On estimation and cross-validation of dynamic treatment regimes with competing risks

Paweł Morzywołek¹ | Johan Steen²,³,⁴ | Wim Van Biesen²,³ | Johan Decruyenaere²,⁴ | Stijn Vansteelandt¹,⁵

¹Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Ghent, Belgium
²Department of Internal Medicine and Pediatrics, Ghent University, Ghent, Belgium
³Renal Division, Ghent University Hospital, Ghent, Belgium
⁴Department of Intensive Care Medicine, Ghent University Hospital, Ghent, Belgium
⁵Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK

Correspondence
Paweł Morzywołek, Department of Applied Mathematics, Computer Science and Statistics, Ghent University, 9000 Ghent, Belgium.
Email: pawel.morzywolek@ugent.be

Funding information
Flemish Research Council (FWO), Grant/Award Number: FWO.OPR.2019.0045.01

The optimal moment to start renal replacement therapy in a patient with acute kidney injury (AKI) remains a challenging problem in intensive care nephrology. Multiple randomized controlled trials have tried to answer this question, but these contrast only a limited number of treatment initiation strategies. In view of this, we use routinely collected observational data from the Ghent University Hospital intensive care units (ICUs) to investigate different prespecified timing strategies for renal replacement therapy initiation based on time-updated levels of serum potassium, pH, and fluid balance in critically ill patients with AKI with the aim to minimize 30-day ICU mortality. For this purpose, we apply statistical techniques for evaluating the impact of specific dynamic treatment regimes in the presence of ICU discharge as a competing event. We discuss two approaches, a nonparametric one — using an inverse probability weighted Aalen-Johansen estimator — and a semiparametric one — using dynamic-regime marginal structural models. Furthermore, we suggest an easy to implement cross-validation technique to assess the out-of-sample performance of the optimal dynamic treatment regime. Our work illustrates the potential of data-driven medical decision support based on routinely collected observational data.

KEYWORDS
Aalen-Johansen estimator, acute kidney injury, competing events, cross-validation, dynamic treatment regimes, marginal structural models, precision medicine, renal replacement therapy, treatment-confounder feedback

1 | INTRODUCTION

Digitalization of patient records and increasing computational power have led to a paradigm shift in the field of medical decision-making from one-size-fits-all interventions to more sophisticated and data-driven intervention strategies optimized for particular subpopulations or individuals. This transition towards precision medicine¹,² has fuelled expectations to successfully tackle some challenging yet unresolved medical problems. One such unresolved question in intensive care medicine is when to initiate renal replacement therapy (RRT, dialysis) in patients with acute kidney injury (AKI). While RRT can avert some of the life-threatening consequences of kidney failure, it is also an invasive treatment and might put patients at increased risk of adverse events. Despite multiple observational studies and clinical trials comparing immediate treatment initiation with strategies that delay treatment initiation until one of the so-called “hard criteria” (sometimes

10.1002/sim.9568
also called “absolute indications”) has been met, the question of optimal timing of RRT initiation remains subject of discussion.3–5

The need for individualized decision support has led causal inference researchers to investigate so-called dynamic treatment regimes (DTR).6,7 A dynamic treatment regime is a set of sequential rules for treatment decisions based on patients’ time-varying covariates, for example, “start RRT in AKI-diagnosed patients if pH fell below 7.1 in the last 24 h”. In contrast, a static treatment regime does not take into account post-baseline information about the patient, for example, “always initiate RRT within 24 h from the AKI diagnosis” or “never initiate RRT in AKI-diagnosed patients.” Some of the early work in the field of dynamic treatment regimes introduced regret-based learning (sometimes referred to as A-learning)8 and g-estimation.9 These methods allow to find flexible optimal decision rules from a rich set of potential treatment strategies. A parallel line of research aimed to constrain the search for optimal treatment rules to a more limited set of simple and well-interpretable treatment strategies. Some of the most prominent methods applied in this context are based on marginal structural models.10–12

The main motivation for our work is to investigate different prespecified treatment strategies for initiation of renal replacement therapy to minimize 30-day ICU mortality in critically ill patients with stage 2 AKI. There are three main complications in the analysis. The first difficulty is that the main event of interest — ICU death within 30 days from AKI diagnosis — is subject to a competing event, namely alive discharge from the ICU (henceforth referred to as ICU discharge). The second challenge relates to the time-varying nature of treatment decisions and of confounding. In practice, past information on patients’ evolving condition and kidney function is used to make a decision on whether to initiate treatment, which in turn influences future kidney function. This induces treatment-confounder feedback that cannot be controlled for via traditional methods to deal with confounding, like stratification, matching, or regression adjustment.13 The third challenge is reliable performance assessment and comparison of considered treatment regimes. The commonly used approach is to evaluate all regimes on the full data, choose the one with the best performance (in our example the one with the lowest 30-day ICU mortality) and report its performance on the same data set. Such in-sample performance evaluation may lead to an overly optimistic assessment.

This article provides guidance on how competing events and treatment-confounder feedback can adequately be taken into account when comparing different dynamic treatment strategies. In particular, we focus on a set of treatment strategies, which involve mild deviations from current clinical practice, and evaluate their comparative effectiveness with the aim to identify an optimal strategy that improves outcomes compared to current clinical practice. We address the question of optimal timing of RRT initiation using a semiparametric estimation approach based on dynamic-regime marginal structural models (DYR-MSM)10–12 in the context of competing risks survival analysis and compare it with a weighted version of the nonparametric Aalen-Johansen estimator14 applied to dynamic treatment regimes. Furthermore, we apply cross-validation for out-of-sample performance assessment of the estimated optimal dynamic treatment regime. Cross-validation methodology for the performance evaluation of statistical procedures is commonly applied in the statistical prediction literature with the aim to avoid overoptimism.15–21

The article is structured as follows. In Section 2, we describe the problem setting and introduce the key concepts related to dynamic treatment regimes and competing risks survival analysis. In Section 3, we describe the use of cross-validation for the out-of-sample performance evaluation of the optimal dynamic treatment regime. In Section 4, we discuss different approaches to estimate cumulative incidence curves under specific dynamic treatment regimes. In Section 5, we present an empirical data application of the discussed methodology for optimizing the time of treatment initiation in the AKI patient population. In Section 6, we provide a discussion of the results.

2 PROBLEM SETTING AND NOTATION

We consider $K \geq 1$ time points at which treatment decisions need to be made. In our setting $K = 30$, since we are interested in 30-day ICU mortality and treatment initiation is considered daily. Furthermore, throughout the main text we assume no censoring due to administrative end of the study or loss to follow-up. However we describe the extension of the methods incorporating censoring due to administrative end of the study or loss to follow-up in Appendix A in the Supplementary Material. In the paragraphs that follow we denote random variables with capital letters and their potential realizations by small letters. In particular, let $A_k, k = 0, \ldots, K − 1$, denote a treatment option at decision time $k$, which takes value $a_k \in \{0, 1\}$, with 1 indicating that RRT has been initiated and 0 otherwise. Let $X_0$ be the information available about the individual at the time of stage 2 AKI diagnosis (which is considered as start of follow-up, or time 0), including baseline characteristics. Similarly, let $X_k, k = 1, \ldots, K$ denote patient information that becomes available between decision times
k – 1 and k, for example, the lowest level of pH in the last 24 h and so forth. Let \( Y_k, k = 0, \ldots \), \( K \) indicate whether or not the event of interest (ICU death) has been observed by time \( k \) (with \( Y_k = 1 \) if it has or \( Y_k = 0 \) otherwise). Similarly, let \( Z_k, k = 0, \ldots \), \( K \) indicate whether or not the competing event (ICU discharge) has been observed by time \( k \) (with \( Z_k = 1 \) if it has or \( Z_k = 0 \) otherwise). Note that only one of two events can be observed for a single patient. We assume the following order of the variables within the \( k \)th time interval \((Z_k, Y_k, X_k, A_k)\). The ordering needs to be aligned with our understanding of the data-generating mechanism and consistently applied across time intervals. Furthermore, define \( O_k \equiv (Z_k, Y_k, X_k) \) and \( H_k \equiv (H_k, \ldots, H_k) \) be the history until and including time \( k \). For all patients in the study population \( A_k = 0 \) for \( k < 0 \), \( Y_k = 0 \) for \( k \leq 0 \) and \( Z_k = 0 \) for \( k \leq 0 \).

A decision rule at time \( k \) is a function that maps the individual’s history to a treatment option \( d_k : (\tilde{O}_k, \tilde{a}_{k-1}) \rightarrow a_k \), for example, “start RRT in AKI-diagnosed patients if pH fell below 7.1 in the last 24 h”. Furthermore, we define \( \tilde{d}_k (\tilde{O}_k, \tilde{a}_{k-1}) \equiv (d_0 (0_0, a_{-1}), \ldots, d_k (\tilde{O}_k, \tilde{a}_{k-1})) \). A dynamic treatment regime is a collection of decision rules, one decision rule corresponding with each of the \( K \) decision points, that is, \( d \equiv \tilde{d}_{K-1} \).

For the purpose of defining and identifying the optimal treatment regime we use the potential outcomes framework from the causal inference literature.\(^{22-24}\) Let \( Y_k^d, k = 1, \ldots, K \) be the potential outcome for the event of interest (ICU death) at time \( k \) that would be observed if all individuals had followed the dynamic treatment regime \( d \) until time \( k – 1 \). Analogously, let \( Z_k^d, k = 1, \ldots, K \) be the potential outcome for the competing event (ICU discharge) at time \( k \) that would be observed if all individuals had followed the dynamic treatment regime \( d \) until time \( k – 1 \). Note that the considered framework merely infers aspects of the distribution under a single, fixed treatment regime, hence alternative frameworks (e.g., decision-theoretic) to infer the effect of dynamic treatment strategies could be used.\(^{25}\)

Let \( D = \{d^{(1)}, \ldots, d^{(m)}\} \) denote a set of prespecified dynamic treatment regimes. We aim to identify the optimal regime \( d_{opt,v} \in D \) for a subpopulation with characteristics \( V = v \) for \( V \subset X_0 \), a possibly empty subset of the baseline covariates. To compare different treatment regimes in the search for the optimal one, we define the value of a regime \( d \) (for the subpopulation with characteristics \( V = v \)) to be 30-day ICU mortality \( V(d; V = v) \equiv P(Y_k^d = 1|V = v) \). Furthermore, we define the optimal regime as the regime that minimizes the value, that is, \( d_{opt,v} \equiv \arg \min_{d \in D} V(d; V = v) \).

To estimate \( P(Y_k^d = 1|V = v) \), for \( k = 1, \ldots, K \), we define the counterfactual event-specific hazard, as the event-specific hazard of an event at time \( k \) under treatment regime \( d \) among patients with characteristics \( V = v \) who have not previously experienced the event of interest nor the competing event. Using the definitions introduced above the counterfactual discrete-time event-specific hazard for the event of interest under the treatment regime \( d \) at time \( k \) can be written as

\[
h_k^{(1)}(d, V) \equiv P(Y_k^d = 1|Y_{k-1}^d = 0, Z_k^d = 0, V).
\] (1)

Similarly, the counterfactual event-specific hazard for the competing event under the treatment regime \( d \) at time \( k \) is defined as

\[
h_k^{(2)}(d, V) \equiv P(Z_k^d = 1|Z_{k-1}^d = 0, Y_{k-1}^d = 0, V).
\] (2)

Even in the absence of censoring due to loss to follow-up, we advocate a hazard-based approach for estimating \( P(Y_k^d = 1|V = v) \), for \( k = 1, \ldots, K \), instead of direct estimation of the risk of the binary outcome \( Y_k^d \). This choice is further motivated in Section 4.4.

It is important to notice that these hazards, while defined in terms of counterfactuals, do not carry a causal interpretation due to a built-in selection bias when comparing two populations defined post-baseline. Indeed, imagine we would like to compare the effect of treatment regime \( d^{(1)} \) relative to \( d^{(2)} \) using time-varying hazard (1). Furthermore, assume that the treatment strategy \( d^{(1)} \) is detrimental to a frail subpopulation of patients. For the sake of the argument we assume that the patient population is randomly assigned to one of the treatment regimes at baseline, however the argument generalizes to observational studies. Due to randomization the proportion of frail patients in both groups will be similar at baseline allowing fair comparison of hazards in both groups at time 0. However, as time progresses the proportion of frail patients in the cohort assigned to \( d^{(1)} \) will decrease due to the detrimental effect of this treatment strategy. This means that at later time points the hazard rate in the group assigned to \( d^{(1)} \) will be computed over a healthier population than in those assigned to \( d^{(2)} \), hence leading to a potentially misleading comparison.\(^{26-28}\)
To avoid such selection bias, we instead compare cumulative incidence functions. These can be expressed as functionals of the counterfactual event-specific hazards by the following formula for $k = 1, \ldots, K$:

$$P(Y_k^d = 1|V) = \sum_{l=1}^{k} h_l^{(1)}(d, V) \left\{ 1 - h_l^{(2)}(d, V) \right\} \prod_{j=1}^{l-1} \left\{ 1 - h_j^{(1)}(d, V) \right\} \left\{ 1 - h_j^{(2)}(d, V) \right\}. \quad (3)$$

This formula expresses that the probability of observing the event of interest by day $k$ equals the sum of probabilities to observe the event of interest by any particular day, but not the competing event on that day (following the assumed order of events within a time interval) and also no event on any of the previous days.

### 3 Cross-Validation for Dynamic Treatment Regimes

The purpose of using cross-validation in our analysis is to assess the best performing dynamic treatment regime on a separate set than the one used to select it. This is encouraged by the TRIPOD (Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis) guidelines for internal model validation. In particular, we wish to evaluate what would be the performance of the optimal DTR from a predefined set of potential treatment regimes $D = \{d^{(1)}, \ldots, d^{(m)}\}$ when deployed in a clinical environment. The commonly used approach is to evaluate all regimes in $D$ on the full data, choose the one with the best performance (in our example the one with the lowest 30-day ICU mortality) and report its performance on the same data set. Such in-sample performance evaluation may lead to an overly optimistic assessment. For instance, imagine we would like to compare two treatment regimes $d^{(1)}$ and $d^{(2)}$. Furthermore, assume both regimes have exactly the same performance. However, once we evaluate and compare them on our data, then merely due to sampling variability, one may seem to perform better than the other. Claiming the optimality of this regime and reporting its in-sample performance would then leave a false optimistic impression of the superiority of one regime.

To perform an out-of-sample evaluation, one could split the data into a training set and a validation set and then evaluate all DTRs on the training set to pick the one with the lowest 30-day ICU mortality. Once this has been done, one could report the out-of-sample performance of this best performing DTR as calculated on the test set. This will result in a less optimistic evaluation than that based on the in-sample performance of the DTR on the training set. It reflects the expected performance of the data-driven procedure to select an optimal DTR when deployed in real-time in a clinical environment. We apply an out-of-sample performance assessment based on cross-validation. This extends the idea of performance evaluation done on the test set by repeating this procedure using different data splits.

We perform $J$-fold cross-validation by splitting the available data set into $J$ mutually exclusive parts of equal size. We combine $J - 1$ folds into a training set and use the remaining $J$th fold as a test set. We permute the role of different folds, so that each fold becomes the test set exactly once. For $j \in \{1, \ldots, J\}$, let $\hat{P}_j^{\text{train}}(Y_{K}^d = 1|V = v)$ be 30-day ICU mortality for the subpopulation with characteristics $V = v$ under the dynamic treatment regime $d$ estimated on the $j$th training set and $\hat{P}_j^{\text{test}}(Y_{K}^d = 1|V = v)$ be 30-day ICU mortality (for the same subpopulation) under the dynamic treatment regime $d$ estimated on the $j$th test set. We estimate the cumulative incidence curve up to time $K$ on the training set for each considered dynamic treatment regime using the methods described in the following sections. For each $j \in \{1, \ldots, J\}$, we choose the best treatment regime $d_{\text{opt},v}^{(j)}$ in the $j$th training set to be the one attaining the lowest 30-day ICU mortality on the $j$th training set

$$d_{\text{opt},v}^{(j)} \equiv \arg\min_{d \in D} \hat{P}_j^{\text{train}}(Y_{K}^d = 1|V = v).$$

This leads to $J$ estimated optimal dynamic treatment regimes $d_{\text{opt},v}^{(1)}, \ldots, d_{\text{opt},v}^{(J)}$, together with the estimated 30-day ICU mortality for the optimal rule on the $j$th test set

$$\hat{P}_j^{\text{test}}(Y_{K}^{d_{\text{opt},v}^{(j)}} = 1|V = v).$$
for \( j = 1, \ldots, J \). We then compute the cross-validated counterfactual 30-day ICU mortality under the optimal regime given by

\[
\frac{1}{J} \sum_{j=1}^{J} \mathbb{P}_{j}^{\text{test}} \left( Y_{k}^{d_{j}^{(\text{opt}, v)}} = 1 \mid V = v \right).
\]

(4)

It is important to note that the optimal \( d_{j}^{(\text{opt}, v)} \) corresponding to different folds \( j = 1, \ldots, J \) might not always be the same regime. In particular, (4) may not correspond to the 30-day ICU mortality of a single DTR, but it can be interpreted as 30-day ICU mortality that would have been obtained by deploying the procedure of choosing the optimal treatment regime out of set \( D \) of possible rules on new data. The benefit of the cross-validated counterfactual 30-day ICU mortality given by (4) is twofold. By comparing it with the cumulative incidence under current standard of care we assess whether deploying the optimal regime is expected to lead to a decrease in the ICU mortality. Furthermore, by comparing the cross-validated mortality with the 30-day ICU mortality under the optimal treatment regime evaluated on the whole data set, we obtain an assessment of the optimism due to selection and evaluation of the optimal treatment regime on the same data set. Note that this assessment may be somewhat conservative (ie, overestimated) due to the reduction in size of the training set.

Cross-validation has also been suggested or considered by others. For instance, Luedtke and van der Laan rely on a cross-validation procedure like the above to obtain data-adaptive estimates of the mean outcome under an estimated optimal regime, along with confidence intervals.

## 4 | ESTIMATION OF THE CUMULATIVE INCIDENCE CURVE

We discuss two approaches to estimate the cumulative incidence curve given by (3) — a nonparametric one, leading to the inverse probability of censoring weighted (IPC weighted) Aalen-Johansen estimator of the cumulative incidence and a semiparametric one, using dynamic-regime marginal structural models (DYR-MSM).

In our motivating example, we are interested in investigating the optimal threshold \( x \) for a decision rule “start RRT in AKI-diagnosed patients if pH fell below \( x \) in the last 24 h”. The IPC weighted Aalen-Johansen estimator targets the cumulative incidence curve for a single treatment regime, independently of the other regimes. The advantage of a dynamic-regime marginal structural model is that it permits to postulate a joint model for different treatment regimes corresponding to different threshold values. This may be appealing when a data set carries little information about certain dynamic treatment regimes and smoothing across similar regimes is warranted.

### 4.1 | Dynamic treatment regime compatibility

Given our aim to identify the best performing treatment strategy \( d_{\text{opt}, v} \) out of the set of prespecified treatment strategies \( D = \{ d^{(1)}, \ldots, d^{(m)} \} \), the favored way to compare different DTRs would be to perform an \( m \)-arm randomized trial in which patients are randomly assigned to one of the considered treatment regimes in \( D \). Unfortunately, once the number of treatment regimes grows, such a randomized trial becomes infeasible. Alternatively, we can evaluate each of the treatment strategies in \( D \) using observational data and compare their performance to identify the best one. The difficulty that we are facing in this case is the fact that we are interested in evaluating retrospectively what would be the effect of applying a hypothetical treatment strategy \( d \in D \) in the study population, whereas the patients included in the observational data set, had been treated based on the treatment strategy followed by their clinicians, which might have been very different from the dynamic treatment strategy \( d \) we wish to evaluate. Therefore, to evaluate performance of the hypothetical treatment strategy \( d \) we need to identify patients in the data set whose observed treatment history coincides with the treatment trajectory that would be observed under the hypothetical treatment regime \( d \) we wish to evaluate, because only these patients can provide information on the effect of hypothetical DTR of interest. We will refer to these patients as compatible with the dynamic treatment regime \( d \) that we consider. Note that all patients are compatible at time 0 with any of the considered regimes \( d \). Now we follow each individual across time and we artificially censor him/her as soon as the observed treatment trajectory deviates from the one that would be observed under the considered DTR \( d \). In other words, we artificially censor a patient, if and as soon as that patient stops being compatible with the considered DTR. We
define the indicator of being compatible with the treatment regime \(d\) at time \(k\) as

\[
C_{d,k} \equiv 1 \left\{ \overline{A}_k = \overline{d}_k \left( \overline{O}_k, \overline{A}_{k-1} \right) \right\},
\]

where

\[
1(X = x) = \begin{cases} 1, & X = x \\ 0, & \text{otherwise} \end{cases}
\]
is an indicator function. Therefore, a patient is compatible with the considered treatment regime \(d\) at time \(k\), if her/his treatment history coincides with the treatment options that would have been assigned by \(d\) for all the past decision points. The idea of estimating the effect of a treatment regime through censoring individuals the first time their treatment differs from the treatment assigned by the considered DTR has been commonly used in the dynamic treatment regimes literature.\(^6,7,10-12,32,34\)

In order to find the optimal treatment regime we need to evaluate each of the DTRs in \(D\) and compare them to identify the one with the lowest resulting mortality. For this purpose we create an extended data set with one copy of each subject in the original observational data set for each considered treatment regime. For each regime we follow the patient population over time and we artificially censor individuals as soon as they are no longer compatible with that given regime. The probability of remaining compatible with the dynamic treatment regime \(d\) at time \(k\) (conditional on being compatible up to time \(k - 1\)) is given by

\[
\pi_{d,k} \equiv \mathbb{P} \left( C_{d,k} = 1 \mid \overline{O}_k, C_{d,k-1} = 1 \right) = \begin{cases} \mathbb{P} \left( C_{d,k} = 1 \mid \overline{X}_k, Y_k = 0, Z_k = 0, C_{d,k-1} = 1 \right), & \text{if } Y_k = 0 \text{ and } Z_k = 0 \\ 1, & \text{otherwise}. \end{cases}
\]

(5)

Therefore, the probability of being artificially censored with respect to the treatment regime \(d\) at time \(k\) (conditional on being compatible up to time \(k - 1\)) is given by \(1 - \pi_{d,k}\). Note that if one of the competing events occurs at time \(k\) (i.e., \(Y_k = 1\) or \(Z_k = 1\)) while an individual is compatible with the dynamic treatment regime \(d\) (i.e., \(C_{d,k-1} = 1\)), then the individual automatically remains compatible for all future time points.

As an example consider a data set containing information on two patients observed over 5 days. Suppose we want to compare two dynamic treatment regimes “start RRT in AKI-diagnosed patients if pH fell below 7.1 in the last 24 h” and “start RRT in AKI-diagnosed patients if pH fell below 7.2 in the last 24 h”. Therefore we make two copies of the original data set — one for each treatment regime that we consider — and stack them on top of each other to obtain an extended data set (see Table 1). The compatibility indicator in Table 1 corresponds to the compatibility status of the patient at the end of the day. Given that a patient is compatible with a considered regime \(d\) at time \(k\), we can determine the compatibility indicator on the following day by checking whether the treatment decision based on the dynamic treatment regime coincides with the observed treatment. For the DTR “start RRT in AKI-diagnosed patients if pH fell below \(x\) in the last 24 h” a given patient would remain compatible if the minimal level of pH in the last 24 h fell below the threshold \(x\) and the treatment has been initiated or if the minimal level of pH did not fall below the threshold \(x\) in the last 24 h and treatment has not been started. Once the patient ceases to be compatible with a particular treatment regime at some time point, then he/she will be incompatible with this regime for all future time points. Furthermore, the patients that are compatible with a particular DTR at the time of experiencing one of the events — ICU death or ICU discharge — remain compatible with the considered dynamic treatment regime until the end of the study.

### 4.2 Identification of the causal estimands

As discussed in the previous section, to compare different dynamic treatment regimes \(d\) we need to estimate \(\mathbb{P} \left( Y_d^k = 1 \mid V = v \right)\). Because this is a counterfactual estimand, it cannot be estimated from the data without invoking common assumptions from the causal inference literature. The consistency assumption is crucial for the application of the potential outcomes framework as it links the observed data to the potential outcomes. It states that for each patient compatible with dynamic treatment regime \(d\) up until time \(k\) the observed variables at time \(k\) correspond to the potential
TABLE 1 Extended data set for comparison of two dynamic treatment regimes “start RRT in AKI-diagnosed patients if pH fell below 7.1 in the last 24 h” and “start RRT in AKI-diagnosed patients if pH fell below 7.2 in the last 24 h”

| ID | Time (days) | DTR | Observed treatment | Min. pH (last 24 h) | Compatibility indicator | PS | IPCW | ICU death | ICU discharge |
|----|-------------|-----|--------------------|---------------------|-------------------------|----|------|-----------|---------------|
| 1  | 1           | 7.1 | 0                  | 7.29                | 1                       | 0.01 | 1.01 | 0         | 0             |
| 1  | 2           | 7.1 | 0                  | 7.24                | 1                       | 0.05 | 1.06 | 0         | 0             |
| 1  | 3           | 7.1 | 1                  | 7.08                | 1                       | 0.4  | 2.67 | 0         | 0             |
| 1  | 4           | 7.1 | 1                  | 7.29                | 1                       | 1.0  | 2.67 | 0         | 0             |
| 1  | 5           | 7.1 | 1                  | 7.29                | 1                       | 1.0  | 2.67 | 0         | 1             |
| 2  | 1           | 7.1 | 0                  | 7.3                 | 1                       | 0.13 | 1.15 | 0         | 0             |
| 2  | 2           | 7.1 | 0                  | 7.29                | 1                       | 0.11 | 1.3  | 0         | 0             |
| 2  | 3           | 7.1 | 1                  | 7.19                | 0                       | 0.48 | 0.0  | 0         | 0             |
| 2  | 4           | 7.1 | 1                  | 7.32                | 0                       | 1.0  | 0.0  | 1         | 0             |
| 2  | 5           | 7.1 | 0                  | 7.32                | 0                       | 1.0  | 0.0  | 1         | 0             |
| 1  | 1           | 7.2 | 1                  | 7.29                | 1                       | 0.01 | 1.01 | 0         | 0             |
| 1  | 2           | 7.2 | 0                  | 7.24                | 1                       | 0.05 | 1.06 | 0         | 0             |
| 1  | 3           | 7.2 | 1                  | 7.08                | 1                       | 0.4  | 2.67 | 0         | 0             |
| 1  | 4           | 7.2 | 1                  | 7.29                | 1                       | 1.0  | 2.67 | 0         | 0             |
| 1  | 5           | 7.2 | 1                  | 7.29                | 1                       | 1.0  | 2.67 | 0         | 1             |
| 2  | 1           | 7.2 | 0                  | 7.3                 | 1                       | 0.13 | 1.15 | 0         | 0             |
| 2  | 2           | 7.2 | 0                  | 7.29                | 1                       | 0.11 | 1.3  | 0         | 0             |
| 2  | 3           | 7.2 | 1                  | 7.19                | 0                       | 0.48 | 2.9  | 0         | 0             |
| 2  | 4           | 7.2 | 1                  | 7.32                | 1                       | 1.0  | 2.9  | 1         | 0             |
| 2  | 5           | 7.2 | 1                  | 7.32                | 1                       | 1.0  | 2.9  | 1         | 0             |

Note: “Min. pH (last 24 h)” is the lowest value of pH measured in the last 24 h for a given patient. “PS” is propensity score given by (7) and “IPCW” is inverse probability of censoring weight given by (6).

outcome variables under treatment regime \( d \). This is expressed for \( k = 0, \ldots, K - 1 \) as

\[
\bar{O}_{k+1}^d = \bar{O}_{k+1} \text{ if } C_{d,k} = 1,
\]

where \( O_{k}^d \equiv (Z_{k}^d, Y_{k}^d, X_{k}^d) \) and \( X_{k}^d \) is the potential outcome for patient information that would become available between decision times \( k - 1 \) and \( k \) under regime \( d \). Another important assumption is sequential randomization, which can be stated as follows

\[
(Z_{k+1}^d, Y_{k+1}^d) \perp A_k | \bar{O}_k, \bar{A}_{k-1} = \bar{d}_{k-1} (\bar{O}_{k-1}, \bar{A}_{k-2}) \text{, for } k = 0, \ldots, K - 1,
\]

where \( Y_{k}^d \equiv (Y_{k}, \ldots, Y_{K}) \) and \( Z_{k}^d \equiv (Z_{k}, \ldots, Z_{K}) \), for \( k = 1, \ldots, K \), denote the future trajectory of the potential outcomes for the event of interest and competing event, respectively, if all individuals had followed the dynamic treatment regime \( d \). This assumption holds by design if treatment initiation is randomly assigned at each decision point \( k \) among patients in whom RRT has not been initiated before time \( k \), with randomization possibly depending on the observed past. However, in observational studies this assumption requires careful consultation with subject matter experts as it essentially states that there are no unmeasured variables that influence both the treatment assignment at a given time point and the final outcome, that is, no unmeasured confounding. Finally, the positivity assumption for the dynamic treatment regimes states
that for \( k = 0, \ldots, K - 1 \)
\[
P \left\{ A_k = d_k \left( \bar{A}_k, \bar{A}_{k-1} \right) | \bar{X}_k, \bar{Y}_k = 0, \bar{Z}_k = 0, C_{d,k-1} = 1 \right\} > 0, \ \text{w.p. 1},
\]
that is, each patient that has remained event-free and compatible with the treatment regime \( d \) up until time \( k \) has a positive probability of receiving treatment prescribed by \( d_k \) at time \( k \).

### 4.3 Inverse probability of censoring weighting

Artificial censoring due to incompatibility with a treatment regime \( d \) at time \( k \) might select a group of patients that is no longer comparable with the risk set under the observed regime in terms of prognostic factors for the outcome. One approach to remedy selection bias due this artificial censoring is the use of inverse probability of censoring weighting (IPC weighting). IPC weights are given by
\[
W_{d,k,l} \equiv \frac{C_{d,k,l}}{\prod_{l=1}^{K} \pi_{d,l,l}}, \tag{6}
\]
where \( \pi_{d,l,l} \) is the probability that an individual \( i \) remains compatible with the DTR \( d \) by the time \( l \) conditional on the observed history and on having remained compatible up to the previous time point given by (5). To obtain the estimates \( \hat{r}_{d,k,l} \), and hence \( \hat{W}_{d,k,l} \), we note first that the upper part of the expression (5) can be written as
\[
P \left\{ A_k = d_k \left( \bar{A}_k, \bar{A}_{k-1} \right) | \bar{X}_k, \bar{Y}_k = 0, \bar{Z}_k = 0, C_{d,k-1} = 1 \right\} = P \left( A_k = 1 | \bar{X}_k, \bar{Y}_k = 0, \bar{Z}_k = 0, C_{d,k-1} = 1 \right) \times \left\{ 1 - P \left( A_k = 1 | \bar{X}_k, \bar{Y}_k = 0, \bar{Z}_k = 0, C_{d,k-1} = 1 \right) \right\}^{1-d_k(\bar{\alpha}, \bar{A}_{k-1})}.
\]
Furthermore, note that for individuals compatible with the DTR \( d \) by the time \( k \) we have
\[
P \left( A_k = 1 | \bar{X}_k, \bar{Y}_k = 0, \bar{Z}_k = 0, C_{d,k-1} = 1 \right) = \begin{cases} 
P \left( A_k = 1 | \bar{X}_k, \bar{Y}_k = 0, \bar{Z}_k = 0, \bar{A}_{k-1} = \bar{\alpha}_{k-1} \right), & \text{if } \bar{\alpha}_{k-1} = \bar{\alpha} \\
1, & \text{otherwise}, \end{cases} \tag{7}
\]
where \( P \left( A_k = 1 | \bar{X}_k, \bar{Y}_k = 0, \bar{Z}_k = 0, \bar{A}_{k-1} = \bar{\alpha}_{k-1} \right) \) is the time-dependent propensity score, that is, the probability of receiving treatment at a given time point given the observed event and covariate history. Note that
\[
P \left( A_k = 1 | \bar{X}_k, \bar{Y}_k = 0, \bar{Z}_k = 0, \bar{A}_{k-1} = \bar{\alpha}_{k-1} \right) = 1 \text{ if } \bar{\alpha}_{k-1} \neq \bar{\alpha}, \text{ because in our analysis we focus on the treatment initiation, hence once the treatment has been initiated at time } k, \text{ that is, } a_k = 1, \text{ for some } k \in \{0, \ldots, K-1\}, \text{ then the indicator for the observed treatment remains 1 from that point onwards, that is, } a_l = 1, \text{ for } l \geq k. \text{ Therefore, we only need to estimate these time-dependent propensity scores in patients in whom the treatment has not been initiated yet, which can be done by fitting a pooled logistic regression model}
\[
P \left( A_k = 1 | \bar{X}_k, \bar{Y}_k = 0, \bar{Z}_k = 0, \bar{A}_{k-1} = \bar{\alpha} \right) = \exp \left( \gamma_0 + \gamma_1 k + \bar{\gamma}_1 \bar{X}_k \right), \text{ for } k = 0, \ldots, K - 1.
\]
In Table 1 the column “PS” contains values of propensity scores given by (7). Once we have the propensity scores we calculate the IPC weights (“IPCW”) using expression (6).

### 4.4 Inverse probability weighted Aalen-Johansen estimator

In this section we let \( V = \emptyset \) or, in other words, we consider the entire patient population. With interest in a specific subpopulation with characteristics \( V \) one may perform exactly the same analysis restricting to the stratum \( V = \nu \) (provided...
that it contains data on a sufficient number of individuals). Under the aforementioned assumptions, the counterfactual event-specific hazards (1) and (2) can be consistently estimated by

\[ \hat{h}_{k}^{(1)}(d) = \frac{\sum_{i=1}^{n} Y_{k,i} (1 - Z_{k,i}) (1 - Y_{k-1,i}) \hat{W}_{d,k-1,i}}{\sum_{i=1}^{n} (1 - Z_{k,i}) (1 - Y_{k-1,i}) \hat{W}_{d,k-1,i}}, \]  

(8)

\[ \hat{h}_{k}^{(2)}(d) = \frac{\sum_{i=1}^{n} Z_{k,i} (1 - Y_{k-1,i}) (1 - Z_{k-1,i}) \hat{W}_{d,k-1,i}}{\sum_{i=1}^{n} (1 - Y_{k-1,i}) (1 - Z_{k-1,i}) \hat{W}_{d,k-1,i}}, \]  

(9)

where \( \hat{W}_{d,k-1,i} \) is given by (6) with \( k \) changed to \( k - 1 \). Expressions (8) and (9) correspond to the inverse probability of censoring weighted estimators of the counterfactual event-specific hazards at time \( k \) for the event of interest and for the competing event, respectively (see Lemma 1 in the Appendix B in the Supplementary Material). Substituting expressions (8) and (9) into formula (3) we obtain the following estimator for the counterfactual cumulative incidence at time \( k = 1, \ldots, K \)

\[ \hat{p}(Y_{k}^{d} = 1) = \sum_{l=1}^{k} \hat{h}_{l}^{(1)}(d) \left\{ 1 - \hat{h}_{l}^{(2)}(d) \right\} \prod_{j=1}^{l-1} \left\{ 1 - \hat{h}_{j}^{(1)}(d) \right\} \left\{ 1 - \hat{h}_{j}^{(2)}(d) \right\}, \]  

(10)

which corresponds to an inverse probability of censoring weighted Aalen-Johansen estimator.

Instead of modeling event-specific hazards (8) and (9), and then combining them into a cumulative incidence curve via the formula (10), one could aim at direct modeling the binary outcome \( Y_{k}^{d} \) to obtain an IPC weighted estimator of \( p(Y_{k}^{d} = 1) \) given by

\[ \hat{p}(Y_{k}^{d} = 1) = \frac{\sum_{i=1}^{n} Y_{k,i} \hat{W}_{d,k-1,i}}{\sum_{i=1}^{n} \hat{W}_{d,k-1,i}}. \]

The estimators that we consider are expected to have better performance. This can be seen by the fact that they generate monotone cumulative incidence curves. The survival curves produced through modeling \( p(Y_{k}^{d} = 1) \) as a binary outcome consecutively for each \( k \in \{1, \ldots, K\} \), would require performing IPC weighting estimation separately at each time point, and this could lead to non-monotonicity of the resulting survival curve resulting from the estimation uncertainty of the IPC weights. In such a case, additional techniques like isotonic regression would be required to restore the monotonicity of the cumulative incidence curves.

### 4.5 Dynamic-regime marginal structural models

An alternative approach to estimate the cumulative incidence curve for different dynamic treatment regimes is to postulate dynamic-regime marginal structural models (DYR-MSM)\(^{10-12} \) for the event-specific hazards (1) and (2). A DYR-MSM is a model for the counterfactual event-specific hazard of the event of interest, for instance, a pooled logistic model

\[ h_{k}^{(1)}(d, V) = \expit(a_0 + a_1 d + a_2 k + a_3 dk + a_4 V), \text{ for } k = 1, \ldots, K, \]  

(11)

where \( d \) is modeled as a categorical variable, that is, it is a vector of dummy variables indicating the dynamic treatment regime. We denote the right hand side of (11) as \( h_{k}^{(1)}(d, V; \alpha) \). Analogously, one may postulate a DYR-MSM for the counterfactual event-specific hazard of the competing event (ICU discharge), for instance

\[ h_{k}^{(2)}(d, V) = \expit(\beta_0 + \beta_1 d + \beta_2 k + \beta_3 dk + \beta_4 V), \text{ for } k = 1, \ldots, K, \]  

(12)

where \( d \) is again modeled as a categorical variable. We denote the right hand side of (12) as \( h_{k}^{(2)}(d, V; \beta) \). Following the developments by Orellana et al.\(^{12} \) and van der Laan and Petersen\(^{13} \) we can consistently estimate the parameters indexing
for $\alpha$ and $\beta$. See Lemma 2 in the Appendix B in the Supplementary Material for technical details. Furthermore, choosing $q_k^{(1)}(d, V) = (1, d, k, dk, V)^\top$ and $q_k^{(2)}(d, V) = (1, d, k, dk, V)^\top$, the estimating Equations (13) and (14) correspond to the weighted score equations of the following logistic regression models

$$\mathbb{P}(Y_k = 1|Y_{k-1} = 0, Z_k = 0, C_{d,k-1} = 1, V) = \expit(\tilde{\alpha}_0 + \tilde{\alpha}_1^d + \tilde{\alpha}_2^k + \tilde{\alpha}_3^d k + \tilde{\alpha}_4^V),$$

$$\mathbb{P}(Z_k = 1|Z_{k-1} = 0, Y_{k-1} = 0, C_{d,k-1} = 1, V) = \expit(\tilde{\beta}_0 + \tilde{\beta}_1^d + \tilde{\beta}_2^k + \tilde{\beta}_3^d k + \tilde{\beta}_4^V),$$

for $k = 1, \ldots, K$ with weights given by (6). The resulting maximum weighted likelihood estimates of the parameters indexing (15) and (16) are consistent for the parameters indexing (11) and (12). Substituting the event-specific hazards $h_k^{(1)}(d, V; \hat{\alpha})$ and $h_k^{(2)}(d, V; \hat{\beta})$ in formula (3) we obtain the following consistent estimator of the cumulative incidence at time $k = 1, \ldots, K$ under regime $d$ in the subpopulation $V$

$$\hat{\mathbb{P}}(Y_k^d = 1|V) = \sum_{i=1}^{K} h_k^{(1)}(d, V; \hat{\alpha}) \left\{ 1 - h_k^{(2)}(d, V; \hat{\beta}) \right\} \prod_{j=1}^{k-1} \left\{ 1 - h_j^{(1)}(d, V; \hat{\alpha}) \right\} \left\{ 1 - h_j^{(2)}(d, V; \hat{\beta}) \right\}.$$

By repeating this calculation for each time $k$, this allows us to compute the cumulative incidence curve for each of the dynamic treatment regimes. Furthermore, by averaging over $V$, that is, obtaining a curve for each patient given his/her observed covariates $V$ and then averaging across all patients, we obtain the marginal cumulative incidence curve for the entire population.

In the example introduced in the Section 4.1, to estimate parameters $\alpha$ and $\beta$ we fit models (15) and (16) to the extended data set restricted to rows with compatibility indicator equal to 1 on the previous day for the patients who are still hospitalized at the intensive care unit, applying the IPC weights given by (6) to account for artificial censoring due to incompatibility with the given dynamic treatment regime.

### 4.5.1 DYR-MSM for heterogeneous treatment effects

The effect of conditioning on $V$ in the marginal structural model is twofold. First, by conditioning on the set of covariates $V$ we permit a different optimal decision rule across different strata $V$, even if regime $d$ is not formally defined as a function of $V$. Indeed, imagine we want to compare dynamic treatment regimes $d^{(1)}$, $d^{(2)}$, and $d^{(3)}$. If we choose $V = \emptyset$ then the optimal rule is one of the three considered rules applied on the whole population. However, if we condition on $V$ then the optimal rule might vary between different strata of the population defined by different levels of $V$, for example, $d^{(3)}$ applied to patients with $V = v_1$ and $d^{(1)}$ applied to patients with $V = v_2$. Therefore, conditional DYR-MSMs allow searching for the optimal rule over a bigger set of rules that, in addition to the aforementioned time-updated physiological parameters, also take into account baseline patient characteristics. Second, conditioning on a set of covariates $V$ in the MSM permits obtaining more stable estimates by applying stabilized IPC weights. If we use stabilized weights $SW_{d,k}$ (to be defined below) we need to make sure that the resulting estimating equations

$$\sum_{i=1}^{n} \sum_{d \in D} \sum_{k=1}^{K} \left\{ Y_{k,i} - h_k^{(1)}(d, V; \alpha) \right\} (1 - Y_{k-1,i}) (1 - Z_{k,i}) q_k^{(1)}(d, V) SW_{d,k-1,i} = 0,$$
\[
\sum_{i=1}^{n} \sum_{d \in D} \sum_{k=1}^{K} \left\{ Z_{k,i} - h_{k,i}^{(2)}(d, V; \beta) \right\} \left( 1 - Z_{k-1,i} \right) (1 - Y_{k-1,i}) q_{k,i}^{(2)}(d, V) SW_{d,k-1} = 0, \tag{19}
\]

remain unbiased. Following Cain et al.\(^{32}\) we argue that the numerator of the stabilized weight \(SW_{d,k}\) can only depend on \(d\) and \(V\), hence \(SW_{d,k} \equiv f_k(d, V) \times W_{d,k}\) for some function \(f_k(d, V)\). If the numerator \(f_k(d, V)\) is only a function of \(d\) and \(V\), then the unbiasedness of the estimating equations (18) and (19) follows from the fact that the estimating equations (13) and (14) are unbiased for all \(q_{k,j}^{(1)}(d, V)\) for \(j = 1, 2\). On the other hand, if the numerator \(f_k\) of the stabilized weight \(SW_{d,k}\) depends on other variables than \(d\) and \(V\), then Equations (18) and (19) are no longer unbiased for all \(q_{k,j}^{(1)}(d, V)\) for \(j = 1, 2\). Therefore, the numerator of the stabilized IPC weights \(f_k\) can only depend on the treatment regime \(d\) and the set of baseline covariates \(V\). In this case a natural choice for the numerator of the stabilized IPC weights is

\[
f_k(d, V) \equiv \prod_{i=1}^{k} P \left\{ A_l = d_l, \left( \bar{O}_{l-1}, \bar{A}_{l-1} \right) \mid \bar{Y}_l = \bar{0}, \bar{Z}_l = \bar{0}, C_{d,l-1} = 1, V \right\}.
\]

Indeed, this choice ensures that when

\[
A_l \perp \bar{O}_l \mid \bar{Z}_{l-1}, \bar{Y}_{l-1}, \bar{A}_{l-1}, V,
\]

that is, the confounding is solely due to baseline confounders or previous exposures, then the resulting estimator is efficient since the weights \(SW_{d,k} = 1\) for all \(k = 1, \ldots, K\) and therefore the resulting estimator is just the maximum likelihood estimator and hence is efficient.

5 | EMPIRICAL EXAMPLE

5.1 | Data

The ICIS (Intensive Care Information System) of the Ghent University Hospital ICUs contains records from all adult patients admitted to the intensive care unit since 2013. Our analysis is restricted to 4769 retrospectively identified stage 2 AKI patients\(^{36}\) (based on the KDIGO — Kidney Disease: Improving Global Outcomes — criteria\(^{37}\)), admitted between January 1, 2013 and December 31, 2017 and who satisfied the following criteria by the time of the inclusion at stage 2 AKI diagnosis: no recorded RRT history, aged 18 years or above, no stage 3 AKI diagnosis and no dialysis restriction.

For each patient we have information on ICU admission time, ICU discharge time, vital status at discharge, timesteps of all dialysis sessions during each ICU episode, baseline covariates (e.g., age, weight, gender, admission category (“No surgery”, “Planned surgery”, “Emergency surgery”), indicator for whether patient has received dialysis prior to current ICU admission, indicator for whether patient had chronic kidney disease diagnosis prior to current ICU admission) and longitudinal measurements (e.g., SOFA scores, indicator for whether KDIGO AKI (stage 1/2/3) creatinine condition has been reached during ICU episode, indicator for whether KDIGO AKI (stage 1/2/3) oliguric condition has been reached during ICU episode, cumulative total fluid intake, cumulative total fluid output, arterial pH values, serum potassium values (in mmol/L), serum ureum values (in mg/dL), serum magnesium values (in mmol/L), fraction of inspired oxygen (FiO\(_2\)), peripheral oxygen saturation (SpO\(_2\)), arterial oxygen concentration (PaO\(_2\)), ratio of arterial oxygen concentration to the fraction of inspired oxygen (P/F ratio), DNR (“Do Not Resuscitate” code) and their timestamps.

5.2 | Results

We consider dynamic treatment regimes that initiate RRT if a patient meets the AKI Stage 2 creatinine criterion and in the last 24 h patient’s pH fell below threshold \(x\) or the oliguric stage 3 AKI criterion had been met. We investigated different pH thresholds \(x \in \{7.0, 7.05, 7.1, 7.15, 7.2\}\) in order to identify the treatment rule resulting in the lowest 30-day ICU mortality.

To perform our analysis we have constructed an extended data set analogous to the example presented in the Table 1. This means we have a copy of the original data for each of the dynamic treatment regimes corresponding to different pH
Cumulative incidence curves for considered DTRs (pH < x) and observed cumulative incidence under the current standard of care (obs). Thick lines correspond to the estimates from the Aalen-Johansen estimator and thin lines correspond to the estimates from the dynamic-regime marginal structural model.

Patients are followed up from the time when the stage 2 AKI criteria are satisfied for the first time (day 0) until ICU death or discharge on day 30, whichever occurs first. We assume that the decision about RRT initiation based on the considered DTRs is performed every 24 h starting from day 0 based on the information available up to a given decision point. If the treatment initiation decision based on the considered DTR does not coincide with the observed treatment, then the patient is censored for this particular DTR from this time point onward, as described in the example in Section 4.1. We applied IPC weighting to deal with potential selection bias due to this artificial censoring. Once we have constructed the extended data set we calculate the cumulative incidence under each prespecified DTR using the IPC weighted Aalen-Johansen estimator given by (10) and marginal structural models given by (17) with \( V = \emptyset \), which means that we do not include any baseline covariates in the DYR-MSMs for the counterfactual event-specific hazards. We have made this choice in our empirical example since the interest of clinicians was in simple rules that do not change across different levels of baseline covariates. Moreover we have investigated the use of stabilized weights in our analysis, but this did not lead to improvement in the stability of weights. Therefore we used unstabilized IPC weights in the final analysis. The resulting MSM-based estimates for the cumulative incidence curve should thus approximate the corresponding weighted Aalen-Johansen estimator. We consider the optimal DTR to be the one with the lowest 30-day ICU mortality. Figure 1 presents the counterfactual cumulative incidence curves under different considered DTRs and the observed cumulative incidence under current standard of care (obs).

Figure 2 presents the cumulative incidence curves for the optimal dynamic treatment regime “initiate RRT if creatinine stage 2 AKI condition has been met and one of the events occurred in the last 24 h: pH < 7.05 or oliguric stage 3 AKI condition has been met,” which resulted in the lowest estimated 30-day ICU mortality among the considered DTRs, computed using the Aalen-Johansen estimator given by (10) and the estimator obtained using marginal structural models given by (17). The suggested treatment rule leads to a 12.6% relative decrease in 30-day ICU mortality from 12.43% for the current standard of care to 10.85%. The confidence intervals for the IPC weighted Aalen-Johansen estimator are based on the infinitesimal jackknife variance estimator, as implemented in the package \texttt{survival} \(^{38}\), and are presented by dashed lines. They are expected to be conservative as a result of ignoring uncertainty in the obtained weights. \(^{39}\) The confidence intervals for the marginal structural model estimator can be obtained using the nonparametric bootstrap. These are not displayed in Figure 2, as these nearly coincide with confidence intervals for the IPC weighted Aalen-Johansen estimator. Note that these confidence intervals have been computed for a particular prespecified treatment regime and as such do not take into account the uncertainty due to the choice of the optimal treatment regime (see Section 6 for further discussion of this issue).
As indicated in Section 3 there is a risk of being overly optimistic about the optimal DTR whenever its selection and its performance assessment are conducted on the same data set. To provide a more realistic out-of-sample performance assessment we computed using 5-fold cross-validation the benchmark for the optimal dynamic treatment regime given by (4). An important question when combining the IPC weights and cross-validation is whether to re-estimate the weights on the training set for each fold. We have chosen to estimate the weights on the complete data prior to performing cross-validation in order to prevent fitting the propensity score on smaller sample sizes. We can see that the cross-validation estimate of the mortality risk (CV-DTR) results in a reduction of 30-day ICU mortality compared to the current standard of care, however the reduction is slightly smaller than suggested by the in-sample estimate. The difference between the cross-validation benchmark (CV-DTR) and the in-sample evaluation of 30-day ICU mortality under the optimal DTR (optDTR) in Figure 2 gives an indication of over-optimism due to choice and performance assessment of the optimal regime on the same data set. There thus remains a suggestion that the estimated optimal treatment regime is expected to lead to a reduction in ICU mortality relative to the standard of care.

Figure 3 presents the unstabilized IPC weights on a number of selected days that have been used to compute the cumulative incidence curves for the best performing dynamic treatment regime. We present plots of the IPC weights on the linear scale in the upper plot and on the logarithmic scale in the lower plot.

Additionally, we investigated the proportion of stage 2 AKI patients who would receive RRT under the best performing DTR and compared it to current practice. This is relevant from a clinical perspective because potentially we could have found a treatment rule that significantly improves ICU survival, but leads to such a high number of additional patients in need of RRT that it would become logistically impossible to provide treatment to all. In Figure 4 we present the comparison of the proportion of stage 2 AKI patients that would receive RRT under the suggested treatment rule (optDTR), which is estimated by the sum of (unstabilized) IPC weights of the treated compatible patients divided by the sum of IPC weights of all compatible patients at a given time point. This follows from the identity

$$E \left( A^d_k \right) = \frac{E \left( A_k W_{d,k-1} \right)}{E \left( W_{d,k-1} \right)},$$

where $A^d_k$ is the potential outcome for treatment under regime $d$, which can be shown following a similar calculation as in the proof of Lemma 1 (see the Appendix B in the Supplementary Material). Based on this estimate, a smaller proportion of patients would receive RRT under the optimal DTR as compared to the current standard of care. 

![Cumulative incidence curves for the optimal DTR (optDTR), 5-fold cross-validation cumulative incidence curve for the optimal DTR (CV-DTR) and the current standard of care (obs). The dashed lines correspond to 95% confidence intervals for the Aalen-Johansen estimators.](image-url)
6 | DISCUSSION

In this article, we illustrated how routinely collected data can be leveraged to inform decisions about RRT initiation among future AKI patients in an intensive care unit. For this purpose we considered two estimators—an inverse probability weighted nonparametric Aalen-Johansen estimator and an estimator based on semiparametric dynamic-regime marginal structural models—to deal with the problem of competing risks and time-varying treatment and confounding. Furthermore, we proposed a cross-validation benchmark to compare the performance of the in-sample selected optimal DTR against, and hence to assess the amount of over-optimism that characterizes the in-sample performance assessment of this optimal DTR.

To the best of our knowledge van der Laan and Luedtke\textsuperscript{21} were the first ones to use cross-validation in context of dynamic treatment regimes. These authors apply cross-validation in order to perform statistical inference for the mean
outcome under the estimated optimal dynamic treatment regime. In contrast, the purpose of cross-validation in our analysis is to compute a cross-validation benchmark for developing insight into the extent to which a different regime may appear optimal on a different subsample, and to obtain a finite-sample evaluation of the degree of optimism due to estimating the optimal regime and performance under that regime from the same sample. This evaluation is less ambitious, and therefore also simpler, than in van der Laan and Luedtke and Andrews et al, who obtain confidence intervals for the mean outcome under the estimated optimal treatment regime that take into account uncertainty related to the choice of the optimal regime. Such less ambitious evaluations are also common when studying the predictive performance of statistical models, where they have become mainstream. We hope that the simplicity of our proposal will likewise encourage its uptake and thereby prevent the winner’s curse.

One of the limitations of our work is the focus on ICU mortality, instead of all-cause mortality. However, we believe this not to induce major concerns in the considered setting since the chance of dying shortly after ICU discharge can be expected to be low, and therefore ICU death to approximate well all-cause mortality in the short term. However, in a general competing events setting, our proposed approach might pose additional challenges. For instance, if the chance of death after ICU discharge were not low, and therefore ICU death would not approximate well all-cause mortality in the short term, then modeling the impact of a given treatment regime on ICU mortality and on ICU discharge could pose interpretative difficulties. For instance, imagine we would like to compare two treatment regimes \( d^{(1)} \) and \( d^{(2)} \). Furthermore, assume that their impact on mortality is the same, however treatment \( d^{(1)} \) leads to longer stays at the ICU. Now, merely due to longer ICU stays we may observe more ICU deaths under treatment regime \( d^{(1)} \), and therefore treatment regime \( d^{(2)} \) may appear beneficial, despite the fact that the impact of both treatment regimes on death is the same. As a remedy, following the lines of Gran et al, one could allow the treatment regime to alter only some of the event-specific
hazards, while keeping the remaining ones unchanged. This means one could model the counterfactual hazard for the event of interest under the considered treatment regime the same way as we did in our analysis, however fixing the hazard for the competing event at that obtained under the current standard of care, therefore assuming no impact of DTR on the competing event mechanism. In such analysis one would be able to compare different DTRs solely based on their impact on the event of interest mechanism and not through potential impact on the competing event mechanism. However, such analysis would necessitate control for all risk factors of ICU discharge that are also associated with survival, which would be difficult to attain given that these processes are so intimately linked.

Another limitation of our analysis is the fact that we rely on correct specification of parametric models—the propensity score model and possibly marginal structural hazard models. One way to avoid these modeling demands would be to use data-adaptive methods via longitudinal minimum loss-based estimation for the 30-day ICU counterfactual mortality.42,43

A final limitation is that medical doctors may be hesitant to base their decision-making on simplified decision rules, which take into account only a limited set of available information. They may arguably be better served with a decision support model, which computes individual patient’s survival chances in either scenario of treatment initiation or its delay. In future work we therefore plan to focus on sequential counterfactual prediction,44,45 which could be used as a framework for such decision support. This permits to incorporate more flexible data-adaptive algorithms into the procedure and to consider treatment rules that are much more tailored to individual patient characteristics. Since the final decision would still rely on the medical doctor, they could incorporate additional information not available to the decision-support model into their decision-making process, which often takes multiple parameters into account in a complex fashion.

ACKNOWLEDGEMENTS
Paweł Morzywołek, Johan Steen, Wim Van Biesen and Stijn Vansteelandt were supported by the Flemish Research Council (FWO Research Project 3G068619—Grant FWO.OPR.2019.0045.01). The authors would like to thank Bram Gadeyne, Veerle Brams and Christian Danneels for technical support.

DATA AVAILABILITY STATEMENT
The data are not publicly available due to confidentiality restrictions. The R code to reproduce our data analysis on synthetic data from the package ltmle43 is available at https://github.com/pmorzywolek/comprisksDTR. The data set has been modified to include also a competing event (see data preparation part in the R code).

ORCID
Paweł Morzywołek https://orcid.org/0000-0002-2553-6675
Johan Steen https://orcid.org/0000-0002-5968-6401
Wim Van Biesen https://orcid.org/0000-0002-4782-5224
Johan Decruyenaere https://orcid.org/0000-0002-3655-4318
Stijn Vansteelandt https://orcid.org/0000-0002-4207-8733

REFERENCES
1. Ashley EA. The precision medicine initiative: a new national effort. JAMA J Am Med Assoc. 2015;313(21):2119-2120.
2. Kosorok M, Laber EB. Precision medicine. Annu Rev Stat Appl. 2019;6(5):263-286.
3. Gaudry S, Hajage D, Schortgen F, et al. Initiation strategies for renal-replacement therapy in the intensive care unit. N Engl J Med. 2016;375(2):122-133.
4. Zarbock A, Kellum JA, Schmidt C, et al. Effect of early vs delayed initiation of renal replacement therapy on mortality in critically ill patients with acute kidney injury: the Elain randomized clinical trial. JAMA J Am Med Assoc. 2016;315(20):2190-2199.
5. Vanmassenhove J, Kielstein J, Jörres A, Biesen WV. Management of patients at risk of acute kidney injury. Lancet. 2017;389(10084):2139-2151.
6. Chakraborty B, Moodie EE. Statistical Methods for Dynamic Treatment Regimes. New York, NY: Springer; 2013.
7. Tsitsi AA, Davidian M, Holloway ST, Laber EB. Dynamic Treatment Regimes: Statistical Methods for Precision Medicine. New York, NY: Chapman & Hall/CRC Press; 2020.
8. Murphy SA. Optimal dynamic treatment regimes. J Royal Stat Soc Ser B Stat Methodol. 2003;65(2):331-366.
9. Robins JM. Optimal structural nested models for optimal sequential decisions. In: Lin DY, Heagerty PJ, eds. Proceedings of the 2nd Seattle Symposium in Biostatistics: Analysis of Correlated Data; New York, NY: Springer; 2004:189-326.
10. Murphy SA, van der Laan MJ, Robins JM, et al. Marginal mean models for dynamic regimes. J Am Stat Assoc. 2001;96(456):1410-1423.
11. van der Laan MJ, Petersen ML. Causal effect models for realistic individualized treatment and intention to treat rules. *Int J Biostat.* 2007;3(1):1-51.
12. Orellana L, Rotnitzky A, Robins JM. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part I: main content. *Int J Biostat.* 2010;6(2):7.
13. Hernán MA, Robins JM. *Causal Inference - What If.* Boca Raton: Chapman & Hall/CRC; 2020.
14. Aalen OO, Johansen S. An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scand J Stat.* 1978;5(3):141-150.
15. Stone M. Cross-validatory choice and assessment of statistical predictions. *J Royal Stat Soc Ser B Stat Methodol.* 1974;36(2):111-147.
16. 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;15:361-387.
17. Steyerberg EW, Harrell FE, Borsboom GI, Eijkemans MJ, Vergouwe Y, Habbema JDF. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol.* 2001;54(8):774-781.
18. Dudoit S, van der Laan MJ. Asymptotics of cross-validated risk estimation in estimator selection and performance assessment. *Stat Methodol.* 2005;2(2):131-154.
19. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning.* New York: Springer; 2009.
20. Smith GC, Seaman SR, Wood AM, Royston P, White IR. Correcting for optimistic prediction in small data sets. *Am J Epidemiol.* 2014;180(3):318-324.
21. van der Laan MJ, Luedtke AR. Targeted learning of the mean outcome under an optimal dynamic treatment rule. *J Causal Infer.* 2014;3(1):61-95.
22. Neyman J. On the application of probability theory to agricultural experiments. *Stat Sci.* 1990;5(4):465-480.
23. Rubin D. Bayesian inference of causal effects. *Ann Stat.* 1978;6(1):34-58.
24. Robins JM. A new approach to causal inference in mortality studies with a sustained exposure period - Application to control of the healthy worker survivor effect. *Math Model.* 1986;7:1393-1512.
25. Dawid AP, Didelez V. Identifying the consequences of dynamic treatment strategies: a decision-theoretic overview. *Stat Surv.* 2010;4:184-231.
26. Hernán MA. The Hazards of Hazard ratios. *Epidemiology.* 2010;21(1):13-15.
27. Martinussen T, Vansteelandt S, Andersen PK. Subtleties in the interpretation of hazard ratios. *Lifetime Data Anal.* 2020;26:833-855.
28. Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernán MA. A causal framework for classical statistical estimands in failure-time settings with competing events. *Stat Med.* 2020;39(8):1199-1236.
29. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC Med.* 2015;13(1):1-10.
30. Chakraborty B, Lamer EB, Zhao YQ. Inference about the expected performance of a data-driven dynamic treatment regime. *Clin Trials.* 2014;11(4):408-417.
31. Luedtke AR, van der Laan MJ. Optimal dynamic treatment rules. In: van der Laan MJ, Rose S, eds. *Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies,* Cham, Switzerland: Springer; 2018:399-418.
32. Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernán MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. *Int J Biostat.* 2010;6(2):18.
33. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol.* 2016;183(8):758-764.
34. Robins JM. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. *Proc Biopharm Sect Am Stat Assoc.* 1993;24-33.
35. Westling T, van der Laan MJ, Carone M. Correcting an estimator of a multivariate monotone function with isotonic regression. *Electron J Stat.* 2020;14(2):3032-3069.
36. Vanmassenhove J, Steen J, Vansteelandt S, et al. The importance of the urinary output criterion for the detection and prognostic meaning of AKI. *Sci Rep.* 2021;11(1):1-9.
37. 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):1-138.
38. Therneau T. Modeling survival data: extending the Cox model, a package for survival analysis in R; 2020.
39. Rotnitzky A, Li L, Li X. A note on over adjustment in inverse probability weighted estimation. *Biometrika.* 2010;97(4):997-1001.
40. Andrews I, Kitagawa T, McCloskey A. Inference on winners. Working paper, National Bureau of Economic Research; 2019.
41. Gran JM, Lie SA, Øyeflaten I, Borgan Ø, Aalen OO. Causal inference in multi-state models-sickness absence and work for 1145 participants after work rehabilitation. *BMC Public Health.* 2015;15(1):1-16.
42. Petersen ML, Schwab J, Gruber S, Blaser N, Schomaker M, van der Laan MJ. Targeted maximum likelihood estimation for dynamic and static longitudinal marginal structural working models. *J Causal Infer.* 2014;2(2):147-185.
43. Lendle SD, Schwab J, Petersen ML, van der Laan MJ. Itmle: an R package implementing targeted minimum loss-based estimation for longitudinal data. *J Stat Softw.* 2017;81(1):1-21.
44. Wager S, Athey S. Estimation and inference of heterogeneous treatment effects using random forests. *J Am Stat Assoc.* 2018;113(S23):1228-1242.
45. Coston A, Kennedy EH, Chouldechova A. Counterfactual predictions under runtime confounding. *Adv Neural Inf Proces Syst.* 2020;33:4150-4162.

**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Morzywołek P, Steen J, Van Biesen W, Decruyenaere J, Vansteelandt S. On estimation and cross-validation of dynamic treatment regimes with competing risks. *Statistics in Medicine.* 2022;41(26):5258-5275. doi: 10.1002/sim.9568