

# Personalizing renal-replacement therapy initiation in the ICU: a statistical reinforcement learning approach

## Introduction

In the last decade, large randomized trials investigated the effect of an early vs a late strategy of renal-replacement therapy (RRT) in ICU patients with acute kidney injury (AKI).<sup>1,2</sup> At the population level, no trial found evidence of any effect. In the static case, heterogeneity of treatment effects was found.<sup>3</sup> Yet, the criteria mandating RRT initiation needs to be refined, accounting for the fact that decisions not to start RRT should be reevaluated every day.

We wished to learn an individualized strategy that recommends whether to start RRT in light of patients' time-evolving characteristics.

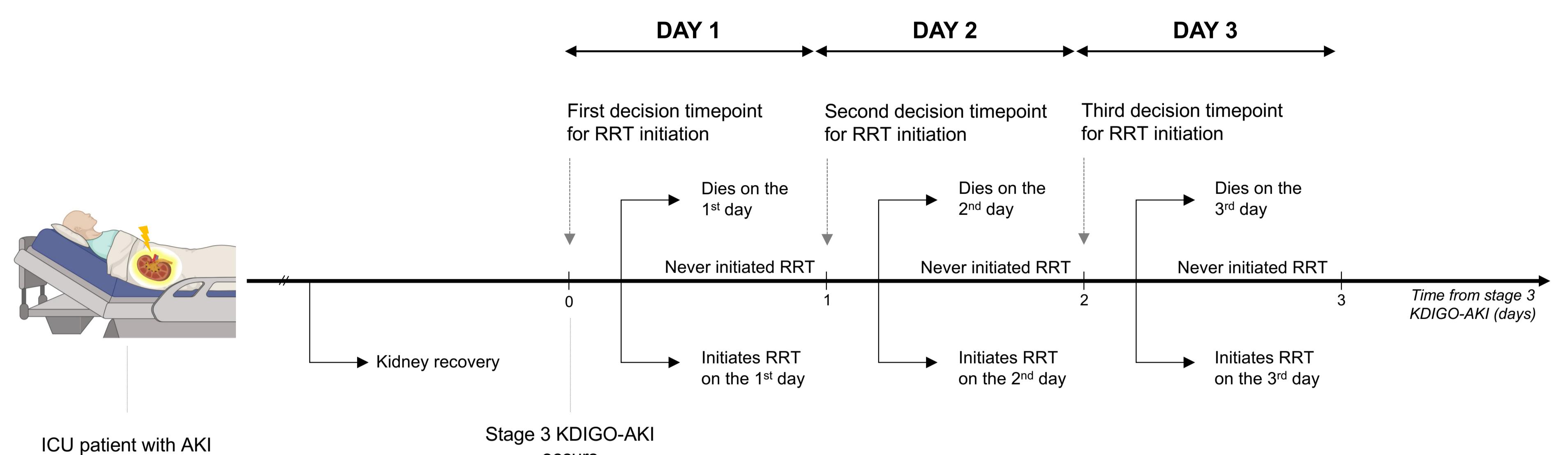


Figure 1. Trajectory of a single patient with acute kidney injury in our learning setup.

## Methods

### Sources of data

- The development sample included participants from the MIMIC-III database which contains routinely collected data from 61 051 distinct ICU admissions of adult patients admitted between 2001 and 2012.
- The validation sample included participants from the AKIKI and AKIKI2 trials,<sup>1,2</sup> two multicenter RCTs conducted in France.

### Population

Eligible patients were adults 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.

### Setup and timepoints for learning RRT initiation strategies

We wished to develop a when-to-treat strategy for RRT initiation in the first 72 hours following a stage 3 KDIGO-AKI. Specifically, we wished to learn a strategy that—for three days in a row after the occurrence of stage 3 KDIGO-AKI—assesses the need to start RRT in light of the history of patients' evolving characteristics. We thus considered three decision timepoints at 0, 24, and 48 hours after the occurrence of stage 3-KDIGO-AKI (Figure 1).

### Outcomes

We used hospital-free days at day 60 (HFD60) as the primary distal outcome. This outcome is unlikely to give rise to the value-misalignment problem in our setting.

### Learning an optimal strategy

To learn an optimal strategy, we used a doubly robust dynamic treatment regimen via weighted least squares.<sup>4</sup> 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. These are defined as  $f_t(h_t) = \mathbb{E}[Y^{\bar{a}_{t-1}, a_{t+1}^{opt}}|H_t = h_t]$  and,  $\gamma_t(a_t, h_t) = \mathbb{E}[Y^{\bar{a}_{t-1}, a_{t+1}^{opt}} - Y^{\bar{a}_{t-1}, 0}|H_t = h_t]$  so that

$$f_t(h_t) + \gamma_t(a_t, h_t) = \mathbb{E}[Y^{\bar{a}_{t-1}, a_{t+1}^{opt}}|H_t = h_t, A_t = a_t].$$

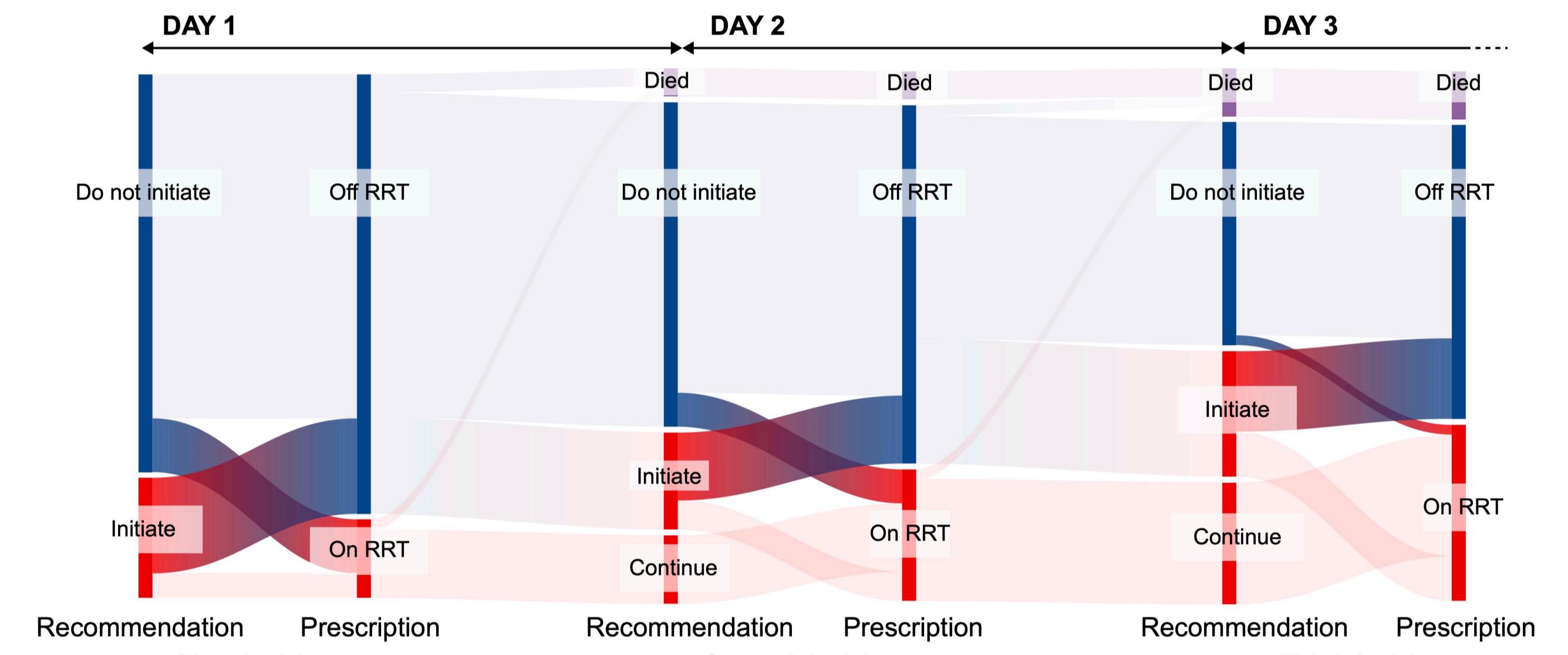
The estimation of  $f_t(\cdot)$  and  $\gamma_t(\cdot)$  starts at  $t = 3$  by regressing  $Y^{\bar{a}_3, a_4^{opt}} = Y$  onto  $(H_3^f, A_3 H_3^Y)$  via weighted least squares with overlap weights  $\tilde{w}_3(H_3) = |A_3 - \hat{e}_3(H_3)|$ . The procedure then follows a backward stepwise approach where we substitute all unobserved potential outcomes with pseudo-outcomes. Specifically, for  $t = \{2,1\}$ , we build pseudo-outcomes  $\tilde{Y}_t = \mathbb{E}[Y^{\bar{a}_t, a_{t+1}^{opt}}|H_t, A_t]$  by taking naïve 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  $(H_t^f, A_t H_t^Y)$  via weighted least squares with weights  $\tilde{w}_t(H_t) = |A_t - \hat{e}_t(H_t)|$ .

### External validation of the learned strategy

To estimate the causal effect of implementing our learned strategy versus following usual care on HFD60, we used the advantage doubly robust (ADR) estimator with terminal state for strategy evaluation.<sup>5</sup>

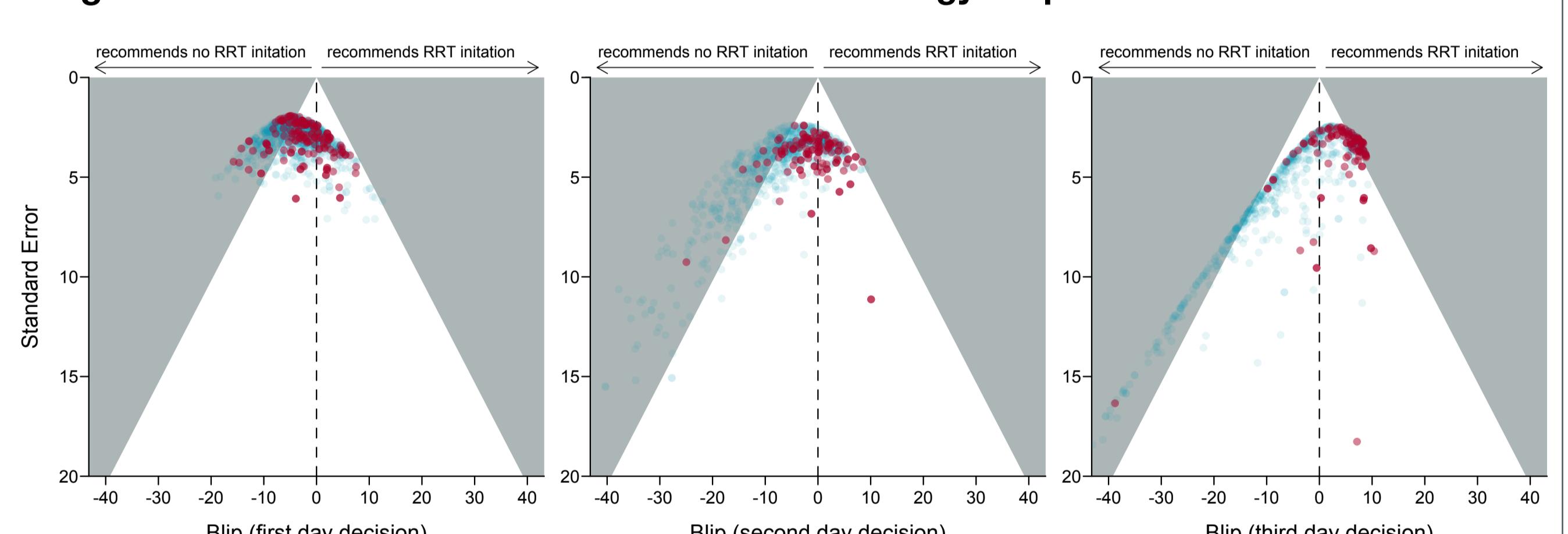
## Results

Figure 2. Comparison of recommendations from our learned strategy and renal-replacement therapy prescriptions observed in the validation set.



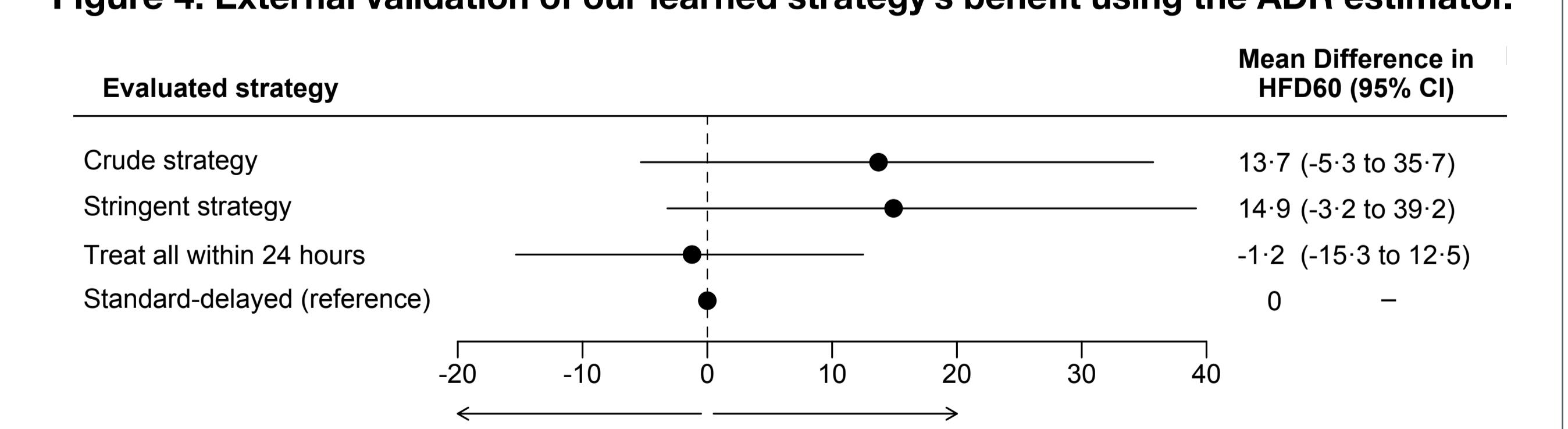
At each decision timepoint, recommendation and prescription to deliver RRT appear in red while recommendation and prescription not to deliver RRT appear in blue. Discrepancies between recommendations and prescriptions are shown in brighter colors.

Figure 3. Recommendations from our learned strategy for patients in the validation set.



Each dot corresponds to a patient for whom a decision whether to initiate RRT needed to be made at the first, second, or third decision timepoint. Dots falling in gray-shaded areas represent patients for whom there is evidence of an effect from RRT initiation at the .05 alpha level.

Figure 4. External validation of our learned strategy's benefit using the ADR estimator.



The “never treat” and “treat all” strategies represent two simpler strategies we also compared to usual care.

## Discussion

Implementing our learned strategy may increase the average number of days that ICU patients with AKI spend alive and outside the hospital. Simpler strategies appeared less promising and potentially harmful.

In the next step, we hope to prospectively test our strategy in ‘silent’ mode. Without informing patient care, off-duty clinicians shall seek to understand why our strategy makes the recommendations it does.

## References

1. Gaudry et al. Initiation strategies for renal replacement therapy in the intensive care unit. NEJM. 2016.
2. Gaudry et al. Comparison of two delayed strategies for renal replacement therapy initiation. Lancet. 2021.
3. Grolleau et al. Personalization of renal replacement therapy initiation. Critical Care. 2022.
4. Wallace et al. Doubly robust dynamic treatment regimen via weighted least squares. Biometrics. 2015.
5. Nie et al. Learning when-to-treat policies. JASA. 2021.

### Check out our clinical decision support tool



### Contact

François Grolleau PhD Candidate  
Hôpital Hôtel-Dieu  
1 Place du Parvis Notre-Dame, Paris, France  
<http://clinicaledempiere.fr/francois-grolleau>  
francois.grolleau@aphp.fr