

# Estimating Treatment Effects for Time-to-Treatment Antibiotic Stewardship in Sepsis

Ruoqi Liu<sup>1</sup>, Katherine H. Buck<sup>2</sup>, Jeffrey M. Caterino<sup>2</sup>, and Ping Zhang<sup>1,3,4,\*</sup>

<sup>1</sup>Department of Computer Science and Engineering. The Ohio State University, 2015 Neil Ave, Columbus OH 43210, USA.

<sup>2</sup>Department of Emergency Medicine. The Ohio State University, 376 W 10th Ave, Columbus OH 43206, USA.

<sup>3</sup>Department of Biomedical Informatics. The Ohio State University, 1800 Cannon Drive, Columbus OH 43210, USA.

<sup>4</sup>Translational Data Analytics institute. The Ohio State University, 1760 Neil Ave, Columbus OH 43210, USA.

\*e-mail: zhang.10631@osu.edu

## ABSTRACT

Sepsis is a life-threatening condition with high in-hospital mortality rate. The timing of antibiotic (ATB) administration poses a critical problem for sepsis management. Existing work studying antibiotic timing either ignores the temporality of the observational data or the heterogeneity of the treatment effects. In this paper, we propose a novel method to estimate Treatment effects for Time-to-Treatment antibiotic stewardship in sepsis (**T4**). **T4** estimates individual treatment effects (ITEs) by recurrently encoding temporal and static variables as potential confounders, and then decoding the outcomes under different treatment sequences. We propose a mini-batch balancing matching that mimics the randomized controlled trial process to adjust the confounding. The model achieves interpretability through a global-level attention mechanism and a variable-level importance examination. Meanwhile, we incorporate **T4** with uncertainty quantification to help prevent overconfident recommendations. We demonstrate that **T4** can identify effective treatment timing with estimated ITEs for antibiotic stewardship on two real-world datasets. Moreover, comprehensive experiments on a synthetic dataset exhibit the outstanding performance of **T4** compared to the state-of-the-art models on ITE estimation.

## Introduction

Sepsis is the body's overwhelming response to infection, which can lead to tissue damage, organ failure, amputations, and death. Sepsis contributes to 6% of hospitalizations and 35% of in-hospital deaths<sup>1</sup>, and costs more than \$27 billion annually in the USA<sup>2</sup>. Based on a recent study on Medicare beneficiaries, approximately 30% of septic patients do not survive for 6 months<sup>3</sup>. Broad-spectrum antibiotics are the first-line medications for sepsis<sup>4,5</sup> because bacterial infection causes most cases<sup>6</sup>.

The current sepsis treatment guideline for antibiotic timing is a one-size-fits-all approach, and when patients with suspected sepsis should receive antibiotics remains controversial<sup>7,8</sup>. The Surviving Sepsis Campaign (SSC) recommends initiating broad-spectrum antibiotics within 1 hour for any patient with suspected sepsis or septic shock<sup>9,10</sup>. While the recommendation is supported by several large observational studies<sup>11,12,13</sup>, there is substantial concern that striving for 1-hour antibiotic delivery for all patients with suspected sepsis may cause serious harm (e.g., antibiotic resistance, *C. difficile* infection)<sup>8,14,15</sup>. Thus determining personalized antibiotic timing at the bedside is urgently needed.

Computational algorithms<sup>16,12,13,17</sup> have been leveraged for examining the optimal antibiotic timing for septic patients using electronic health records (EHRs). EHRs contain irregularly sampled temporal data, including patient's lab test results, vital signs and demographics. The main idea is to estimate the treatment effects of different timings of antibiotics on septic outcomes (e.g., in-hospital mortality). However, most

studies have either ignored the temporality of EHRs<sup>12,13</sup> or the heterogeneity of treatment effects<sup>16,17</sup>. These two issues are crucial for identifying effective and precise therapy for septic patients.

Some other literature<sup>18,19,20,21,22</sup> of treatment design for sepsis treat the problem as an off-policy evaluation using observational data. For example, Komorowski et al.,<sup>21</sup> propose a model to learn treatment policy based on patient trajectories (i.e., states, actions and observations) by optimizing a reward determined by patient survival. However, our method derives optimal treatment options by estimating individual potential outcomes for future timestamps. Our problem setting is more challenging that 1) we need to do counterfactual reasoning based on only observed data and 2) we need to adjust time-varying confounding and estimate unbiased individual causal effects.

In this paper, we study the problem of identifying the most effective timing for antibiotic administration in septic patients using EHR data. As shown in Figure 1, the patient's information are extracted and compiled from EHRs and then used to build the model for antibiotic administration timing recommendation. To address the aforementioned challenges, we propose a novel framework to estimate Trustworthy Treatment effects for Time-to-Treatment antibiotic stewardship in sepsis (**T4**). The proposed model **T4** first estimates individual treatment effects (ITEs) of receiving antibiotics by recurrently encoding temporal and static information obtained before the current timestamp (baseline period), and then decoding the potential outcomes under different treatment sequences after

the current timestamp (follow-up period). We apply balancing matching for each mini-batch via treatment propensity as balancing scores to construct a *pseudo* balanced mini-batch, thus adjust the influence of confounders. We also provide the model interpretability of treatment recommendation by analyzing: 1) the contribution of each timestamp in the baseline to treatment recommendation with attention mechanism and 2) the contribution of each variable to treatment recommendation via variable importance examination that excludes each variable in evaluating the influence on model loss. Meanwhile, we adopt MC Dropout<sup>23</sup> to estimate uncertainty and quantify the confidence behind the ITE estimation and treatment recommendation.

We evaluate the effectiveness of treatment recommendation on two nonoverlapping real-world EHR datasets: Medical Information Mart for Intensive Care version III (MIMIC-III)<sup>24</sup> and AmsterdamUMCdb<sup>25</sup>. The results show that the mortality rate of patients who receive the antibiotics at the time we recommend is notably lower than the patients who do not, indicating that our model offers effective timings of antibiotic administration that help to reduce the mortality rate. We demonstrate the application of our model on time-to-treatment recommendation using a concrete patient example. We also analyze model interpretability by visualizing the global and variable-level contribution to treatment recommendation via a concrete case study. Moreover, we conduct comparison experiments on ITE estimation using a synthetic dataset, and our model outperforms the state-of-the-art ITE estimation methods.

The contributions of this paper include the following:

- We propose an end-to-end treatment timing recommendation framework that seamlessly integrates the treatment effect estimation model, uncertainty quantification, and model interpretability for making transparent treatment recommendation.
- We develop a new ITE estimation method that can model time-varying information and adjust the influence of temporal confounding variables via balancing matching that mimics the randomized controlled trial process.
- We incorporate the ITE estimation with uncertainty quantification and interpretable analysis to achieve reliable treatment recommendation.
- We illustrate the usage of the proposed model in two real-world EHR datasets. The results show that our model can successfully identify effective timing of treatment and thus pave the way for personalized and precision medicine. We further conduct comprehensive comparison experiments on a synthetic dataset for ITE estimation.

## Overall framework

**T4** recurrently encodes the patient’s temporal covariates extracted from the baseline period, then decodes the potential outcomes with different treatment sequences in the follow-up period. Both encoder and decoder are built based on long short-term memory (LSTM)<sup>26</sup>, a deep recurrent neural network that is widely used for modeling time series data. **T4** adjusts the influence of confounders via *balancing matching* (Fig. 3) to generate balanced mini-batches. Each patient in the mini-batch is matched with the corresponding counterfactual outcomes using the observed outcomes of his or her similar patients in different treatment groups. The similarity of patients are estimated using propensity scores<sup>27</sup> of receiving the current treatment. Figure 2 illustrates the framework of the proposed method. The training procedure of **T4** is shown in Algorithm 1.

---

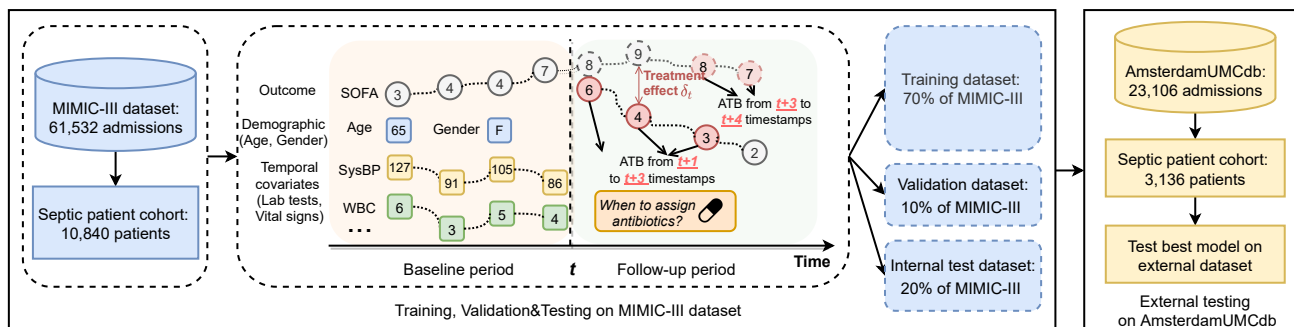
### Algorithm 1: T4 Training Procedure

---

**Input:** Treatment assignments  $A$ , temporal covariates  $X$ , outcomes  $Y$ , static covariates  $d$   
**Output:** Potential outcome  $Y'(A_{t+1,t+\zeta})$ , treatment recommendation  $\pi'$

- 1 Pre-train **T4** as propensity estimator according to Eq. (16);
- 2 **for**  $epoch = 1, \dots, EPOCH$  **do**
- 3     Obtain the similar patients of each mini-batch for balancing matching according to Eq. (14);
- 4     **for**  $t = 1, \dots, current\_timestamp$  **do**
- 5         Obtain the embedding of temporal and static covariates as  $e_t^x \in \mathcal{R}^{K_x^e}$  and  $e^d \in \mathcal{R}^{K_d \times K_d^e}$ ;
- 6         Encode  $e_t^x, e^d, a_t$  with *LSTM* and obtain  $h_t$ ;
- 7         Obtain the attention weight  $\alpha$  according to Eq. (5);
- 8     **end**
- 9     Compute the potential outcome  $y'_t$  according to Eq. (6);
- 10    **for**  $t = current\_timestamp, \dots, current\_timestamp + \zeta$  **do**
- 11         Decode  $e^d, a_t, y'_t$  using *LSTM* and obtain  $h_t$ ;
- 12         Obtain the attention weight  $\beta$  according to Eq. (8);
- 13     **end**
- 14     Compute the potential outcome  $y'_t$  according to Eq. (9);
- 15     Compute the treatment effect  $\delta'$  according to Eq. (10);
- 16     Compute the treatment strategy  $\pi'$  according to Eq. (11);
- 17     Compute outcome prediction loss according to Eq. (17);
- 18 **end**

---



**Figure 1.** Overall data flow of **T4** framework. The data from MIMIC-III is randomly split into training, validation and testing datasets with percentages of 70%, 10% and 20%, respectively. The validation dataset is used to select the best model parameters and the testing dataset is used as an internal evaluation dataset. **T4** framework is used to estimate individual treatment effects for ATB administration timing recommendation. An external dataset obtained from AmsterdamUMCdb is used as an external test set.

## Results

### Datasets

**MIMIC-III.** MIMIC-III<sup>24</sup> is a large, freely-available database comprising de-identified health-related data associated with over 40,000 patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012. It contains patients' demographics, vital signs, lab tests and treatment assignments.

**AmsterdamUMCdb.** AmsterdamUMCdb<sup>25</sup> is the first freely accessible European intensive care database. It is endorsed by the European Society of Intensive Care Medicine (ESICM) and its Data Science Section. It contains de-identified health data related to tens of thousands of intensive care unit admissions, including demographics, vital signs, laboratory tests and medications.

In both datasets, we included adult septic patients fulfilling the international consensus sepsis-3 criteria<sup>28</sup>. We extracted 10,840 patients and 3,136 patients from MIMIC-III and AmsterdamUMCdb after applying exclusion criteria respectively. Here, the causal inference problem we studied is the treatment effects of antibiotic therapy among sepsis patients given the observed confounding variables. There are three essential components that should be identified from the patient data: (1) **Treatments**: we consider multiple kinds of antibiotic therapy during the ICU stays. At each timestamp, a binary treatment indicator will denote whether the patient is assigned with antibiotics or not; (2) **Confounders**: we obtain 22 temporal covariates (i.e., vital signs: temperature, heart rate, etc.; lab tests: potassium, sodium, etc.) and 4 static covariates (i.e., age, gender, etc.) as potential confounders; (3) **Outcomes**: we compile the 24-hour Sepsis-related Organ Failure Assessment (SOFA)<sup>29</sup> score as the primary outcome, which is computed based on the degree of dysfunction of six organ systems. The definition of sepsis-3 patient cohort is in study design of Methods, the list of antibiotics is in Supplementary Table 1, the list of patient's covariates is in Supplementary Table 3 and

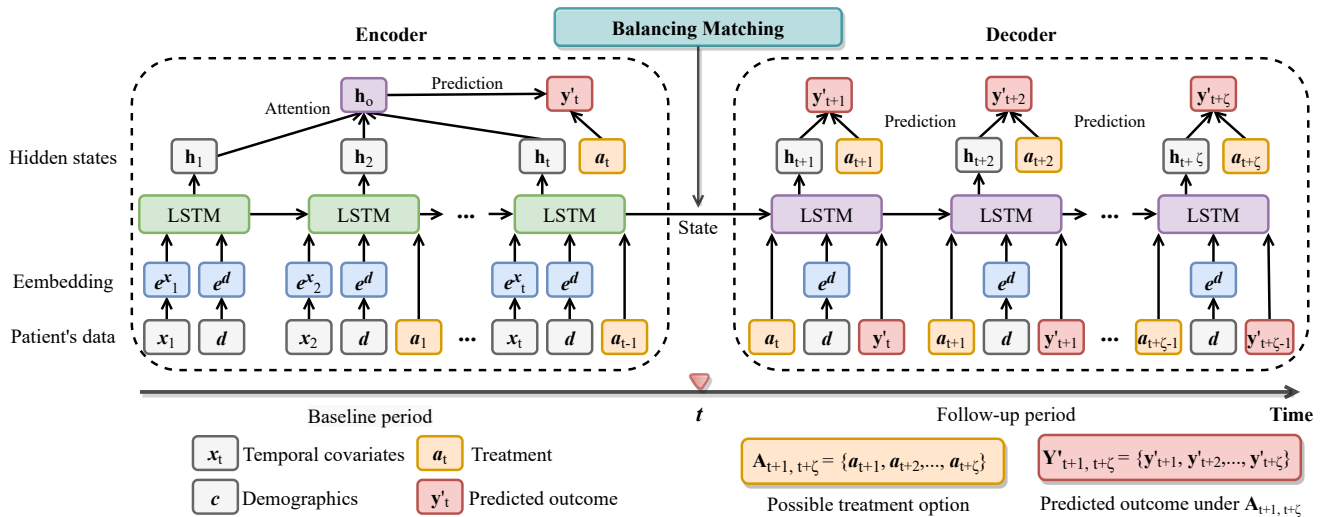
computation of SOFA scores is in Supplementary Table 2.

### Model performance

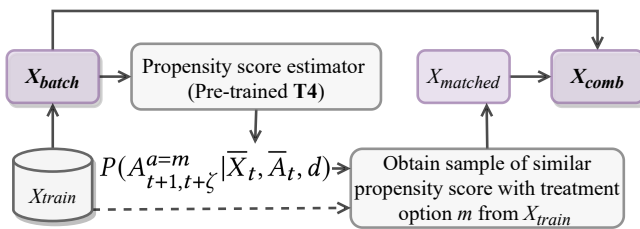
#### Population level analysis

As we have no access to the counterfactual outcomes in the real-world dataset, so that we are not able to directly evaluate the model performance in terms of counterfactual prediction. Thus, we evaluate the model performance by comparing the treatment effects of recommended timing of administration (determined by estimated ITE according to Eq. (11)) and the observed timing of administration on patients' mortality rate. Specifically, we first obtain a target group of patients whose observed timing of ATB administration is different from the model recommendation. Then we derive a compared group to the target group by involving the most similar patients whose observed timing of ATB administration matches the model recommendation. We use the variables obtained from the baseline period (time window before the follow-up period) to perform patient similarity. We use Euclidean distance as the similarity measurement. Finally we compare the mortality rate within these two groups and expect that the mortality rate of patients whose observed treatments match our recommendation would be much lower than the patients whose observed treatments are different from recommendation.

As shown in Fig. 4, we compute and compare 30 days mortality rate and 60 days mortality rate on two datasets, respectively. The black dashed lines in the two plots denote the average mortality rate among the population, which are the baselines for the comparison. We find that the mortality rate of patients who receive treatments at different timestamps as our recommendation is higher than the average mortality rate baseline, while the mortality rate of patients who receive the treatments at the same timestamps as our recommendation is lower than the baseline. We evaluate the model concerning different lengths of the follow-up period (i.e.,  $\zeta \in \{3, 4, 5, 6, 7\}$ ) and the mortality rates for patients with the same treatments are consistently lower than the mortality rates for patients



**Figure 2.** The framework of **T4**. **T4** consists of three main components: 1) the encoder network recurrently encodes the patient’s baseline information, including temporal and static covariates, and treatment assignments via the LSTM network; 2) the decoder network is initialized with encoder outputs and predicts the outcomes under different treatment sequences; 3) the balancing matching constructs balanced mini-batches via propensity as balancing scores during the training process. The details of balancing matching are shown in Fig. 3.



**Figure 3.** Illustration of the balancing matching. During the training process, each patient in the batch is matched with the corresponding counterfactual outcomes using the observed outcomes of his/her nearest neighbors in other treatment groups in the training data. A pre-trained **T4** is used to estimate the propensity scores for computing the patient’s distance. The matched batch and original batch are combined together for the training.

with different treatments. Results show that our model recommends effective treatment strategies (reflecting on lower mortality rate), and provides potential clinical insights for doctors to decide the timing of antibiotic administration for the septic patients. We also observe that model performs consistently on the external testing set from AmsterdamUMCdb, which demonstrates the robustness of our model when applying to a different dataset with different feature distribution.

### Individual level analysis

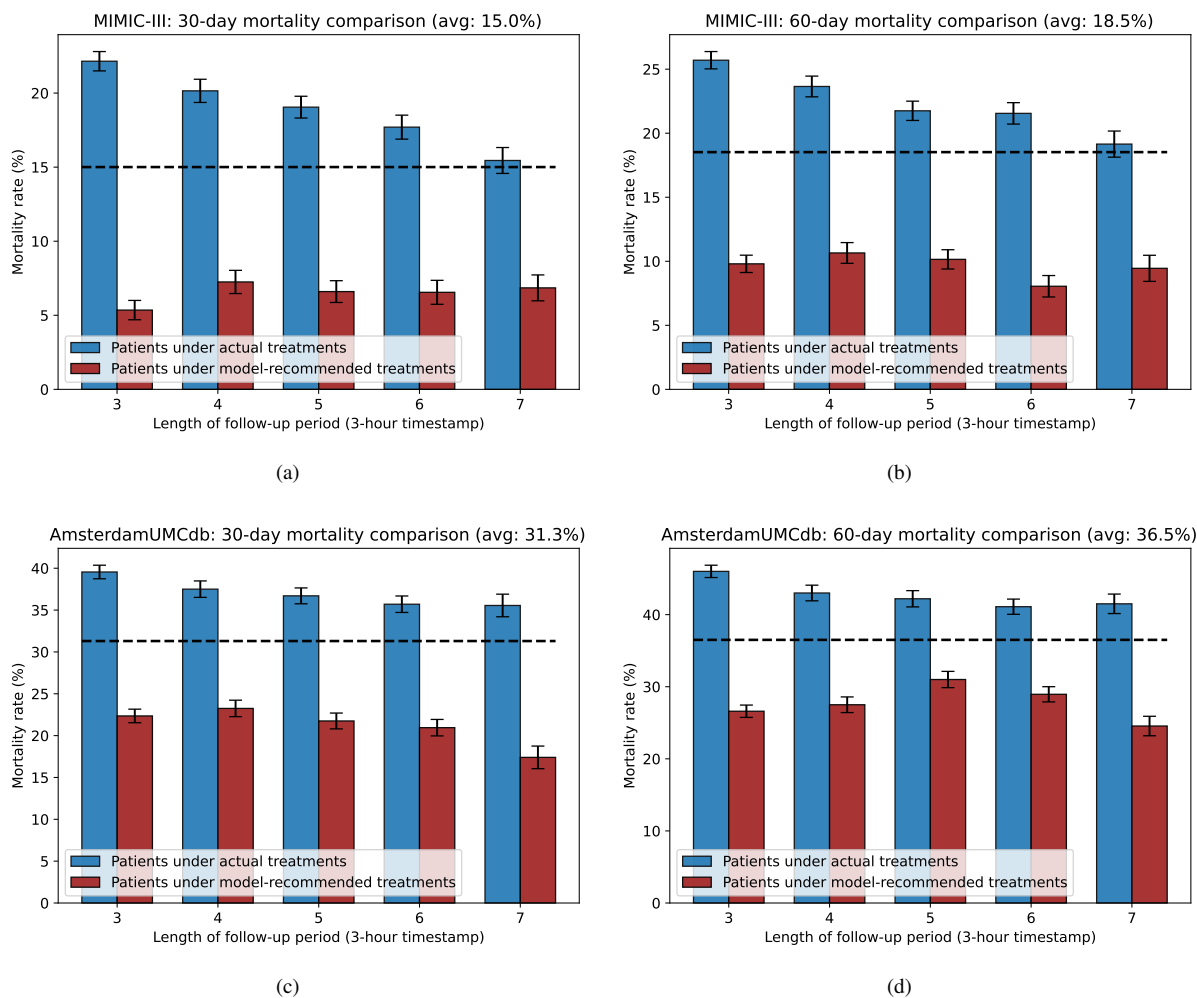
To further demonstrate how our model recommends antibiotics based on the estimated ITEs with uncertainty quantification, we utilize a real-world patient case. As shown in Fig. 5(a), we use the predicted ITEs (red line) equipped with

uncertainty estimates (red shadowed area) as the criteria to recommend the timing of antibiotic administration at each timestamp. Specifically, at each timestamp, we recommend the antibiotics if the upper bound of the predicted ITE is lower than zero, and we will not recommend if the lower bound of the predicted ITE is higher than zero. Here, zero serves as the baseline, denoting no difference between having a treatment and not. In this example, the upper bound of the predicted ITE is lower than zero for all the timestamps in the follow-up period. Thus, the optimal antibiotic recommendation provided by our model is to continuously take antibiotics during the follow-up period.

We also compare true outcomes under the actually received antibiotics and predicted outcomes under the recommended antibiotics. Here, we use the SOFA score as the outcomes, where higher values are associated with worse status and higher mortality rate<sup>30,31</sup>. From Fig. 5(a), we observe that the patient receives antibiotics since the 4-*th* timestamp, which results in large SOFA scores during the follow-up period. Conversely, our model recommends the patient to take antibiotics earlier (from the 1-*th* to the 6-*th* timestamp), and the predicted SOFA scores under recommended timing of antibiotics are much lower than the true SOFA scores. The results demonstrate that our model can identify effective timing of antibiotic administration for septic patients to help improve their disease condition and reduce the mortality rate.

### A Case Study for Model Interpretability

We demonstrate the model interpretability of treatment recommendation using a concrete case study. We visualize both global and variable level contribution in Fig. 5(b). We also



**Figure 4.** Mortality rate comparison of two datasets. Within the bar-chart, the error bars denote 95% confidence interval with  $n=30$  bootstrap samples. Blue and red bars denote patients under actual treatments and patients under model-recommended treatments, respectively. The total mortality rate of two groups of patients is plotted using the black dashed line, which serves as the baseline. Fig. 4(a) and 4(b) are mortality comparison on MIMIC-III dataset. Fig. 4(c) and Fig. 4(d) are mortality comparison on AmsterdamUMCdb dataset.

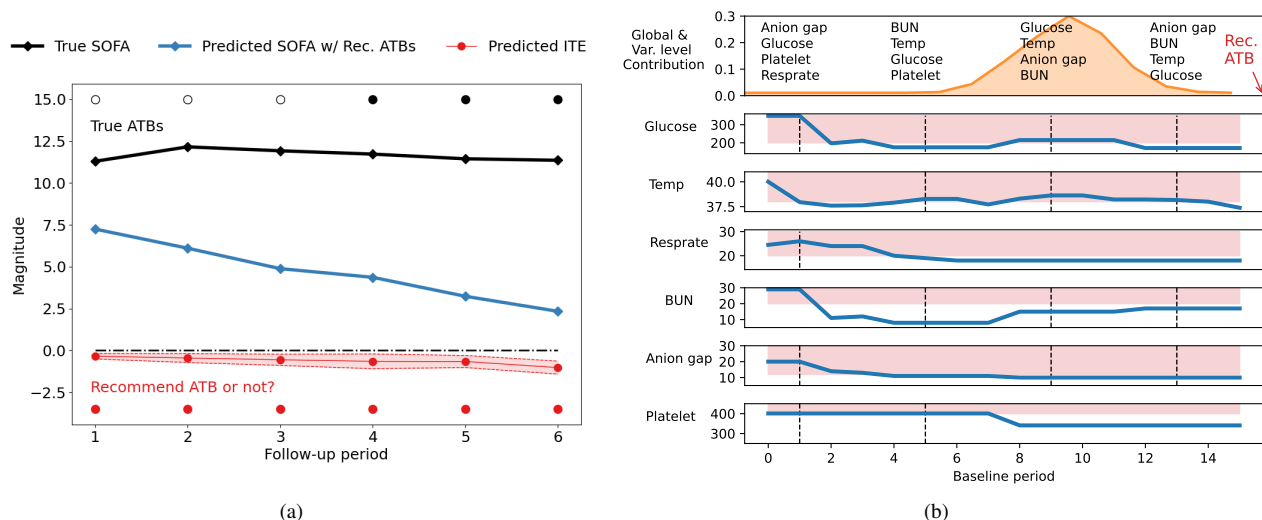
plot in this figure the dynamics of each important variable. Here, the vital signs include temperature, and respiratory rate (resprate); the lab tests include glucose, blood urea nitrogen (BUN), anion gap, and platelet. We use the dashed black lines to denote the timestamps with high contribution to the treatment recommendation. We observe that the values of most variables are within or close to the abnormal range at those high contribution timestamps. Our model recommends the patient to take antibiotics at the end of the baseline period with regards to these warning signals. Taking the glucose as an example, the normal range should be lower than 140 mg/dL, and a reading of more than 200 mg/dL indicates diabetes<sup>1</sup>. We observe that the values of glucose are far from the nor-

mal range, which maintain above 200 mg/dL and even reach 300 mg/dL at early stage. This case study shows that our model achieves the transparent treatment recommendation via visualizing important timestamps and variables contributing to the recommendation, paving the way for interpretable and precise treatment recommendation.

## Discussion

In this study we propose a novel framework to estimate treatment effects for treatment recommendation. The proposed model **T4** first estimates ITEs by recurrently encoding historical temporal patient information and static information, and then decoding the potential outcomes under different treatment sequences. We apply balancing matching for each mini-batch to construct a balanced mini-batch and adjust the influ-

<sup>1</sup><https://www.mayoclinic.org/diseases-conditions/diabetes/diagnosis-treatment/drc-20371451>



**Figure 5.** The case study of a patient for treatment recommendation and model interpretability. In Fig. 5(a), the predicted values of ITEs (red line with shadowed area denoting the uncertainty estimates) for antibiotic (ATB) recommendation. An ATB will be recommended to the patient if the upper bound of predicted ITE is lower than zero and will not be recommended if the lower bound of predicted ITE is higher than zero, where zero is the threshold for determining whether to recommend ATBs. In Fig. 5(b), the most important global baseline timestamps (orange area) and variables contributing to the treatment recommendation are denoted at the top subplot. An ATB is recommended at the end of the baseline period.

ence of confounders. We also provide the model interpretability of treatment recommendation to analyze both global-level and variable-level contribution via attention mechanism and variable importance analysis, respectively. Meanwhile, the model uncertainty quantification helps to avoid overconfident treatment recommendation. We illustrate the usage of the proposed model in two real-world EHR datasets, showing that our model can successfully identify effective treatment strategies and thus pave the way for personalized and precision medicine.

## Individual treatment effect estimation

### Comparison experiments

Several studies have proposed estimating individual treatment effects (ITEs) using causal inference techniques on observational data, such as matching-based methods (e.g., propensity score matching<sup>27</sup>), forest-based methods (e.g., causal forest<sup>32</sup>), or representation learning-based methods (e.g., counterfactual regression<sup>33</sup>). These methods are mainly designed for static data and are not satisfied for estimating ITEs on EHRs, while the proposed **T4** fully considers the time-varying information and adjust the temporal confounding via balancing matching operation.

To illustrate the model performance on ITE estimation and treatment recommendation, we design experiments on a synthetic dataset. We simulate 5000 patients with 50 timestamps, 20 temporal covariates and 5 static covariates. We use the first 40 timestamps as baseline period and the remaining as follow-up period. We use Precision in Estimation of Heterogeneous Effect (PEHE) and the error of Average Treat-

ment Effect ( $\epsilon$ AATE) to evaluate the model performance. We conduct comparison experiments against the state-of-the-art methods of ITE estimation: (1) **Classical methods:** Linear Regression (LR)<sup>34</sup>, Random Forest (RF)<sup>35</sup> and support vector machine (SVM)<sup>36</sup>; (2) **Forest-based methods:** Causal Forest (CF)<sup>32</sup> and Bayesian Additive Regression Trees (BART)<sup>37</sup>; (3) **Representation learning-based methods:** Counterfactual Regression (CFR)<sup>33</sup>, GANITE<sup>38</sup> and Dragonnet<sup>39</sup>. (4) **Time-varying based methods:** Recurrent Marginal structural Network (RMSM)<sup>40</sup> and Counterfactual Recurrent Network (CRN)<sup>41</sup>.

Results in Supplementary Table 8 show that our model outperforms the state-of-the-art ITE estimation methods.

### Ablation study on balancing matching

We evaluate the influence of different percentages of balancing matching samples on the model performance. We vary the percentages from 0 to 100% and show the performance change in Supplementary Fig. 8. We observe that the error of ITE estimation in terms of both PEHE and  $\epsilon$ AATE decreases as including more matching samples in a mini-batch during the training process. Specifically, performance highly increases when only a small number of matching samples (around 20%) are provided, and the curve tends to gently slope downward as the percentage of matching samples exceeds 30%. The results demonstrate that the balancing matching improves the model performance on ITE estimation by largely decreasing the estimation error.

## Controversies in antibiotics for septic patients

The timing of antibiotic treatment is controversial. While many studies suggest early antibiotic regimens for any patients with suspected sepsis or septic shock, there is substantial concern that early antibiotic assignments may cause serious harm including higher mortality rates<sup>7</sup>. Several recent cohort studies and RCTs<sup>8,15,42,43,44</sup> suggest that the timing of antibiotic treatment should depend on the severity of illness (i.e., sepsis, severe sepsis or septic shock) and the likelihood of true infection. They point out that immediate antibiotic regimens benefit patients with severe illness (e.g., septic shock) while in less critically ill patients, immediate antibiotic regimens may lead to overprescribing and potential harm. For example, a recent study<sup>45</sup> shows that an overdose of antibiotics is associated with a 20% increase in the odds of death in patients who received adequate therapy. According to the study<sup>45</sup>, the morbidity of overdose antibiotics may be more obvious in less critically ill patients compared to the patients with septic shock as the morbidity of other acute severe illness surpasses the possible morbidity comes from an antibiotic overdose. There are also some general explanations for the association between antibiotic overdose and higher mortality. Besides the antibiotic resistance, antibiotics themselves also cause harm (e.g., organ injury, mitochondrial dysfunction, the impact on the microbiome, and overgrowth by fungi and *Clostridium difficile*)<sup>46</sup>. Specifically, the study<sup>45</sup> shows that unnecessarily broad empiric therapy was associated with a 26% increased risk of *Clostridium difficile* infection. Thus determining personalized antibiotic timing at the bedside is urgently needed.

## Limitation of public clinical data

### Unavailability of counterfactual outcomes

As the ground truth counterfactual outcomes are not available in the real-world data, we evaluate the effectiveness of our model in two ways: 1) mortality rate comparison as shown in Fig. 4 and 2) factual prediction on SOFA scores as shown in Supplementary Table 7.

### Type of antibiotic treatment

The choice of the type of antibiotic treatment is based on suspected infection sites according to empirical antibiotic studies and guidelines<sup>47,48</sup>. However, the sites of suspected infection are not available in our public clinical datasets (MIMIC-III and AmsterdamUMCdb), especially during the first 48 hours since ICU admission as the determination of true infection sites is complicated and takes time to obtain the results. In the future, we will incorporate suspected infection sites to provide recommendations for a specific type of antibiotic, orthogonal etc.

### Blood cultures

Blood cultures are deemed as the gold standard for antibiotic treatment regimens (e.g., initiation or de-escalation of antibiotics). However, in our public datasets, most blood cultures

are taken 12 hours within the ICU admission and usually the results are available after 2-3 days. According to the proposed framework as illustrated in 2, we only leverage the first 48 hours data since ICU admission and therefore the results of blood cultures may not be available in this period.

In future work, we can develop a more practical and precise antibiotic recommendation system that combines the model's general recommendations with real patient conditions (e.g., suspected infection sites, blood cultures and other concomitant therapies) if available.

In summary, we propose a computational framework to estimate treatment effects for time-to-treatment antibiotic stewardship in sepsis. Experiments on both real-world and synthetic datasets have demonstrated the superiority of the model performance compared to the state-of-the-art models, while also providing interpretability of the treatment recommendation.

## Methods

In this section, we first introduce the study design, then we present the proposed model for estimating treatment effects.

### Study design

We evaluate the proposed treatment recommendation framework through a retrospective study on two large real-world EHR datasets (MIMIC-III and AmsterdamUMCdb) with recorded patients' demographics, vital signs, lab tests, medications and diagnosis.

### Definition of Sepsis in two datasets

We obtain the septic patients according to the recent sepsis-3 criteria<sup>28</sup> with respect to **1**)  $t_{susp}$ : time of clinical suspicion of infection (i.e., earlier timestamp of antibiotics and blood cultures within a specified duration) and **2**)  $t_{SOFA}$ : two-point deterioration in SOFA score<sup>29</sup> within a 6-hour period. The patient is diagnosed with sepsis when these two events happen close to each other. Specifically,  $t_{SOFA}$  happens 24 hours earlier than  $t_{susp}$  or 12 hours later than  $t_{susp}$ . We exclude patients whose age is under 18 years old at the time of ICU admission. We also exclude patients whose ICU stay is less than 9 hours or longer than 20 days.

We extract 22 temporal covariates (i.e., vital signs and lab tests) and 4 static covariates (Supplementary Table 3). We encode the patients' time series data into discrete 3-hour time steps. The covariates with multiple records within a single time step are averaged. The missing data are imputed using the values obtained from the last time step.

### Preliminary

We extract patient information from longitudinal observational data. For each patient, let  $\bar{A}_T = \{a_1, a_2, \dots, a_T\} \in \mathcal{N}^T$  be the treatment assignments with  $a_t = 1$  if the patient receive the treatment at  $t$ -th timestamp and  $a_t = 0$  otherwise. Let  $\bar{X}_T = \{x_1, x_2, \dots, x_T\} \in \mathcal{R}^{T \times K_x}$  be the temporal covariates and  $\bar{Y}_T = \{y_1, y_2, \dots, y_T\} \in \mathcal{R}^T$  be the outcomes of  $T$  timestamps.

The patient has static covariates  $d \in \mathcal{R}^{K_d}$ , such as gender and age. The observational data for the patient can be represented as  $\mathcal{D} = \{A_T, \bar{X}_T, d, \bar{Y}_T\}$ .

Our goal is to estimate the treatment effects with temporal and static covariates via predicting the potential outcomes under different treatment sequences. We adopt the potential outcomes framework<sup>49,50</sup> to examine the causal effects under the treatments. The potential outcome is the outcome that would have been observed if the patient had received treatment. We extend the potential outcome framework in our application scenario. Given the observational data up to  $t$ -th timestamp and treatment assignments  $\mathbf{A}_{t+1,t+\zeta}$ , the patient has potential outcomes  $\mathbb{E}[Y(\mathbf{A}_{t+1,t+\zeta})] = \{\mathbb{E}[Y(\mathbf{A}_j | \bar{X}_t, \bar{A}_t, d)]\}_{j=t+1}^{j=t+\zeta}$  in the following  $\zeta$  time period. Specifically,  $\mathbf{A}_{t+1,t+\zeta} = \{a_{t+1}, a_{t+2}, \dots, a_{t+\zeta}\}$  denotes any treatment assignments from  $t+1$  to  $t+\zeta$  timestamp. For each timestamp  $j$ , there are two potential outcomes  $\mathbb{E}[Y(\mathbf{A}_j^{a=1} | \bar{X}_t, \bar{A}_t, d)]$ ,  $\mathbb{E}[Y(\mathbf{A}_j^{a=0} | \bar{X}_t, \bar{A}_t, d)]$ , which are corresponding to different treatment assignments.

To estimate the treatment effect of a given treatment assignment during the following  $\zeta$  time period, we define the individual treatment effect (ITE),  $\delta_j$  on  $(t+j)$ -th timestamp as follows,

$$\delta_j = \mathbb{E}[Y(\mathbf{A}_j^{a=1} | \bar{X}_t, \bar{A}_t, \hat{A}_{(t+1,t+j-1)}, d)] - \mathbb{E}[Y(\mathbf{A}_j^{a=0} | \bar{X}_t, \bar{A}_t, \hat{A}_{(t+1,t+j-1)}, d)] \quad (1)$$

where  $\hat{A}_{(t+1,t+j-1)}$  is the learned optimal treatments between timestamp  $t+1$  and  $t+j-1$ . The treatment effects of  $\zeta$  time period is  $\Delta = [\delta_{t+1}, \dots, \delta_{t+\zeta}]$ . In this paper, we use the Sepsis-related Organ Failure Assessment (SOFA) scores<sup>29</sup> (i.e., range from 0 to 24, larger values associated with severe disease status and higher mortality) as outcomes. The computation of SOFA scores can be found in Supplementary Table 2. The recommended treatment assignments are as follows,

$$\pi^* = [\mathbb{1}(\delta_{t+1} < 0), \dots, \mathbb{1}(\delta_{t+\zeta} < 0)] \quad (2)$$

where  $\mathbb{1}(\cdot)$  equals to 1 if the inside expression is true otherwise 0.

## Assumptions

Our ITE estimation is based on the standard causal assumptions<sup>51,52</sup> as follow,

**Assumption 1 (Consistency)** *The potential outcome under treatment history  $\bar{A}_t$  equals to the observed outcome if the actual treatments history is  $\bar{A}_t$ .*

**Assumption 2 (Positivity)** *Given the observational data of the history, if the the probability  $P(A_t = 1 | \bar{X}_t, \bar{A}_{t-1}, d) \neq 0$ , then the probability of receiving treatment 0 or 1 is positive, i.e.,  $0 < P(A_t = 1 | \bar{X}_t, \bar{A}_{t-1}, d) < 1$ , for all  $\bar{A}_t$ .*

**Assumption 3 (Sequential Strong Ignorability)** *Given the observational data of the history, the treatment assigned for time  $t$  is independent of the potential outcome of time  $t$ , i.e.,  $\mathbf{Y}(A_t) \perp\!\!\!\perp A_t | \bar{X}_t, \bar{A}_{t-1}, d$ , for all treatment sequences  $\mathbf{A}_t$ .*

## Treatment effect estimation with T4

**T4** recurrently encodes the patient's temporal covariates via Long short-term memory (LSTM)<sup>26</sup>, then decodes the potential outcomes with different treatment sequences. **T4** adjusts the influence of confounders via balancing matching to generate balanced mini-batches. Figure 2 illustrates the framework of the proposed method.

### Encoder for baseline period

We convert the initial high-dimensional covariates  $x_t \in \mathcal{R}^{K_x}$  into a lower dimensional and continuous data embedding  $e_t^x \in \mathcal{R}^{K_x^e}$  as,

$$e_t^x = W_e x_t + b_e \quad (3)$$

where  $W_e \in \mathcal{R}^{K_x^e \times K_x}$  is the weight matrix,  $b_e \in \mathcal{R}^{K_x^e}$  is the bias vector and  $K_x^e$  is the dimension of the embedded temporal vectors. That is, we have embedding of temporal covariates  $E^x = \{e_1^x, e_2^x, \dots, e_t^x\} \in \mathcal{R}^{t \times K_x^e}$ . Similarly, we convert the static covariates (demographics) into embedding as  $E^d \in \mathcal{R}^{K_d \times K_d^e}$ , where  $K_d$  is the number of static covariates and  $K_d^e$  is the dimension of embedded static vectors.

Given the embedding of temporal and static covariates, and the treatment assignments at each timestamp, the encoder builds upon the LSTM as follows,

$$h_1, h_2, \dots, h_t = LSTM([e^{x_1}, e_d], [e^{x_2}, e_d, a_1], \dots, [e^{x_t}, e_d, a_{t-1}]) \quad (4)$$

where  $h_t \in \mathcal{R}^{K_h}$  is the hidden state at  $t$ -th timestamp and  $K_h$  is the dimension of hidden vectors. The last hidden state  $h_t$  is used to initialize the decoder. We aggregated all the hidden states via attention mechanism for automatically focusing on important historical timestamps. We calculate the attention weight  $\alpha_{t,s}$  using a method that concatenates each previous hidden state  $h_s$  with the current state  $h_t$ , and the product of two states. That is,

$$\alpha_{t,s} = \text{score}(h_t, h_s) = \Phi(W_\alpha^\top [h_t, h_s, h_t \odot h_s]) \quad (5)$$

$$\alpha_t = \text{softmax}(\alpha_{t,1}, \alpha_{t,2}, \dots, \alpha_{t,t-1})$$

where  $\Phi$  is hyperbolic tangent function,  $W_\alpha \in \mathbb{R}^{3K_h}$  is learnable parameter matrix. Using the generated attention energies, we calculate the context vector  $h_o$  for each patient up to  $t$  time stamp as  $h_o = \sum_{s=1}^{t-1} \alpha_{t,s} h_s$ .

We predict the potential outcome  $y_t$  using the attentively aggregated vector  $h_o$ , current hidden state  $h_t$  and the treatment  $a_t$ . The prediction serves as the input to the initial state of the decoder,

$$y_t' = W_p([h_o, h_t, a_t]) + b_p \quad (6)$$

where  $W_p \in \mathcal{R}^{K_y \times (2K_h+1)}$  and  $b_p \in \mathcal{R}^{K_y}$  are parameters to learn.



### Decoder for follow-up period

Initializing with the last hidden state of the encoder and true/predicted outcomes, the decoder recurrently predicts the potential outcome at each timestamp with different treatment sequences. We obtain the hidden states of decoder as,

$$h_{t+1}, \dots, h_{t+\zeta} = LSTM([a_t, e^d, y_t], \dots, [a_{t+\zeta-1}, e^d, y_{t+\zeta-1}]) \quad (7)$$

We integrate the encoder outputs and current hidden state of decoder via an attention layer. We generate the aggregated context vector  $c_{t+j}$  at each timestamp as,

$$\begin{aligned} \beta_{t+j,u} &= \Phi(W_\beta^\top [h_{t+j}, h_u]) \\ \beta_{t+j,u} &= \text{softmax}(\beta_{t+j,1}, \beta_{t+j,2}, \dots, \beta_{t+j,t}) \\ c_{t+j} &= \sum_{u=1}^t \beta_{t+j,u} h_u \end{aligned} \quad (8)$$

where  $\Phi$  is hyperbolic tangent function,  $\{h_u\}_{u=1}^t$  are the encoder outputs and  $W_\beta \in \mathcal{R}^{2K_h}$  is the parameter to learn.

We predict the outcomes  $y'_{t+j}$  by combining the learned context vector with current treatment as

$$y'_{t+j} = W_q [h_{t+j}, c_{t+j}, a_{t+j}] + b_q \quad \text{for } j = 1, 2, \dots, \zeta \quad (9)$$

where  $W_q \in \mathcal{R}^{K_y \times (2K_h+1)}$  and  $b_q \in \mathcal{R}^{K_y}$  are parameters to learn. During the training, we use teach forcing technique with ratio equals to 0.5 to train the model with ground truth treatments and outcomes. In the inference/testing, we feed the decoder's predictions (both outcomes and treatments) back to itself for each step. The current predictions are based on the previous predictions during the inference, which is consistent to the practical application scenario.

We identify the optimal treatment sequence via a greedy-style strategy instead of checking every possible treatment trajectory. We select the best treatment option according to the predicted outcomes at each step and use it for next prediction. Compared to permutation of all possible combination (up to  $2^\zeta$ ) of treatments, our strategy is more time efficient with the increasing of  $\zeta$ .

Then we can compute the treatment effect for  $(t+j)$ -th timestamp using Eq.(1) as,

$$\delta'_j = y'^{a=1}_{t+j} - y'^{a=0}_{t+j} \quad \text{for } j = 1, 2, \dots, \zeta \quad (10)$$

where  $y'^{a=1}_{t+j}$  is the predicted outcome when receiving the treatment at  $(t+j)$ -th timestamp, and  $y'^{a=0}_{t+j}$  is the predicted outcome when not receiving the treatment. Thus we determine the optimal treatment assignments among all  $\zeta$  time period using Eq.(2) as,

$$\pi' = [\mathbb{1}(\delta'_{t+1} < 0), \dots, \mathbb{1}(\delta'_{t+\zeta} < 0)] \quad (11)$$

### Balancing matching

To adjust the inherent treatment selection bias in the data, we adopt the idea of balancing scores<sup>53</sup> to construct *pseudo* mini-batches that mimic the corresponding randomized controlled

trial (RCT) process (i.e., the treatment groups are randomly split and the patient distribution in each group is balanced.) We illustrate the process of balancing matching in Fig. 3. Specifically, we match, for each patient in the original mini-batch, the unobserved counterfactual outcomes (i.e., the potential outcomes under other possible treatment options except the observed one), with the observed outcomes of nearest neighbors in the training data. There are several methods to obtain the nearest neighbor by computing the distance among individuals. Here, we estimate the distance via the propensity score<sup>27</sup>, which is defined as the conditional probability receiving the treatments  $m_*$  given historical information up to the current timestamp:

$$PS^{m_*} = P(A_{t+1,t+\zeta}^{a=m_*} | \bar{X}_t, \bar{A}_t, d) \quad (12)$$

Here,  $m_*$  denote a possible treatment sequence during  $\zeta$ . We use a pre-trained **T4** as a propensity score estimator to calculate the propensity scores for each patient in the training set.

The distance between patient  $i$  with treatments  $m_*$ , and the patient  $j$  with treatments  $n_*$  is defined using absolute distance as,

$$\Delta_n(i, j) = |PS_i^{m_*} - PS_j^{n_*}| \quad (13)$$

where  $PS_i^{m_*}$  and  $PS_j^{n_*}$  denote the estimated propensity scores for two respective patients. We then obtain the nearest neighbours of patient  $i$  in treatment group  $n_*$  as,

$$\kappa_n(i) = \arg \min_j \Delta_n(i, j) \quad (14)$$

Finally the matched mini-batch is combined with the original mini-batch as a whole for the following training process.

### Objective function

We first pre-train our model **T4** to estimate the propensity scores for balancing matching. We obtain the treatment prediction using a linear layer and sigmoid function as,

$$a'_{t+j} = \text{sigmoid}(W_a^\top [h_{t+j}, c_{t+j}] + b_a) \quad (15)$$

where  $W_a \in \mathcal{R}^{2K_h}$  and  $b_a \in \mathcal{R}$  are parameters to learn. We use cross-entropy loss for the treatment prediction as,

$$\mathcal{L}_a = -\frac{1}{N} \frac{1}{\zeta} \sum_{i=1}^N \sum_{j=1}^{\zeta} (a'_{t+j} \log a^i_{t+j} + (1 - a'_{t+j}) \log (1 - a^i_{t+j})) \quad (16)$$

The training objective function for the outcome prediction is the mean squared error between the predicted potential outcomes and factual outcomes as,

$$\mathcal{L}_y = \frac{1}{N} \sum_{i=1}^N \frac{1}{\zeta} \sum_{j=1}^{\zeta} (y'^i_{t+j} - y^i_{t+j})^2 \quad (17)$$

The overall training procedure of **T4** is demonstrated in 1.

## Model interpretability

Interpretability is a very desiring property in treatment effect estimation and treatment timing recommendation problems. In this paper, we realize the interpretability of treatment recommendation by analyzing both global-level contribution and variable-level contribution.

### Global-level contribution

The global-level contribution is the contribution of each timestamp in the baseline period to the treatment recommendation given in the follow-up period. The outputs of the encoder are sent to the decoder and integrated together with the hidden states of decoder through the attention layer. We obtain the learned attention weights  $\beta_{t+j,u}$  as the contribution of  $u$ -th timestamp to the treatment recommendation given at  $(t+j)$ -th timestamp according to Eq. 9.

### Variable-level contribution

Each timestamp contains a number of temporal variables (e.g., lab tests, vital signs, etc.), and based solely on the contribution at the global level, we are unable to identify the impact of each individual variable. We then examine the contribution of each variable via a variable importance analysis. Specifically, given the temporal covariates  $x_u$ , we first predict outcomes  $y'_{x_u-i}$  when excluding all the information from the  $i$ -th dimension of  $x_t$ . Here, we mask the corresponding information by replacing them with the mean value of  $i$ -th variable across in the dataset. We compute the prediction loss  $\mathcal{L}_y(y'_{x_u-i}, y)$  using Eq. (17) except that the predicted outcomes are replaced with  $y'_{x_u-i}$ . Finally, the contribution of each variable  $i$  at  $u$ -th timestamp is computed as,

$$\begin{aligned} \omega_{u,i} &= \mathcal{L}_y(y'_{x_u-i}, y) - \mathcal{L}_y(y', y) \\ \omega_{u,i} &= \text{softmax}(\omega_{u,1}, \omega_{u,2}, \dots, \omega_{u,K_x}) \end{aligned} \quad (18)$$

where  $\mathcal{L}_y(y', y)$  is the prediction loss when all features of  $x_t$  are included in the loss computation. We multiply the global-level contribution and the variable-level contribution ( $\beta_{t+j,u} \omega_{u,i}$ ) to obtain the contribution of each variable at each timestamp.

## Uncertainty quantification

The uncertainty quantification of the estimated treatment effects is also important for treatment recommendation. In this paper, we adopt MC Dropout<sup>23,54</sup> to quantify the model uncertainty by applying dropout during both training and testing process. Specifically, with the dropout enabled during the testing, the model generates a different output every forward pass for the same input. Suppose we have  $K$  iterations, and for iteration  $k$ , we obtain the estimated effect  $\delta'_{j,k}$ . Then the model uncertainty  $\eta(\delta'_j)$  is computed as,

$$\eta(\delta'_j) = \frac{1}{K} \sum_{k=1}^K (\delta'_{j,k})^2 - \left( \frac{1}{K} \sum_{k=1}^K (\delta'_{j,k}) \right)^2 \quad (19)$$

In this way, each ITE  $\delta'_j$  is equipped with according uncertainty estimates  $\eta(\delta'_j)$ . We use the estimated uncertainty for 1) quantifying the confidence associated with the estimated ITEs and provided recommendation. If the estimated uncertainty exceeds a certain threshold (i.e.,  $\eta(\delta'_j) > \eta_0$ ), our model will alert the doctors that the provided recommendation are not reliable; 2) determining whether to assign a treatment at each timestamp. We derive the standard deviation from the variance and then calculate the 95% confidence intervals of ITE estimator. Our treatment recommendation strategy is that, at each timestamp in follow-up period, the treatment will be assigned to the patient if the upper bound of  $\delta'_j$  is less than zero, and the treatment will not be assigned if the lower bound of  $\delta'_j$  is larger than or equal to zero. The estimated uncertainty here is to enhance the robustness of prediction and guarantee the effectiveness of treatment recommendation.

## Implementation Details

The proposed model is implemented using Python 3.6 and PyTorch 1.4<sup>2</sup> and trained on Ubuntu 20.04 with NVIDIA GeForce RTX 2080 Ti. We train our model using the adaptive moment estimation (Adam) algorithm. Dropout<sup>23</sup> is enabled during the training and testing for uncertainty estimation. The data is randomly split into training, validation and test sets with percentages of 70%, 10%, 20%, and the validation set is used to improve the models and select the best model hyper-parameters (Supplementary Table 4). We report the performance on the test sets for all methods. The final results are averaged on five random realizations. The values of both temporal and static covariates are normalized as follows,

$$x_{t,i} = \frac{x_{t,i} - \text{mean}(x_{t,i})}{\text{std}(x_{t,i})} \quad (20)$$

where  $\text{mean}(x_{t,i})$  and  $\text{std}(x_{t,i})$  are the mean and standard deviation of  $i$ -th variable in  $x_t$  over the entire dataset.

## Data availability

MIMIC-III dataset is publicly available from PhysioNet<sup>3</sup>. AmsterdamUMCdb is publicly available from Amsterdam Medical Data Science website<sup>4</sup>.

## Code availability

The source code for this paper can be downloaded from the Github repository at <https://github.com/ruoqi-liu/T4> or the Zenodo repository at <https://doi.org/10.5281/zenodo.7003982>.

## References

1. Rhee, C. *et al.* Incidence and trends of sepsis in us hospitals using clinical vs claims data, 2009-2014. *Jama* **318**, 1241–1249 (2017).

<sup>2</sup><https://pytorch.org/>

<sup>3</sup><https://mimic.physionet.org/>

<sup>4</sup><https://amsterdammedicaldatascience.nl/>

2. Arefian, H. *et al.* Hospital-related cost of sepsis: a systematic review. *J. Infect.* **74**, 107–117 (2017).
3. Buchman, T. G. *et al.* Sepsis among medicare beneficiaries: 1. the burdens of sepsis, 2012–2018. *Critical care medicine* **48**, 276 (2020).
4. Treatment for sepsis. <https://www.sepsis.org/sepsis-basics/treatment/> (2021).
5. Dellinger, R. P. *et al.* Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock, 2012. *Intensive care medicine* **39**, 165–228 (2013).
6. Bacterial-infections in sepsis. <https://www.sepsis.org/sepsisand/bacterial-infections/> (2021).
7. Moss, S. R. & Prescott, H. C. Current controversies in sepsis management. In *Seminars in respiratory and critical care medicine*, vol. 40, 594–603 (Thieme Medical Publishers, 2019).
8. Klompas, M. & Rhee, C. Current sepsis mandates are overly prescriptive, and some aspects may be harmful. *Critical care medicine* **48**, 890–893 (2020).
9. Rhodes, A. *et al.* Campaña para sobrevivir a la sepsis: recomendaciones internacionales para el tratamiento de la sepsis y el choque séptico: 2016. *Crit Care Med [internet]. mar* **45** (2017).
10. Levy, M. M., Evans, L. E. & Rhodes, A. The surviving sepsis campaign bundle: 2018 update. *Intensive care medicine* **44**, 925–928 (2018).
11. Kalil, A. C., Johnson, D. W., Lisco, S. J. & Sun, J. Early goal-directed therapy for sepsis: a novel solution for discordant survival outcomes in clinical trials. *Critical care medicine* **45**, 607–614 (2017).
12. Liu, V. X. *et al.* The timing of early antibiotics and hospital mortality in sepsis. *Am. journal respiratory critical care medicine* **196**, 856–863 (2017).
13. Seymour, C. W. *et al.* Time to treatment and mortality during mandated emergency care for sepsis. *New Engl. J. Medicine* **376**, 2235–2244 (2017).
14. Force, I. S. T. Infectious diseases society of america (idsa) position statement: why idsa did not endorse the surviving sepsis campaign guidelines. *Clin. infectious diseases: an official publication Infect. Dis. Soc. Am.* **66**, 1631 (2018).
15. Rhee, C., Strich, J. R., Klompas, M., Yealy, D. M. & Masur, H. Sep-1 has brought much needed attention to improving sepsis care. . . but now is the time to improve sep-1. *Critical Care Medicine* **48**, 779–782 (2020).
16. Zhang, D., Micek, S. T. & Kollef, M. H. Time to appropriate antibiotic therapy is an independent determinant of postinfection icu and hospital lengths of stay in patients with sepsis. *Critical care medicine* **43**, 2133–2140 (2015).
17. Shashikumar, S. P., Josef, C., Sharma, A. & Nemati, S. Deepaise—an end-to-end development and deployment of a recurrent neural survival model for early prediction of sepsis. *arXiv preprint arXiv:1908.04759* (2019).
18. Tsoukalas, A., Albertson, T. & Tagkopoulos, I. From data to optimal decision making: a data-driven, probabilistic machine learning approach to decision support for patients with sepsis. *JMIR medical informatics* **3**, e3445 (2015).
19. Raghu, A. *et al.* Deep reinforcement learning for sepsis treatment. *arXiv preprint arXiv:1711.09602* (2017).
20. Raghu, A., Komorowski, M. & Singh, S. Model-based reinforcement learning for sepsis treatment. *arXiv preprint arXiv:1811.09602* (2018).
21. Komorowski, M., Celi, L. A., Badawi, O., Gordon, A. C. & Faisal, A. A. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nat. medicine* **24**, 1716–1720 (2018).
22. Utomo, C. P., Li, X. & Chen, W. Treatment recommendation in critical care: A scalable and interpretable approach in partially observable health states. In *International Conference On Information Systems* (2018).
23. Gal, Y. & Ghahramani, Z. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. *ICML'16 (JMLR.org)*, 2016).
24. Johnson, A. E. *et al.* Mimic-iii, a freely accessible critical care database. *Sci. data* **3**, 1–9 (2016).
25. Thorat, P. J. *et al.* Sharing icu patient data responsibly under the society of critical care medicine/european society of intensive care medicine joint data science collaboration: The amsterdam university medical centers database (amsterdamumcdb) example. *Critical care medicine* **49**, e563 (2021).
26. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural computation* **9**, 1735–1780 (1997).
27. Rosenbaum, P. R. & Rubin, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**, 41–55 (1983).
28. Singer, M. *et al.* The third international consensus definitions for sepsis and septic shock (sepsis-3). *Jama* **315**, 801–810 (2016).
29. Vincent, J.-L. *et al.* The sofa (sepsis-related organ failure assessment) score to describe organ dysfunction/failure (1996).
30. Vincent, J.-L. *et al.* Use of the sofa score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. *Critical care medicine* **26**, 1793–1800 (1998).
31. Ferreira, F. L., Bota, D. P., Bross, A., Mélot, C. & Vincent, J.-L. Serial evaluation of the sofa score to predict outcome in critically ill patients. *Jama* **286**, 1754–1758 (2001).
32. Wager, S. & Athey, S. Estimation and inference of heterogeneous treatment effects using random forests. *J. Am. Stat. Assoc.* **113**, 1228–1242 (2018).
33. Shalit, U., Johansson, F. D. & Sontag, D. Estimating individual treatment effect: generalization bounds and algorithms. In *International Conference on Machine Learning*, 3076–3085 (PMLR, 2017).
34. Seber, G. A. & Lee, A. J. *Linear regression analysis*, vol.

- 329 (John Wiley & Sons, 2012).
35. Liaw, A., Wiener, M. *et al.* Classification and regression by randomforest. *R news* **2**, 18–22 (2002).
  36. Wang, L., Zhang, Z. & Design, C. X. R. C. Theory and applications. *Support. Vector Mach. Springer-Verlag, Berlin Heidelberg* (2005).
  37. Hill, J. L. Bayesian nonparametric modeling for causal inference. *J. Comput. Graph. Stat.* **20**, 217–240 (2011).
  38. Yoon, J., Jordon, J. & van der Schaar, M. Ganite: Estimation of individualized treatment effects using generative adversarial nets. In *International Conference on Learning Representations* (2018).
  39. Shi, C., Blei, D. & Veitch, V. Adapting neural networks for the estimation of treatment effects. In *NeurIPS'19*, 2503–2513 (2019).
  40. Lim, B. Forecasting treatment responses over time using recurrent marginal structural networks. In *NeurIPS'18*, 7483–7493 (2018).
  41. Bica, I., Alaa, A. M., Jordon, J. & van der Schaar, M. Estimating counterfactual treatment outcomes over time through adversarially balanced representations. *arXiv preprint arXiv:2002.04083* (2020).
  42. Dupuis, C. & Timsit, J.-F. Antibiotics in the first hour: is there new evidence? *Expert. Rev. Anti-infective Ther.* **19**, 45–54 (2021).
  43. Im, Y. *et al.* Time-to-antibiotics and clinical outcomes in patients with sepsis and septic shock: a prospective nationwide multicenter cohort study. *Critical Care* **26**, 1–10 (2022).
  44. Alam, N. *et al.* Prehospital antibiotics in the ambulance for sepsis: a multicentre, open label, randomised trial. *The Lancet Respir. Medicine* **6**, 40–50 (2018).
  45. Rhee, C. *et al.* Prevalence of antibiotic-resistant pathogens in culture-proven sepsis and outcomes associated with inadequate and broad-spectrum empiric antibiotic use. *JAMA Netw. Open* **3**, e202899–e202899 (2020).
  46. Singer, M. Antibiotics for sepsis: does each hour really count, or is it incestuous amplification? (2017).
  47. Strich, J. R., Heil, E. L. & Masur, H. Considerations for empiric antimicrobial therapy in sepsis and septic shock in an era of antimicrobial resistance. *The J. Infect. Dis.* **222**, S119–S131 (2020).
  48. Severe sepsis and septic shock antibiotic guide - stanford medicine.
  49. Rubin, D. B. Causal inference using potential outcomes: Design, modeling, decisions. *J. Am. Stat. Assoc.* **100**, 322–331 (2005).
  50. Robins, J. M. & Hernán, M. A. Estimation of the causal effects of time-varying exposures. *Longitud. data analysis* **553**, 599 (2009).
  51. Robins, J. M., Hernan, M. A. & Brumback, B. Marginal structural models and causal inference in epidemiology (2000).
  52. Hernan, M. A. & Robins, J. M. Causal inference (2010).
  53. Schwab, P., Linhardt, L. & Karlen, W. Perfect match: A simple method for learning representations for counterfactual inference with neural networks. *arXiv preprint arXiv:1810.00656* (2018).
  54. Jesson, A., Mindermann, S., Shalit, U. & Gal, Y. Identifying causal-effect inference failure with uncertainty-aware models. *Adv. Neural Inf. Process. Syst.* **33** (2020).
  55. Bangaru, S. P., Suhas, J. & Ravindran, B. Exploration for multi-task reinforcement learning with deep generative models. *arXiv preprint arXiv:1611.09894* (2016).
  56. Xiao, T. & Kesineni, G. Generative adversarial networks for model based reinforcement learning with tree search. *Univ. California, Berkeley, Tech. report* (2016).
  57. Andersen, P.-A., Goodwin, M. & Granmo, O.-C. The dreaming variational autoencoder for reinforcement learning environments. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, 143–155 (Springer, 2018).
  58. Bica, I., Alaa, A. M. & van der Schaar, M. Time series deconfounder: Estimating treatment effects over time in the presence of hidden confounders. *arXiv preprint arXiv:1902.00450* (2019).
  59. Li, R. *et al.* G-net: a recurrent network approach to g-computation for counterfactual prediction under a dynamic treatment regime. In *Machine Learning for Health*, 282–299 (PMLR, 2021).
  60. Austin, P. C., Grootendorst, P. & Anderson, G. M. A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a monte carlo study. *Stat. medicine* **26**, 734–753 (2007).
  61. Imbens, G. W. & Rubin, D. B. *Causal inference in statistics, social, and biomedical sciences* (Cambridge University Press, 2015).
  62. Chandra, J. *et al.* A novel vascular leak index identifies sepsis patients with a higher risk for in-hospital death and fluid accumulation. *Critical Care* **26**, 1–10 (2022).

## Acknowledgements

This work was funded in part by the National Center for Advancing Translational Research of the National Institutes of Health under award number CTSA Grant UL1TR002733. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## Author contributions

PZ conceived the project. RL and PZ developed the method. RL conducted the experiments. RL and PZ analyzed the results. RL, KB, JC and PZ wrote the manuscript. All authors read and approved the final manuscript.

## **Competing interests**

The authors declare no competing interests.

## **Additional information**

Correspondence and requests for materials should be addressed to PZ.