1}$ et $Y^{t=0}$ par "sélection" des résultats potentiels élémentaires compatibles avec la strate (C, A, N, D) et l'allocation Z d'un patient :

$$\begin{aligned} Y^{t=1} &\stackrel{\text{def}}{=} CY^{s=c, z=1, t=1} + AZY^{s=a, z=1, t=1} + A(1 - Z)Y^{s=a, z=0, t=1} + DY^{s=d, z=0, t=1}, \\ Y^{t=0} &\stackrel{\text{def}}{=} CY^{s=c, z=0, t=0} + NZY^{s=n, z=1, t=0} + N(1 - Z)Y^{s=n, z=0, t=0} + DY^{s=d, z=1, t=0}. \end{aligned}$$

Remarquez que cette définition pour $Y^{t=1}$ et $Y^{t=0}$ n'est pas restrictive. En particulier, elle implique la possibilité qu'un *alway-taker* donné ait un résultat potentiel $Y^{t=1}$ différent s'il est randomisé avec $Z = 0$ ou avec $Z = 1$; c'est à dire, il est possible que $Y^{s=a, z=0, t=1} \neq Y^{s=a, z=1, t=1}$. De la même manière, un *never-taker* donné peut avoir un résultat potentiel $Y^{t=0}$ différent s'il est randomisé avec $Z = 0$ ou avec $Z = 1$; c'est à dire, il est possible que $Y^{s=n, z=0, t=0} \neq Y^{s=n, z=1, t=0}$. En d'autres termes, on ne fait pas d'hypothèse de type exclusion-restriction.

On introduit enfin l'hypothèse suivante

$$Y = TY^{t=1} + (1 - T)Y^{t=0}. \quad (\text{Consistency})$$

Pour la clarté de l'exposé, nous nous focaliserons sur le cas où Y est binaire. L'extension au cas où Y est continu est naturelle ; celle-ci est énoncée en annexe.

²Pour des raisons similaires il n'est pas nécessaire d'introduire les variables $Y^{s=a, z=1, t=0}$, $Y^{s=a, z=0, t=0}$, $Y^{s=c, z=1, t=0}$, $Y^{s=c, z=0, t=0}$, $Y^{s=d, z=1, t=1}$, et $Y^{s=d, z=0, t=0}$.

3 Notations

Par soucis de clarté pour la suite de l'exposé, définissons

$$\begin{aligned} P_{c11}(X) &\stackrel{\text{def}}{=} \mathbb{E}(C|Z=1, T=1, X), \\ P_{a11}(X) &\stackrel{\text{def}}{=} \mathbb{E}(A|Z=1, T=1, X), \\ P_{c00}(X) &\stackrel{\text{def}}{=} \mathbb{E}(C|Z=0, T=0, X), \\ P_{n00}(X) &\stackrel{\text{def}}{=} \mathbb{E}(N|Z=0, T=0, X). \end{aligned}$$

Définissons également la généralisation du score de propension dans un contexte de compliance imparfaite à la randomisation

$$\pi(X, Z) \stackrel{\text{def}}{=} \mathbb{E}(T|X, Z)$$

ainsi que les probabilités conditionnelles de prise du traitement et d'allocation au traitement respectivement $e(X) \stackrel{\text{def}}{=} \mathbb{E}[T|X]$ et $\eta(X) \stackrel{\text{def}}{=} \mathbb{E}[Z|X] = \alpha$.

Définissons maintenant

$$\begin{aligned} q_{c11}(X) &\stackrel{\text{def}}{=} \mathbb{E}(Y^{s=c, z=1, t=1}|X), \\ q_{a11}(X) &\stackrel{\text{def}}{=} \mathbb{E}(Y^{s=a, z=1, t=1}|X), \\ q_{c00}(X) &\stackrel{\text{def}}{=} \mathbb{E}(Y^{s=c, z=0, t=0}|X), \\ q_{n00}(X) &\stackrel{\text{def}}{=} \mathbb{E}(Y^{s=a, z=0, t=0}|X). \end{aligned}$$

Définissons aussi,

$$\begin{aligned} q_{11}(X) &\stackrel{\text{def}}{=} \mathbb{E}(Y|Z=1, T=1, X), \\ q_{00}(X) &\stackrel{\text{def}}{=} \mathbb{E}(Y|Z=0, T=0, X), \end{aligned}$$

et enfin $q(X) \stackrel{\text{def}}{=} \mathbb{E}(Y|X)$.

4 Complier average causal effect

Notre target estimand, le “CACE” est

$$\Delta \stackrel{\text{def}}{=} \mathbb{E}(Y^{t=1} - Y^{t=0}|C=1).$$

La méthode d'estimation de CACE proposée dans ce travail s'appuie sur le résultat suivant.

Lemme 1. *Le CACE peut être représenté de la manière suivante :*

$$\Delta = \mathbb{E} \left[\{q_{c11}(X) - q_{c00}(X)\} \rho_c(X) \right] / \mathbb{E}[\rho_c(X)].$$

Preuve.

$$\begin{aligned}
\Delta &\stackrel{\text{def}}{=} \mathbb{E}[Y^{t=1} - Y^{t=0}|C = 1] \\
&= \mathbb{E}[Y^{s=c, z=1, t=1} - Y^{s=c, z=0, t=0}|C = 1] \\
&= \mathbb{E}\left[\mathbb{E}(Y^{s=c, z=1, t=1} - Y^{s=c, z=0, t=0}|X)|C = 1\right] \\
&= \mathbb{E}\left[\mathbb{E}(Y^{s=c, z=1, t=1}|X) - \mathbb{E}(Y^{s=c, z=0, t=0}|X)|C = 1\right] \\
&= \mathbb{E}\left[q_{c11}(X) - q_{c00}(X)|C = 1\right] \\
&= \int_{\mathcal{X}} dx \{q_{c11}(x) - q_{c00}(x)\} f_{X|C}(x|1) \\
&= \int_{\mathcal{X}} dx \{q_{c11}(x) - q_{c00}(x)\} \frac{f_{C|X}(1|x)}{f_C(1)} f_X(x) \\
&= \mathbb{E}\left[\{q_{c11}(X) - q_{c00}(X)\} \frac{\mathbb{E}(C|X)}{\mathbb{E}(C)}\right] \\
&= \mathbb{E}\left[\{q_{c11}(X) - q_{c00}(X)\} \mathbb{E}(C|X)\right] / \mathbb{E}(C) \\
&= \mathbb{E}\left[\{q_{c11}(X) - q_{c00}(X)\} \mathbb{E}(C|X)\right] / \mathbb{E}[\mathbb{E}(C|X)] \\
&= \mathbb{E}\left[\{q_{c11}(X) - q_{c00}(X)\} \rho_c(X)\right] / \mathbb{E}[\rho_c(X)]
\end{aligned}$$

□

Le but de la procédure d’inférence proposée dans ce travail est d’estimer les fonctions $q_{c11}(\cdot)$, $q_{c00}(\cdot)$, et $\rho_c(\cdot)$ afin de dériver un estimateur “plug-in” pour Δ .

5 Inférence

On considère une expérience $(X_i, C_i, A_i, N_i, D_i, Z_i, T_i, Y_i)_{1 \leq i \leq n} \stackrel{\text{iid}}{\sim} \mathcal{P}$ où seules les données $(X_i, Z_i, T_i, Y_i)_{1 \leq i \leq n}$ sont observables, les variables $(C_i, A_i, N_i, D_i)^T$ étant considérées comme latentes.

5.1 Etape 1 : estimation simultanée des fonctions ρ_c , ρ_a , ρ_n , ρ_d

Lemme 2. *Notre procédure d’estimation va crucialement utiliser le fait que*

$$\pi(X, Z) = \sum_{k \in \{c, a, n, d\}} \rho_k(X) \mu_k(Z)$$

où on note $\mu_k(Z) \stackrel{\text{def}}{=} \mathbb{E}(T^{s=k}|Z)$.

Preuve.

$$\begin{aligned}
\pi(X, Z) &\stackrel{\text{def}}{=} \mathbb{E}(T|X, Z) \\
&= \mathbb{E}(CT^{s=c} + AT^{s=a} + NT^{s=n} + DT^{s=d}|X, Z) \\
&= \mathbb{E}(C|X, Z) \mathbb{E}(T^{s=c}|X, Z) + \mathbb{E}(A|X, Z) \mathbb{E}(T^{s=a}|X, Z) \\
&\quad + \mathbb{E}(N|X, Z) \mathbb{E}(T^{s=n}|X, Z) + \mathbb{E}(D|X, Z) \mathbb{E}(T^{s=d}|X, Z) \\
&= \mathbb{E}(C|X) \mathbb{E}(T^{s=c}|Z) + \mathbb{E}(A|X, Z) \mathbb{E}(T^{s=a}|Z) \\
&\quad + \mathbb{E}(N|X, Z) \mathbb{E}(T^{s=n}|Z) + \mathbb{E}(D|X, Z) \mathbb{E}(T^{s=d}|Z) \\
&= \rho_c(X)\mu_c(Z) + \rho_a(X)\mu_a(Z) + \rho_n(X)\mu_n(Z) + \rho_d(X)\mu_d(Z) \\
&= \sum_{k \in \{c, a, n, d\}} \rho_k(X)\mu_k(Z)
\end{aligned}$$

La troisième égalité vient du fait que $(C, A, N, D) \perp\!\!\!\perp T^{s=k}|X, Z, \forall k \in \{c, a, n, d\}$. \square

Le lemme 2 implique que $\pi(\cdot)$ peut être vue comme un mélange des fonctions connues $\mu_k(\cdot), k \in \{c, a, n, d\}$ pour laquelle $\rho_c(\cdot), \rho_a(\cdot), \rho_n(\cdot), \rho_d(\cdot)$, les composantes respectives du mélange, sont des fonctions inconnues.

En annexe de ce travail (Théorème 7), nous montrons que, sous hypothèses paramétriques pour $\rho_n(\cdot), k = \{c, a, n, d\}$, le modèle $\pi(x, z; \delta) = \sum_{k \in \{c, a, n, d\}} \rho_k(X; \delta)\mu_k(Z)$ est identifiable quand les covariables $x \in \mathcal{X}$ sont continues (voir remarque 8 pour le détail des conditions techniques). Nous proposons d'estimer simultanément $\rho_c(\cdot; \delta), \rho_a(\cdot; \delta), \rho_n(\cdot; \delta), \rho_d(\cdot; \delta)$ par l'algorithme d'EM No. 1. Une alternative d'estimation non paramétrique de ces fonctions est donnée dans l'algorithme EM-like No. 4.

5.2 Etape 2 : estimation distincte des fonctions q_{c11} et q_{c00}

Lemme 3. *Notre procédure d'estimation va cruciallement utiliser le fait que*

$$\begin{aligned}
(i) \quad q_{11}(X) &= P_{c11}(X)q_{c11}(X) + P_{a11}(X)q_{a11}, \\
(ii) \quad q_{00}(X) &= P_{c00}(X)q_{c00}(X) + P_{n11}(X)q_{n00}.
\end{aligned}$$

Preuve.

$$\begin{aligned}
q_{11}(X) &\stackrel{\text{def}}{=} \mathbb{E}(Y|Z = 1, T = 1, X) \\
&= \mathbb{E}(Y^{t=1}|Z = 1, T = 1, X) \\
&= \mathbb{E}(CY^{s=c, z=1, t=1} + AY^{s=a, z=1, t=1}|Z = 1, T = 1, X) \\
&= \mathbb{E}(C|Z = 1, T = 1, X) \mathbb{E}(Y^{s=c, z=1, t=1}|Z = 1, T = 1, X) \\
&\quad + \mathbb{E}(A|Z = 1, T = 1, X) \mathbb{E}(Y^{s=a, z=1, t=1}|Z = 1, T = 1, X) \\
&= \mathbb{E}(C|Z = 1, T = 1, X) \mathbb{E}(Y^{s=c, z=1, t=1}|X) + \mathbb{E}(A|Z = 1, T = 1, X) \mathbb{E}(Y^{s=a, z=1, t=1}|X) \\
&= P_{c11}(X)q_{c11}(X) + P_{a11}(X)q_{a11}
\end{aligned}$$

La deuxième égalité vient de *consistency*, la quatrième vient de *principal ignorability* et la cinquième

vient du fait que la formulation de notre problème implique

$$Y^{s=k, z=l, t=m} \perp\!\!\!\perp (Z, T) | X \quad \forall (k, l, m)$$

(ceci est vérifiable formellement par argument de d-separation sur le graphe acyclique correspondant à notre modèle probabiliste). Le même argument s'applique pour $q_{00}(X)$. \square

Lemme 4. *Les fonctions inconnues P_{c11} , P_{a11} , P_{c00} , P_{n00} peuvent être réexprimées selon les fonctions ρ_c , ρ_a , ρ_n , estimées à l'étape 1.*

- (i) $P_{c11}(X) = \rho_c(X)/\{\rho_c(X) + \rho_a(X)\},$
- (ii) $P_{a11}(X) = \rho_a(X)/\{\rho_c(X) + \rho_a(X)\},$
- (iii) $P_{c00}(X) = \rho_c(X)/\{\rho_c(X) + \rho_n(X)\},$
- (iv) $P_{n00}(X) = \rho_n(X)/\{\rho_c(X) + \rho_n(X)\}.$

Preuve.

$$\begin{aligned} P_{c11}(X) &\stackrel{\text{def}}{=} \mathbb{E}(C|Z=1, T=1, X) \\ &= \mathbb{P}(C=1|Z=1, T=1, X) \\ &= \mathbb{P}(T=1|C=1, Z=1, X) \frac{\mathbb{P}(C=1|Z=1, X)}{\mathbb{P}(T=1|Z=1, X)} \\ &= \frac{\mathbb{P}(C=1|Z=1, X)}{\mathbb{P}(T=1|Z=1, X)} \\ &= \frac{\mathbb{P}(C=1|X)}{\mathbb{E}(T|Z=1, X)} \\ &= \frac{\mathbb{E}(C|X)}{\mathbb{E}(C|X) + \mathbb{E}(A|X)} \\ &= \rho_c(X)/\{\rho_c(X) + \rho_a(X)\}. \end{aligned}$$

Un argument similaire s'applique pour P_{a11} , P_{c00} , et P_{n00} . \square

Le lemme 3 (i) implique que $q_{11}(\cdot)$ peut être vue comme un mélange de $q_{c11}(\cdot)$ et $q_{a11}(\cdot)$ (fonctions inconnues). Le lemme 4 (i) et (ii) implique que les composantes du mélange $P_{c11}(\cdot)$ et $P_{a11}(\cdot)$ peuvent être facilement estimées dès la fin de l'étape 1 par $\hat{P}_{c11}(\cdot) = \rho_c(\cdot; \hat{\delta})/\{\rho_c(\cdot; \hat{\delta}) + \rho_a(\cdot; \hat{\delta})\}$ et $\hat{P}_{a11}(\cdot) = \rho_a(\cdot; \hat{\delta})/\{\rho_c(\cdot; \hat{\delta}) + \rho_a(\cdot; \hat{\delta})\}$. Dans ces conditions, quand Y est binaire, nous proposons, d'estimer $q_{c11}(\cdot; \beta)$ par l'algorithme d'EM No. 2 en utilisant le sous échantillon $(X_i, Y_i)_{i:Z_i=1, T_i=1}$ avec les poids $(P_{c11}(X_i))_{i:Z_i=1, T_i=1}$. De la même manière nous proposons d'estimer $q_{c00}(\cdot; \gamma)$ par l'algorithme d'EM No. 2 en utilisant le sous échantillon $(X_i, Y_i)_{i:Z_i=0, T_i=0}$ avec les poids $(P_{c00}(X_i))_{i:Z_i=0, T_i=0}$.

Quand Y est continue, nous proposons, d'utiliser l'algorithme d'EM No. 3. Une alternative d'estimation non paramétrique des fonctions q_{c11} et q_{c00} est donnée dans l'algorithme EM-like No. 5 quand Y est binaire et dans l'algorithme EM-like No. 6 quand Y est continue.

5.3 Etape 3: estimation de CACE

En nous appuyant sur le lemme 1 nous proposons un estimateur “*plug-in*” ou “*principal ignorability*” pour Δ

$$\hat{\Delta}_{PI} = \frac{\sum_{i=1}^n \{q_{c11}(X_i; \hat{\beta}) - q_{c00}(X_i; \hat{\gamma})\} \rho_c(X_i; \hat{\delta})}{\sum_{i=1}^n \rho_c(X_i; \hat{\delta})}. \quad (1)$$

6 Propriétés asymptotiques

En réexprimant l'équation (1), on remarque que $\hat{\Delta}_{PI}$ est solution d'une équation de la forme

$$\sum_{i=1}^n \psi_0(X_i; \Delta_{PI}, \delta, \beta, \gamma) = 0$$

où

$$\psi_0(x; \Delta_{PI}, \delta, \beta, \gamma) = \{q_{c11}(x; \beta) - q_{c00}(x; \gamma) - \Delta_{PI}\} \rho_c(x; \delta).$$

On remarque par ailleurs que $\hat{\beta}$, $\hat{\gamma}$ et $\hat{\delta}$ sont les paramètres de modèles de mélange. Chaque paramètre est la solution d'une équation de “score” pour le modèle de mélange correspondant. En particulier, les paramètres $\hat{\delta}$ sont solution d'une équation de la forme

$$\sum_{i=1}^n \psi_1(X_i, Z_i, T_i; \delta) = 0$$

où

$$\psi_1(x, z, t; \delta) = \frac{\partial}{\partial \delta} \ln \left\{ \sum_{s \in \{c, a, n, d\}} \frac{\exp \delta_s^T x}{\sum_{k \in \{c, a, n, d\}} \exp \delta_k^T x} \mu_s(z)^t (1 - \mu_s(z))^{1-t} \right\}.$$

Dans le cas où Y est binaire,³ les paramètres $\hat{\beta}, \hat{\gamma}$ sont solutions d'équations de la forme

$$\begin{aligned} \sum_{i=1}^n \psi_2(X_i, Z_i, T_i, Y_i; \delta, \beta) &= 0, \\ \sum_{i=1}^n \psi_3(X_i, Z_i, T_i, Y_i; \delta, \gamma) &= 0 \end{aligned}$$

où

$$\begin{aligned} \psi_2(x, z, t, y; \delta, \beta) &= \mathbb{1}(z = 1) \mathbb{1}(t = 1) \frac{\partial}{\partial \beta} \ln \left\{ P_{c11}(x; \delta) \text{expit}(\beta_c^T x)^y (1 - \text{expit}(\beta_c^T x))^{1-y} \right. \\ &\quad \left. + P_{a11}(x; \delta) \text{expit}(\beta_{nc}^T x)^y (1 - \text{expit}(\beta_{nc}^T x))^{1-y} \right\}, \\ \psi_3(x, z, t, y; \delta, \gamma) &= \mathbb{1}(Z_i = 0) \mathbb{1}(T_i = 0) \frac{\partial}{\partial \gamma} \ln \left\{ P_{c00}(X_i; \delta) \text{expit}(\gamma_c^T X_i)^{Y_i} (1 - \text{expit}(\gamma_c^T X_i))^{1-Y_i} \right. \\ &\quad \left. + P_{n00}(X_i; \delta) \text{expit}(\gamma_{nc}^T X_i)^{Y_i} (1 - \text{expit}(\gamma_{nc}^T X_i))^{1-Y_i} \right\}. \end{aligned}$$

³Des ré-expressions similaires sont possibles dans le cas où Y est continu. Par souci de clarté, nous nous focalisons ici sur le cas où Y est binaire.

Ces ré-expressions permettent de définir une fonction d'estimation sans biais

$$\psi(x, z, t, y; \Delta_{PI}, \delta, \beta, \gamma) = \left(\psi_0^T(x; \Delta_{PI}, \delta, \beta, \gamma), \psi_1^T(x, z, t; \delta), \psi_2^T(x, z, t, y; \delta, \beta), \psi_3^T(x, z, t, y; \delta, \gamma) \right)^T$$

permettant de voir $\hat{\Delta}_{PI}$ comme un M-estimateur partiel. Par la théorie de la M-estimation [12], sous conditions de régularité et en supposant que les solutions $(\Delta_{PI}, \delta, \beta, \gamma)$ existent et sont bien définies, on a

$$n^{1/2} \begin{pmatrix} \hat{\Delta}_{PI} - \Delta \\ \hat{\delta} - \delta \\ \hat{\beta} - \beta \\ \hat{\gamma} - \gamma \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma),$$

où la matrice de covariance Σ peut être estimée par l'estimateur sandwich empirique (voir Corollaire 1, section 2.2.1 de cette thèse). Celui-ci est implémenté en R numériquement dans le package **geex** ou algorithmiquement dans le package **Mestim**.

7 Cas particuliers où monotonie et/ou exclusion restriction sont respectées

Dans cette section, nous dérivons de nouveaux estimateurs pour CACE pour les situations où les hypothèses de monotonie et/ou d'exclusion restriction sont vérifiées. Nous anticipons que les estimateurs décrits ci-dessous auront empiriquement des variances inférieures à celle de l'estimateur $\hat{\Delta}_{PI}$, dans les conditions où les hypothèses correspondantes sont correctes.

7.1 Cas particulier où monotonie est vérifiée

Dans la situation où monotonie est vérifiée, il n'y a pas de *defiers*, c'est à dire on a $D \equiv 0$. On peut alors ré-exprimer le traitement pris par un patient de la manière suivante :

$$\begin{aligned} T &\stackrel{\text{def}}{=} CT^{s=c} + AT^{s=a} + NT^{s=n} && (\text{Monotonicity}) \\ &= C \times Z + A \times 1 + N \times 0. \end{aligned}$$

En l'absence de *defiers*, une adaptation directe du Lemme 2 conduit à

$$\pi(X, Z) = \sum_{k \in \{c, a, n\}} \rho_k(X) \mu_k(Z)$$

et l'estimation de $\rho_c(\cdot)$ peut être obtenue par des adaptations similaires dans l'algorithme d'EM No. 1 (cas paramétrique) ou l'algorithme d'EM No. 4 (cas non paramétrique). On obtient ainsi $\hat{\rho}_c^{\text{no defiers}}(\cdot)$ que l'on peut facilement substituer à $\hat{\rho}_c(\cdot)$ dans l'estimateur $\hat{\Delta}_{PI}$ pour facilement obtenir l'estimateur $\hat{\Delta}_{PI}^{\text{no defiers}}$.

7.2 Cas particulier où exclusion restriction est vérifiée

Dans la situation où exclusion restriction est vérifiée, on peut ré-exprimer les résultats potentiels élémentaires de la manière suivante :

$$\begin{aligned} Y^{s=a,z=0,t=1} &= Y^{s=a,z=1,t=1} \stackrel{\text{def}}{=} Y^{s=a,t=1}, \\ Y^{s=n,z=0,t=0} &= Y^{s=n,z=1,t=0} \stackrel{\text{def}}{=} Y^{s=n,t=0}. \end{aligned} \quad (\text{Exclusion restriction})$$

Dans ces conditions, les équations pour $Y^{t=1}$ et $Y^{t=0}$ se simplifient et on obtient

$$\begin{aligned} Y^{t=1} &= CY^{s=c,z=1,t=1} + AY^{s=a,t=1} + DY^{s=d,z=0,t=1}, \\ Y^{t=0} &= CY^{s=c,z=0,t=0} + NY^{s=n,t=0} + DY^{s=d,z=1,t=0}. \end{aligned}$$

Lemme 5. *Sous l'hypothèse d'exclusion restriction on a*

$$\Delta = \mathbb{E} \left[q(X) \frac{\eta(X)\rho_c(X)}{e(X)\mathbb{E}[\rho_c(X)|T=1]} \middle| T=1 \right] - \mathbb{E} \left[q(X) \frac{\eta(X)\rho_c(X)}{e(X)\mathbb{E}[\rho_c(X)|T=0]} \middle| T=0 \right].$$

Corollaire 6. *Sous l'hypothèse d'exclusion restriction, il n'est pas nécessaire d'estimer les fonctions $q_{c11}(\cdot)$ et $q_{c00}(\cdot)$ pour estimer CACE ; l'estimation des fonctions $\rho_c(\cdot)$, $q(\cdot)$, $e(\cdot)$ et $\eta(\cdot)$ est suffisante et on peut utiliser l'estimateur*

$$\hat{\Delta}_{PI}^{ER} = \frac{\sum_{i:T_i=1} \hat{q}(X_i) \frac{\hat{\eta}(X_i)}{\hat{e}(X_i)} \hat{\rho}_c(X_i)}{\sum_{i:T_i=1} \hat{\rho}_c(X_i)} - \frac{\sum_{i:T_i=0} \hat{q}(X_i) \frac{\hat{\eta}(X_i)}{\hat{e}(X_i)} \hat{\rho}_c(X_i)}{\sum_{i:T_i=0} \hat{\rho}_c(X_i)}.$$

Notez que contrairement aux fonctions $q_{c11}(X) = \mathbb{E}(Y^{s=c,z=1,t=1}|X)$ et $q_{c00}(X) = \mathbb{E}(Y^{s=c,z=0,t=0}|X)$, les fonctions $q(X) = \mathbb{E}(Y|X)$, $e(X) = \mathbb{E}(T|X)$ et $\eta(X) = \mathbb{E}(Z|X)$ sont d'estimation simple ; en particulier, elles ne requièrent pas l'utilisation de modèles de mélange.

Preuve. Sous l'hypothèse d'exclusion restriction on a

$$\begin{aligned}
& \mathbb{E}[Y^{t=1}|C = 1] = \mathbb{E}[Y^{t=1}|C = 1, T = 1] && \text{(exclusion restriction)} \\
& = \mathbb{E}[Y|C = 1, T = 1] && \text{(consistency)} \\
& = \mathbb{E}[\mathbb{E}(Y|X)|C = 1, T = 1] \\
& = \mathbb{E}[q(X)|C = 1, T = 1] \\
& = \int_{\mathcal{X}} dx q(x) f_{X|C,T}(x|1, 1) \\
& = \int_{\mathcal{X}} dx q(x) \frac{f_{C|X,T}(1|x, 1)}{f_{C|T}(1|1)} f_{X|T}(x|1) && \text{(bayes)} \\
& = \mathbb{E} \left[q(X) \frac{\mathbb{P}(C = 1|X, T = 1)}{\mathbb{P}(C = 1|T = 1)} \middle| T = 1 \right] \\
& = \mathbb{E} \left[q(X) \frac{\mathbb{P}(T = 1|X, C = 1) \mathbb{P}(C = 1|X)}{\mathbb{P}(T = 1|X) \mathbb{P}(C = 1|T = 1)} \middle| T = 1 \right] && \text{(bayes)} \\
& = \mathbb{E} \left[q(X) \frac{\mathbb{P}(Z = 1|X) \mathbb{P}(C = 1|X)}{\mathbb{P}(T = 1|X) \mathbb{E}(C|T = 1)} \middle| T = 1 \right] && \text{(probabilités totales)} \\
& = \mathbb{E} \left[q(X) \frac{\mathbb{E}(Z|X) \mathbb{E}(C|X)}{\mathbb{E}(T|X) \mathbb{E}[\mathbb{E}(C|X)|T = 1]} \middle| T = 1 \right] \\
& = \mathbb{E} \left[q(X) \frac{\eta(X) \rho_c(X)}{e(X) \mathbb{E}[\rho_c(X)|T = 1]} \middle| T = 1 \right]
\end{aligned}$$

La première égalité vient du fait que sous l'hypothèse d'exclusion restriction on a

$$\{Y^{t=0}, Y^{t=1}\} \perp\!\!\!\perp T|C$$

; ceci est vérifiable formellement par argument de d-separation sur le graphe acyclique correspondant à notre modèle probabiliste où, en présence d'exclusion restriction, il n'y plus de flèche allant de Z vers $\{Y^{t=0}, Y^{t=1}\}$. La neuvième égalité vient du fait que

$$\begin{aligned}
\mathbb{P}(T = 1|X, C = 1) &= \mathbb{P}(Z = 1|X, C = 1) \overbrace{\mathbb{P}(T = 1|X, C = 1, Z = 1)}^{=1} \\
&\quad + \mathbb{P}(Z = 0|X, C = 1) \overbrace{\mathbb{P}(T = 1|X, C = 1, Z = 0)}^{=0}.
\end{aligned}$$

Le même argument s'applique pour $\mathbb{E}[Y^{t=0}|C = 1]$ et la substitution du résultat ci-dessus dans la définition de Δ prouve le Lemme 5. \square

8 Discussion

La suite de ce travail consistera en la réalisation de simulations de Monte Carlo pour comparer nos estimateurs de CACE aux estimateurs usuels dans différentes situations où les hypothèses de monotonie, exclusion restriction, ignorabilité principale sont respectées ou en partie violées. Nous prévoyons également d'une application de nos estimateurs sur les données de l'essai randomisé

RESICARD [13]. Enfin, l'implémentation de nos estimateurs dans le package R CACEmix et la librairie Python CACEmixPy est en cours.

Bibliographie

1. International Council For Harmonisation of Technical Requirements For Pharmaceuticals For Human Use (ICH). *E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials* 2019.
2. Little, R. J. & Rubin, D. B. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annual Review of Public Health* **21**, 121–145 (2000).
3. Frangakis, C. E. & Rubin, D. B. Principal stratification in causal inference. *Biometrics* **58**, 21–29 (2002).
4. Imbens, G. W. & Angrist, J. D. Identification and estimation of local average treatment effects. *Econometrica: Journal of the Econometric Society*, 467–475 (1994).
5. Heckman, J. J. & Vytlacil, E. Policy-relevant treatment effects. *American Economic Review* **91**, 107–111 (2001).
6. Angrist, J. D., Imbens, G. W. & Rubin, D. B. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* **91**, 444–455 (1996).
7. Stuart, E. A. & Jo, B. Assessing the sensitivity of methods for estimating principal causal effects. *Statistical Methods in Medical Research* **24**, 657–674 (2015).
8. Mansournia, M. A., Higgins, J. P., Sterne, J. A. & Hernán, M. A. Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology* **28**, 54 (2017).
9. Jo, B. & Stuart, E. A. On the use of propensity scores in principal causal effect estimation. *Statistics in Medicine* **28**, 2857–2875 (2009).
10. Jordan, M. I. & Jacobs, R. A. Hierarchical mixtures of experts and the EM algorithm. *Neural Computation* **6**, 181–214 (1994).
11. Imbens, G. W. & Rubin, D. B. Bayesian inference for causal effects in randomized experiments with noncompliance. *The Annals of Statistics*, 305–327 (1997).
12. Stefanski, L. A. & Boos, D. D. The calculus of M-estimation. *The American Statistician* **56**, 29–38 (2002).
13. Assyag, P. *et al.* RESICARD: East Paris network for the management of heart failure: absence of effect on mortality and rehospitalization in patients with severe heart failure admitted following severe decompensation. *Archives of Cardiovascular Diseases* **102**, 29–41 (2009).

Identifiability

We assume that the space of covariable \mathcal{X} is an open subset of \mathbb{R}^p and set $\Theta = \mathbb{R}^p \setminus \{0\} \times \mathbb{R}^p \times \mathbb{R}^p$; we denote $\mathcal{Z} = \{0, 1\}$ the space of allocated treatment. For $\beta = (\beta_c, \beta_a, \beta_n) \in \Theta$, we consider functions $\pi(\cdot; \beta): \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}$ of the form

$$\pi(x, z; \beta) = z\rho_c(x; \beta) + \rho_a(x; \beta) + 0 \times \rho_n(x; \beta) + (1 - z)\rho_d(x; \beta)$$

where

$$\begin{aligned}\rho_k(x; \beta) &= \frac{\exp(\beta_k^T x)}{1 + \sum_{l \in \{c, a, n\}} \exp(\beta_l^T x)}, \quad \text{for } k \in \{c, a, n\}, \quad \text{and} \\ \rho_d(x; \beta) &= \frac{1}{1 + \sum_{l \in \{c, a, n\}} \exp(\beta_l^T x)}.\end{aligned}$$

Theorem 7. *The statistical model $\{\pi(\cdot; \beta)\}_{\beta \in \Theta}$ is identifiable.*

Remark 8. In the above model, we assumed that $\beta_c \neq 0$. In view of our application, this assumption that $\beta_c \neq 0$ is very mild. Indeed, letting $\beta_c = 0$ implies that $\rho_c = \rho_d$. In plain words, this means that, for every patient, the probability of being a complier and the probability of being a defier are the same. Thus, our assumption holds as soon as a single patient has a probability of being a complier that is different from their probability of being a defier.

Proof of Theorem 7. Consider $\pi(\cdot; \beta)$ and $\pi(\cdot; \gamma)$ with β and $\gamma \in \Theta$. Assume that $\pi(\cdot; \beta) = \pi(\cdot; \gamma)$ on $\mathcal{X} \times \mathcal{Z}$. Specializing in $z = 0$ and $z = 1$, we get the following equations on \mathcal{X} :

$$\begin{cases} \rho_c(x; \beta) + \rho_a(x; \beta) = \rho_c(x; \gamma) + \rho_a(x; \gamma) \\ \rho_a(x; \beta) + \rho_d(x; \beta) = \rho_a(x; \gamma) + \rho_d(x; \gamma). \end{cases}$$

This system is equivalent to

$$\begin{cases} \left(1 + \sum_{l \in \{c, a, n\}} \exp(\gamma_l^T x)\right) \left(1 + \exp(\beta_a^T x)\right) = \left(1 + \sum_{l \in \{c, a, n\}} \exp(\beta_l^T x)\right) \left(1 + \exp(\gamma_a^T x)\right) \\ \left(1 + \sum_{l \in \{c, a, n\}} \exp(\gamma_l^T x)\right) \left(\exp(\beta_c^T x) + \exp(\beta_a^T x)\right) = \left(1 + \sum_{l \in \{c, a, n\}} \exp(\beta_l^T x)\right) \left(\exp(\gamma_a^T x) + \exp(\gamma_c^T x)\right) \end{cases} \quad (2)$$

Using the expansion of \exp in power series, the first order term of the expansion of the System (2) provides the following identities

$$\begin{cases} \beta_c + \beta_a + \gamma_n = \gamma_c + \gamma_a + \beta_n \\ \gamma_c + \gamma_n + \beta_a = \beta_c + \beta_n + \gamma_a \end{cases}$$

which leads to

$$\begin{cases} \beta_c = \gamma_c \\ \gamma_n + \beta_a = \beta_n + \gamma_a. \end{cases} \quad (3)$$

Using again the expansion of \exp in power series, the second order term of the expansion of the

System (2) provides the following identities

$$\begin{cases} [\gamma_c^T x + \beta_a^T x]^2 + [\gamma_n^T x + \beta_a^T x]^2 + [\gamma_c^T x]^2 + [\gamma_n^T x]^2 = [\beta_c^T x + \gamma_a^T x]^2 + [\beta_n^T x + \gamma_a^T x]^2 + [\beta_c^T x]^2 + [\beta_n^T x]^2 \\ [\gamma_c^T x + \beta_a^T x]^2 + [\beta_n^T x + \gamma_c^T x]^2 + [\gamma_c^T x]^2 + [\gamma_a^T x]^2 = [\beta_c^T x + \gamma_a^T x]^2 + [\beta_c^T x + \gamma_n^T x]^2 + [\beta_c^T x]^2 + [\beta_a^T x]^2. \end{cases}$$

Using the relations provided by the System (3), the above equations simplify to

$$\begin{cases} [\gamma_c^T x + \beta_a^T x]^2 + [\gamma_n^T x]^2 = [\beta_c^T x + \gamma_a^T x]^2 + [\beta_n^T x]^2 \\ [\beta_n^T x + \gamma_c^T x]^2 + [\gamma_a^T x]^2 = [\beta_c^T x + \gamma_n^T x]^2 + [\beta_a^T x]^2. \end{cases}$$

Expanding and simplifying the above expression, we get

$$\begin{cases} [\beta_a^T x]^2 + 2[\gamma_c^T x][\beta_a^T x] + [\gamma_n^T x]^2 = [\beta_n^T x]^2 + 2[\beta_c^T x][\gamma_a^T x] + [\gamma_a^T x]^2 \\ [\beta_a^T x]^2 + 2[\beta_c^T x][\gamma_n^T x] + [\gamma_n^T x]^2 = [\gamma_a^T x]^2 + 2[\beta_n^T x][\gamma_c^T x] + [\beta_n^T x]^2. \end{cases}$$

Substracting the second equation to the first one in the above system provides the relation

$$2[\beta_c^T x]([\beta_a^T x] - [\gamma_n^T x]) = 2[\beta_c^T x]([\gamma_a^T x] - [\beta_n^T x]).$$

That is

$$[\beta_c^T x]([\beta_a - \gamma_n - \gamma_a + \beta_n]^T x) = 0.$$

Consider the function $\phi: x \mapsto \beta_c^T x$ and $\psi: x \mapsto [\beta_a - \gamma_n - \gamma_a + \beta_n]^T x$. Assuming $\beta_c \neq 0$, we get that $\psi = 0$ on the open set $\mathcal{X} \setminus \ker \phi$. Since ψ is linear, it is zero on all of \mathbb{R}^p . It follows that

$$\beta_a - \gamma_n = \gamma_a - \beta_n.$$

This relation together with the System of equations (3) implies that

$$\beta_c = \gamma_c, \quad \beta_a = \gamma_a, \quad \beta_n = \gamma_n.$$

□

EM Algorithms

Algorithm 1 The EM procedure for estimating $\rho_s(\cdot; \delta_s)$ where $s \in \{c, a, n, d\}$.

Input: Data $(X_i^\rho, Z_i, T_i)_{1 \leq i \leq n}$ where X_i^ρ is a relevant subset of the variables contained in X_i .
Initialize the prior probabilities associated with the nodes of the tree as

$$g_{c,i} \leftarrow 1/4, \quad g_{a,i} \leftarrow 1/4, \quad g_{n,i} \leftarrow 1/4, \quad \text{and} \quad g_{d,i} \leftarrow 1/4.$$

Compute individual contributions to each expert's likelihood as

$$\begin{aligned} L_{c,i} &\leftarrow Z_i T_i + (1 - Z_i)(1 - T_i) \\ L_{a,i} &\leftarrow T_i \\ L_{n,i} &\leftarrow 1 - T_i \\ L_{d,i} &\leftarrow Z_i(1 - T_i) + (1 - Z_i)T_i \end{aligned}$$

Iterate until convergence on the parameters $\delta = (\delta_c^T, \delta_a^T, \delta_n^T, \delta_d^T)^T$:

Compute the posterior probabilities associated with the nodes of the tree as

▷ E-step

$$\begin{aligned} h_{c,i} &\leftarrow g_{c,i} L_{c,i} / \sum_{s \in \{c, a, n, d\}} g_{s,i} L_{s,i} \\ h_{a,i} &\leftarrow g_{a,i} L_{a,i} / \sum_{s \in \{c, a, n, d\}} g_{s,i} L_{s,i} \\ h_{n,i} &\leftarrow g_{n,i} L_{n,i} / \sum_{s \in \{c, a, n, d\}} g_{s,i} L_{s,i} \\ h_{d,i} &\leftarrow g_{d,i} L_{d,i} / \sum_{s \in \{c, a, n, d\}} g_{s,i} L_{s,i} \end{aligned}$$

For the gating network $\rho(\cdot)$ estimate parameters δ by solving the IRLS problem

▷ M-step

$$\delta \leftarrow \arg \max_{\delta} \sum_{i=1}^n \sum_{s \in \{c, a, n, d\}} h_{s,i} \ln \left(\exp \delta_s^T X_i^\rho / \sum_{k \in \{c, a, n, d\}} \exp \delta_k^T X_i^\rho \right)$$

as a multinomial logistic regression with features $(X_i^\rho)_{1 \leq i \leq n}$, and labels $(h_{1,i}, h_{2,i}, h_{3,i}, h_{4,i})_{1 \leq i \leq n}$.
Update the prior probabilities associated with the nodes of the tree as

$$\begin{aligned} g_{c,i} &\leftarrow \exp \delta_c^T X_i^\rho / \sum_{k \in \{c, a, n, d\}} \exp \delta_k^T X_i^\rho \\ g_{a,i} &\leftarrow \exp \delta_a^T X_i^\rho / \sum_{k \in \{c, a, n, d\}} \exp \delta_k^T X_i^\rho \\ g_{n,i} &\leftarrow \exp \delta_n^T X_i^\rho / \sum_{k \in \{c, a, n, d\}} \exp \delta_k^T X_i^\rho \\ g_{d,i} &\leftarrow \exp \delta_d^T X_i^\rho / \sum_{k \in \{c, a, n, d\}} \exp \delta_k^T X_i^\rho \end{aligned}$$

Return: $\rho_s(x; \hat{\delta}) = \exp \delta_s^T x / \sum_{k \in \{c, a, n, d\}} \exp \delta_k^T x$.

Algorithm 2 The EM procedure for estimating $q_c(\cdot; \zeta)$ where q_c is either q_{c11} or q_{c00} depending on the data subset considered (similarly, $q_{nc}(\cdot; \zeta)$ is either q_{a11} or q_{n00} depending on the data subset considered).

Input: Data $(X_i^\zeta, Y_i)_{i:Z_i=1, T_i=1}$ or $(X_i^{\zeta^T}, Y_i)_{i:Z_i=0, T_i=0}$ where X_i^ζ is a relevant subset of the variables contained in X_i , and P_c .*

Initialize the prior probabilities associated with the nodes of the tree as

$$g_{c,i} \leftarrow P_c, \quad g_{nc,i} \leftarrow 1 - P_c.$$

Initialize the parameters (ζ_c, ζ_{nc}) of the experts $(q_c(\cdot), q_{nc}(\cdot))$ at random e.g.,

$$\zeta_c \sim \mathcal{N}(0, D) \quad \zeta_{nc} \sim \mathcal{N}(0, D)$$

with D a diagonal matrix.

Compute individual predictions from the initiated expert networks $(q_c(\cdot), q_{nc}(\cdot))$ as

$$q_{c,i} \leftarrow \text{expit}(\zeta_c^T X_i^\zeta) \quad q_{nc,i} \leftarrow \text{expit}(\zeta_{nc}^T X_i^\zeta)$$

Iterate until convergence on the parameters $\zeta = (\zeta_c^T, \zeta_{nc}^T)^T$:

Compute individual contributions to each expert's likelihood as

$$\begin{aligned} L_{c,i} &\leftarrow q_{c,i}^{Y_i} (1 - q_{c,i})^{1-Y_i} \\ L_{nc,i} &\leftarrow q_{nc,i}^{Y_i} (1 - q_{nc,i})^{1-Y_i} \end{aligned}$$

Compute the posterior probabilities associated with the nodes of the tree as

▷ E-step

$$\begin{aligned} h_{c,i} &\leftarrow g_{c,i} L_{c,i} / (g_{c,i} L_{c,i} + g_{nc,i} L_{nc,i}) \\ h_{nc,i} &\leftarrow 1 - h_{c,i} \end{aligned}$$

For the expert network $q_c(\cdot)$ estimate parameters ζ_c by solving the IRLS problem

▷ M-step

$$\zeta_c \leftarrow \arg \max_{\zeta_c} \sum_{i=1}^{n'} h_{c,i} \left[Y_i \ln \{ \text{expit}(\zeta_c^T X_i^\zeta) \} + (1 - Y_i) \ln \{ 1 - \text{expit}(\zeta_c^T X_i^\zeta) \} \right]$$

as a weighted logistic regression with features X_i^ζ , labels Y_i and weights $h_{c,i}$.

For the expert network $q_{nc}(\cdot)$ estimate parameters ζ_{nc} by solving the IRLS problem

$$\zeta_{nc} \leftarrow \arg \max_{\zeta_{nc}} \sum_{i=1}^{n'} h_{nc,i} \left[Y_i \ln \{ \text{expit}(\zeta_{nc}^T X_i^\zeta) \} + (1 - Y_i) \ln \{ 1 - \text{expit}(\zeta_{nc}^T X_i^\zeta) \} \right]$$

as a weighted logistic regression with features X_i^ζ , labels Y_i and weights $h_{nc,i}$.

Update the predictions from the expert networks $(q_c(\cdot), q_{nc}(\cdot))$ as

$$q_{c,i} \leftarrow \text{expit}(\zeta_c^T X_i^\zeta) \quad q_{nc,i} \leftarrow \text{expit}(\zeta_{nc}^T X_i^\zeta)$$

Return: $q_c(x; \hat{\zeta}) = \text{expit}(\zeta_c^T x)$

* P_c represents $(P_{c11}(X_i))_{i:Z_i=1, T_i=1}$ or $(P_{c00}(X_i))_{i:Z_i=0, T_i=0}$ depending on the considered data subset. Inside the algorithm, the whole procedure is run on either the $i:Z_i=1, T_i=1$ or $i:Z_i=0, T_i=0$ data subset. We drop these subscripts to avoid clutter. As a reminder of this, we note sums over n' rather than n .

Algorithm 3 The EM procedure for estimating $q_c(\cdot; \zeta)$ when Y is continuous.

Input: Data $(X_i^\zeta, Y_i)_{i:Z_i=1, T_i=1}$ or $(X_i^{\zeta^T}, Y_i)_{i:Z_i=0, T_i=0}$ where d -dimensional X_i^ζ is a relevant subset of the variables contained in X_i , and P_c .^{*}

Initialize the prior probabilities associated with the nodes of the tree as

$$g_{c,i} \leftarrow P_c, \quad g_{nc,i} \leftarrow 1 - P_c.$$

Initialize the parameters $((\zeta_c, \sigma_c^2), (\zeta_{nc}, \sigma_{nc}^2))$ of the experts $(q_c(\cdot), q_{nc}(\cdot))$ at random e.g.,

$$\begin{aligned} \zeta_c &\sim \mathcal{N}(0, D) \quad \sigma_c^2 \leftarrow 1 \\ \zeta_{nc} &\sim \mathcal{N}(0, D) \quad \sigma_{nc}^2 \leftarrow 1 \end{aligned}$$

with D a diagonal matrix.

Compute individual predictions from the initiated expert networks $(q_c(\cdot), q_{nc}(\cdot))$ as

$$q_{c,i} \leftarrow \zeta_c^T X_i^\zeta \quad q_{nc,i} \leftarrow \zeta_{nc}^T X_i^\zeta$$

Iterate until convergence on the parameters $\zeta = (\zeta_c^T, \zeta_{nc}^T)^T$:

Compute individual contributions to each expert's likelihood as

$$\begin{aligned} L_{c,i} &\leftarrow \mathcal{N}_{\mathcal{L}}(Y_i | \mu = \zeta_c^T X_i^\zeta, \sigma^2 = \sigma_c^2) \\ L_{nc,i} &\leftarrow \mathcal{N}_{\mathcal{L}}(Y_i | \mu = \zeta_{nc}^T X_i^\zeta, \sigma^2 = \sigma_{nc}^2) \end{aligned}$$

Compute the posterior probabilities associated with the nodes of the tree as

▷ E-step

$$\begin{aligned} h_{c,i} &\leftarrow g_{c,i} L_{c,i} / (g_{c,i} L_{c,i} + g_{nc,i} L_{nc,i}) \\ h_{nc,i} &\leftarrow 1 - h_{c,i} \end{aligned}$$

For the expert network $q_c(\cdot)$ estimate parameters ζ_c by solving the weighted least-squares problem

▷ M-step

$$\zeta_c \leftarrow \arg \min_{\zeta_c} \sum_{i=1}^{n'} h_{c,i} (Y_i - \zeta_c^T X_i^\zeta)^2$$

as a weighted linear regression with features X_i^ζ , labels Y_i and weights $h_{c,i}$.

For the expert network $q_{nc}(\cdot)$ estimate parameters ζ_{nc} by solving the weighted least-squares problem

$$\zeta_{nc} \leftarrow \arg \min_{\zeta_{nc}} \sum_{i=1}^{n'} h_{nc,i} (Y_i - \zeta_{nc}^T X_i^\zeta)^2$$

as a weighted linear regression with features X_i^ζ , labels Y_i and weights $h_{nc,i}$.

Update the variance parameter for each expert $(q_c(\cdot), q_{nc}(\cdot))$ as

$$\sigma_c^2 \leftarrow \frac{1}{n' - d} \sum_{i=1}^{n'} h_{c,i} (Y_i - \zeta_c^T X_i^\zeta)^2, \quad \sigma_{nc}^2 \leftarrow \frac{1}{n' - d} \sum_{i=1}^{n'} h_{nc,i} (Y_i - \zeta_{nc}^T X_i^\zeta)^2$$

Update the predictions from the expert networks $(q_c(\cdot), q_{nc}(\cdot))$ as

$$q_{c,i} \leftarrow \zeta_c^T X_i^\zeta \quad q_{nc,i} \leftarrow \zeta_{nc}^T X_i^\zeta$$

Return: $q_c(x; \hat{\zeta}) = \zeta_c^T x$

* P_c represents $(P_{c11}(X_i))_{i:Z_i=1, T_i=1}$ or $(P_{c00}(X_i))_{i:Z_i=0, T_i=0}$ depending on the considered data subset. Inside the algorithm, the whole procedure is run on either the $i:Z_i=1, T_i=1$ or $i:Z_i=0, T_i=0$ data subset. We drop these subscripts to avoid clutter. As a reminder of this, we note sums over n' rather than n .

Algorithm 4 The nonparametric EM-like procedure for estimating $\rho_s(\cdot)$ where $s \in \{c, a, n, d\}$.

Input: Data $(X_i^\rho, Z_i, T_i)_{1 \leq i \leq n}$ where X_i^ρ is a relevant subset of the variables contained in X_i .

Initialize the prior probabilities associated with the nodes of the tree as

$$g_{c,i} \leftarrow 1/4, \quad g_{a,i} \leftarrow 1/4, \quad g_{n,i} \leftarrow 1/4, \quad \text{and} \quad g_{d,i} \leftarrow 1/4.$$

Compute individual contributions to each expert's likelihood as

$$\begin{aligned} L_{c,i} &\leftarrow Z_i T_i + (1 - Z_i)(1 - T_i) \\ L_{a,i} &\leftarrow T_i \\ L_{n,i} &\leftarrow 1 - T_i \\ L_{d,i} &\leftarrow Z_i(1 - T_i) + (1 - Z_i)T_i \end{aligned}$$

Iterate until convergence:

Compute the posterior probabilities associated with the nodes of the tree as

▷ E-step

$$\begin{aligned} h_{c,i} &\leftarrow g_{c,i} L_{c,i} / \sum_{s \in \{c, a, n, d\}} g_{s,i} L_{s,i} \\ h_{a,i} &\leftarrow g_{a,i} L_{a,i} / \sum_{s \in \{c, a, n, d\}} g_{s,i} L_{s,i} \\ h_{n,i} &\leftarrow g_{n,i} L_{n,i} / \sum_{s \in \{c, a, n, d\}} g_{s,i} L_{s,i} \\ h_{d,i} &\leftarrow g_{d,i} L_{d,i} / \sum_{s \in \{c, a, n, d\}} g_{s,i} L_{s,i} \end{aligned}$$

For the gating network fit $\hat{\rho}_s(\cdot)$, $s \in \{c, a, n, d\}$ as a multiclass classification

▷ M-step

problem with features $(X_i^{\hat{\rho}})_{1 \leq i \leq n}$, and labels $(h_{1,i}, h_{2,i}, h_{3,i}, h_{4,i})_{1 \leq i \leq n}$.

Update the prior probabilities associated with the nodes of the tree as

$$g_{c,i} \leftarrow \hat{\rho}_c(X_i \rho), \quad g_{a,i} \leftarrow \hat{\rho}_a(X_i \rho), \quad g_{n,i} \leftarrow \hat{\rho}_n(X_i \rho), \quad g_{d,i} \leftarrow \hat{\rho}_d(X_i \rho)$$

Return: $\hat{\rho}_c(\cdot)$

Algorithm 5 The non parametric EM-like procedure for estimating $q_c(\cdot)$ where q_c is either q_{c11} or q_{c00} depending on the data subset considered (similarly, $q_{nc}(\cdot)$ is either q_{a11} or q_{n00}).

Input: Data $(X_i^\zeta, Y_i)_{i:Z_i=1, T_i=1}$ or $(X_i^{\zeta T}, Y_i)_{i:Z_i=0, T_i=0}$ where X_i^ζ is a relevant subset of the variables contained in X_i , and P_c .*

Initialize the prior probabilities associated with the nodes of the tree as

$$g_{c,i} \leftarrow P_c, \quad g_{nc,i} \leftarrow 1 - P_c.$$

Initialize the individual predictions from the expert networks $(q_c(\cdot), q_{nc}(\cdot))$ as

$$q_{c,i} \sim \mathcal{U}_{[0,1]} \quad q_{nc,i} \sim \mathcal{U}_{[0,1]}$$

Iterate until convergence:

Compute individual contributions to each expert's likelihood as

$$\begin{aligned} L_{c,i} &\leftarrow q_{c,i}^{Y_i} (1 - q_{c,i})^{1 - Y_i} \\ L_{nc,i} &\leftarrow q_{nc,i}^{Y_i} (1 - q_{nc,i})^{1 - Y_i} \end{aligned}$$

Compute the posterior probabilities associated with the nodes of the tree as

▷ E-step

$$\begin{aligned} h_{c,i} &\leftarrow g_{c,i} L_{c,i} / (g_{c,i} L_{c,i} + g_{nc,i} L_{nc,i}) \\ h_{nc,i} &\leftarrow 1 - h_{c,i} \end{aligned}$$

For the expert network $q_c(\cdot)$ fit $\hat{q}_c(\cdot)$ via a weighted nonparametric classifier with features X_i^ζ , labels Y_i and weights $h_{c,i}$.
▷ M-step

For the expert network $q_{nc}(\cdot)$ fit $\hat{q}_{nc}(\cdot)$ via a weighted nonparametric classifier with features X_i^ζ , labels Y_i and weights $h_{nc,i}$.

Update the predictions from the expert networks $(q_c(\cdot), q_{nc}(\cdot))$ as

$$q_{c,i} \leftarrow \hat{q}_c(X_i^\zeta) \quad q_{nc,i} \leftarrow \hat{q}_{nc}(X_i^\zeta)$$

Return: $\hat{q}_c(\cdot)$

* P_c represents $(P_{c11}(X_i))_{i:Z_i=1, T_i=1}$ or $(P_{c00}(X_i))_{i:Z_i=0, T_i=0}$ depending on the considered data subset. Inside the algorithm, the whole procedure is run on either the $i:Z_i=1, T_i=1$ or $i:Z_i=0, T_i=0$ data subset.

Algorithm 6 The nonparametric EM-like procedure for estimating $q_c(\cdot)$ when Y is continuous.

Input: Data $(X_i^\zeta, Y_i)_{i:Z_i=1, T_i=1}$ or $(X_i^{\zeta T}, Y_i)_{i:Z_i=0, T_i=0}$ where X_i^ζ is a relevant subset of the variables contained in X_i , and P_c .*

Initialize the prior probabilities associated with the nodes of the tree as

$$g_{c,i} \leftarrow P_c, \quad g_{nc,i} \leftarrow 1 - P_c.$$

Initialize the variance parameters $(\sigma_c^2, \sigma_{nc}^2)$

$$\sigma_c^2 \leftarrow 1 \quad \sigma_{nc}^2 \leftarrow 1$$

Initialize the individual predictions from the expert networks $(q_c(\cdot), q_{nc}(\cdot))$ as

$$q_{c,i} \sim \mathcal{N}(0, 1) \quad q_{nc,i} \sim \mathcal{N}(0, 1)$$

Iterate until convergence:

Compute individual contributions to each expert's likelihood as

$$\begin{aligned} L_{c,i} &\leftarrow \mathcal{N}_{\mathcal{L}}(Y_i | \mu = q_{c,i}, \sigma^2 = \sigma_c^2) \\ L_{nc,i} &\leftarrow \mathcal{N}_{\mathcal{L}}(Y_i | \mu = q_{nc,i}, \sigma^2 = \sigma_{nc}^2) \end{aligned}$$

Compute the posterior probabilities associated with the nodes of the tree as

▷ E-step

$$\begin{aligned} h_{c,i} &\leftarrow g_{c,i} L_{c,i} / (g_{c,i} L_{c,i} + g_{nc,i} L_{nc,i}) \\ h_{nc,i} &\leftarrow 1 - h_{c,i} \end{aligned}$$

For the expert network $q_c(\cdot)$ fit $\hat{q}_c(\cdot)$ via weighted nonparametric regression

▷ M-step

with features X_i^ζ , labels Y_i and weights $h_{c,i}$.

For the expert network $q_{nc}(\cdot)$ fit $\hat{q}_{nc}(\cdot)$ via weighted nonparametric regression

with features X_i^ζ , labels Y_i and weights $h_{nc,i}$.

Update the variance parameter for each expert $(q_c(\cdot), q_{nc}(\cdot))$ as

$$\sigma_c^2 \leftarrow \frac{1}{n'} \sum_{i=1}^{n'} h_{c,i} (Y_i - \hat{q}_c(X_i^\zeta))^2, \quad \sigma_{nc}^2 \leftarrow \frac{1}{n'} \sum_{i=1}^{n'} h_{nc,i} (Y_i - \hat{q}_{nc}(X_i^\zeta))^2$$

Update the predictions from the expert networks $(q_c(\cdot), q_{nc}(\cdot))$ as

$$q_{c,i} \leftarrow \hat{q}_c(X_i^\zeta) \quad q_{nc,i} \leftarrow \hat{q}_{nc}(X_i^\zeta)$$

Return: $\hat{q}_c(\cdot)$

* P_c represents $(P_{c11}(X_i))_{i:Z_i=1, T_i=1}$ or $(P_{c00}(X_i))_{i:Z_i=0, T_i=0}$ depending on the considered data subset. Inside the algorithm, the whole procedure is run on either the $i:Z_i=1, T_i=1$ or $i:Z_i=0, T_i=0$ data subset. We drop these subscripts to avoid clutter. As a reminder of this, we note sums over n' rather than n .

Annexe B : Calculs et définitions

1 Annexe à la section “Inférence causale”

Dans le cas où Y est binaire l’expression de l’estimateur TMLE est :

$$\hat{\Delta}_{TMLE} = n^{-1} \sum_{i=1}^n q^*(1, X_i; \hat{\beta}, \hat{\gamma}, \hat{\epsilon}) - q^*(0, X_i; \hat{\beta}, \hat{\gamma}, \hat{\epsilon})$$

où

$$\begin{aligned} q^*(1, X_i; \hat{\beta}, \hat{\gamma}, \hat{\epsilon}) &= \text{expit} \left[\text{logit}\{q(1, X_i; \hat{\gamma})\} + \hat{\epsilon} H(1, X_i; \hat{\beta}) \right] \\ q^*(0, X_i; \hat{\beta}, \hat{\gamma}, \hat{\epsilon}) &= \text{expit} \left[\text{logit}\{q(0, X_i; \hat{\gamma})\} + \hat{\epsilon} H(0, X_i; \hat{\beta}) \right] \end{aligned}$$

et la “variable astucieuse” (*clever covariate*) est toujours dénotée

$$H(A_i, X_i; \hat{\beta}) = \frac{A_i}{\pi(X_i; \hat{\beta})} - \frac{1 - A_i}{1 - \pi(X_i; \hat{\beta})}.$$

Le paramètre ϵ est estimé lors de l’étape ciblage via

$$\hat{\epsilon} = \arg \max_{\epsilon} \prod_{i=1}^n \text{expit} \left[\text{logit}\{q(A_i, X_i; \hat{\gamma})\} + \epsilon H(A, X_i; \hat{\beta}) \right]^{Y_i} \left\{ 1 - \text{expit} \left[\text{logit}\{q(A_i, X_i; \hat{\gamma})\} + \epsilon H(A, X_i; \hat{\beta}) \right] \right\}^{1-Y_i}.$$

2 Annexe à la section “Théorie de la M-estimation”

Théorème 2 (Variance asymptotique sandwich – notation fonctionnelle). *Si $T(\cdot)$ est un M-estimateur de type ψ avec $\psi(z, \delta)$ dérivable selon δ , z -(presque partout) alors $T(\cdot)$ est \sqrt{n} -convergent et*

$$\sqrt{n}(T(F_n) - T(F)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma)$$

où la matrice de variance-covariance Σ a pour expression

$$\Sigma = \left[\mathbb{E} \left\{ \frac{\partial \psi(Z, \delta)}{\partial \delta^T} \Big|_{\delta=T(F)} \right\} \right]^{-1} \mathbb{E} \left\{ \psi(Z, T(F)) \psi^T(Z, T(F)) \right\} \left[\left[\mathbb{E} \left\{ \frac{\partial \psi(Z, \delta)}{\partial \delta^T} \Big|_{\delta=T(F)} \right\} \right]^{-1} \right]^T.$$

Corollaire 2 (Variance asymptotique sandwich empirique – notation fonctionnelle). *Si $T(\cdot)$ est un M-estimateur de type ψ avec $\psi(z, \delta)$ dérivable selon δ , z -(presque partout) alors la distribution asymptotique de $T(F_n)$ peut être estimée par*

$$T(F_n) \stackrel{\text{d}}{\sim} \mathcal{N}(0, n^{-1} \Sigma(F_n))$$

où $\Sigma(\cdot)$ est l'estimateur asymptotique sandwich empirique :

$$\Sigma(F_n) = A^{-1} B (A^{-1})^T$$

avec

$$A = n^{-1} \sum_{i=1}^n \left\{ \frac{\partial \psi(Z_i, \delta)}{\partial \delta^T} \Big|_{\delta=T(F_n)} \right\},$$

$$B = n^{-1} \sum_{i=1}^n \left\{ \psi(Z_i, T(F_n)) \psi^T(Z_i, T(F_n)) \right\}.$$

3 Annexe à la section “Mixtures d’experts et algorithme de maximisation d’espérance”

On montre ci-dessous que la représentation énoncée pour $p_{Y,S|X}$ induit bien une densité de probabilité :

$$\begin{aligned}
& \int_{\mathcal{Y}} \sum_{s \in \mathcal{S}} p_{Y,S|X}(y, s|x; \theta) dy \\
&= \int_{\mathcal{Y}} \sum_{s=0}^1 \left[p_{Y|X,S}(y|x, 1; \alpha_1) \text{expit}(\beta^T x) \right]^s \left[p_{Y|X,S}(y|x, 1; \alpha_0) \{1 - \text{expit}(\beta^T x)\} \right]^{1-s} dy \\
&= \int_{\mathcal{Y}} p_{Y|X,S}(y|x, 1; \alpha_1) \text{expit}(\beta^T x) dy + \int_{\mathcal{Y}} p_{Y|X,S}(y|x, 1; \alpha_0) \{1 - \text{expit}(\beta^T x)\} dy \\
&= \text{expit}(\beta^T x) \int_{\mathcal{Y}} p_{Y|X,S}(y|x, 1; \alpha_1) dy + \{1 - \text{expit}(\beta^T x)\} \int_{\mathcal{Y}} p_{Y|X,S}(y|x, 1; \alpha_0) dy \\
&= \text{expit}(\beta^T x) + \{1 - \text{expit}(\beta^T x)\} = 1
\end{aligned}$$

et il apparaît par ailleurs trivialement que $p_{Y,S|X}(y, s|x; \theta) > 0, \forall y \in \mathcal{Y}, s \in \{0, 1\}, x \in \mathcal{X}, \theta \in \Theta$.

4 Annexe à la section “Apprentissage par renforcement de stratégies dynamiques optimales”

Hypothèse 7 (Positivité séquentielle dans le contexte d’individus entrants dans un état terminal). Pour tout $h_k \in \mathcal{H}_k$, $k = 1, \dots, K$, les relations suivantes sont vérifiées. Si $h_k = \Phi$:

$$\forall a \in \mathcal{A}_k \setminus \{0\}, e_{k,a}(h_k) = 0, \quad \text{et} \quad e_{k,0}(h_k) = 1,$$

sinon : il existe des constantes $0 < \eta$, et $\eta_0 < 1$ telles que

$$\forall a \in \mathcal{A}_k \setminus \{0\}, e_{k,a}(h_k) > \eta, \quad \text{et} \quad e_{k,0}(h_k) > 1 - \eta_0.$$

Causal inference methods for personalized medicine: an application to the timing of renal replacement therapy initiation

Résumé : La médecine personnalisée représente l'aspiration de cliniciens et de chercheurs à proposer au temps opportun, des traitements adaptés aux caractéristiques de chaque patient. Traditionnellement, l'élaboration de règles personnalisées de traitements s'appuie sur la connaissance clinique et biologique. Ces approches centrées sur la connaissance experte ont cependant des limites, en particulier en réanimation où les maladies sont complexes et hétérogènes. Dans ce travail de thèse, nous adoptons une démarche différente, centrée sur l'apprentissage statistique causal appliquée à des données cliniques, pour produire des règles de médecine personnalisée. Nous focalisons notre attention sur une application en réanimation : le recours à l'épuration extra-rénale dans l'insuffisance rénale aiguë. Nos contributions sont les suivantes : (i) Nous illustrons le principe des essais émulés pour identifier à partir de données observationnelles des sous-groupes de patients bénéficiant de l'hémodialyse intermittente plutôt que de l'hémofiltration continue. (ii) Nous utilisons des données d'essais randomisés contrôlés pour estimer l'effet traitement individuel du recours à une stratégie d'initiation précoce versus tardive de l'épuration extra-rénale. (iii) Nous développons un cadre général pour l'évaluation de règles personnalisées de traitement à partir de données observationnelles. (iv) Nous estimons et validions une stratégie dynamique optimale pour l'initiation de l'épuration extra-rénale dans les trois jours suivant l'insuffisance rénale aiguë. A la lumière de nos résultats, nous discutons la place des méthodes causales d'apprentissage statistique pour développer, à partir de données structurées courantes, des règles de médecine personnalisée pertinentes pour la pratique clinique.

Mots-clés : Inférence causale, médecine personnalisée, apprentissage statistique par renforcement.

Abstract: Personalized medicine is the clinician's and researcher's long-held ideal of providing timely treatments tailored to the characteristics of each individual patient. Historically, the development of individualized treatment rules has been based on clinical and biological knowledge. However, these expert-knowledge-based approaches have their limitations, particularly in intensive care, where diseases are complex and heterogeneous. In this thesis, we use a different approach, centered on causal statistical learning applied to clinical data, to produce individualized treatment rules. We focus on an application in intensive care: the initiation of renal replacement therapy in acute kidney injury. Our contributions are as follows: (i) We illustrate the principle of emulated trials to identify from observational data subgroups of patients benefiting from intermittent hemodialysis rather than continuous hemofiltration. (ii) We use data from randomized controlled trials to estimate the individual treatment effect of using an early versus a late strategy of renal replacement therapy initiation. (iii) We develop a comprehensive framework for the evaluation of individualized treatment rules based on observational data. (iv) We estimate and validate an optimal dynamic strategy for the initiation of renal replacement therapy within three days of acute kidney injury. In the light of our results, we discuss the role of causal inference and statistical learning methods in developing clinically relevant individualized treatment rules from structured electronic health record data.

Keywords: Causal inference, Personalized medicine, & Statistical reinforcement learning.
