Medical Dead-ends and Learning to Identify High-risk States and Treatments

Mehdi Fatemi, Taylor W. Killian, Jayakumar Subramanian, Marzyeh Ghassemi

2021 NeurIPS

Presenter: Wei-Chun Tsai

National Cheng Kung University

國立成功大學
National Cheng Kung University
1931

OUTLINE

- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

OUTLINE

- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

Introduction - Off-policy Reinforcement Learning (RL)

- Off-policy Reinforcement Learning (RL)
  - Isolate behavioural policies from the target policy
  - Important in safety-critical domains
  - Significant advances made possible by off-policy RL combined with DNNs
- Pitfalls
  - The performance degrade drastically in fully offline settings
  - Significantly overfit to data-collection artifacts
  - RL estimates of optimal policies are largely unreliable in healthcare due to legal and ethical implications

Introduction - Dead-end Discovery (DeD)

- Dead-end Discovery (DeD)
  - Paradigm Shift: identify **treatments to avoid** as opposed to what treatment to select
  - Goal: avoid future **dead-ends**, which negative outcomes are inevitable
- Validation DeD in a carefully constructed toy domain
- Evaluation - Septic
  - Septic: highly prevalent, physiologically severe, costly, poorly understood.
  - DeD confirms the existence of dead-ends, and demonstrate that 12% of treatments administered to terminally ill patients reduce their chances of survival

OUTLINE

- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

Related Work

- RL in Health
  - Recent work: seeking to develop optimal treatment recommendation policies
  - An **optimal policy** that maximizes a patient's chance of recovery is both **computationally and experimentally infeasible**
- Safety in RL
  - Recent work: evaluated in online settings, where data can be acquired or models can be tested against new cases
- Dead-ends
  - Proposed by Fatemi et al. in the context of exploration
  - Adapting this approach and expanding the theoretical results to an offline RL setting

OUTLINE

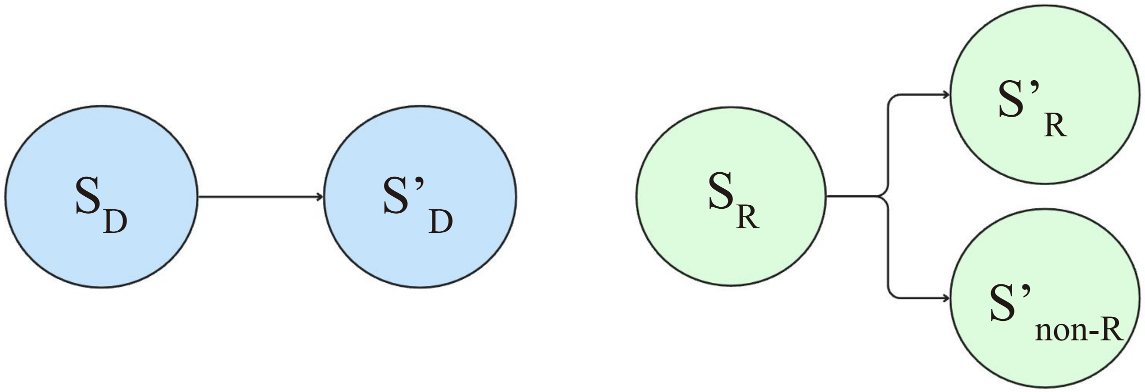
- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

Methods - Math Framework

- Markov Decision Processes (MDP)
  - 2 independent MDPs: M_D and M_R
  - Optimal State-treatment Value Function: Q^*_D and Q^*_R
  - Optimal State Value Function: V^*_D and V^*_R
  - Reward Function
    - M_D returns -1 with any transition to a negative terminal state (0 otherwise)
    - M_R returns $+1$ with any transition to a positive terminal state (0 otherwise)
  - $Q^*_D(s,a) \in [-1,0]$, $Q^*_R(s,a) \in [0,1]$

Methods - Math Framework

- Special States
  - Terminal States S_T: the final observation of any recorded trajectory
  - Dead-end S_D: negative outcomes are unavoidable (happening w.p.1)
  - Rescue S_R: positive outcome is reachable (with probability 1)



Dead-end states

Rescue states

Methods - Math Framework

- Mathematical Proof Conclusion
  - V^*_D of all dead-end states will be precisely -1.
  - $Q^*_D(s, a) = -1$ for all treatments a at state s if and only if s is a dead-end.
- Summary
  - Treatment Security: abiding by the maximum hope of a positive outcome
  - Connecting the RL concept of value functions to dead-end discovery
  - V^*_D enables detecting dead-end states
  - Q^*_D enables further treatment avoidance

OUTLINE

- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

Methods - Neural Network Based State Construction and Identification

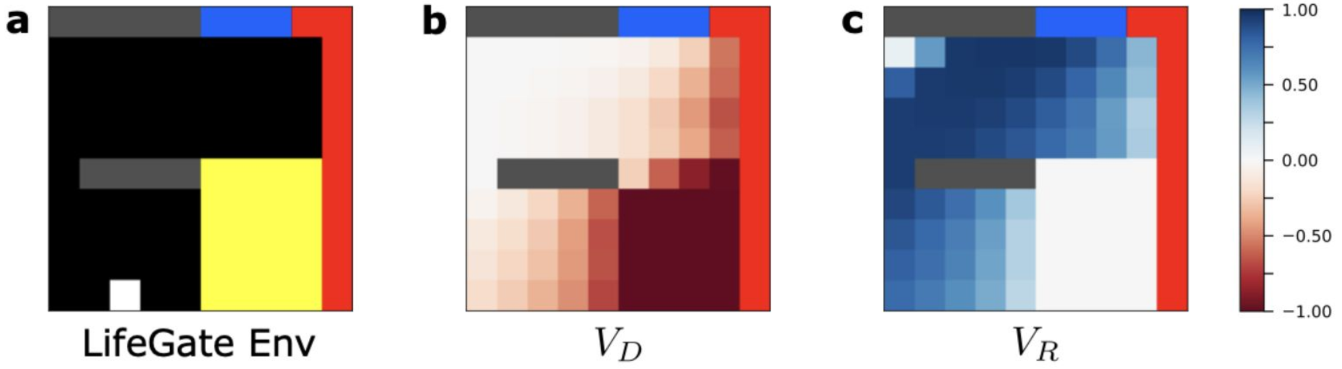
- State Construction (SC-Network)
  - Constructing states of patients
  - Transforming a single or possible sequence of observations into a fixed embedding
- Identification (D-Network and R-Network)
  - Trained using Double DQN algorithm
  - Computing Q_D and Q_R for all treatment of given state

OUTLINE

- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

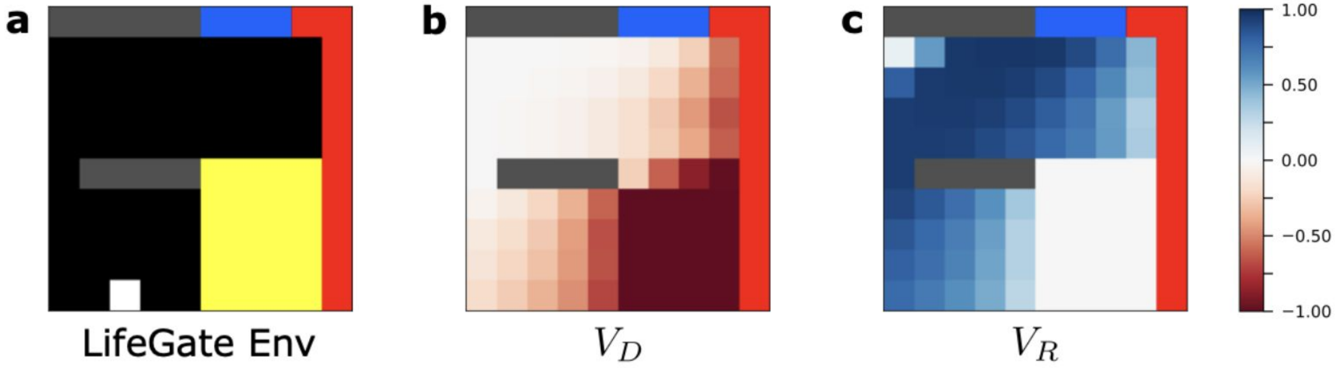
Methods - Toy Problem Validation: Life-Gate

- Provide a tabular toy-example (Life-Gate)
- Set up
    - White: Agent
    - Gray: Obstacle (neutral)
    - Yellow: Dead-ends
    - Black
    - Blue - Life gate
    - Red - Death gate
- Actions: moving up, down, left, right, and doing nothing (no-up)



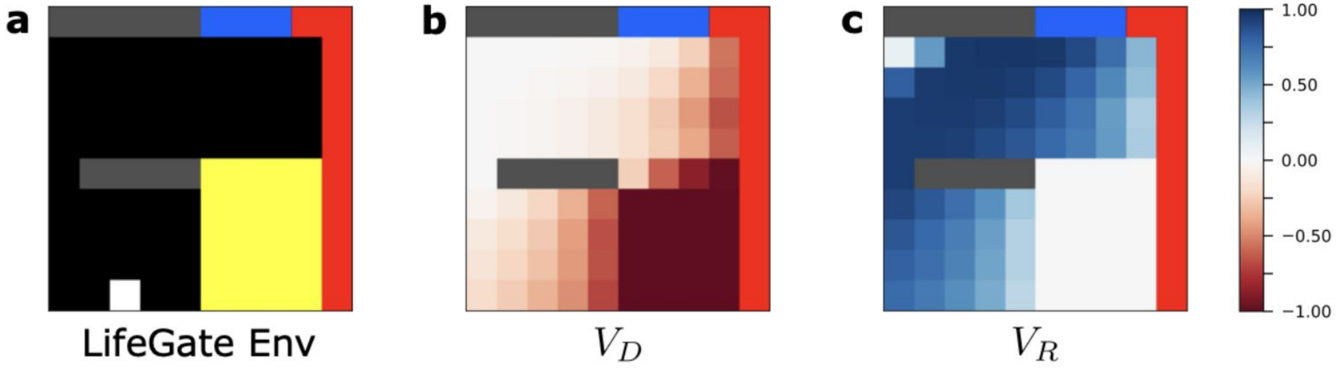
Methods - Toy Problem Validation: Life-Gate

- Black Area
  - DEATH-DRIFT = 40%
- Yellow Area (dead-end states)
  - Right: 70%
  - No actions: 30%
- Adjacent states to dead-ends are possibly the most critical to alert



Methods - Toy Problem Validation: Life-Gate

- Conclusion
  - $\delta_D = -0.7$ and $\delta_R = 0.7$ seem to clearly set the boundary for most states
  - Only for all yellow area (aside from the few erroneous states), $V_D = -1$
  - No dead-end state can be a rescue, as seen by $V_R = 0$ for the yellow area



OUTLINE

- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

Empirical Setup for Dead-end Analysis

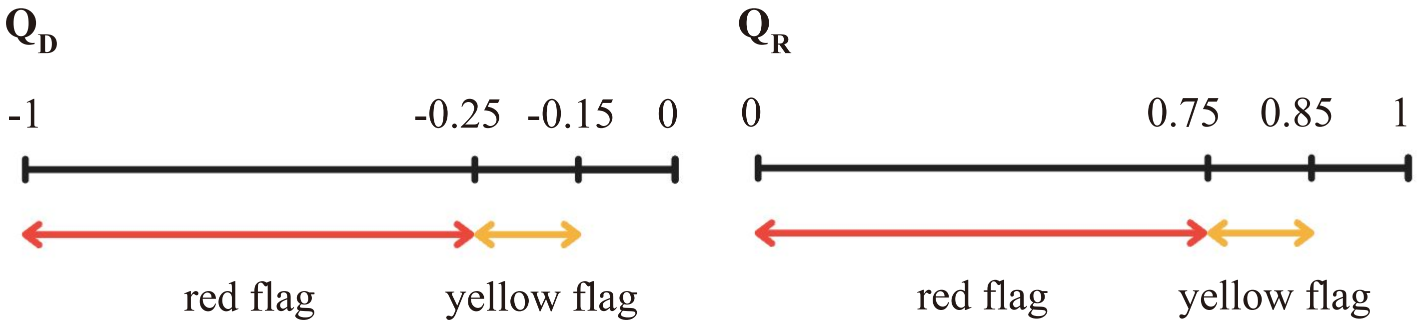
- Data
  - MIMIC (Medical Information Mart for Intensive Care) - III dataset
- Training
  - training the SC-, D-, and R- networks in an offline manner using retrospective data
  - Train: 75%, Validation: 5%, Test: 20%
  - Imbalance of Data: additional data buffer
    - Store the last transition of nonsurvivors trajectories
    - Minibatch (size 64): main data (size 62) + data buffer (size 2)

OUTLINE

- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

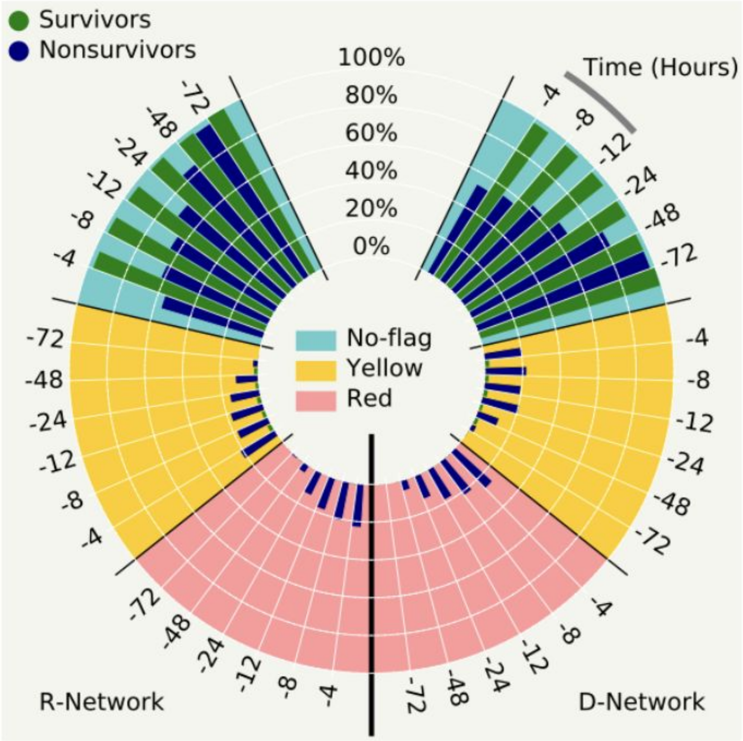
Empirical Results - Septic Dead-End State Prediction

- Experiment
  - To flag potentially non-secure treatments
  - Examine if Q_D and Q_R of each treatment at a given state pass certain thresholds δ_D and δ_R
  - Red flag: δ_D = -0.25 and δ_R = 0.75, minimize both false positives and false negatives
  - Yellow flag: δ_D = -0.15 and δ_R = 0.85 for higher sensitivity and early indication



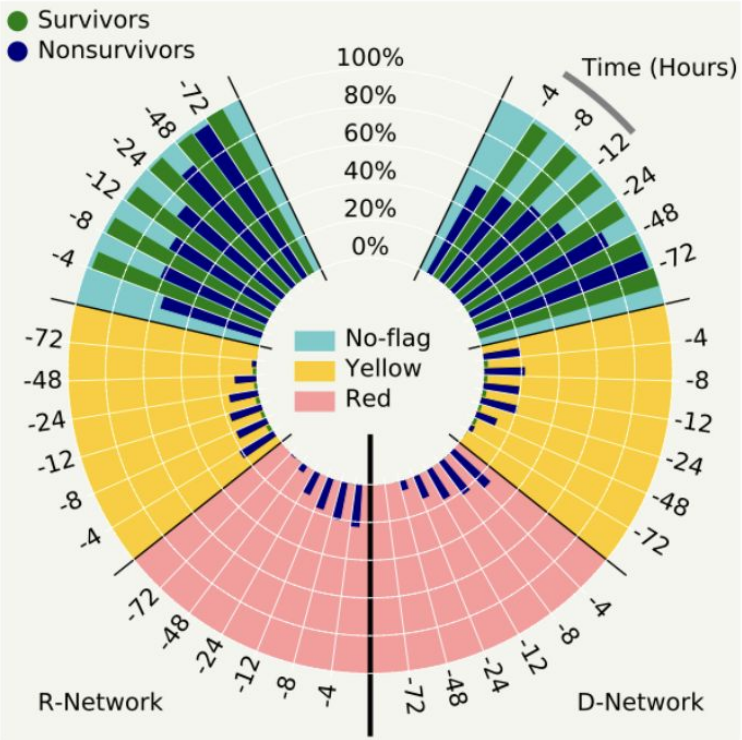
Empirical Results - Septic Dead-End State Prediction

- Results
  - As nonsurvivors approach death, DeD identifies **increasing percentages** of patients raising **fatal flags**
  - Flag emergence for ICU patients
    - A clear worsing trend of state values for non-surviving patients as they approach their terminal state



Empirical Results - Septic Dead-End State Prediction

- Results
  - **Distinctive difference** between the trend of values in **survivors and nonsurvivors**
    - survivors: raise nearly no red flag
    - non-survivors: a steep reduction in no-flag zone with increasing numbers of patients flagged in the Red zone
  - red-flag membership for **long periods** strongly correlates with mortality

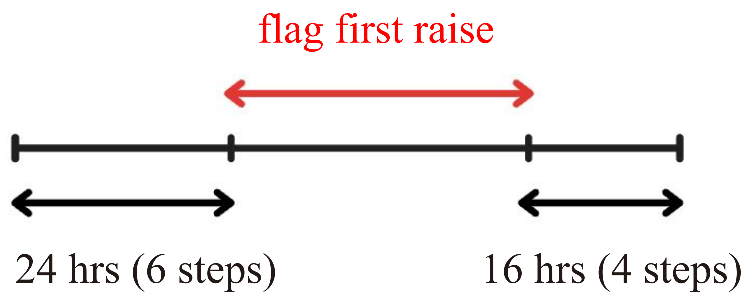


OUTLINE

- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

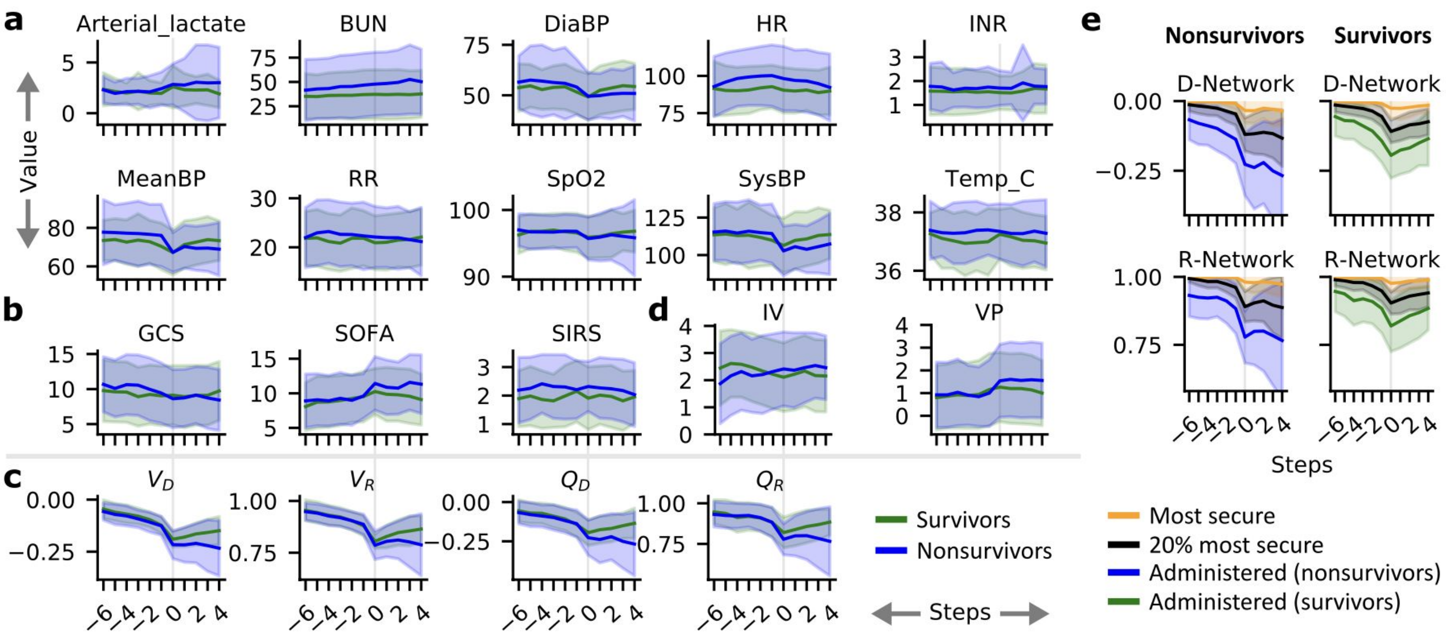
First Flag Analysis

- Experiment
  - To further support our hypothesis that dead-end states exist among septic patients and may be preventable
  - Patient alignment: point select all trajectories in the test data with at least 24 hours (6 steps) prior to the first flag and at least 16 hours (4 steps) afterwards
  - Excluding patients with flags that occur either too early or too late



First Flag Analysis

- Result



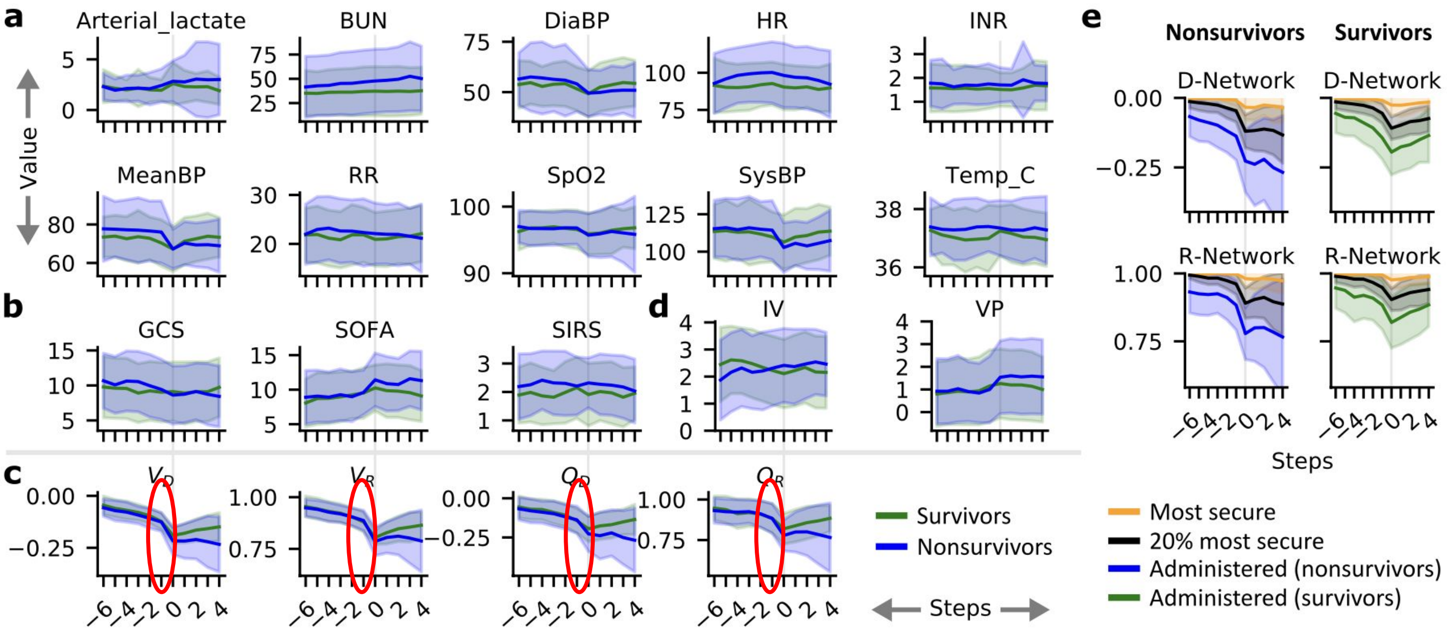
Trend of measures around the first raised flag

First Flag Analysis

- Result
  - Prior to the first flag: V and Q values have similar behavior in survivors and nonsurvivors
  - After the flag is raised: a similar diverging trend among various clinical measures
    - Slight improvement in all value estimates
    - Values of non-surviving patient trajectories quickly collapse
    - Survivors continue to improve

First Flag Analysis

- Main Points
  - DeD identifies a clear **critical point** in the care timeline where non-surviving patients



d

OUTLINE

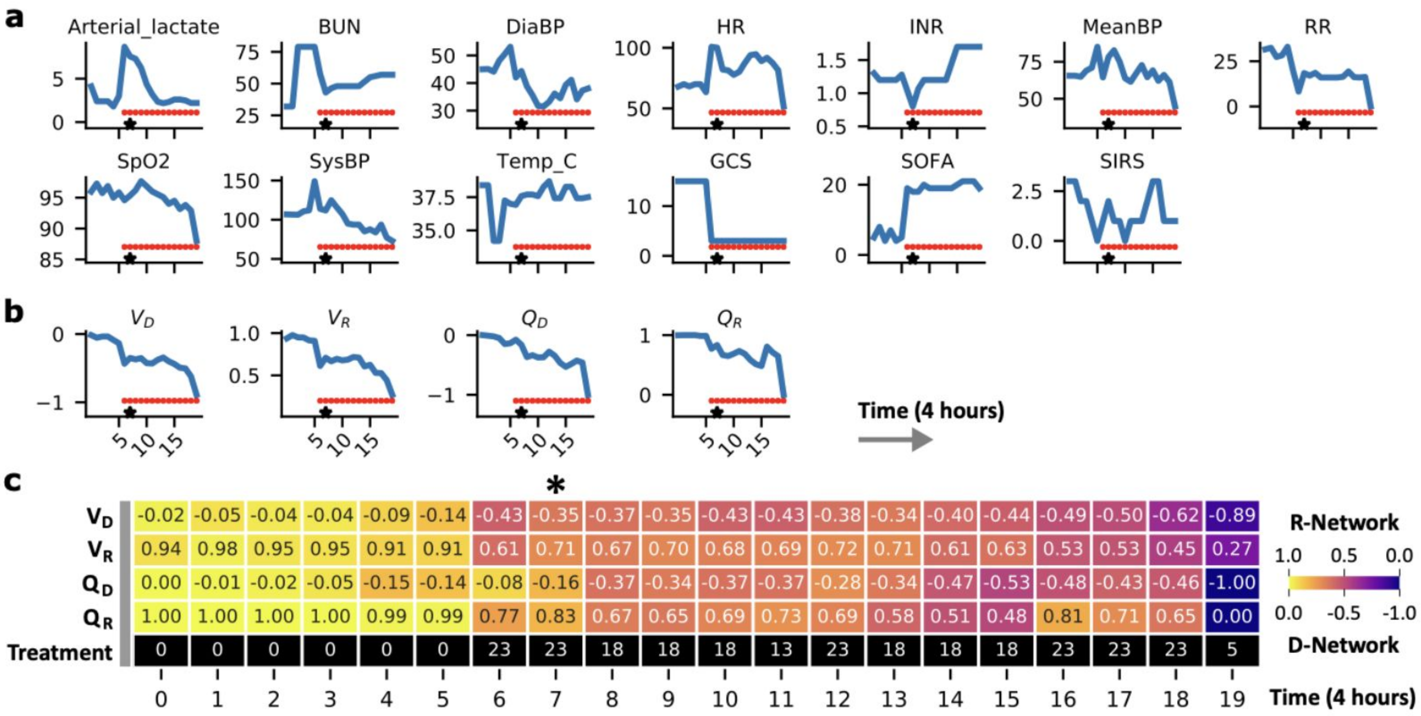
- Introduction
- Related Work
- Methods
  - Math Framework
  - Neural Network Based State Construction and Identification
  - Toy Problem Validation: Life-Gate
- Empirical Setup for Dead-end Analysis
- Empirical Results
  - Septic Dead-End State Prediction
  - First Flag Analysis
  - Individual Trajectories
- Discussion

Individual Trajectories

- Experiment
  - Extracting relevant information from the electronic health record data (EHR)
  - Projecting the state representations of the patient’s trajectory using t-SNE
- Result
  - Certain areas in the t-SNE projection of observed patient states appear to correspond with dead-end states
  - Clinically established measures (SOFA, GCS) closely follow the decrease in DeD estimated values
  - Qualitative analysis suggests that the estimates of Q_D and Q_R are reliable and informative

Individual Trajectories

- Result



OUTLINE

Discussion

- **Key Contributions**
  - Introduced an RL-based method to **avoid harmful treatments**
  - Targeted **dead-ends**, where negative outcomes are inevitable
  - Applied to **sepsis**, a major cause of death, aiming to improve outcomes
- **Impact and Novelty**
  - First RL approach to **flag bad treatments**, not just optimal ones
  - Generic algorithm with **security guarantees**
  - Provides **insights for ICU** interventions, focusing on risky treatments

Discussion

- Applications
  - **DeD** is suited for **safety-critical settings** with limited data
  - Relevant to fields like robotics and industrial control
- Limitations
  - **Extrapolation risk** despite median value use
  - Lack of analysis on **demographic sensitivity**
  - No **external validation** from other hospitals or clinicians

Discussion

- Applications
  - **DeD** is suited for **safety-critical settings** with limited data
  - Relevant to fields like robotics and industrial control
- Limitations
  - **Extrapolation risk** despite median value use
  - Lack of analysis on **demographic sensitivity**
  - No **external validation** from other hospitals or clinicians