

Estimating treatment effects for time-to-treatment antibiotic stewardship in sepsis

Received: 13 August 2021

Accepted: 2 March 2023

Published online: 06 April 2023

 Check for updates

Ruoqi Liu¹, Katherine M. Hunold², Jeffrey M. Caterino² & Ping Zhang^{1,3,4}  

Sepsis is a life-threatening condition with a high in-hospital mortality rate. The timing of antibiotic 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. Here we propose a novel method (called T4) to estimate treatment effects for time-to-treatment antibiotic stewardship in sepsis. T4 estimates individual treatment effects by recurrently encoding temporal and static variables as potential confounders, and then decoding the outcomes under different treatment sequences. We propose 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 equip T4 with an uncertainty quantification to help prevent overconfident recommendations. We demonstrate that T4 can identify effective treatment timing with estimated individual treatment effects for antibiotic stewardship on two real-world datasets. Moreover, comprehensive experiments on a synthetic dataset exhibit the outstanding performance of T4 compared with the state-of-the-art models on estimation of individual treatment effect.

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 deaths¹, and costs more than US\$27 billion annually in the USA². Based on a recent study of Medicare beneficiaries, approximately 30% of patients with sepsis do not survive for 6 months³. Broad-spectrum antibiotics are the first-line medications for sepsis^{4,5} because bacterial infection causes most cases⁶.

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^{7,8}. The Surviving Sepsis Campaign recommends initiating broad-spectrum antibiotics

within 1 hour for any patient with suspected sepsis or septic shock^{9,10}. Although the recommendation is supported by several large observational studies^{11–13}, there is substantial concern that striving for 1 hour antibiotic delivery for all patients with suspected sepsis may cause serious harm (for example, antibiotic resistance and *Clostridium difficile* infection)^{8,14,15}. Therefore, determining personalized antibiotic timing at the bedside is urgently needed.

Computational algorithms using electronic health records (EHRs)^{12,13,16,17} have been leveraged to examine the optimal antibiotic timing for patients with sepsis. EHRs contain irregularly sampled temporal data, including patients' lab results, vital signs and demographics. The aim is to estimate the treatment effects of different

¹Department of Computer Science and Engineering, The Ohio State University, Columbus, OH, USA. ²Department of Emergency Medicine, The Ohio State University, Columbus, OH, USA. ³Department of Biomedical Informatics, The Ohio State University, Columbus, OH, USA. ⁴Translational Data Analytics institute, The Ohio State University, Columbus, OH, USA. ✉ e-mail: zhang.10631@osu.edu

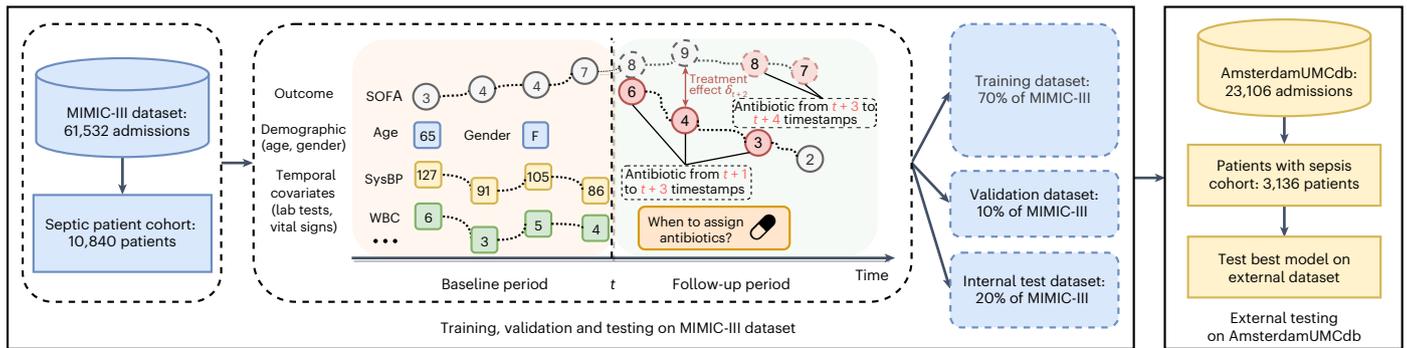


Fig. 1 | Overall data flow of T4 framework. Data from MIMIC-III are randomly split into training (70%), validation (10%) and testing (20%) datasets. The validation dataset is used to select the best model parameters and the testing dataset is used for internal evaluation. The T4 framework is used to estimate

ITEs for antibiotic administration timing recommendation. An external dataset obtained from AmsterdamUMCdb is used as an external test set. (SysBP: systolic blood pressure. WBC: white blood cell).

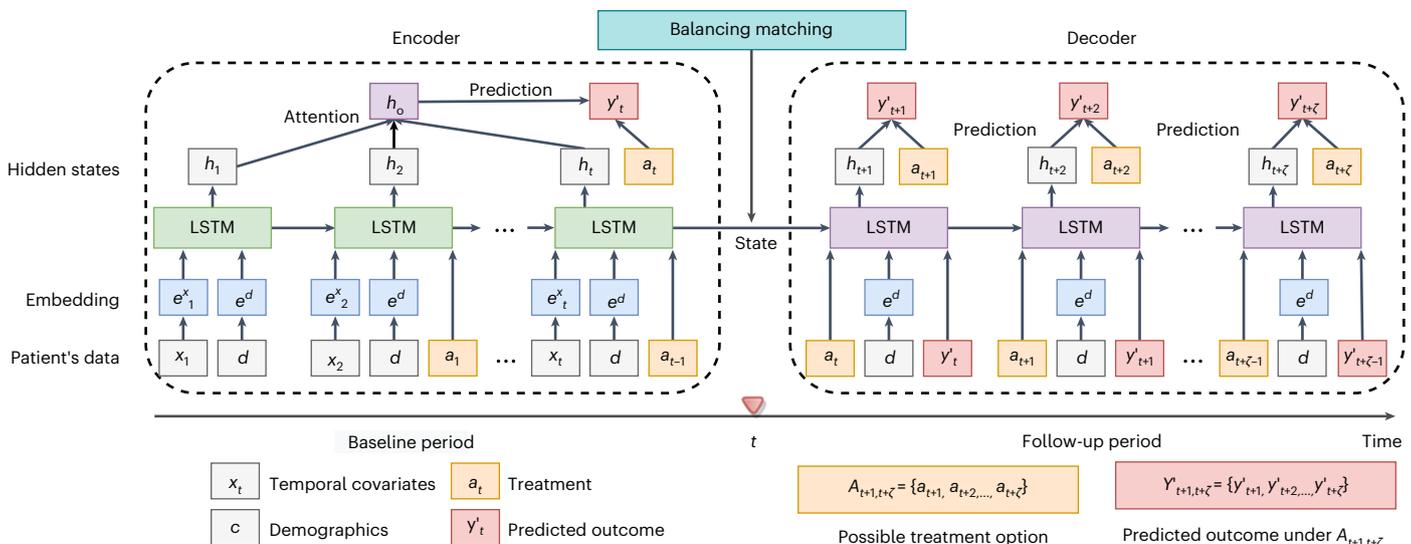


Fig. 2 | The T4 framework. T4 consists of three main components: (1) the encoder network recurrently encodes the patient's baseline information, including temporal and static covariates, and treatment assignments via the LSTM network; (2) the decoder network is initialized with encoder outputs and

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.

timings of antibiotics on septic outcomes (for example, in-hospital mortality). However, most studies have either ignored the temporality of EHRs^{12,13} or the heterogeneity of treatment effects^{16,17}. These two issues are crucial for identifying effective and precise therapy for patients with sepsis.

Previous studies^{18–22} of treatment design for sepsis regard the problem as an off-policy evaluation using observational data. For example, a model²¹ to learn treatment policy based on patient trajectories (that is, 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: (1) we need to do counterfactual reasoning based only on observed data and (2) we need to adjust time-varying confounding and estimate unbiased individual causal effects.

Herein, we study the problem of identifying the most effective timing for antibiotic administration in patients with suspected sepsis using EHR data. As shown in Fig. 1, the patients' information is extracted and compiled from EHRs and then used to build the model for antibiotic administration timing recommendation. We propose a

framework, T4, to estimate treatment effects for time-to-treatment antibiotic stewardship in sepsis. T4 first estimates the 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 pseudobalanced mini-batch, thus adjusting the influence of confounders. We also provide model interpretability of treatment recommendations by analysing: (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 Dropout²³ 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

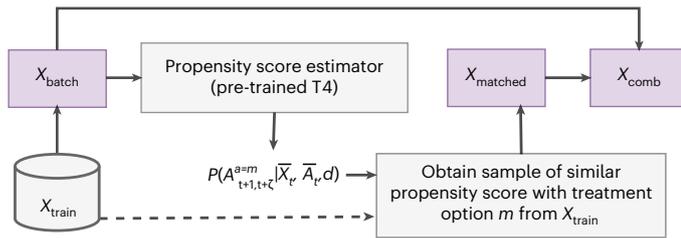


Fig. 3 | Illustration of the balancing matching. During the training process, each patient in the batch is matched with the corresponding counterfactual outcomes using the observed outcomes of their nearest neighbours in other treatment groups in the training data. A pre-trained T4 is used to estimate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

Mart for Intensive Care version III (MIMIC-III)²⁴ and AmsterdamUMCdb²⁵. 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 analyse model interpretability by visualizing the global and variable-level contributions to treatment recommendation via a concrete case study. Moreover, we conduct comparison experiments on ITE estimation using synthetic and semi-synthetic datasets, 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 uncertainty quantification and interpretable analysis into the ITE estimation 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 synthetic and semi-synthetic datasets for ITE estimation.

Overall framework

As illustrated in Fig. 2, 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)²⁶, a deep recurrent neural network that is widely used for modelling 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 scores²⁷ of receiving the current treatment. 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 π'

- 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 α 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 β according to Eq. (8);
- 13 **end**
- 14 Compute the potential outcome y_t' according to Eq. (9);
- 15 Compute the treatment effect δ' according to Eq. (10);
- 16 Compute the treatment strategy π' according to Eq. (11);
- 17 Compute outcome prediction loss according to Eq. (17);
- 18 **end**

Results

Datasets

MIMIC-III²⁴ 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.

AmsterdamUMCdb²⁵ is the first freely accessible European intensive care database. It is endorsed by the European Society of Intensive Care Medicine 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 patients with sepsis fulfilling the international consensus Sepsis-3 criteria²⁸. We extracted data for 10,840 patients and 3,136 patients from MIMIC-III and AmsterdamUMCdb, respectively, after applying exclusion criteria. The causal inference problem we studied is the treatment effects of antibiotic therapy among patients with sepsis given the observed confounding variables. There are three essential components that should be identified from the patient data. (1) Treatments: we considered multiple kinds of antibiotic therapy during ICU stays. At each timestamp, a binary treatment indicator will denote whether the patient is assigned antibiotics or not. (2) Confounders: we obtained 22 temporal covariates

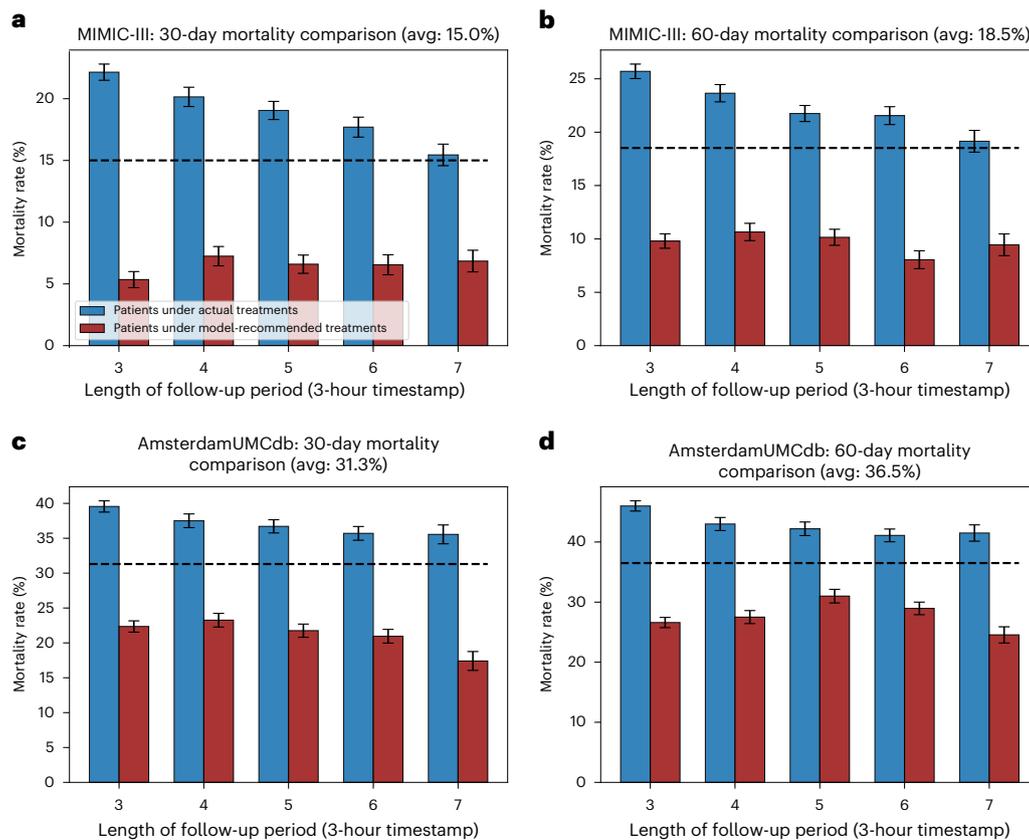


Fig. 4 | Mortality rate comparison of two datasets. **a, b**, Mortality comparisons on the MIMIC-III dataset over 30 days (**a**) and 60 days (**b**). **c, d**, Mortality comparisons on the AmsterdamUMCdb dataset over 30 days (**c**) and 60 days (**d**). Error bars denote 95% confidence intervals with $n = 30$ bootstrap samples. The

total mortality rate of two groups of patients is plotted using the black dashed line, which serves as the baseline. More data statistics of two groups of patients are in Supplementary Table 6.

(vital signs including temperature and heart rate; and lab tests including potassium and sodium levels) and 4 static covariates (age, gender and so on) as potential confounders. (3) Outcomes: we compiled the 24 hour Sepsis-related Organ Failure Assessment (SOFA)²⁹ score as the primary outcome, which is computed based on the degree of dysfunction of six organ systems. The definition of the Sepsis-3 patient cohort is in the study design section of Methods; a list of antibiotics is in Supplementary Table 1, computation of SOFA scores is in Supplementary Table 2 and the list of patients' covariates is in Supplementary Table 3.

Experiment on real-world data

Population-level analysis. PLACEHOLDER TEXT.

As we have no access to the counterfactual outcomes in the real-world dataset, 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 equation (11)) and the observed timing of administration on patients' mortality rate. Specifically, we first obtain a target group of patients whose observed timing of antibiotic 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 antibiotic 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 (see Supplementary Fig. 2 for patient similarity evaluation). 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 the recommendation.

As shown in Fig. 4, we compute and compare 30 days' mortality rate and 60 days' mortality rate on two datasets. The black dashed lines in the two plots denote the average mortality rates among the population, which are the baselines for 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 (that is, $\zeta \in \{3, 4, 5, 6, 7\}$) and the mortality rates for patients with the same treatments are consistently lower than the mortality rates for patients 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 patients with sepsis. We also observe that the model performs consistently on the external testing set from AmsterdamUMCdb, which demonstrates the robustness of our model when applied to a different dataset with different feature distributions. We further evaluate our model on suspected septic patients and report the results in Supplementary Fig. 3.

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. 5a, we use the predicted ITEs (red line) equipped with uncertainty

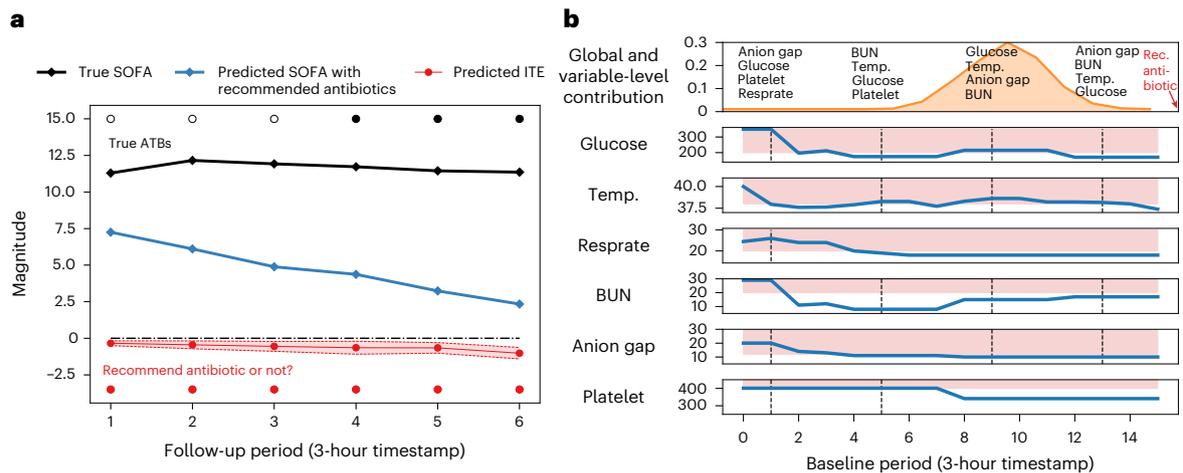


Fig. 5 | Case study of a patient for treatment recommendation and model interpretability. **a**, Predicted values of ITEs (red line with shadowed area denoting the uncertainty estimates) for antibiotic recommendation. An antibiotic will be recommended to the patient if the upper bound of predicted ITE 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 antibiotics. **b**, The most important global baseline timestamps (orange area) and variables contributing to the treatment recommendation are denoted at the top subplot. An antibiotic is recommended at the end of the baseline period.

estimates (red shadowed area) as the criteria to recommend the timing of antibiotic administration at each timestamp (see Supplementary Table 7 for average standard deviation of estimated treatment effect in 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 them if the lower bound of the predicted ITE is higher than zero. Here, zero denotes 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 antibiotics that were actually received and predicted outcomes under the recommended antibiotics. Here, we use the SOFA score as the outcome, where higher values are associated with worse status and higher mortality rate^{30,31}. From Fig. 5a, we observe that the patient received antibiotics from the fourth timestamp, which results in large SOFA scores during the follow-up period. Conversely, our model recommends the patient take antibiotics earlier (from the first to the sixth 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 patients with sepsis 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 contributions in Fig. 5b. We also plot 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 platelets. 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 take antibiotics at the end of the baseline period with regard to these warning signals. Taking glucose as an example, the normal range should be lower than 140 mg dl⁻¹, and a reading of more than 200 mg dl⁻¹ indicates diabetes³². We observe that the values of glucose are far from the normal range, which remains above 200 mg dl⁻¹ and

even reaches 300 mg dl⁻¹ at the early stage. This case study shows that our model achieves a transparent treatment recommendation via visualizing important timestamps and variables contributing to the recommendation, paving the way for interpretable and precise treatment recommendation. More case studies are provided in Supplementary Figs. 5 and 6.

Experiments on synthetic data

Comparison experiments. PLACEHOLDER TEXT.

Several studies have proposed estimating ITEs using causal inference techniques on observational data, including matching-based methods (for example, propensity score matching²⁷), forest-based methods (for example, causal forest³³) and representation-learning-based methods (for example, counterfactual regression³⁴). These methods are mainly designed for static data and are not suitable for estimating ITEs on EHRs, whereas T4 fully considers time-variable information and adjusts temporal confounding using balancing matching.

To illustrate the model's performance on ITE estimation and treatment recommendation, we performed experiments on a synthetic dataset. We simulated 5,000 patients with 50 timestamps, 20 temporal covariates and 5 static covariates. We used the first 40 timestamps as the baseline period and the remaining timestamps as the follow-up period. We use Precision in Estimation of Heterogeneous Effect (PEHE) and the error of Average Treatment Effect (eATE) to evaluate the model's performance. We conducted comparison experiments against state-of-the-art methods of ITE estimation: (1) classical methods: linear regression³⁵, random forest³⁶ and support vector machine³⁷; (2) forest-based methods: causal forest³³ and Bayesian additive regression trees³⁸; (3) representation-learning-based methods: counterfactual regression³⁴, GANITE³⁹ and Dragonnet⁴⁰; (4) time-varying-based methods: recurrent marginal structural network⁴¹, counterfactual recurrent network⁴² and G-Net⁴³. Results in Supplementary Table 10 show that our model outperforms state-of-the-art ITE estimation methods. More details of data simulation and results analysis can be found in the Supplementary Information.

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 ϵ ATE decreases as more matching samples are included 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 balancing matching improves the model performance on ITE estimation by substantially decreasing the estimation error.

Discussion

In this study, we propose a novel framework to estimate treatment effects for treatment recommendation. Our 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 influence of confounders. We also provide the model interpretability of treatment recommendation to analyse both global and variable-level contributions via an attention mechanism and variable importance analysis, respectively. Meanwhile, the model uncertainty quantification helps to avoid overconfident treatment recommendations. 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.

Controversies in antibiotics for patients with sepsis

The timing of antibiotic treatment is contentious. 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 rates⁷. Several recent cohort studies and randomized controlled trials^{8,15,44–46} suggest that the timing of antibiotic treatment should depend on the severity of illness (that is, sepsis, severe sepsis or septic shock) and the likelihood of true infection. They point out that immediate antibiotic regimens benefit patients with severe illness (for example, septic shock) while in less critically ill patients, immediate antibiotic regimens may lead to overprescribing and potential harm. For example, a recent study⁴⁷ showed 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 study⁴⁷, the morbidity of overdose antibiotics may be more obvious in less critically ill patients compared with 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 antibiotic resistance, antibiotics themselves also cause harm (for example, organ injury, mitochondrial dysfunction, the impact on the microbiome, and overgrowth by fungi and *C. difficile*)⁴⁸. Specifically, the study⁴⁷ showed that unnecessarily broad empiric therapy was associated with a 26% increased risk of *C. difficile* infection. Thus determining personalized antibiotic timing at the bedside is urgently needed.

Limitations of synthetic patient data

We evaluate the effectiveness of our model through comprehensive comparison experiments on synthetic and semi-synthetic datasets for treatment effect estimation (Supplementary Tables 9 and 10). The results demonstrate that the proposed method achieves a more accurate estimation of treatment effect than the state-of-the-art methods. Though outstanding performance has been achieved on the synthetic patient data, there are still some potential limitations.

Fully synthetic data. The fully synthetic data generated by statistical models from scratch is flexible as the entire data generation process

and specify the data scale (for example, the number of patients or the number of features) can be predefined according to the experimental design. However, such a predefined data generation process is merely to mimic the real data generation and may not fully represent the complex and heterogeneous patient EHR data⁴⁹. Additionally, rare cases and outliers that are present in the real patient data may not be accurately captured by the fully synthetic data.

Semi-synthetic data. Semi-synthetic data is simulated based on real patient data where we obtain the real patient covariates and simulate the treatment assignments and outcomes based on the covariates. Thus, compared with fully synthetic data, semi-synthetic data contain more realistic patients and thus can better capture rare cases and outliers. Whereas, semi-synthetic data are highly constrained by the original data. If the original data are limited or of poor quality, it is challenging to generate high-quality semi-synthetic data.

Thus, we further conduct comprehensive experiments on two real-world patient datasets that contain more diverse and realistic patient trajectories than the synthetic data.

Limitations of real-world patient data

Unavailability of counterfactual outcomes. PLACEHOLDER TEXT.

The ground truth counterfactual outcomes are not available in real-world patient data thus we can not directly evaluate the performance of treatment effect estimation on real-world data. Instead, we demonstrate in Supplementary Table 8 that the proposed method performs better than the state-of-the-art methods in factual outcome prediction of SOFA scores. Moreover, we evaluate the effectiveness of learned optimal timing of treatment and demonstrate in Fig. 4 that the mortality rate of the patients with recommended timing of treatment is lower than the patients with the observed timing of treatment.

Type of antibiotic treatment. The type of antibiotic treatment choice is based on suspected infection sites according to empirical antibiotic studies and guidelines^{50,51}. 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 and so on.

Blood cultures. Blood cultures are deemed the gold standard for antibiotic treatment regimens (for example, initiation or de-escalation of antibiotics). However, in our public datasets, most blood cultures are taken 12 hours before or after the ICU admission and usually the results are available after 2–3 days. According to the proposed framework as illustrated in Fig. 2, we only leverage the first 48 hours of data since ICU admission; 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 (for example, suspected infection sites, blood cultures and other concomitant therapies) if available.

Methods

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

Study design

We evaluated 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 defined the patients with sepsis according to the recent Sepsis-3 criteria²⁸ with respect to (1) t_{susp} : time of clinical suspicion of infection (that is, earlier timestamp of antibiotics and blood cultures within a specified duration) and (2) t_{SOFA} : two-point deterioration in SOFA score²⁹ 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 excluded patients 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 (that is, 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 were averaged. The missing data were imputed using 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 $\mathcal{D} = \{\bar{A}_T, \bar{X}_T, d, \bar{Y}_T\}$.

Our goal is to estimate the treatment effects with temporal and static covariates by predicting the potential outcomes under different treatment sequences. We adopt the potential outcomes framework^{52,53} 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 ζ 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 correspond to different treatment assignments.

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

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

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 ζ time period is $\Delta = [\delta_{t+1}, \dots, \delta_{t+\zeta}]$. In this paper, we use SOFA scores²⁹ (that is, range from 0 to 24, with larger values associated with more 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,

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

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

Assumptions

Our ITE estimation is based on the standard causal assumptions^{54,55} as follows,

Assumption 1. Consistency. The potential outcome under treatment history \bar{A}_t equals the observed outcome if the actual treatment history is A_t .

Assumption 2. Positivity. Given the observational data of the history, if 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, that is, $0 < P(A_t = 1 | \bar{X}_t, \bar{A}_{t-1}, d) < 1$, 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 , that is, $\mathbf{Y}(\mathbf{A}_t) \perp A_t | \bar{X}_t, \bar{A}_{t-1}, d$, for all treatment sequences \mathbf{A}_t .

Treatment effect estimation with T4

T4 recurrently encodes the patient's temporal covariates via LSTM²⁶, 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 $e_t^x \in \mathcal{R}^{K_e^x}$ as,

$$e_t^x = W_e x_t + b_e \tag{3}$$

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

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

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 an 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,

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

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

where Φ 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,

$$y_t^* = W_p([h_o, h_t, a_t]) + b_p \tag{6}$$

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 the decoder as,

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

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

$$\begin{aligned} \beta_{t+j,u} &= \Phi \left(W_{\beta}^T [h_{t+j}, h_u] \right) & \kappa_n(i) &= \arg \min_j \Delta_n(i, j) \quad (14) \\ \beta_{t+j,u} &= \text{softmax} (\beta_{t+j,1}, \beta_{t+j,2}, \dots, \beta_{t+j,t}) \\ c_{t+j} &= \sum_{u=1}^t \beta_{t+j,u} h_u \end{aligned} \quad (8)$$

where Φ 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 the teach-forcing technique with a ratio of 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 with 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 the next prediction. Compared with the permutation of all possible combinations (up to 2^{ζ}) of treatments, our strategy is more time efficient with the increasing of ζ . Then we can compute the treatment effect for $(t + j)$ -th timestamp using equation (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 ζ time period using equation (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 scores⁵⁶ to construct pseudo mini-batches that mimic the corresponding randomized controlled trial process (that is, 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 (that is, the potential outcomes under other possible treatment options except the observed one), with the observed outcomes of nearest neighbours in the training data. There are several methods to obtain the nearest neighbour by computing the distance among individuals. Here, we estimate the distance via the propensity score²⁷, which is defined as the conditional probability receiving the treatments m , given historical information up to the current timestamp:

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

Here, m_s denotes a possible treatment sequence during ζ . 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_s , and the patient j with treatments n_s is defined using absolute distance as,

$$\Delta_n(i, j) = |PS_i^{m_s} - PS_j^{n_s}| \quad (13)$$

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

Finally the matched mini-batch is combined with the original mini-batch as a whole for the following training process. Our balancing matching incorporates the balancing matching with representation learning as a whole and thus inherits the advantages of both the matching-based methods and the representation-learning-based methods. The confounding bias will be mitigated within each mini-batch and the proposed model achieves better performance than the state-of-the-art methods in treatment effect estimation (see results in Supplementary Tables 9 and 10).

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,

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

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

$$\mathcal{L}_a = -\frac{1}{N} \frac{1}{\zeta} \sum_{i=1}^N \sum_{j=1}^{\zeta} (a^i_{t+j} \log a^i_{t+j} + (1 - a^i_{t+j}) \log (1 - a^i_{t+j})) \quad (16)$$

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 Algorithm 1.

Model interpretability

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

Global contribution. The global 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 the 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 equation (9).

Variable-level contribution. Each timestamp contains a number of temporal variables (for example, lab tests and vital signs), 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_u . Here, we mask the corresponding information by replacing them with the mean value of i -th variable across the dataset. We compute the prediction loss $\mathcal{L}_y(y'_{x_u-i}, y)$ using equation (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_u are included in the loss computation. We multiply the global-level contribution and

the variable-level contribution ($\beta_{r;u} \omega_{u,i}$) 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^{23,57} to quantify the model uncertainty by applying dropout during both the training and testing process. Specifically, with the dropout enabled during the testing, the model generates a different output for 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 δ'_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 (that is, $\eta(\delta'_j) > \eta_0$), our model will alert the doctors that the provided recommendation is 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 the ITE estimator. Our treatment recommendation strategy is that, at each timestamp in the follow-up period, the treatment will be assigned to the patient if the upper bound of δ'_j is less than zero, and the treatment will not be assigned if the lower bound of δ'_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. Our experiments show that the adopted uncertainty quantification method is more efficient in monitoring estimation error when compared to propensity calibration and random exclusion (see Supplementary Fig. 7).

Implementation details

The proposed model is implemented using Python 3.6 and PyTorch 1.4 and trained on Ubuntu 20.04 with NVIDIA GeForce RTX 2080 Ti. We train our model using the adaptive moment estimation (Adam) algorithm. Dropout²³ is enabled during the training and testing for uncertainty estimation. The data is randomly split into training, validation and test sets (70%, 10%, 20%, respectively) and the validation set is used to improve the models and select the best model hyper-parameters (see Supplementary Tables 4 and 5 for hyper-parameter tuning of the proposed model and baseline models, respectively). The maximum length of the baseline period is 48-hour (additional analysis of the maximum length of the baseline period is in Supplementary Fig. 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. The detailed model configuration is in Supplementary Fig. 1.

Data availability

The MIMIC-III dataset is publicly available from PhysioNet (<https://mimic.physionet.org/>). AmsterdamUMCdb is publicly available from the Amsterdam Medical Data Science website (<https://amsterdam-medicaldatascience.nl/>).

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.7683025>.

References

- Rhee, C. et al. Incidence and trends of sepsis in us hospitals using clinical vs claims data, 2009-2014. *JAMA* **318**, 1241–1249 (2017).
- Arefian, H. et al. Hospital-related cost of sepsis: a systematic review. *J. Infect.* **74**, 107–117 (2017).
- Buchman, T. G. et al. Sepsis among medicare beneficiaries: 1. the burdens of sepsis, 2012–2018. *Crit. Care Med.* **48**, 276 (2020).
- Treatment for sepsis. *Sepsis Alliance* <https://www.sepsis.org/sepsis-basics/treatment/> (2021).
- Dellinger, R. P. et al. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock, 2012. *Intensive Care Med.* **39**, 165–228 (2013).
- Bacterial-infections in sepsis. *Sepsis Alliance* <https://www.sepsis.org/sepsisand/bacterial-infections/> (2021).
- Moss, S. R. & Prescott, H. C. Current controversies in sepsis management. *Semin. Respir. Crit. Care Med.* **40**, 594–603 (2019).
- Klompas, M. & Rhee, C. Current sepsis mandates are overly prescriptive, and some aspects may be harmful. *Crit. Care Med.* **48**, 890–893 (2020).
- Rhodes, A. et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. *Intensive care medicine.* **43**, 304–377 (2017).
- Levy, M. M., Evans, L. E. & Rhodes, A. The Surviving Sepsis Campaign bundle: 2018 update. *Intensive Care Med.* **44**, 925–928 (2018).
- 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. *Crit. Care Med.* **45**, 607–614 (2017).
- Liu, V. X. et al. The timing of early antibiotics and hospital mortality in sepsis. *Am. J. Resp. Crit. Care Med.* **196**, 856–863 (2017).
- Seymour, C. W. et al. Time to treatment and mortality during mandated emergency care for sepsis. *N. Engl. J. Med.* **376**, 2235–2244 (2017).
- IDSA Sepsis Task Force. Infectious Diseases Society of America (IDSA) position statement: why IDSA did not endorse the Surviving Sepsis Campaign guidelines. *Clin. Infect. Dis.* **66**, 1631–1635 (2018).
- 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. *Crit. Care Med.* **48**, 779–782 (2020).
- 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. *Crit. Care Med.* **43**, 2133–2140 (2015).
- Shashikumar, S. P., Josef, C., Sharma, A. & Nemati, S. DeepAISE-an interpretable and recurrent neural survival model for early prediction of sepsis. *Artificial intelligence in medicine* **113**, 102036 (2019).
- 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 Med. Inform.* **3**, e3445 (2015).
- Raghu, A. et al. Deep reinforcement learning for sepsis treatment. Preprint at <https://doi.org/10.48550/arXiv.1711.09602> (2017).
- Raghu, A., Komorowski, M. & Singh, S. Model-based reinforcement learning for sepsis treatment. Preprint at <https://doi.org/10.48550/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. Med.* **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 *Int. Conf. Information Systems* (2018).
23. Gal, Y. & Ghahramani, Z. Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In *Int. Conf. Machine Learning '16* (2016).
24. Johnson, A. E. et al. MIMIC-III, a freely accessible critical care database. *Sci. Data* **3**, 160035 (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. *Crit. Care Med.* **49**, e563 (2021).
26. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **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. **22**, 707–710 (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. *Crit. Care Med.* **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. Diabetes-diagnosis and treatment. *Mayo Clinic* <https://www.mayoclinic.org/diseases-conditions/diabetes/diagnosis-treatment/drc-20371451> (2021).
33. Wager, S. & Athey, S. Estimation and inference of heterogeneous treatment effects using random forests. *J. Am. Stat. Assoc.* **113**, 1228–1242 (2018).
34. Shalit, U., Johansson, F. D. & Sontag, D. Estimating individual treatment effect: generalization bounds and algorithms. In *Int. Conf. Machine Learning* 3076–3085 (PMLR, 2017).
35. Seber, G. A. & Lee, A. J. *Linear Regression Analysis* Vol. 329 (John Wiley and Sons, 2012).
36. Liaw, A. & Wiener, M. et al. Classification and regression by randomforest. *R News* **2**, 18–22 (2002).
37. Wang, L. (ed.) *Support Vector Machines: Theory and Applications* (Springer, 2005).
38. Hill, J. L. Bayesian nonparametric modeling for causal inference. *J. Comput. Graph. Stat.* **20**, 217–240 (2011).
39. Yoon, J., Jordon, J. & van der Schaar, M. GANITE: estimation of individualized treatment effects using generative adversarial nets. In *Int. Conf. Learning Representations* (2018).
40. Shi, C., Blei, D. & Veitch, V. Adapting neural networks for the estimation of treatment effects. In *NeurIPS'19* 2503–2513 (2019).
41. Lim, B. Forecasting treatment responses over time using recurrent marginal structural networks. In *NeurIPS'18* 7483–7493 (2018).
42. Bica, I., Alaa, A. M., Jordon, J. & van der Schaar, M. Estimating counterfactual treatment outcomes over time through adversarially balanced representations. In *Int. Conf. Learning Representations* (2020).
43. 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).
44. Dupuis, C. & Timsit, J.-F. Antibiotics in the first hour: is there new evidence? *Expert Rev. Anti Infect. Ther.* **19**, 45–54 (2021).
45. Im, Y. et al. Time-to-antibiotics and clinical outcomes in patients with sepsis and septic shock: a prospective nationwide multicenter cohort study. *Crit. Care* **26**, 19 (2022).
46. Alam, N. et al. Prehospital antibiotics in the ambulance for sepsis: a multicentre, open label, randomised trial. *Lancet Resp. Med.* **6**, 40–50 (2018).
47. 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 (2020).
48. Singer, M. Antibiotics for sepsis: does each hour really count, or is it incestuous amplification? *Am. J. Respir. Crit. Care Med.* **196**, 800–802 (2017).
49. Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F. & Mahmood, F. Synthetic data in machine learning for medicine and healthcare. *Nat. Biomed. Eng.* **5**, 493–497 (2021).
50. Strich, J. R., Heil, E. L. & Masur, H. Considerations for empiric antimicrobial therapy in sepsis and septic shock in an era of antimicrobial resistance. *J. Infect. Dis.* **222**, S119–S131 (2020).
51. *Severe Sepsis and Septic Shock Antibiotic Guide* (Stanford Health, 2017).
52. Rubin, D. B. Causal inference using potential outcomes: design, modeling, decisions. *J. Am. Stat. Assoc.* **100**, 322–331 (2005).
53. Robins, J. M. & Hernán, M. A. in *Longitudinal Data Analysis* (eds. Fitzmaurice, G. et al.) 553–599 (Chapman and Hall, 2009).
54. Robins, J. M., Hernán, M. A. & Brumback, B. Marginal structural models and causal inference in epidemiology. *Epidemiology* **11**, 550–560 (2000).
55. Hernán, M. A. & Robins, J. M. *Causal Inference* (2010).
56. Schwab, P., Linhardt, L. & Karlen, W. Perfect match: a simple method for learning representations for counterfactual inference with neural networks. Preprint at <https://doi.org/10.48550/arXiv.1810.00656> (2018).
57. Jesson, A., Mindermann, S., Shalit, U. & Gal, Y. Identifying causal-effect inference failure with uncertainty-aware models. *Adv. Neural Inf. Process. Syst.* **33**, 11637–11649 (2020).

Acknowledgements

This work was funded in part by the National Science Foundation under award number IIS-2145625 and by the National Institutes of Health under award number UL1TR002733. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation or the National Institutes of Health.

Author contributions

P.Z. conceived the project. R.L. and P.Z. developed the method. R.L. conducted the experiments. R.L. and P.Z. analysed the results. R.L., K.M.H., J.M.C. and P.Z. wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42256-023-00638-0>.

Correspondence and requests for materials should be addressed to Ping Zhang.

Peer review information *Nature Machine Intelligence* thanks Deepti Gurdasani and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Estimating treatment effects for time-to-treatment antibiotic stewardship in sepsis

In the format provided by the authors and unedited

Supplementary Material

Additional related work

On-policy learning Existing work of on-policy learning with reinforcement learning is proposed to first represent the actual environment by using a generative model and then learn optimal policies based on the generative environment. Bangaru et al.,¹³ introduce an adaptive exploration signal as a pseudo-reward from a deep generative model in order to deduce the Markov Decision Process (MDP). Xiao et al.,¹⁴ propose to apply generative adversarial networks (GAN) to learn the dynamics of the environment for model-based reinforcement learning. Andersen et al.,¹⁵ propose to create a generative environment using variational autoencoder (VAE) and learn optimal policies based on the generated samples. Baucum et al.,¹⁶ propose transitional variational autoencoders (tVAE), a generative model that produces realistic patient trajectories with few distributional assumptions, and can learn effective treatment policies.

Time-varying ITE More recently, transformers have become state-of-the-art in modeling sequential data and achieve better performance than LSTMs. A recent work¹⁷ proposes a transformer-based model for time-varying ITE with a novel counterfactual domain confusion loss to address confounding bias.

Additional details on experimental setup

Real-world data

Supplementary Table 1. The list of antibiotics in MIMIC-III Dataset. There are 18 kinds of antibiotics in total.

Category	Name
Antibiotic	Cefazolin, Cefepime, Ceftazidime, Ciprofloxacin, Clindamycin, Erythromycin, Gentamicin, Levofloxacin, Metronidazole, Moxifloxacin, Piperacillin, Rifampin, Tobramycin, Vancomycin, Amikacin, Ampicillin, Azithromycin, Aztreonam

Supplementary Table 2. The definition of SOFA score and its components across six organ systems. Each SOFA component score ranges from 0 (normal) to 4 (most abnormal). The total SOFA score ranges from 0 (normal) to 24 (most abnormal).

SOFA score	1	2	3	4
Respiration PaO ₂ /FiO ₂ , mmHg	< 400	< 300	< 200	< 100
Coagulation Platelets × 10 ³ /mm ³	< 150	< 100	< 50	< 20
Liver Bilirubin, mg/dl (μmol/l)	1.2 - 1.9 (20 - 32)	2.0 - 5.9 (33 - 101)	6.0 - 11.9 (102 - 204)	> 12.0 (> 204)
Cardiovascular Hypotension	MAP < 70 mmHg	Dopamine ≤ 5 or dobutamine (any dose)	Dopamine > 5 or epinephrine ≤ 0.1 or norepinephrine ≤ 0.1	Dopamine > 15 or epinephrine > 0.1 or norepinephrine > 0.1
Central nervous system (CNS) Glasgow Coma Score (GCS)	13 - 14	10 - 12	6 - 9	<6
Renal Creatinine, mg/dl (μmol/l) or urine output	1.2 - 1.9 (110 - 170)	2.0 - 3.4 (171 - 299)	3.5-4.9 (300 - 440) or < 500 ml/day	> 5.0 (> 440) or <200 ml/day

Supplementary Table 3. The list of variables in MIMIC-III and AdmsterdamDB. There are 22 temporal covariates and 4 demographics and static variables. PT: Prothrombin Time; BUN: Blood Urea Nitrogen; WBC: White Blood Cells count;

Category	MIMIC-III		AmsterdamDB		
	Mean	Std.	Mean	Std.	
Demographics	Age	65.55	16.44	61.30	17.90
	Gender	43% Female	-	42% Female	-
	Weight	81.67	25.50	79.83	13.61
	Height	169.30	11.17	175.15	8.44
Lab test	Anion gap	13.35	3.80	8.70	4.62
	Bicarbonate	25.65	5.27	25.63	6.35
	Bilirubin	3.36	6.41	3.15	6.85
	Creatinine	1.50	1.46	1.28	1.03
	Chloride	104.00	6.60	108.60	46.31
	Glucose	134.00	66.83	133.9	45.74
	Hematocrit	29.96	5.13	38.98	1.67
	Hemoglobin	10.09	1.79	12.57	1.64
	Lactate	2.44	2.14	2.40	2.95
	Platelet	235.05	155.28	220.82	171.65
	Potassium	4.08	0.63	5.58	602.56
	PT	17.76	8.95	1.59	10.12
	Sodium	138.84	5.32	140.88	43.45
	BUN	29.85	23.54	14.15	9.80
	WBC	11.23	7.64	14.56	11.80
Vital signs	Heart Rate	87.81	18.30	92.70	23.65
	SysBP	120.92	23.28	126.05	139.59
	DiasBP	61.41	14.55	60.77	31.11
	Meanbp	78.70	16.88	82.12	47.34
	Respratory	20.48	5.90	21.99	7.71
	Temperature	36.96	0.85	36.73	21.14
	SpO2	97.00	3.27	96.09	7.43

```
Seq2Seq(
  (encoder): Encoder(
    (embedding): Linear(in_features=22, out_features=32, bias=True)
    (embedding_static): Linear(in_features=4, out_features=32, bias=True)
    (rnn): LSTM(33, 128, num_layers=2, batch_first=True)
    (fc_out): Linear(in_features=161, out_features=1, bias=True)
    (dropout): Dropout(p=0.3, inplace=False)
    (attention_encoder): Sequential(
      (0): Linear(in_features=128, out_features=1, bias=True)
      (1): Tanh()
    )
    (ps_out): Sequential(
      (0): Linear(in_features=160, out_features=128, bias=True)
      (1): ReLU()
      (2): Linear(in_features=128, out_features=1, bias=False)
    )
  )
)

(decoder): AttentionDecoder(
  (embedding): Linear(in_features=1, out_features=32, bias=True)
  (embedding_static): Linear(in_features=4, out_features=32, bias=True)
  (attn_f): Attn(
    (attn): Linear(in_features=128, out_features=128, bias=True)
  )
  (rnn): LSTM(65, 128, num_layers=2, batch_first=True, dropout=0.3)
  (fc_out_1): Sequential(
    (0): Linear(in_features=256, out_features=128, bias=True)
    (1): ReLU()
    (2): Linear(in_features=128, out_features=1, bias=True)
  )
  (fc_out_0): Sequential(
    (0): Linear(in_features=256, out_features=128, bias=True)
    (1): ReLU()
    (2): Linear(in_features=128, out_features=1, bias=True)
  )
  (ps_out): Sequential(
    (0): Linear(in_features=256, out_features=128, bias=True)
    (1): ReLU()
    (2): Linear(in_features=128, out_features=1, bias=True)
  )
  (dropout): Dropout(p=0.3, inplace=False)
)
)
```

Supplementary Figure 1. Model configuration.

Supplementary Table 4. Hyperparameter search range and optimal hyperparameters for **T4**.

	Hyperparameter range	Optimal hyperparameters
Augmentation ratio	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 1.0	0.4
Learning rate	5e-3, 1e-3, 5e-4, 1e-4, 5e-5, 1e-5	5e-5
Embedding size	16, 32, 64, 128	32
Hidden layer size	32, 64, 128, 256	128
Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5	0.3
Batch size	8, 16, 32, 64	32

Synthetic Data

To illustrate the model performance on ITE estimation and treatment recommendation, we design experiments on a synthetic dataset. We first simulate temporal covariates x_t as a weighted sum of historical covariates \bar{x}_{t-1} and treatment assignments \bar{a}_{t-1} as follows,

$$x_t | \bar{x}_{t-1}, \bar{a}_{t-1} \sim \frac{1}{\sum_{j=1}^{t-1} w_j} \sum_{j=1}^{t-1} w_j (x_j + \lambda a_j) \quad (21)$$

where the weight $w_j = w_{j-1} * 2$ forms a geometric sequence with ratio 2, $\lambda \sim \mathcal{N}(0, 0.1)$ controls the influence of historical treatment assignments. The initial covariates are simulated by a multi-variable Gaussian distribution as $x_1 \sim \mathcal{N}(0^{k_t}, 0.1 \cdot (\Sigma \cdot \Sigma^\top))$, where $\Sigma \sim \mathcal{U}((-1, 1)^{k_t \times k_t})$ is simulated by a uniform distribution and k_t is the dimension of the temporal covariates. We also simulate static covariates c with the same distribution. We then simulate the treatment assignment a_t at each timestamp as,

$$a_t | x_t, c \sim \text{Bernoulli}(\sigma(s^\top [x_t, c] + m)) \quad (22)$$

where $s \sim \mathcal{N}(0^k, 0.1 \cdot I)$, k is the dimension of all covariates, I is the identity matrix, and $m \sim \mathcal{N}(0, 0.1)$. $\sigma(\cdot)$ is the sigmoid function and $[\cdot]$ concatenates two vectors as a whole. The outcome is simulated as a function of temporal covariates, static covariates and treatments as follows,

$$y_t | x_t, a_t \sim q^\top [x_t, c, \beta a_t] + \varepsilon \quad (23)$$

where $q \sim \mathcal{N}(0^{k+1}, 0.05 \cdot (\Sigma \cdot \Sigma^\top))$, $\Sigma \sim \mathcal{U}((-1, 1)^{(k+1) \times (k+1)})$, $\beta \sim \mathcal{N}(0, 0.5)$ controls the influence of treatments, and $\varepsilon \sim \mathcal{N}(0, 0.1)$.

In this paper, we simulate 5000 patients with 50 timestamps, $k_t = 20$ temporal covariates and 5 static covariates. We use the first 40 timestamps as the baseline period and the remaining as the follow-up period. Note that the timestamp here corresponds to the hour in the physical time unit without any time interval aggregation as in real-world data. The synthetic dataset and codes for simulation are available at <https://github.com/ruoqi-liu/T4>.

Semi-synthetic data based on MIMIC-III

To further evaluate our model on real patient trajectories while having available counterfactual ground truth for evaluation, we generate a semi-synthetic dataset based on MIMIC-III data. Specifically, we simulate treatment assignment a_t and potential outcome y_t given real patient covariates (i.e., 22 time-varying covariates and 4 static covariates). The treatment assignment is simulated by $a_t | x_t, c \sim \text{Bernoulli}(\sigma(s^\top [x_t, c] + m))$, where $s \sim \mathcal{N}(0^k, 0.1 \cdot I)$, k is the dimension of all covariates, I is the identity matrix, and $m \sim \mathcal{N}(0, 0.1)$. $\sigma(\cdot)$ is the sigmoid function and $[\cdot]$ concatenates two vectors as a whole. The outcome is simulated as a function of temporal covariates, static covariates and treatments as $y_t | x_t, a_t \sim q^\top [x_t, c, \beta a_t] + \varepsilon$, where $q \sim \mathcal{N}(0^{k+1}, 0.05 \cdot (\Sigma \cdot \Sigma^\top))$, $\Sigma \sim \mathcal{U}((-1, 1)^{(k+1) \times (k+1)})$, $\beta \sim \mathcal{N}(0, 0.5)$ controls the influence of treatments, and $\varepsilon \sim \mathcal{N}(0, 0.1)$.

Baseline Methods

We conduct comparison experiments against the state-of-the-art methods of ITE estimation in the following categories: (1) **Classical methods:** Linear Regression (LR)⁴, Random Forest (RF)⁵ and support vector machine (SVM)⁶ are included as the basic machine learning models for comparison. They directly regard the treatment assignment as an additional feature and predict potential outcomes based on the patient’s covariates under different treatments; (2) **Forest-based methods:** Causal Forest (CF)² and Bayesian Additive Regression Trees (BART)⁷ are two commonly used tree-based models for causal effect estimation. CF extends the classical model RF in that instead of minimizing prediction error, it split the data for maximizing the difference of treatment effects across splits. BART is a non-parametric Bayesian regression tree model that each tree is a learner constrained by a regularization prior; (3) **Representation learning-based methods:** Counterfactual Regression (CFR)³ construct balanced representations in the hidden space via deep neural network. We compare two variants of CFR: one is equipped with Wasserstein (WASS) as distance metrics for distribution balance and the other is the vanilla version without any distribution balance (TARNET). GANITE⁸ estimates the ITEs via a generative adversarial network by generating and discriminating counterfactuals. Dragonnet⁹ jointly optimize propensity prediction and potential outcome prediction for ITE estimation. (4) **Time-varying based methods: Recurrent Marginal structural Network (RMSM)**¹⁰ adopts recurrent marginal structural network for predicting the patient’s potential response to a series of treatments. Counterfactual Recurrent Network (CRN)¹¹ adopts adversarial training techniques to balance the historical confounding variables. G-Net¹² is a g-computation based deep sequential modeling framework that provides estimates of treatment effects under dynamic and time-varying treatment strategies. We further adapt existing time-varying methods in our setting by estimating treatment effects and obtaining optimal treatment policies at each timestamp.

Hyperparameter tuning for time-varying based ITE models.

Supplementary Table 5. Ranges for hyperparameter tuning for time-varying based ITE models.

Model	Module	Hyperparameter	Search range
RMSN	Propensity networks	Learning rate	5e-3, 1e-3, 5e-4, 1e-4, 5e-5, 1e-5
		Embedding size	16, 32, 64, 128
		Hidden layer size	32, 64, 128, 256
		Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
		Batch size	8, 16, 32, 64
		Random search iterations	50
	Encoder	Learning rate	5e-3, 1e-3, 5e-4, 1e-4, 5e-5, 1e-5
		Embedding size	16, 32, 64, 128
		Hidden layer size	32, 64, 128, 256
		Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
		Batch size	8, 16, 32, 64
		Random search iterations	50
	Decoder	Learning rate	5e-3, 1e-3, 5e-4, 1e-4, 5e-5, 1e-5
		Embedding size	16, 32, 64, 128
		Hidden layer size	32, 64, 128, 256
		Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
		Batch size	8, 16, 32, 64
		Random search iterations	20
CRN	Encoder	Learning rate	5e-3, 1e-3, 5e-4, 1e-4, 5e-5, 1e-5
		Embedding size	16, 32, 64, 128
		Hidden layer size	32, 64, 128, 256
		Balancing representation size (d_r)	32, 64, 128, 256
		FC hidden units	$0.5d_r, 1d_r, 2d_r$
		Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
		Batch size	8, 16, 32, 64
	Random search iterations	50	
	Decoder	Learning rate	5e-3, 1e-3, 5e-4, 1e-4, 5e-5, 1e-5
		Embedding size	16, 32, 64, 128
		Hidden layer size	Balancing representation size of encoder
		Balancing representation size (d_r)	32, 64, 128, 256
		FC hidden units	$0.5d_r, 1d_r, 2d_r$
		Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
Batch size		8, 16, 32, 64	
Random search iterations	30		
G-Net	-	Learning rate	5e-3, 1e-3, 5e-4, 1e-4, 5e-5, 1e-5
		Embedding size	16, 32, 64, 128
		Hidden layer size	32, 64, 128, 256
		Output size (d_r)	32, 64, 128, 256
		FC hidden units	$0.5d_r, 1d_r, 2d_r$
		Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
		Batch size	8, 16, 32, 64
		Random search iterations	50

Performance Measurement

We use Precision in Estimation of Heterogeneous Effect (PEHE) to evaluate the model performance on ITE estimation. Specifically, PEHE computes the mean squared error (MSE) between the values of ground truth ITE δ_j^i and estimated ITE $\hat{\delta}_j^i$ as follows,

$$\text{PEHE} = \frac{1}{\zeta} \frac{1}{N} \sum_{j=1}^{\zeta} \sum_{i=1}^N (\delta_j^i - \delta_j'^i)^2 \quad (24)$$

Besides the individual-level evaluation, we are also interested in the causal effect over the entire population. We use the error of Average Treatment Effect (ε_{ATE}) to evaluate the model performance, which is computed as the mean absolute error (MAE) between the ground truth and estimated ATE as,

$$\varepsilon_{\text{ATE}} = \frac{1}{\zeta} \sum_{j=1}^{\zeta} \left| \frac{1}{N} \sum_{i=1}^N \delta_j^i - \frac{1}{N} \sum_{i=1}^N \delta_j'^i \right| \quad (25)$$

Two metrics are all averaged on ζ follow-up timestamps and PEHE is regarded as the primary evaluation metric.

Additional experimental results

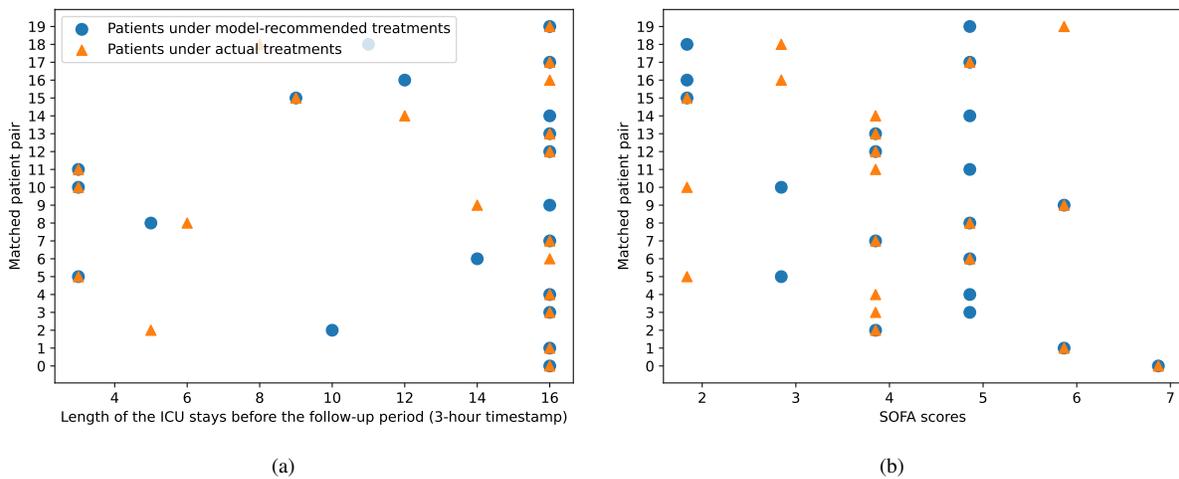
Real-world data

Additional results of the population-level analysis. In addition to the population-level analysis in Fig. 4, we further calculate the percentage of patients in the target group (i.e., the patients under actual treatments that are different from the model's recommendation) among all the patients in Supplementary Table 6.

Supplementary Table 6. The percentage of patients who received treatments different from the model's recommendation (median, %).

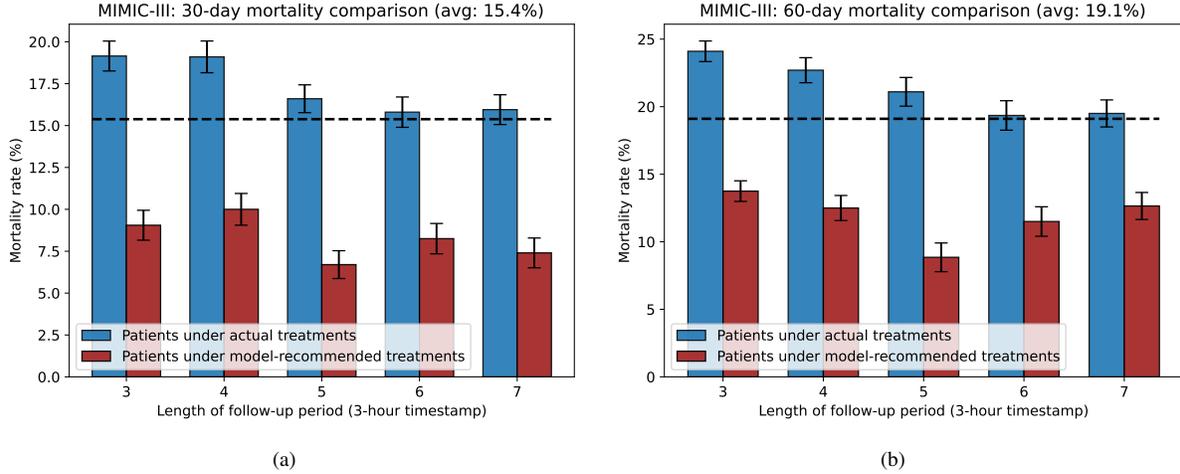
Follow-up period	30-day mortality					60-day mortality				
	3	4	5	6	7	3	4	5	6	7
MIMIC-III	50.10	53.40	48.81	36.37	23.90	50.41	53.24	50.38	36.39	23.69
AmsterdamUMCdb	54.55	64.28	67.46	64.72	59.33	54.76	64.34	68.17	63.00	58.58

To demonstrate the similarity between two groups of patients for mortality comparison, we show the distribution of the length of ICU stays before the follow-up period and SOFA scores in Supplementary Fig. 2. We randomly select 20 patients under actual treatments and their matched patients under model-recommended treatments. We observe that the length of the ICU stays and the SOFA scores are similar within each matched pair (i.e., most blue circles and yellow triangles are overlapped or close to each other).



Supplementary Figure 2. Similarity evaluation of 20 randomly selected patients. Fig.2(a) shows the length of the ICU stays of matched pairs. Fig. 2(b) shows the SOFA scores of matched pairs.

Besides experiments on septic patients, we also evaluate our model on patients with suspected sepsis who will or will not develop sepsis eventually. We conduct experiments on the suspected sepsis cohort obtained from the MIMIC-III dataset. As shown in Supplementary Fig. 3, the mortality rate of patients with treatments received at different timestamps as our recommendation (under actual treatments) is higher than the average mortality rate baseline, while the mortality rate of patients with treatments received at the same timestamps as our recommendation is lower than the baseline. This demonstrates that the model can also provide effective treatment strategies for suspected septic patients.



Supplementary Figure 3. Mortality rate comparison of suspected septic patients obtained from MIMIC-III dataset. 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 a black dashed line, which serves as the baseline.

Quantitative analysis of the width of confidence intervals. In Supplementary Table 7, we show a quantitative analysis of the relationship between the standard deviations and the length of the follow-up period. We calculate the average standard deviation of all the patients in the test set for each follow-up step. We observe that the standard deviation increases with the increase of the follow-up period.

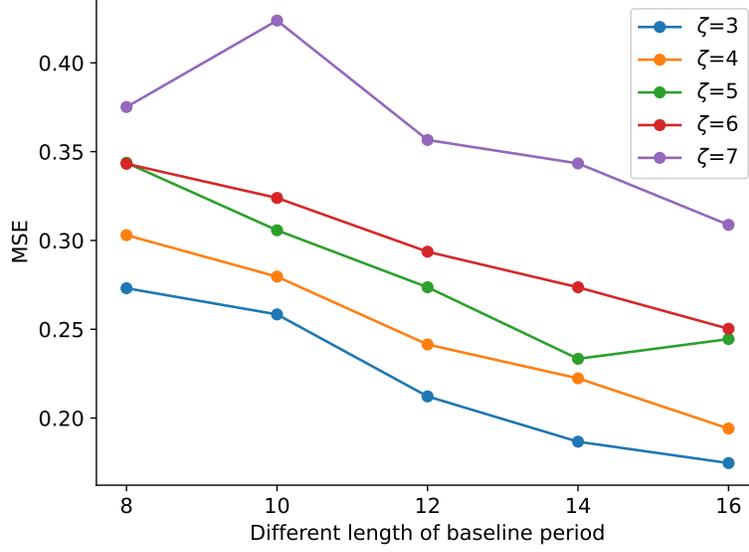
Supplementary Table 7. The average standard deviation of estimated treatment effect δ in the follow-up period.

ζ	1	2	3	4	5	6
Standard deviation of δ	0.03843	0.03950	0.03975	0.03996	0.03992	0.04011

Influence of the length of the baseline period. We set the maximum length of the baseline period as 16 timestamps. Here, we aggregate 3-hour as 1 timestamp as the real-world EHR data is commonly sparse and irregularly sampled. Thus 16 timestamps denote 48-hour in physical time. We aim to use the early data collected during the follow-up period for capturing the personalized characteristics of septic patients and then recommend optimal timing for ATBs in the follow-up period. We also explore the influence of the length of the baseline period and follow-up period using the MIMIC-III dataset. We evaluate the performance of factual prediction on SOFA scores over different lengths of the baseline period (8 to 16 timestamps). As shown in Supplementary Fig. 4, the model achieves the best performance with the length of baseline period equals 16 timestamps.

Comparison experiments on factual prediction. We conduct comprehensive comparison experiments for the factual prediction of SOFA scores on the MIMIC-III dataset. In Supplementary Table 8, we report the mean and standard deviation of mean squared error (MSE) of each method among different lengths of the follow-up period. The results show that our method consistently outperforms all the baselines in the factual prediction of SOFA scores.

Additional individual-level analysis. We showcase the situation that our recommended treatment timing is later than the factual treatment timing (Supplementary Fig. 5), indicating that the recommended timing of antibiotics depends on the

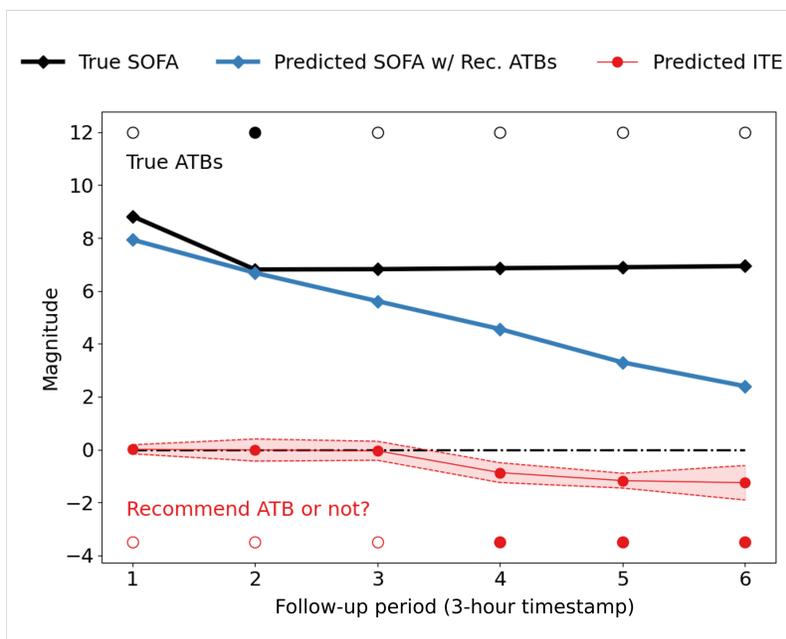


Supplementary Figure 4. The influence of the different lengths of baseline period to factual prediction performance on MIMIC-III dataset.

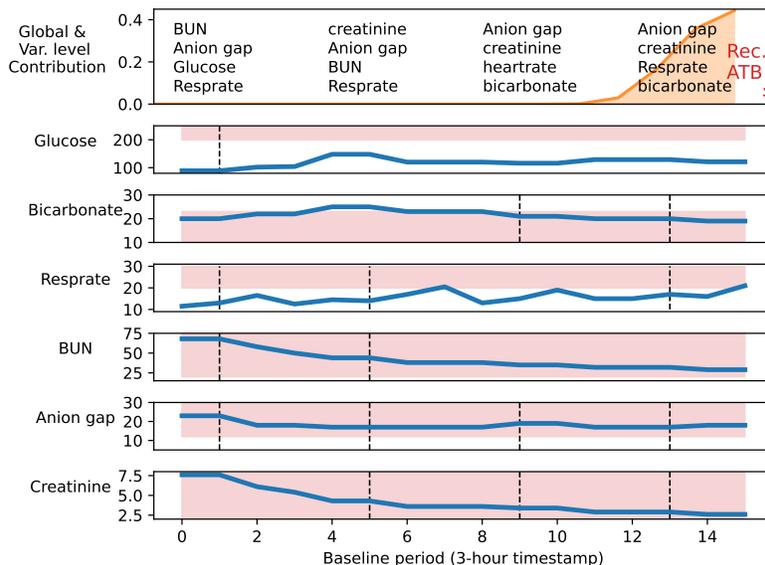
Supplementary Table 8. Performance comparison of factual prediction of SOFA scores on MIMIC-III dataset. Here, we report the estimated mean squared error (MSE) of each method among five different lengths of the follow-up period.

Method		$\zeta = 3$	$\zeta = 4$	$\zeta = 5$	$\zeta = 6$	$\zeta = 7$
Base model	LR	0.739 ± 0.094	0.749 ± 0.099	0.768 ± 0.122	0.803 ± 0.171	0.841 ± 0.233
	RF	0.613 ± 0.030	0.624 ± 0.029	0.634 ± 0.031	0.638 ± 0.032	0.635 ± 0.028
	SVM	0.639 ± 0.029	0.646 ± 0.030	0.652 ± 0.031	0.656 ± 0.030	0.660 ± 0.031
Representation learning based	CFR WASS ³	0.646 ± 0.017	0.657 ± 0.019	0.656 ± 0.015	0.665 ± 0.018	0.674 ± 0.015
	TARNET ³	0.663 ± 0.012	0.682 ± 0.022	0.697 ± 0.014	0.687 ± 0.017	0.688 ± 0.024
	GANITE ⁸	0.877 ± 0.016	0.884 ± 0.019	0.892 ± 0.013	0.896 ± 0.014	0.904 ± 0.013
	Dragonnet ⁹	0.642 ± 0.012	0.640 ± 0.009	0.650 ± 0.015	0.665 ± 0.022	0.666 ± 0.017
Forest based	Causal Forest ²	0.657 ± 0.022	0.666 ± 0.022	0.676 ± 0.020	0.682 ± 0.021	0.681 ± 0.020
	BART ⁷	0.579 ± 0.028	0.594 ± 0.024	0.612 ± 0.028	0.617 ± 0.022	0.619 ± 0.024
Time-varying based	RMSN ¹⁰	0.247 ± 0.020	0.233 ± 0.007	0.260 ± 0.005	0.321 ± 0.017	0.378 ± 0.017
	CRN ¹¹	0.225 ± 0.011	0.235 ± 0.008	0.263 ± 0.007	0.319 ± 0.018	0.391 ± 0.015
	G-Net ¹²	0.221 ± 0.010	0.230 ± 0.006	0.261 ± 0.009	0.313 ± 0.008	0.372 ± 0.012
Ours	T4	0.173 ± 0.013	0.212 ± 0.011	0.248 ± 0.007	0.275 ± 0.013	0.335 ± 0.019

individual patient’s health status. We also provide the model interpretability results in terms of global-level and variable-level contributions to treatment recommendation in Supplementary Fig. 6.



Supplementary Figure 5. An example to illustrate the treatment recommendation process. 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. The patient takes ATB in the early stage of the follow-up period, while the model recommends taking ATB later. The predicted SOFA score under recommended ATBs (blue) is much lower than the true SOFA score (black line).



Supplementary Figure 6. An example to illustrate model interpretability on treatment recommendation. 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. **T4** provides transparent antibiotics recommendation based on the time-varying variables in the baseline period.

Semi-synthetic data

The comparison results on the semi-synthetic dataset are shown in Supplementary Table 9. The proposed model yields the best performance among all the baselines.

Supplementary Table 9. Performance comparison of treatment effect estimation on semi-synthetic datasets based on MIMIC-III. Here, we report the PEHE of each method among five different lengths of the follow-up period. The results are the average and standard deviation over 10 random runs.

Method		$\zeta = 3$	$\zeta = 4$	$\zeta = 5$	$\zeta = 6$	$\zeta = 7$
Base model	LR	1.707 ± 0.239	1.723 ± 0.236	1.741 ± 0.237	1.750 ± 0.242	1.772 ± 0.246
	RF	1.167 ± 0.217	1.236 ± 0.242	1.227 ± 0.221	1.257 ± 0.21	1.278 ± 0.201
	SVM	0.894 ± 0.128	0.930 ± 0.127	0.941 ± 0.119	0.953 ± 0.115	0.966 ± 0.112
Representation learning based	CFR WASS ³	0.782 ± 0.022	0.791 ± 0.037	0.809 ± 0.031	0.831 ± 0.028	0.864 ± 0.041
	TARNET ³	0.814 ± 0.033	0.837 ± 0.021	0.839 ± 0.042	0.852 ± 0.031	0.881 ± 0.053
	GANITE ⁸	0.821 ± 0.041	0.832 ± 0.026	0.842 ± 0.024	0.867 ± 0.037	0.889 ± 0.039
	Dragonnet ⁹	0.758 ± 0.021	0.762 ± 0.018	0.787 ± 0.073	0.795 ± 0.026	0.823 ± 0.016
Forest based	Causal Forest ²	1.014 ± 0.127	1.034 ± 0.128	1.043 ± 0.126	1.050 ± 0.123	1.087 ± 0.129
	BART ⁷	0.732 ± 0.163	0.744 ± 0.228	0.792 ± 0.137	0.829 ± 0.182	0.858 ± 0.205
Time-varying based	RMSN ¹⁰	0.483 ± 0.023	0.502 ± 0.012	0.537 ± 0.017	0.561 ± 0.026	0.579 ± 0.023
	CRN ¹¹	0.480 ± 0.012	0.506 ± 0.024	0.532 ± 0.021	0.552 ± 0.021	0.571 ± 0.024
	G-Net ¹²	0.475 ± 0.013	0.497 ± 0.011	0.534 ± 0.021	0.548 ± 0.023	0.569 ± 0.027
Ours	T4	0.320 ± 0.027	0.348 ± 0.012	0.372 ± 0.013	0.384 ± 0.022	0.395 ± 0.024

Synthetic data

The performance of the proposed **T4** and existing baseline models on different lengths of follow-up period is shown in Supplementary Table 10. We vary $\zeta = 3$ to 7 and observe that **T4** demonstrates better performance compared to the baselines in terms of PEHE and ϵ ATE. We first examine the performance of base machine learning models including LR, RF and SVM. They directly regard the treatment assignment as an additional feature without considering the influence of confounders. Thus, the overall performance is lower than the other two categories which adjusts the influence of confounders and selection bias. Among the three base models, SVM performs better than others for different lengths of the follow-up period.

Representation learning based approaches achieve better performance than the base models. They take advantage of deep neural network to learn the representations of confounders. GANITE, which adopts the Generative Adversarial Nets (GANs) for estimating ITEs by modeling the counterfactual distributions, performs better than other representation learning models. TARNET and CFR WASS share the same fundamental framework excepts that CFR WASS is equipped with Wasserstein (WASS) as distance measurement for balancing the distribution of different treatment groups in the hidden space. Thus, the performance of CFR WASS is better than TARNET. Dragonnet achieves comparative performance with TARNET. The performance of forest based models is relatively better than representation learning based models and BART achieves better scores than the CF. Time-varying based methods (RMSN and CRN) are designed to estimate ITE with time-varying data. They generally outperform other baselines which are mainly designed for static data.

Our proposed **T4** achieves outstanding performance than all baselines. The model fully considers the temporality of data by modeling the time-varying confounders via a recurrent neural network, and adjusting the influence of confounders by constructing *pseudo* balanced training batches via balancing matching operation. To evaluate the effectiveness of balancing matching, we design two ablations of the original model. 1) **T4-w/o BM**: we remove the entire balancing matching component from the model; 2) **T4-w/ Rand.**: we augment the training batches with a portion of randomly selected samples. From the Supplementary Table 10, we observe the performance of **T4-w/o BM** is lower than the complete **T4** model, which demonstrates that balancing matching improves the model performance by constructing balanced batches and adjusting confounders. We also find that the performance of **T4-w/ Rand.** is better than **T4-w/o BM** but far from **T4**. The results show that random mini-batch augmentation only yields limited improvement. Our proposed balancing matching is specifically designed to adjust the influence of confounding via matching the samples based on the computed balancing scores, and therefore fundamentally facilitates the ITE estimation. Thus, compared to **T4-w/o BM** and **T4-w/ Rand.**, **T4** achieves the best performance for ITE estimation. We also consider an enhancing MC-dropout with variational dropout for recurrent networks¹ as a replacement for the original dropout layer (i.e., dropout is applied on the embedding layer). The performance of our model with variational dropout (**T4** (Variat. MC)) is comparable to the model with original dropout layer.

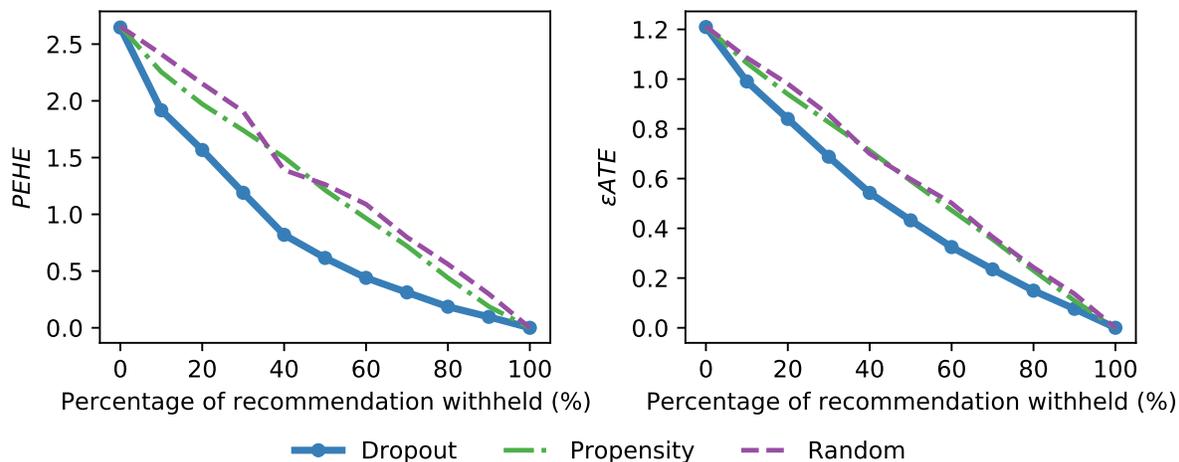
Influence of uncertainty quantification to treatment recommendation

We use the estimated uncertainty to quantify the confidence of estimated ITEs and provided treatment recommendation. In Supplementary Fig. 7, we demonstrate that our model equipped with uncertainty estimates $\eta(\delta')$ achieves lower estimation error when excluding fewer individuals (withholding recommendation) with uncertain estimation compared to other uncertainty quantification approaches. Here, the percentage of recommendation withheld in the figure denotes the percentage of excluded individuals with low confidence during the evaluation. The approach of propensity¹⁸ computes and rank the propensity scores (Eq. (12)) for each individual. Then it excludes the individuals with computed propensity scores either close to 0 or 1. These individuals are likely to violate the overlap assumption¹⁹ in causal inference and lead to inaccurate ITE estimation. The random approach randomly excludes the individuals and serves as the baseline.

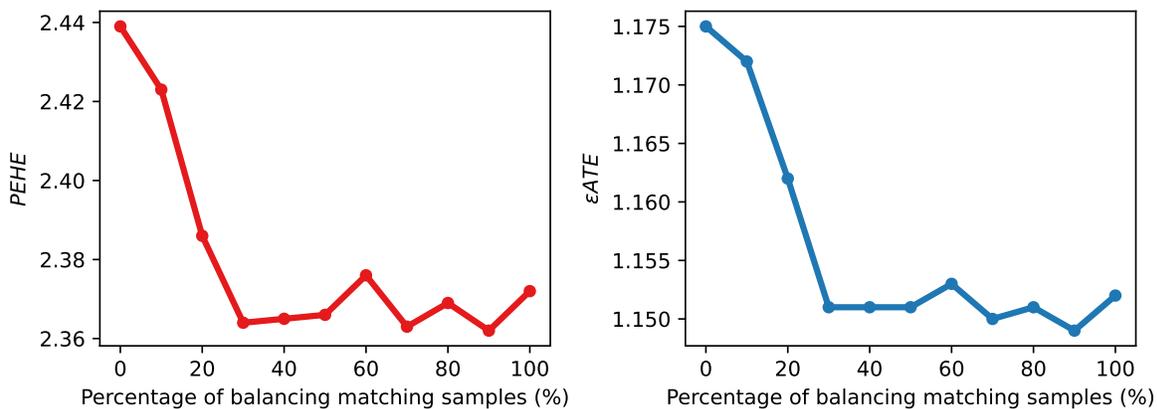
We vary the percentage of recommendation withheld and show the PEHE and ϵ ATE on the remaining data samples in the figure. We observe that estimated PEHE and ϵ ATE for three approaches all decline with the increasing percentage of recommendation withheld because samples with high uncertainty and inaccurate ITE estimation are excluded sequentially. The uncertainty quantification method $\eta(\delta')$ we adopt in our model achieves the best performance among others in terms of fast decreasing estimation error when the percentage of recommendation withheld increases. The results demonstrate that uncertainty quantification helps treatment recommendation by explicitly examining the confidence associated with ITE estimation and treatment recommendation.

Supplementary Table 10. Performance comparison of ITE estimation on synthetic datasets. Here, we report the estimated PEHE and ϵ ATE of each method among five different lengths of the follow-up period.

Method	$\zeta = 3$		$\zeta = 4$		$\zeta = 5$		$\zeta = 6$		$\zeta = 7$		
	PEHE	ϵ ATE	PEHE	ϵ ATE	PEHE	ϵ ATE	PEHE	ϵ ATE	PEHE	ϵ ATE	
Base model	LR	3.958 \pm 0.072	1.500 \pm 0.026	3.987 \pm 0.072	1.505 \pm 0.025	4.017 \pm 0.07	1.511 \pm 0.024	4.048 \pm 0.064	1.517 \pm 0.023	4.086 \pm 0.059	1.524 \pm 0.023
	RF	3.989 \pm 0.077	1.506 \pm 0.027	4.021 \pm 0.077	1.512 \pm 0.026	4.052 \pm 0.076	1.518 \pm 0.026	4.084 \pm 0.069	1.523 \pm 0.025	4.124 \pm 0.062	1.53 \pm 0.024
	SVM	3.947 \pm 0.075	1.498 \pm 0.027	3.979 \pm 0.074	1.504 \pm 0.026	4.007 \pm 0.074	1.509 \pm 0.026	4.038 \pm 0.070	1.515 \pm 0.025	4.078 \pm 0.064	1.522 \pm 0.025
Representation learning based	CFR WASS ³	3.794 \pm 0.199	1.447 \pm 0.032	3.823 \pm 0.188	1.451 \pm 0.030	3.851 \pm 0.175	1.455 \pm 0.029	3.884 \pm 0.178	1.459 \pm 0.029	3.914 \pm 0.178	1.464 \pm 0.029
	TARNET ³	3.848 \pm 0.099	1.469 \pm 0.020	3.898 \pm 0.108	1.478 \pm 0.021	3.937 \pm 0.103	1.484 \pm 0.021	3.967 \pm 0.102	1.489 \pm 0.021	3.998 \pm 0.116	1.494 \pm 0.023
	GANITE ⁸	3.786 \pm 0.212	1.460 \pm 0.035	3.808 \pm 0.195	1.463 \pm 0.033	3.842 \pm 0.195	1.471 \pm 0.034	3.869 \pm 0.208	1.473 \pm 0.035	3.911 \pm 0.204	1.480 \pm 0.035
	Dragonnet ⁹	3.830 \pm 0.084	1.466 \pm 0.018	3.872 \pm 0.096	3.872 \pm 0.096	3.922 \pm 0.091	1.480 \pm 0.019	3.945 \pm 0.098	1.484 \pm 0.020	3.980 \pm 0.112	1.491 \pm 0.022
Forest based	Causal Forest ²	3.755 \pm 0.201	1.441 \pm 0.033	3.785 \pm 0.192	1.446 \pm 0.031	3.810 \pm 0.180	1.449 \pm 0.027	3.843 \pm 0.182	1.454 \pm 0.022	3.875 \pm 0.182	1.459 \pm 0.030
	BART ⁷	3.750 \pm 0.205	1.44 \pm 0.034	3.782 \pm 0.193	1.446 \pm 0.031	3.806 \pm 0.176	1.449 \pm 0.030	3.841 \pm 0.180	1.454 \pm 0.030	3.871 \pm 0.180	1.459 \pm 0.029
Time-varying based	RMSN ¹⁰	2.480 \pm 0.022	1.195 \pm 0.011	2.510 \pm 0.017	1.201 \pm 0.009	2.599 \pm 0.032	1.218 \pm 0.012	2.569 \pm 0.027	1.211 \pm 0.012	2.661 \pm 0.027	1.230 \pm 0.018
	CRN ¹¹	2.470 \pm 0.030	1.187 \pm 0.010	2.539 \pm 0.024	1.202 \pm 0.011	2.627 \pm 0.028	1.219 \pm 0.011	2.631 \pm 0.031	1.219 \pm 0.014	2.675 \pm 0.029	1.235 \pm 0.019
	G-Net ¹²	2.476 \pm 0.009	1.193 \pm 0.014	2.507 \pm 0.011	1.200 \pm 0.009	2.591 \pm 0.013	1.213 \pm 0.007	2.558 \pm 0.013	1.213 \pm 0.011	2.651 \pm 0.015	1.225 \pm 0.012
Ours	T4-w/o BM	2.439 \pm 0.012	1.175 \pm 0.008	2.495 \pm 0.019	1.188 \pm 0.010	2.516 \pm 0.017	1.192 \pm 0.007	2.555 \pm 0.021	1.200 \pm 0.009	2.582 \pm 0.024	1.207 \pm 0.014
	T4-w/ Rand.	2.406 \pm 0.016	1.165 \pm 0.011	2.474 \pm 0.015	1.183 \pm 0.011	2.491 \pm 0.013	1.186 \pm 0.012	2.510 \pm 0.019	1.189 \pm 0.010	2.548 \pm 0.020	1.197 \pm 0.012
	T4 (Variat. MC)	2.372 \pm 0.014	1.160 \pm 0.012	2.416 \pm 0.020	1.168 \pm 0.011	2.442 \pm 0.021	1.167 \pm 0.009	2.466 \pm 0.019	1.174 \pm 0.013	2.557 \pm 0.026	1.200 \pm 0.015
	T4	2.362 \pm 0.020	1.149 \pm 0.012	2.384 \pm 0.019	1.156 \pm 0.011	2.430 \pm 0.022	1.164 \pm 0.014	2.473 \pm 0.019	1.174 \pm 0.007	2.476 \pm 0.024	1.176 \pm 0.010



Supplementary Figure 7. The estimated PEHE and ϵ ATE of different uncertainty quantification approaches when we vary the percentage of recommendation withheld. The evaluation metrics are reported on the remaining samples with the provided recommendation.



Supplementary Figure 8. The performance change with regards to PEHE and ϵ ATE when the percentage of balancing matching samples in each training batch increases.

References

1. Gal, Y. & Ghahramani, Z. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. ICML'16 (JMLR.org, 2016).
2. Wager, S. & Athey, S. Estimation and inference of heterogeneous treatment effects using random forests. *J. Am. Stat. Assoc.* **113**, 1228–1242 (2018).
3. 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).
4. Seber, G. A. & Lee, A. J. *Linear regression analysis*, vol. 329 (John Wiley & Sons, 2012).
5. Liaw, A., Wiener, M. *et al.* Classification and regression by randomforest. *R news* **2**, 18–22 (2002).
6. Wang, L., Zhang, Z. & Design, C. X. R. C. Theory and applications. *Support. Vector Mach. Springer-Verlag, Berlin Heidelberg* (2005).
7. Hill, J. L. Bayesian nonparametric modeling for causal inference. *J. Comput. Graph. Stat.* **20**, 217–240 (2011).
8. 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).
9. Shi, C., Blei, D. & Veitch, V. Adapting neural networks for the estimation of treatment effects. In *NeurIPS'19*, 2503–2513 (2019).
10. Lim, B. Forecasting treatment responses over time using recurrent marginal structural networks. In *NeurIPS'18*, 7483–7493 (2018).
11. Bica, I., Alaa, A. M., Jordon, J. & van der Schaar, M. Estimating counterfactual treatment outcomes over time through adversarially balanced representations. *Int. Conf. on Learn. Represent.* (2020).
12. 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).
13. Bangaru, S. P., Suhas, J. & Ravindran, B. Exploration for multi-task reinforcement learning with deep generative models. *arXiv preprint arXiv:1611.09894* (2016).
14. Xiao, T. & Kesineni, G. Generative adversarial networks for model based reinforcement learning with tree search. *Univ. California, Berkeley, Tech. report* (2016).
15. 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).
16. Baucum, M., Khojandi, A. & Vasudevan, R. Improving deep reinforcement learning with transitional variational autoencoders: A healthcare application. *IEEE J. Biomed. Heal. Informatics* **25**, 2273–2280 (2020).
17. Melnychuk, V., Frauen, D. & Feuerriegel, S. Causal transformer for estimating counterfactual outcomes. In *International Conference on Machine Learning*, 15293–15329 (PMLR, 2022).
18. 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).
19. Imbens, G. W. & Rubin, D. B. *Causal inference in statistics, social, and biomedical sciences* (Cambridge University Press, 2015).