# Individualized Fluid and Vasopressor Dosing: A Machine Learning Approach

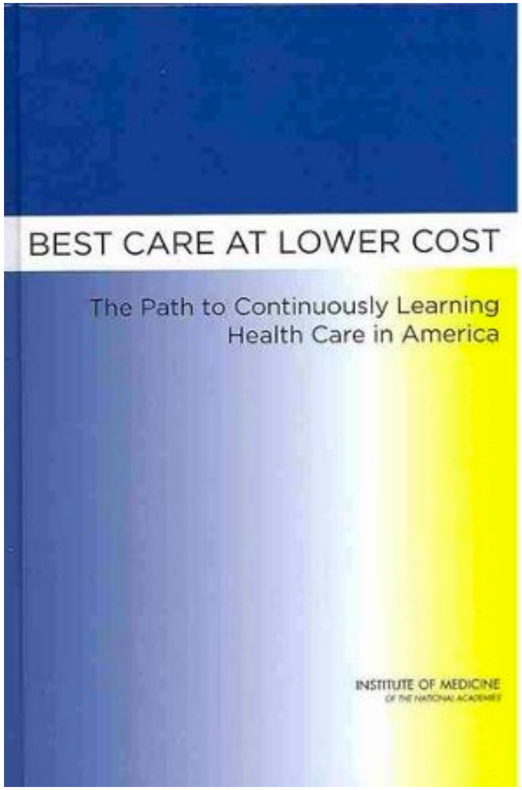
Shamim Nemati[1], Matthew D. Stanley[1], Timothy G. Buchman[1], Fereshteh Razmi[1], Li-wei Lehman[2]

[1]Emory University School of Medicine,

[2]Massachusetts Institute of Technology

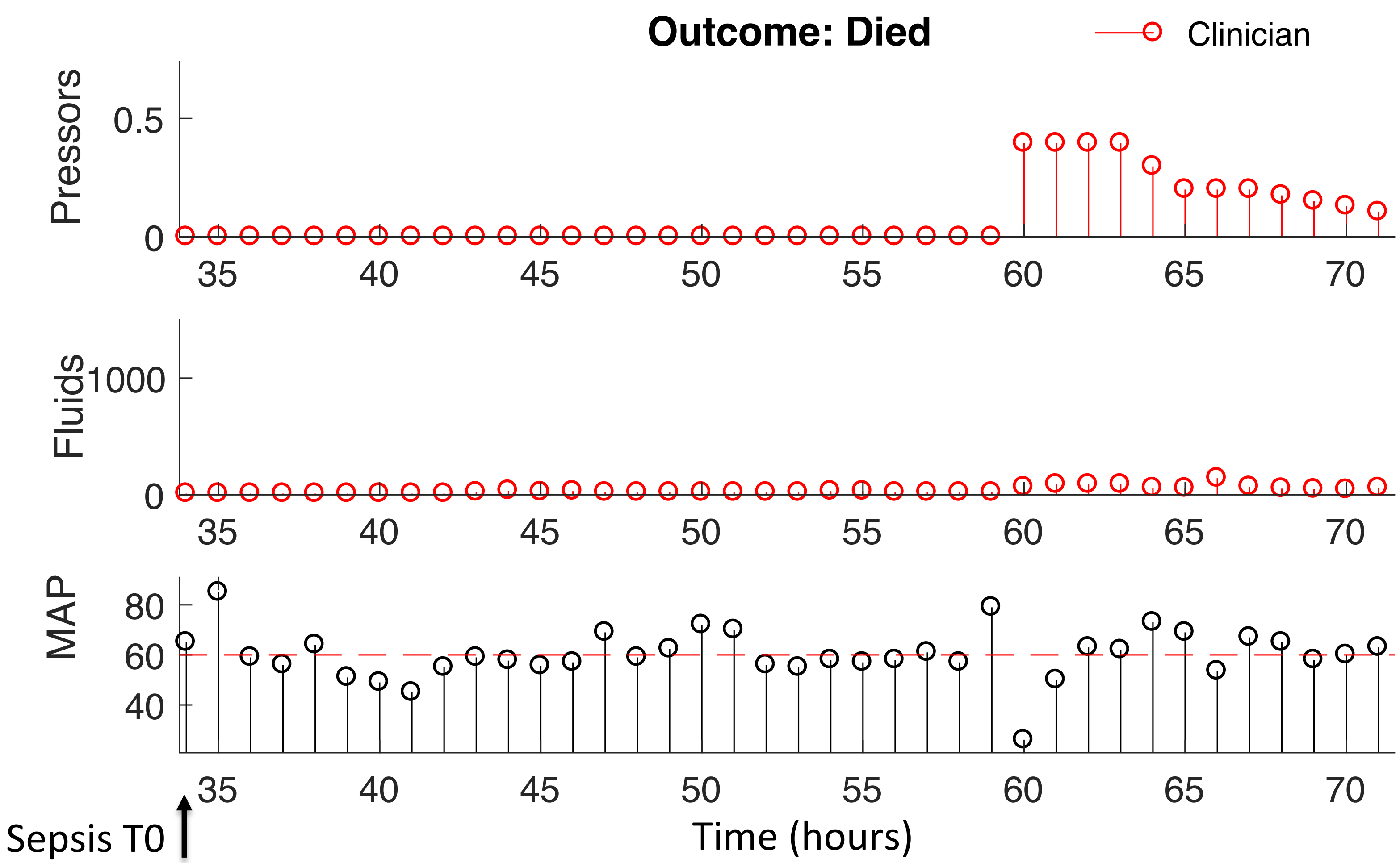
# Continuously Learning Healthcare System

"Achieving higher quality care at lower cost will require fundamental commitments to the incentives, culture, and leadership that foster continuous learning, as the lessons from research and each care experience are systematically captured, assessed, and translated into reliable care." – IOM (2012)




- Learning (fast) requires 'Intelligence'
  - Intelligent systems can learn from experience.
- What is it so special about "each care experience"?
  - Patients are different
  - Variations in care: 'innovation' or 'error'?

# An Example of “Sequential Decision Making”



# Pressors and Fluids Dosing in the ICU

- How effective are our treatments for sepsis?

Seymour CW, Gesten F, Prescott HC, et al: Time to Treatment and Mortality during Mandated Emergency Care for Sepsis. *N Engl J Med* 2017. **376**(23): p. 2235-2244.

> "More rapid completion of a 3-hour bundle of sepsis care and rapid administration of antibiotics, but not rapid completion of an initial bolus of intravenous fluids, were associated with lower risk-adjusted in-hospital mortality."

- Call for personalization ("Precision Medicine")
- Evidence-based medicine: randomized clinical trials (RCT), more randomized clinical trials, meta-analysis
  - Expensive
  - Time-consuming
  - Limitations due to ethical considerations
- Can we instead learn from "care experiences" (clinical 'big data')
  - Sample trajectories of medication dosing and patient response

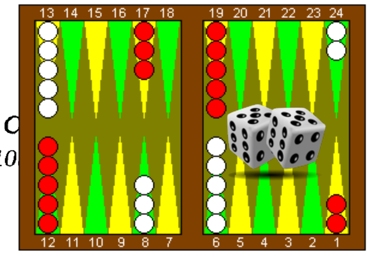
# Taking Inspiration from Gamers and AI

## Examples of Sequential Decision Making

**Neurogammon: A Neural-Network Backgammon Program**

1990s

**Gerald Tesauro**
*IBM Thomas J. Watson Research C*
*PO Box 704, Yorktown Heights, NY 10*




**Abstract**

Neurogammon 1.0 is a complete backgammon program which uses multi-layer neural networks to make move decisions and doubling decisions. The networks were trained by back-propagation on large expert data sets. Neurogammon appears to play backgammon at a substantially higher level than conventional programs. At the recently held First Computer Olympiad in London, Neurogammon won the backgammon competition with a perfect record of 5 wins and no losses, thereby becoming the first learning program ever to win any tournament.

nature

Menu 2015

archive ▸ volume 518 ▸ issue

*NATURE* | LETTER

日本語要約

**Human-level control through deep reinforcement learning**

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg & Demis Has

- Complex Systems
  - **Dimensionality:** Number of poss ations of the pieces is very large.
  - **Stochasticity:** No two g tly the same.
  - **Credit Assign** ment does not always provide immediate feedback, so the value of an s action pair is not always obvious.

# Sequential Decision Making in Medicine

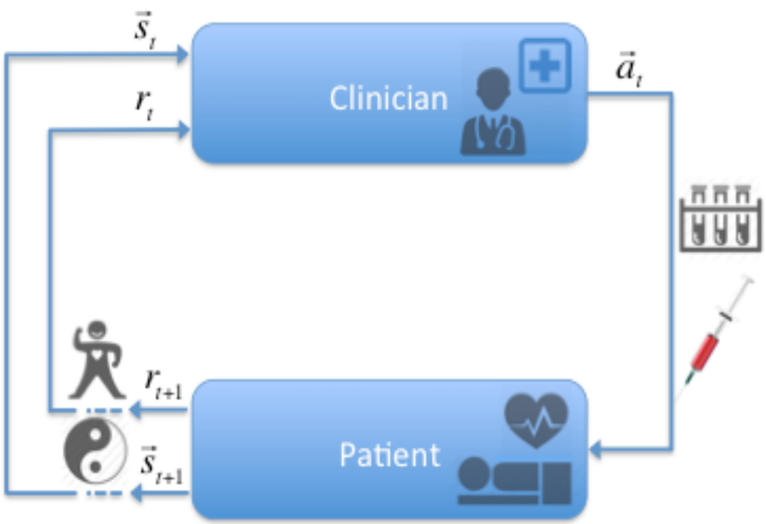
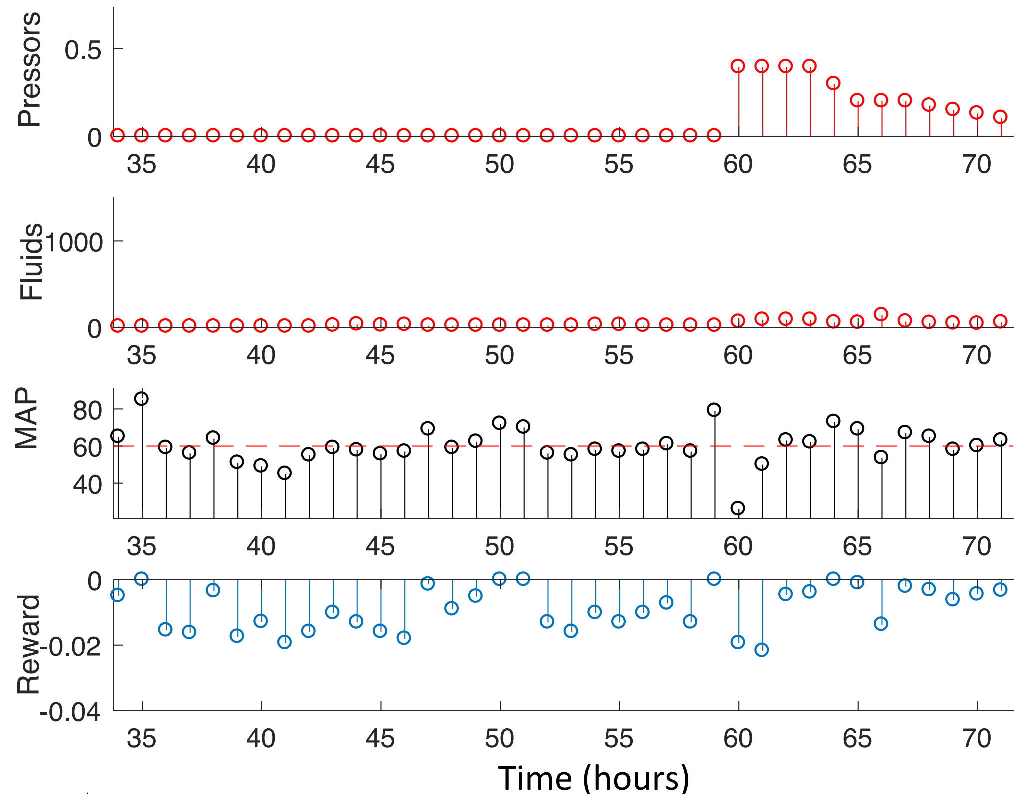
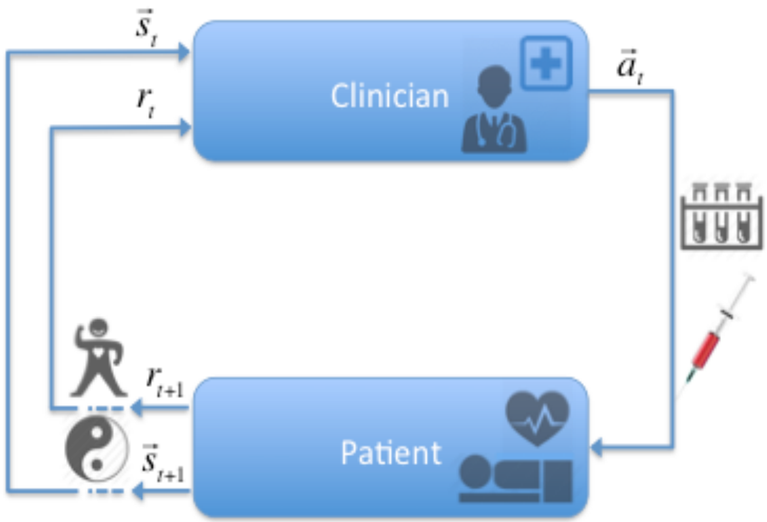



Fig. Generic Sequential Decision Making in Medicine

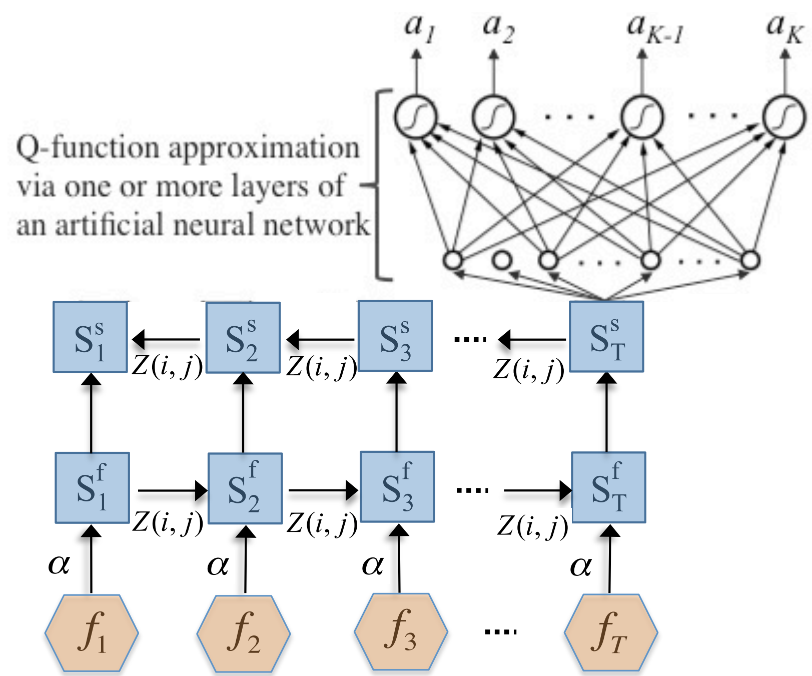



- **State ($s_t$)**: Patient's dynamic (evolving) phenotype
    - Location of objects on a computer screen
- **Actions ($a_t$)**: Increase/Decrease Vasopressor and Fluid Dose by x%
    - Move left, right, up, down
- **Reward ($r_t$)**: Stable MAP (short-term), healthy end-organ (e.g., Kidney) function and survival (long-term).
- **Objective**: Find a (dosing) policy that maximizes accumulated long-term rewards.

# End-to-end State Estimation and Reinforcement Learning




Medication Dosing in the ICU: Case of Heparin



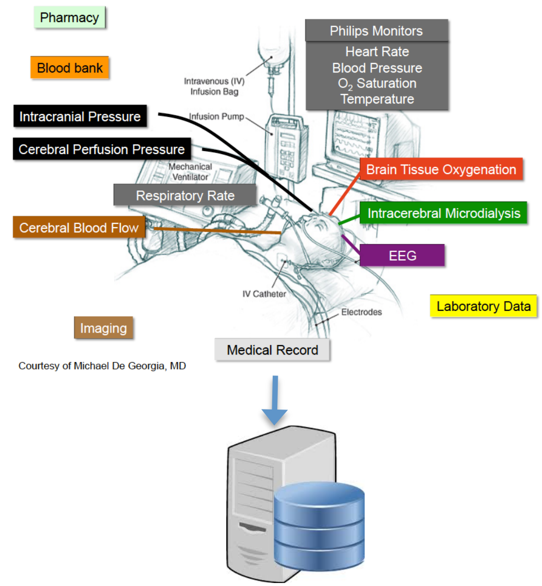

## Continuous Action-Spaces:

- Consider $n$ ($n$=1..$N$) sample trajectories of a patient, define: $R_t^{(n)} = \sum_{t'=t}^{T} \gamma^{t'-t} r_{t'}^{(n)}$
- William's REINFORCE Algorithm: $J(\theta) = \sum_{n=1}^{N} \sum_{t=1}^{T} E_{p(s_{t:T};\theta)}[R_t^{(n)}]$
- We can get a sample approximation to the gradient:

$$\nabla J(\theta) = \frac{1}{N} \sum_{n=1}^{N} \sum_{t=1}^{T} [\nabla_\theta \log \pi(a_t^{(n)} \mid s_{1:t}^{(n)};\theta)] R_t^{(n)}$$

# Data and Study Design

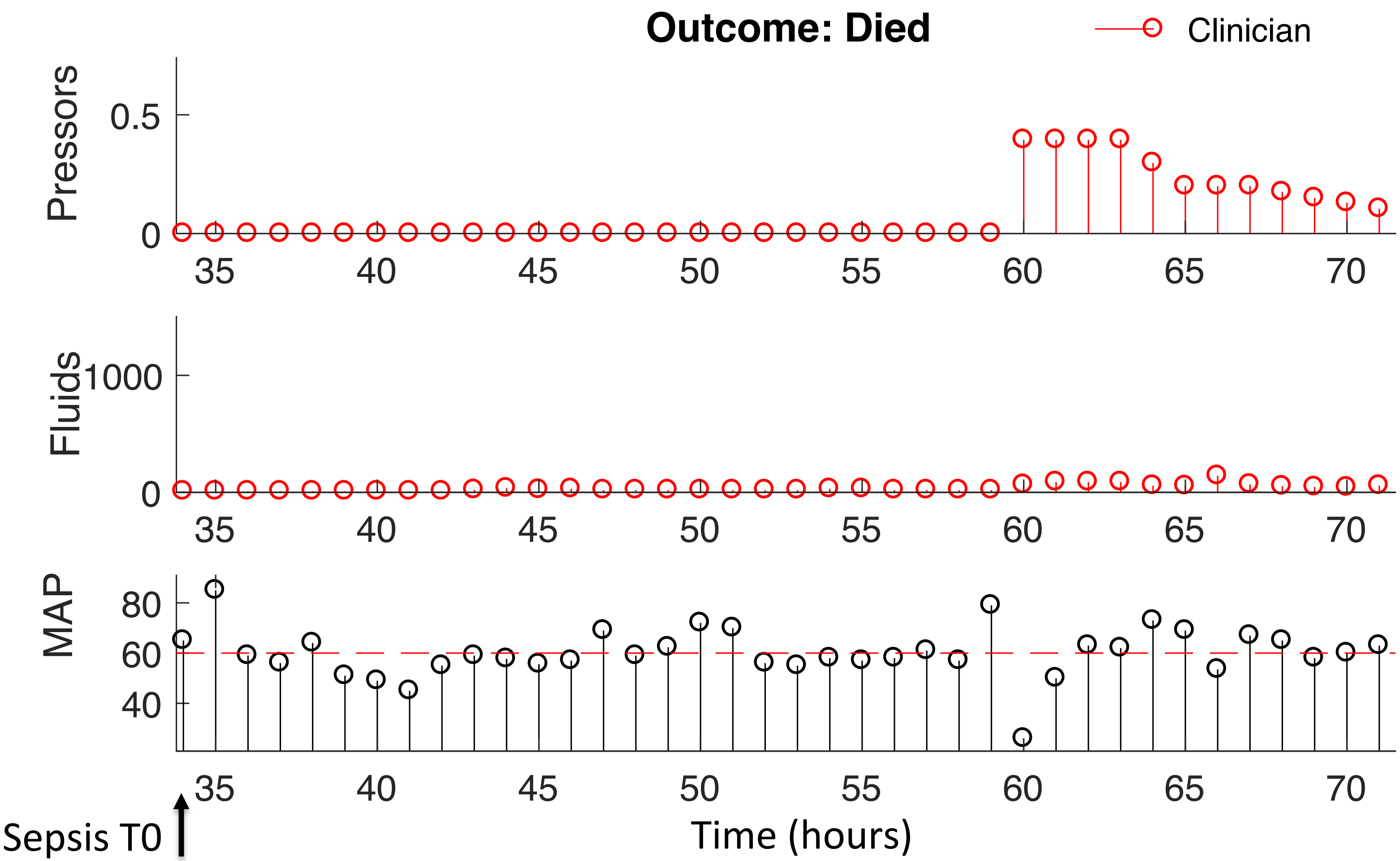
- Over 5,500 ICU patients from the MIMIC-III database with ICU-onset sepsis (sepsis-3 def.)
- **Input**: Clinical and physiological variables (from sepsis onset and the proceeding 72-hour):
  - HR, MAP, SBP, RR, SpO2, Temp, GCS, FiO2, Potassium, Sodium, Chloride, Glucose, BUN, Creatinine, Magnesium, Calcium, Total Bilirubin, Albumin, Hemoglobin, WBC Count, Platelets Count, Arterial pH, PaO2, PaCO2, Base Excess, Lactate,HCO3, PaO2/FiO2, Vasopressor Dose (median), Fluid Balance Input (4 hours), Fluid Balance Output (4 hours), Cumulated Balance, SOFA, MAP(t-1), MAP(t-2), MAP(t-3), Vasopressor Dose(t-1), Vasopressor Dose(t-2), Vasopressor Dose(t-3), Fluid (t-1), Fluid (t-2), Fluid (t-3)
  - Charlson Comorbidity Index (CCI), Age, Gender, Admission SOFA, SAPS II, etc. to adjust for severity of illness of each patient.
    - Over 40 features → mapped to 10D state



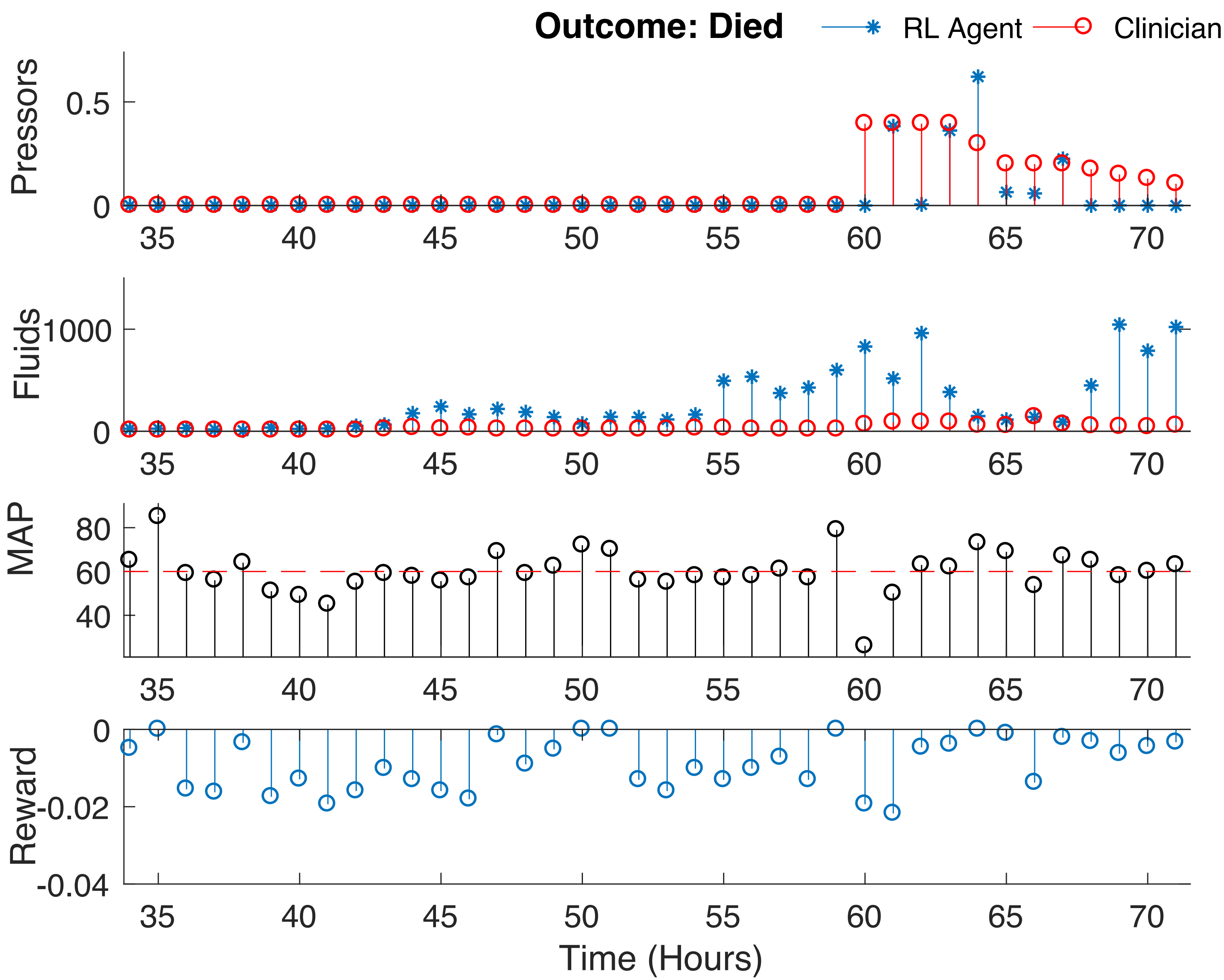

MIMIC-III
(Medical Information Mart for Intensive Care III)
*N*=Thousands

- **Action**: Continuous dose of fluids and pressors (normalized & standardized based on relative strength to norepinephrine)
- **Reward**: stable MAP, low kidney SOFA, survival → Extremely Naïve!
- **Objective:** find a dosing policy that maximizes the expected total reward over proceeding 72 hours.

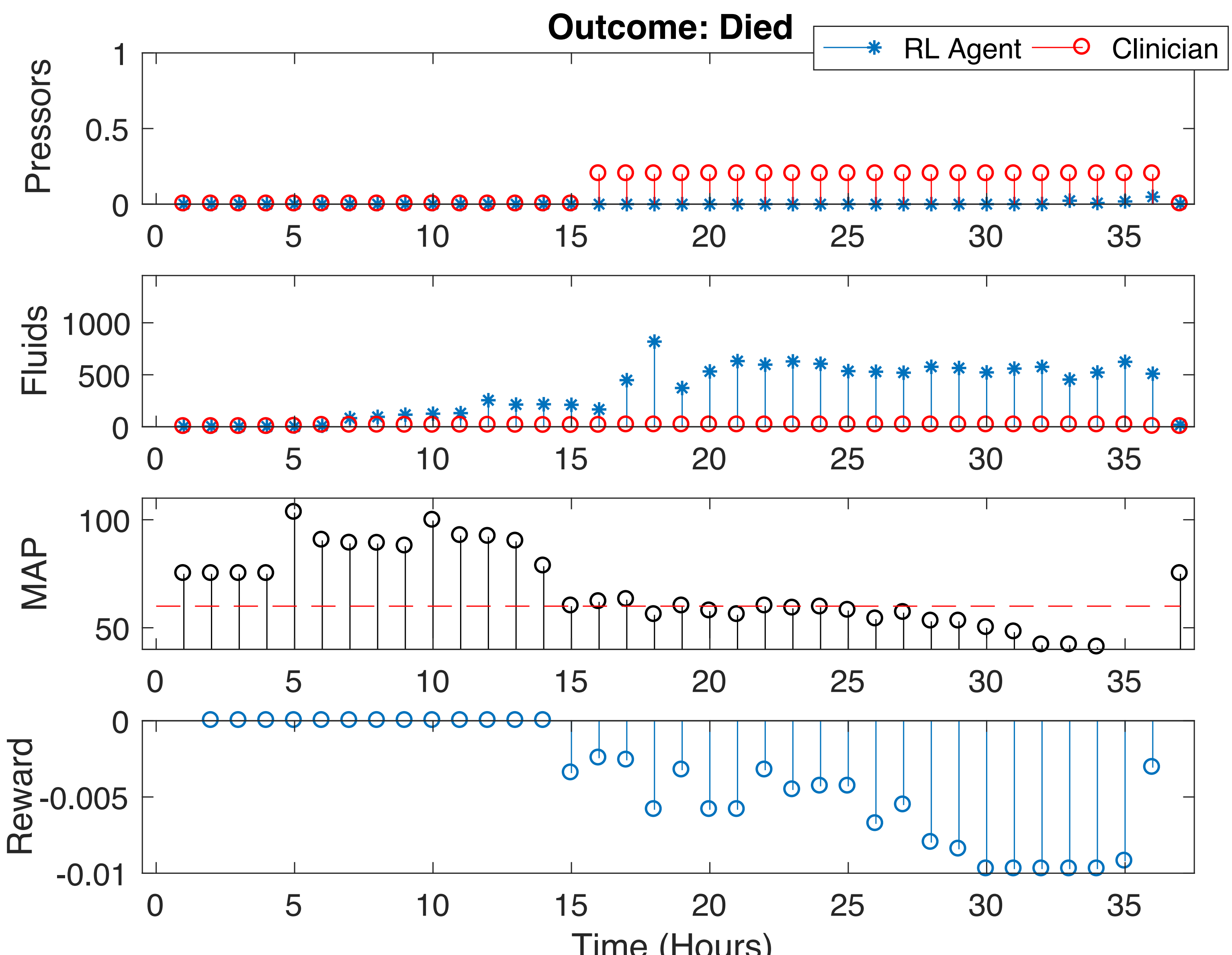
# Example 1:Pressors and Fluid Dosing



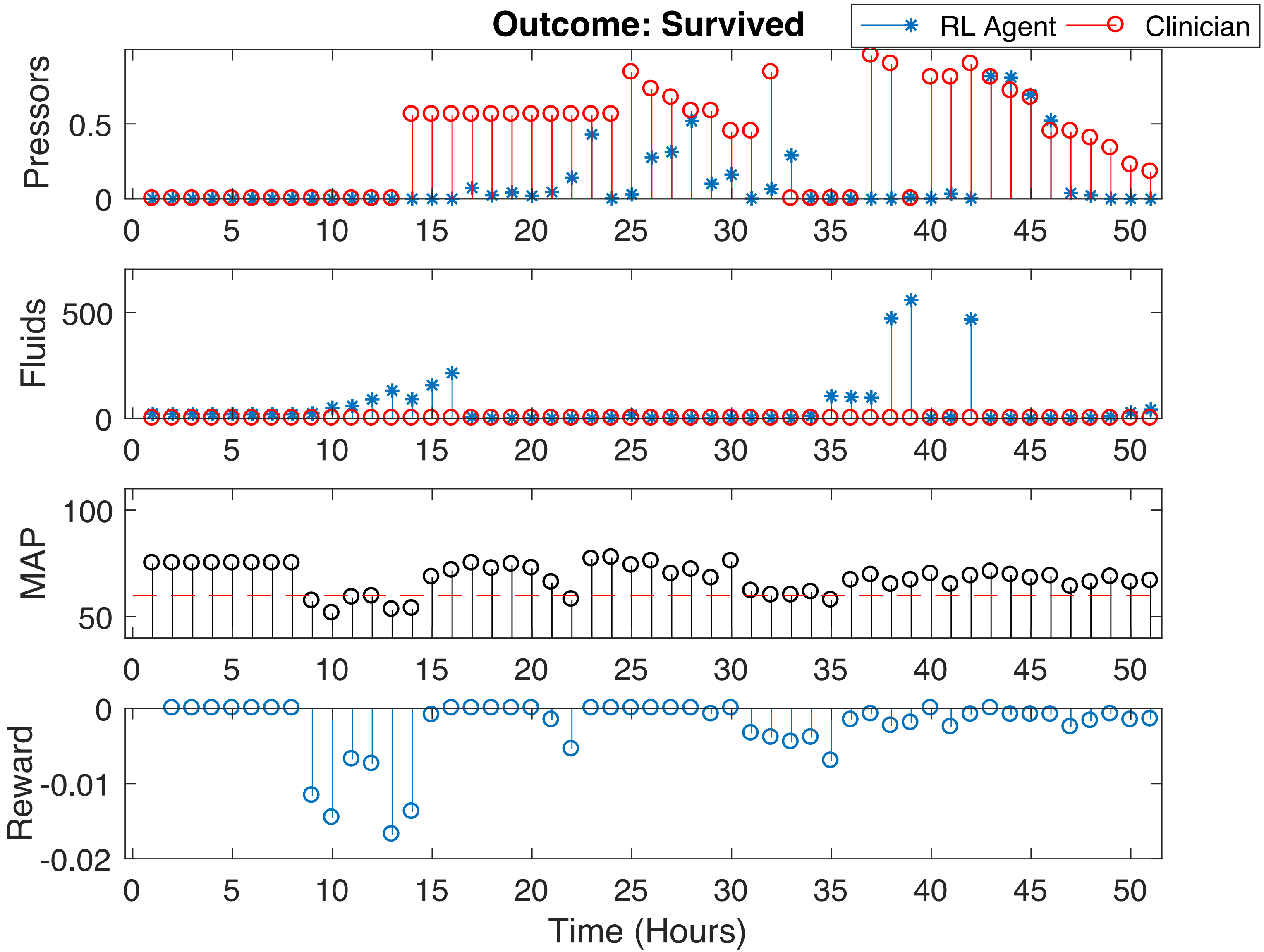
After adjusting for confounding factors both "Time to 1st dose of pressors" and "Volume of 1st dose of pressors" are significantly associated with death.



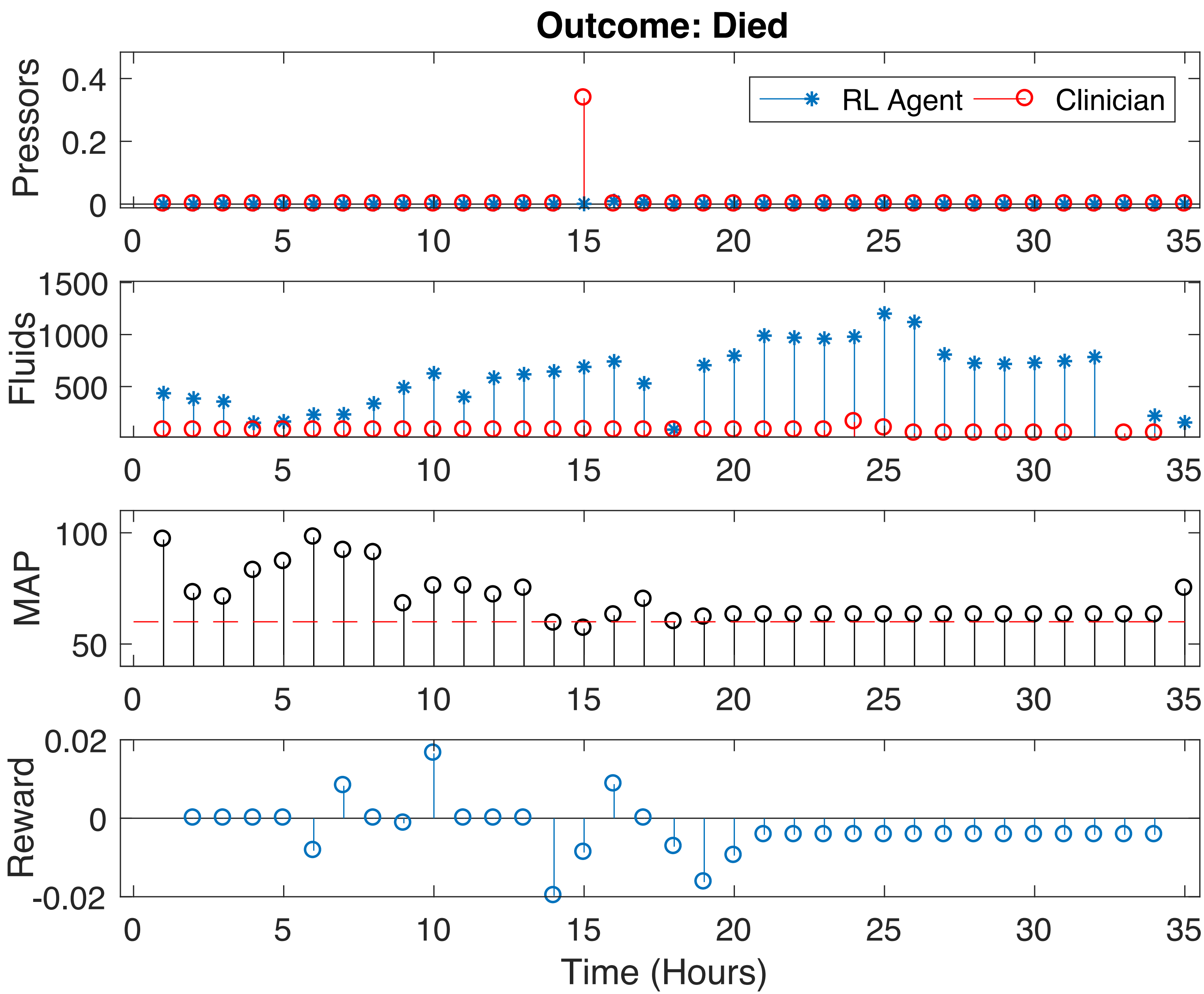
# Example 2:Pressors and Fluid Dosing



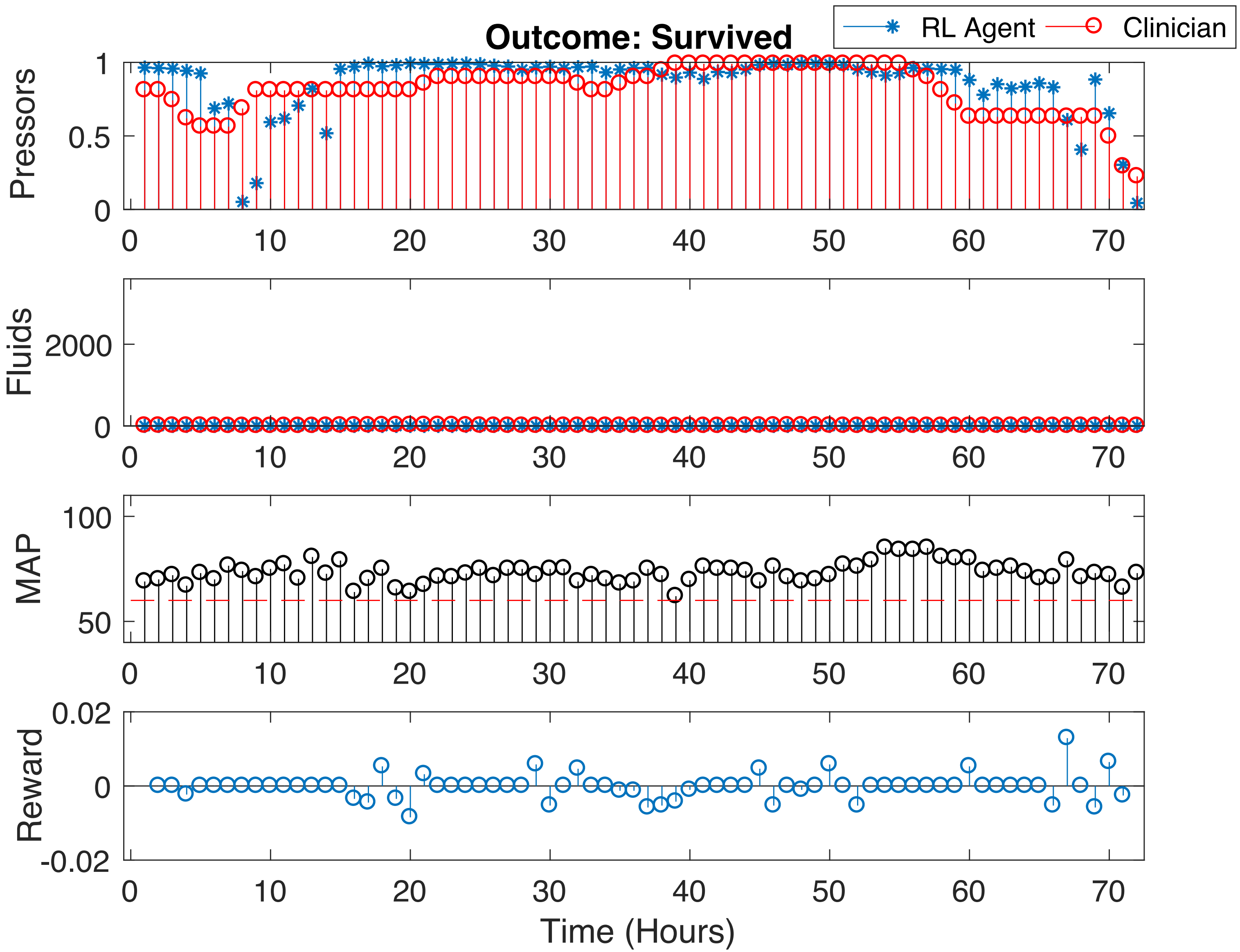
# Example 4:Pressors and Fluid Dosing



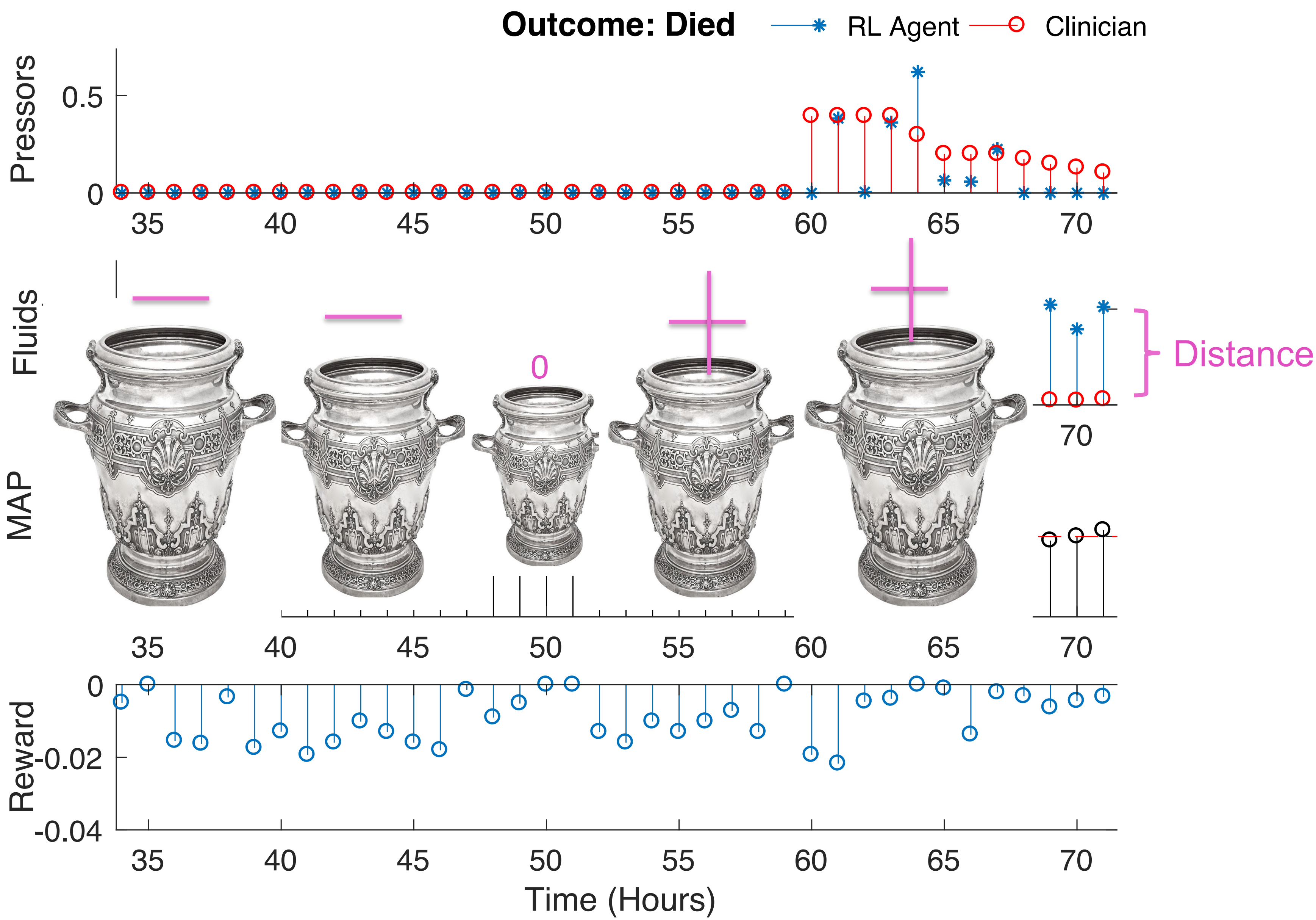
# Example 5:Pressors and Fluid Dosing



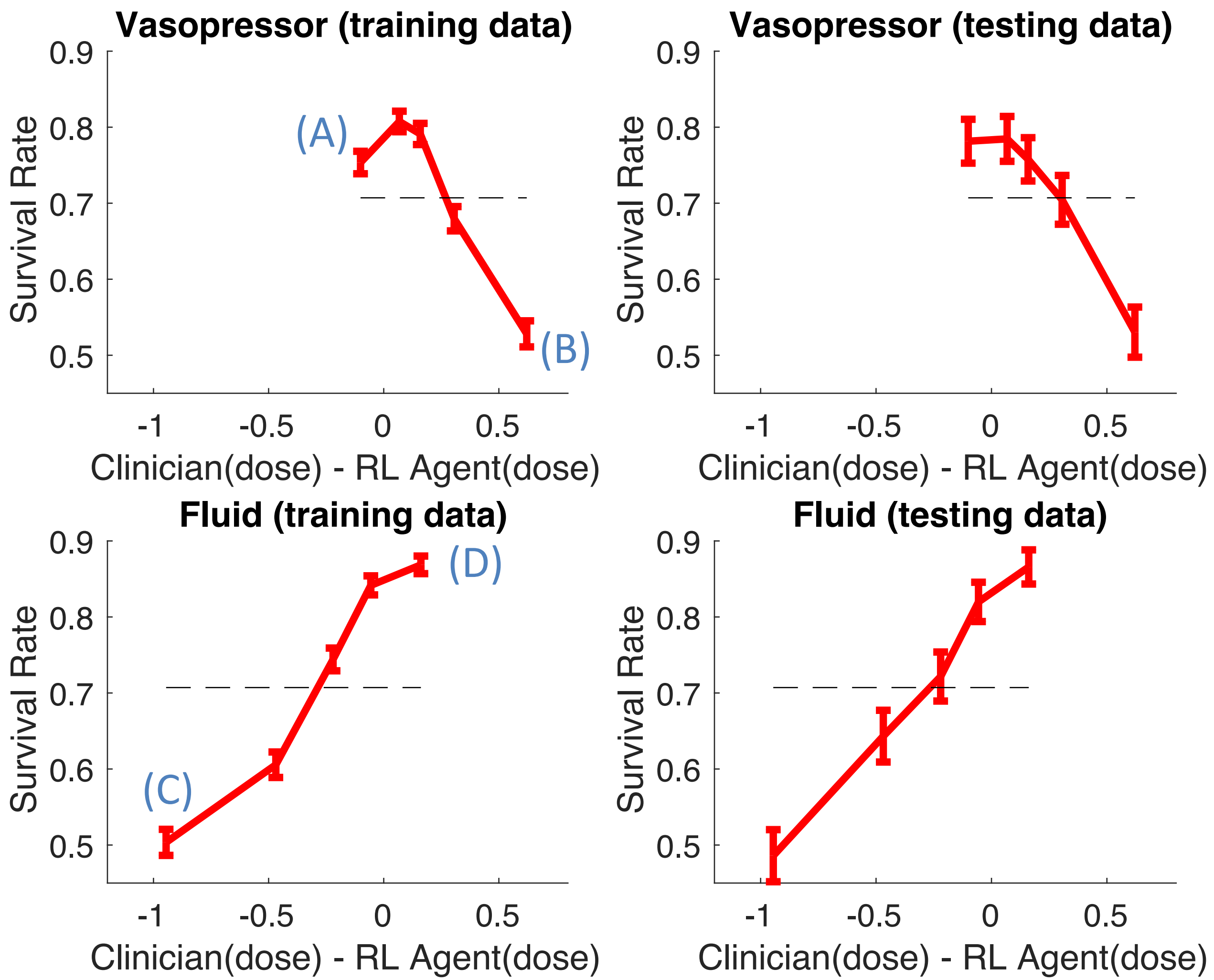
# Example 6:Pressors and Fluid Dosing



# Quantifying Model Performance Across the Population



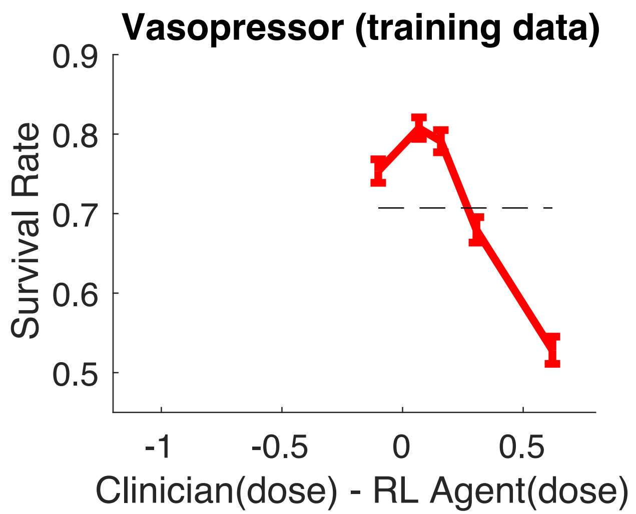
# Assessing Model Performance Across the Population



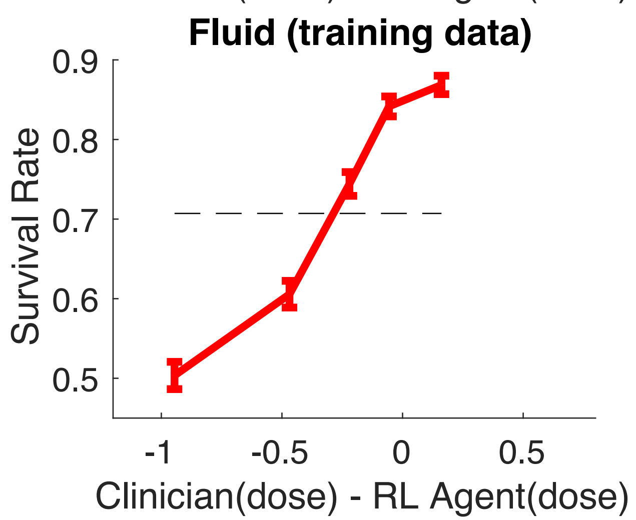

- (A) Patients who are given <u>Pressors</u> at slightly lower rate than the recommendation of the RL agent.
- (B) Patients who are given <u>Pressors</u> at much higher rate than the recommendation of the RL agent.
- (C) Patients who are given <u>Fluids</u> at much lower dose than the recommendation of the RL agent.
- (D) Patients who are given <u>Fluids</u> at slightly higher rate than the recommendation of the RL agent.

# Assessing Model Performance Across the Population

| Features | Survival odds | $p$ (adjusted) |
|---|---|---|
| Bin1 (-0.24<=d<0.03) | **1.26** | 0.002 |
| Bin2 (0.03<=d<0.10) | **1.6** | 0 |
| Bin3 (0.10<=d<0.22) | **1.45** | 0 |
| Bin4 (0.22<=d<0.40) | **0.82** | 0.004 |
| Gender | **1.19** | 0.016 |
| Age | **0.99** | 0 |
| Elixhauser | 1.03 | 0.13 |
| CCI | **0.83** | 0 |



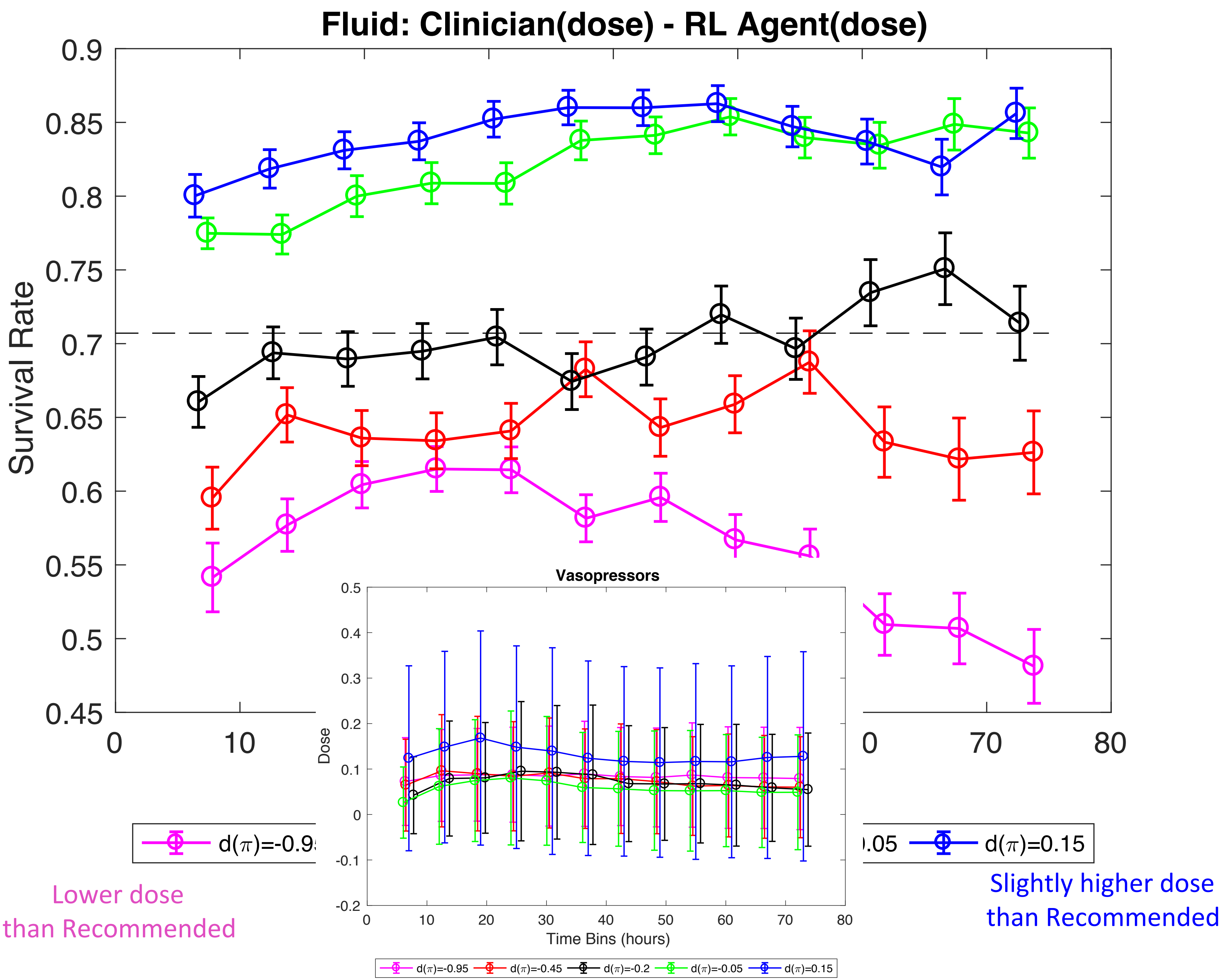

| Features | Survival odds | $p$ (adjusted) |
|---|---|---|
| Bin1 (-1.27<=d<-0.62) | **0.31** | 0 |
| Bin2 (-0.62<=d<-0.32) | **0.53** | 0 |
| Bin3 (-0.32<=d<-0.12) | 1.07 | 0.36 |
| Bin4 (-0.12<=d<0.01) | **2.06** | 0 |
| Gender | **1.57** | 0 |
| Age | 1 | 0.30 |
| Elixhauser | **1.13** | 0 |
| CCI | **0.93** | 0.0001 |

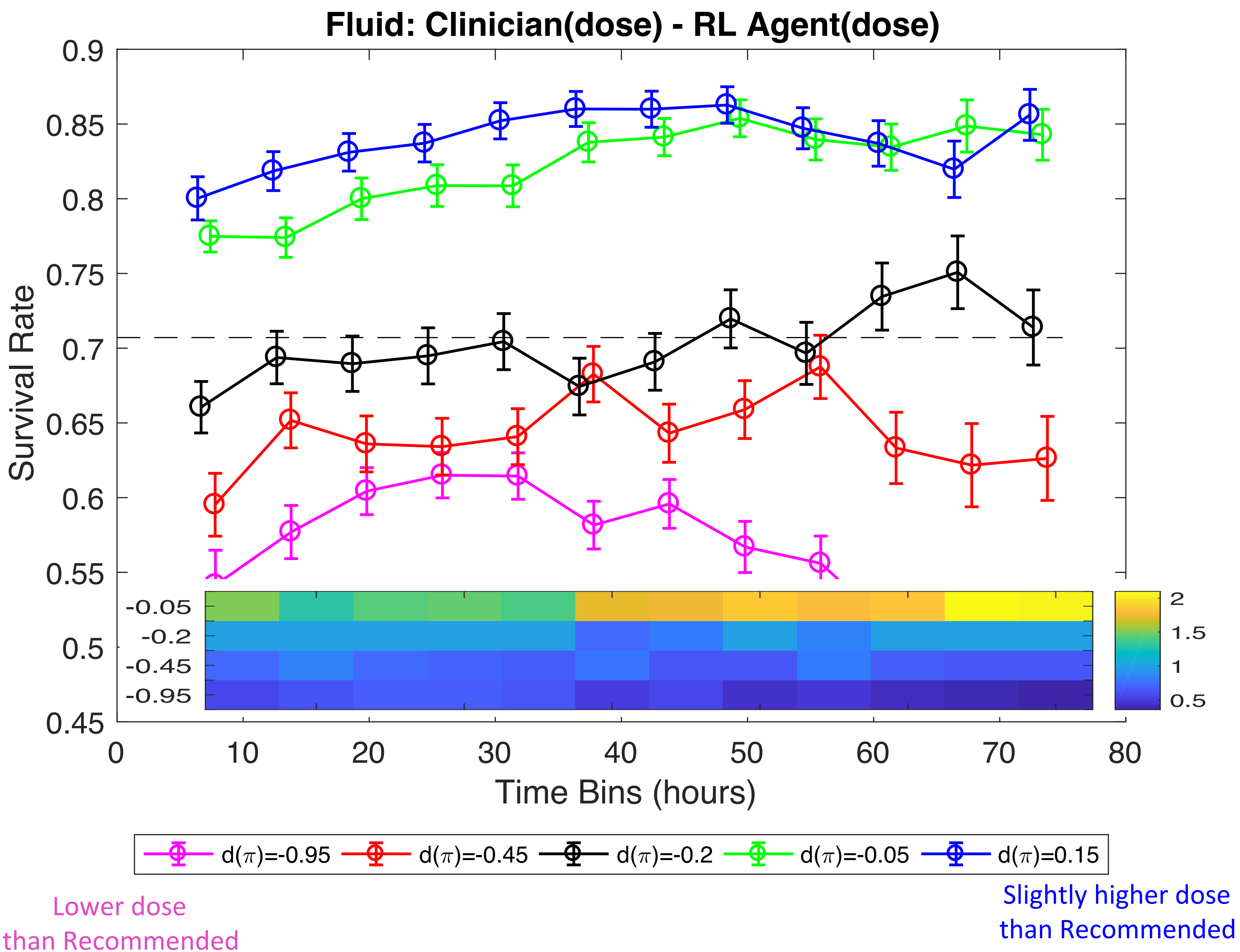


# Summary of Observations

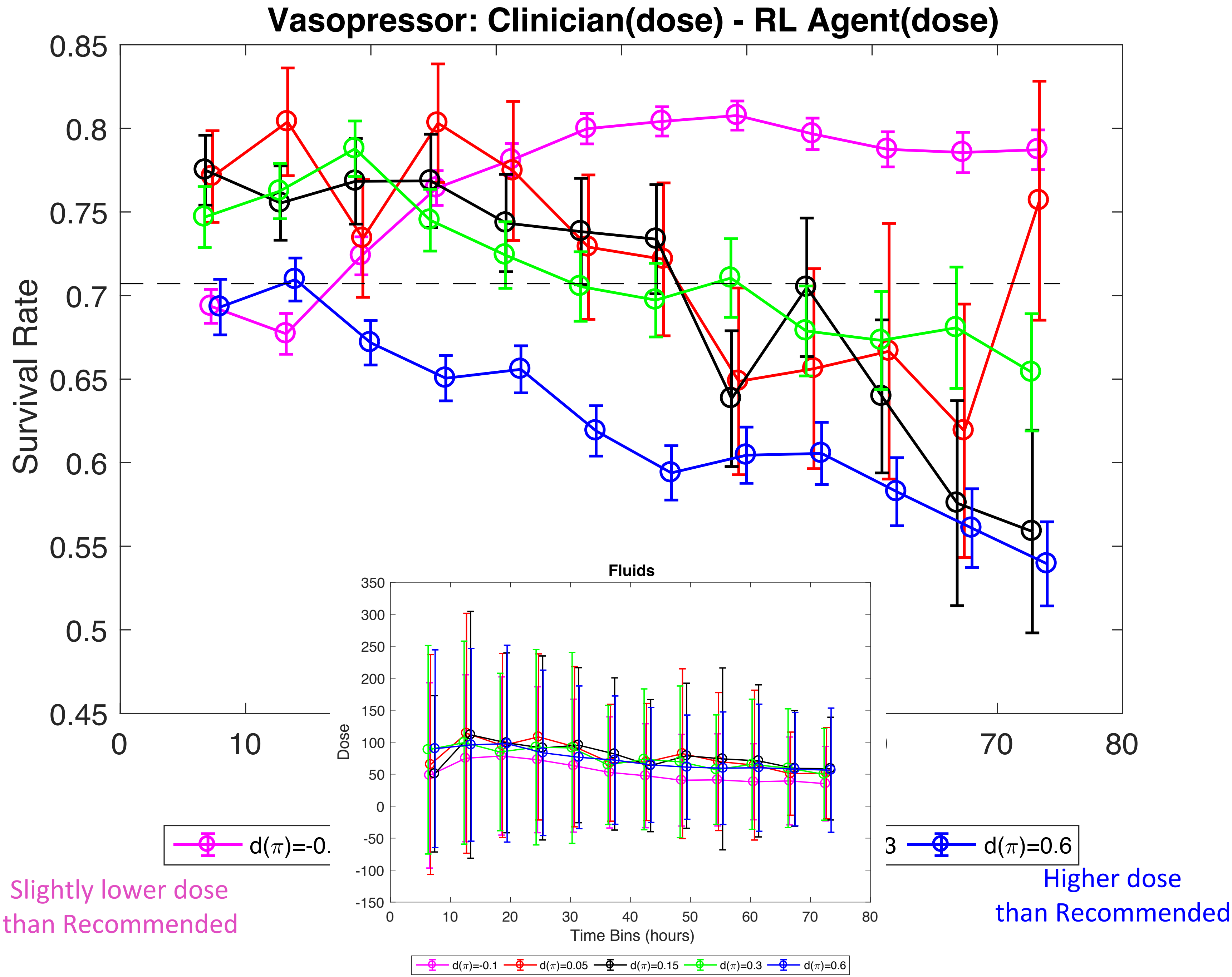
- The RL Agent (AI) learns to personalize dosing:
  1) Compared to a clinician, AI learns to start early with fluids, and gives more fluids over time (than pressors).
  2) It learns to dose pressors at a lower rate than clinicians.
  3) Adherence to the AI's personalized dosing policy is positively associated with Survival, even after adjusting for confounding factors.

- What can we say about the dosing trajectory?
  - We can break down each patient record into <u>6 hours segments</u> (72/6=12 bin) & repeat the previous analysis.

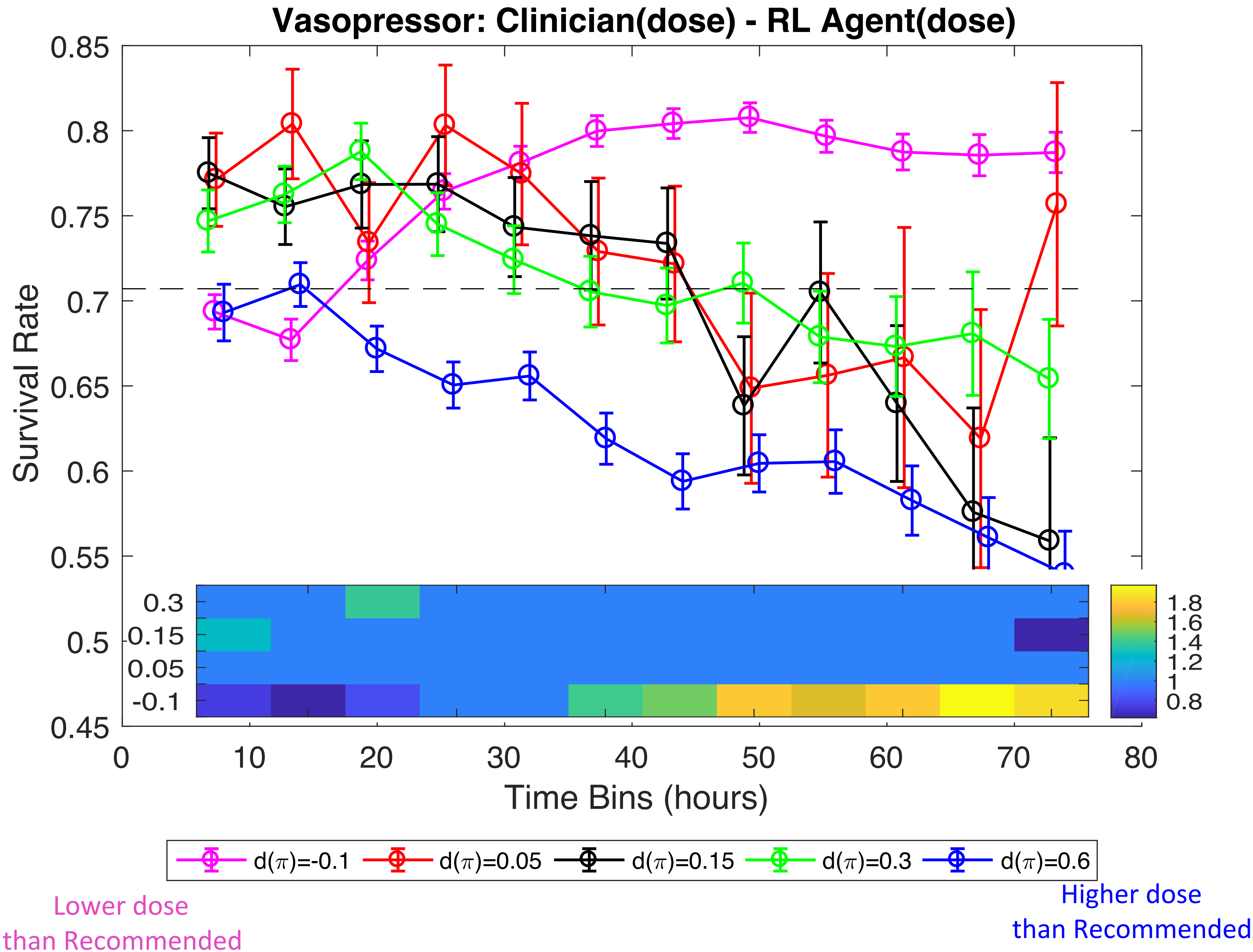

Fluid: Clinician(dose) - RL Agent(dose)
Survival Rate
d(π)=-0.9
.05
d(π)=0.15
Vasopressors
Dose
Time Bins (hours)
d(π)=-0.95
d(π)=-0.45
d(π)=-0.2
d(π)=-0.05
d(π)=0.15
Lower dose
than Recommended
Slightly higher dose
than Recommended



Fluid: Clinician(dose) - RL Agent(dose)
Survival Rate
Time Bins (hours)
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0
10
20
30
40
50
60
70
80
-0.05
-0.2
-0.45
-0.95
2
1.5
1
0.5
d(π)=-0.95
d(π)=-0.45
d(π)=-0.2
d(π)=-0.05
d(π)=0.15
Lower dose
than Recommended
Slightly higher dose
than Recommended

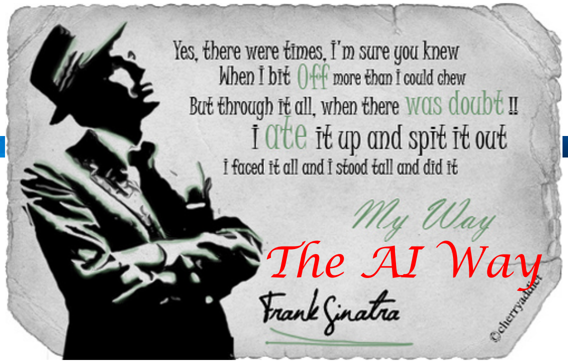


Vasopressor: Clinician(dose) - RL Agent(dose)
Survival Rate
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0
10
70
80
d(π)=-0.
3
d(π)=0.6
Fluids
Dose
Time Bins (hours)
d(π)=-0.1
d(π)=0.05
d(π)=0.15
d(π)=0.3
d(π)=0.6
Slightly lower dose than Recommended
Higher dose than Recommended



Vasopressor: Clinician(dose) - RL Agent(dose)
Survival Rate
Time Bins (hours)
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0
10
20
30
40
50
60
70
80
0.3
0.15
0.05
-0.1
1.8
1.6
1.4
1.2
1
0.8
d(π)=-0.1
d(π)=0.05
d(π)=0.15
d(π)=0.3
d(π)=0.6
Lower dose
than Recommended
Higher dose
than Recommended

# Conclusion

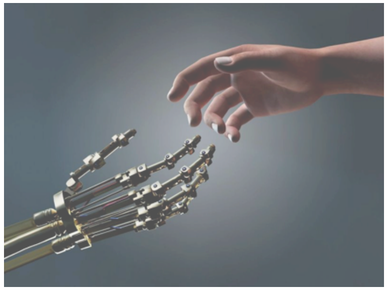



## Observations:

- RL Agent (AI) learns to dose pressors at a lower rate than clinicians
- AI learns to start fluids early and tends to give more fluids over time
- Adherence to AI's personalized dosing is positively associated with Survival, even after adjusting for confounding factors.

## Perspectives:

- "Doctors are Geniuses" → dosing protocols are too simplistic
- Clinician-in-the-Loop Control (GPS systems!) & "Continuous Learning" (IOM)



## Challenges for the future:

- Clinical 'big data' typically includes missing observations/covariates
- Very high-dimensional observation and action spaces (many drugs)
  - Representation Learning, state-estimation, modular and hierarchical design

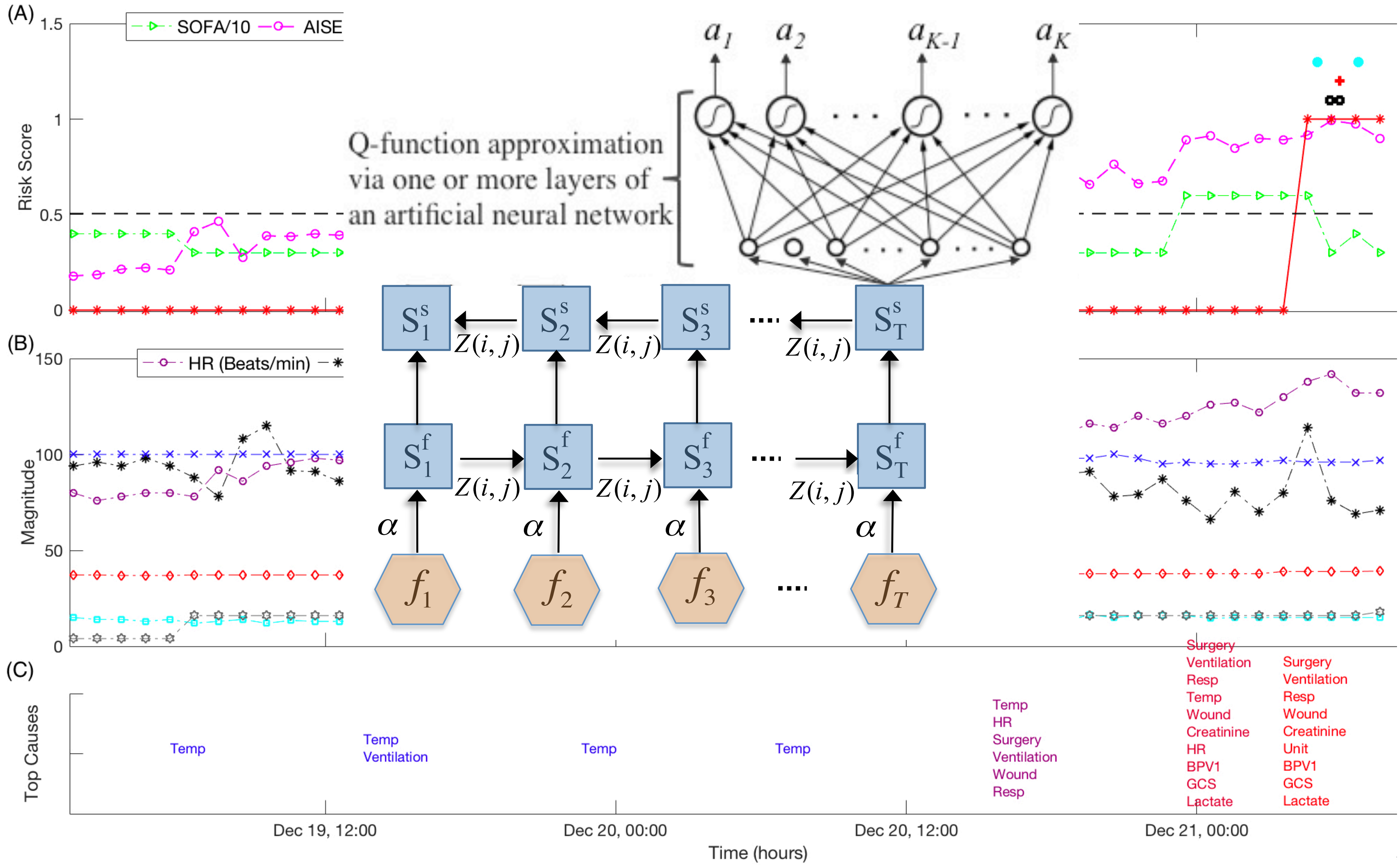
# *Thank You! Questions?*

**Shamim Nemati**
Assistant Professor, Biomedical Informatics, Emory University
Email: shamim.nemati@alum.mit.edu
Website: http://NematiLab.info

**Fundings and Acknowledgements:**

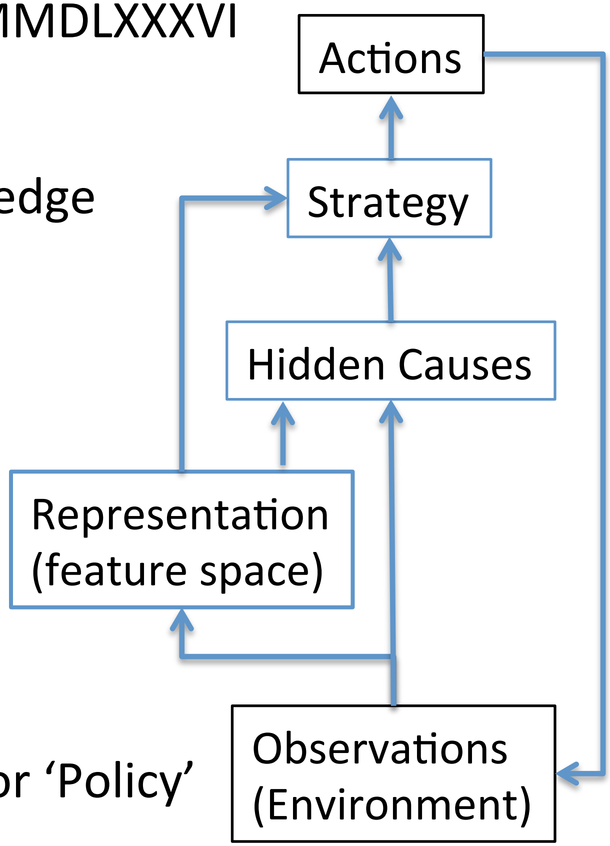
- Ruth L. Kirschstein National Research Service Award (2010-2012)
- James S. McDonnell Postdoctoral Fellowship (2013-2014)
- NIH BD2K Career Development Award (K01) (2015-2020)
- Healthcare Innovation Seed Grant, Emory/Georgia Tech (2016-2017)
- Emory Radiology-BMI Seed Grant (2016-2017)

# Not a blackbox: RL Agent Can Reason!



# Evolution of Intelligence (David Krakauer)

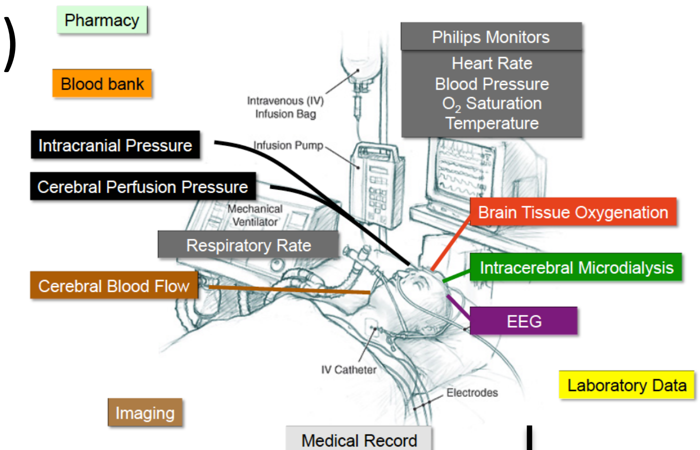
- Representation
  - 3998 in Arabic numerals versus Roman numerals: MMMCMXCVIII
  - 42 x 133 = 5586 vs. XLII x CXXXIII = MMMMMDLXXXVI
- Inference (AI / Machine Learning)
  - Going from known knowledge to unknown knowledge
  - Hidden or latent causes, nechanisms
- Strategy (Control System)
  - From causes/mechanism to action/therapy
  - Sequential Decision Making
- Competition
  - Survival of the fittest strategy
  - A form of optimization: what is the best strategy or 'Policy'



# Intelligent Systems in Health Care

## Rule-based Expert Systems (80s-90s)

- Simulate the thinking processes of the human medical expert
- MYCIN (infectious diseases), ONCOCIN (oncology), Quick Medical Reference
- Diagnosis/<u>directly assist decision-making</u>




Courtesy of Michael De Georgia, MD

## Data-driven Machine Learning (2000-Current)

- Better representation (features) or visualization
- Prediction/classification/phenotype discovery
- <u>Indirectly assist with decision-making</u>

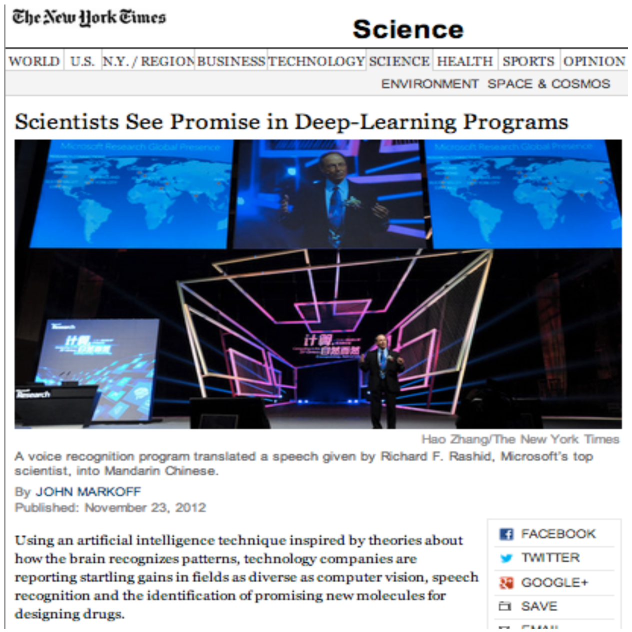
Modern Clinical Databases
(*N* of Thousands)

## Data-driven Expert Systems

- Derives structures (features) and rules from data
- <u>Directly assists with decision-making</u>, but
- A Global Positioning System (GPS)
- Moreover, can learn from “every ride” or “every care experience”

# How can we Build a Data-driven Continuously Learning HealthCare System?

- Big Data
  - Volume and Variety, Velocity and Veracity
  - Lots and lots of examples
- Deep Learning
  - Representation Learning
  - Finding 'Regularities' across examples
  - Latent state estimation / Perception
- Reinforcement Learning
  - Sequential decision making / Control


The New York Times — Science

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION

ENVIRONMENT SPACE & COSMOS

Scientists See Promise in Deep-Learning Programs

Hao Zhang/The New York Times

A voice recognition program translated a speech given by Richard F. Rashid, Microsoft's top scientist, into Mandarin Chinese.

By JOHN MARKOFF

Published: November 23, 2012

Using an artificial intelligence technique inspired by theories about how the brain recognizes patterns, technology companies are reporting startling gains in fields as diverse as computer vision, speech recognition and the identification of promising new molecules for designing drugs.

FACEBOOK

TWITTER

GOOGLE+

SAVE

➢ What do <u>Atari games</u>, <u>Backgammon</u>, and <u>dosing of Vasopressors and Fluids</u> have in common?
  - ➢ Sequential decision making
  - ➢ We can learn optimal decision rules from clinical examples ('big data')

# You can ignore this slide … but if you pay attention you may increase your long-term reward!

- Optimal action-value function:

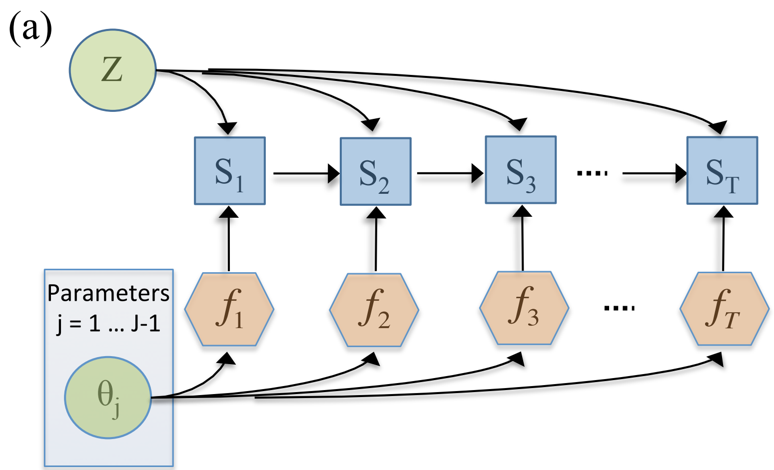
$$Q^*(s,a) = \max_{\pi} E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots \mid s_t = s, a_t = a, \pi]$$

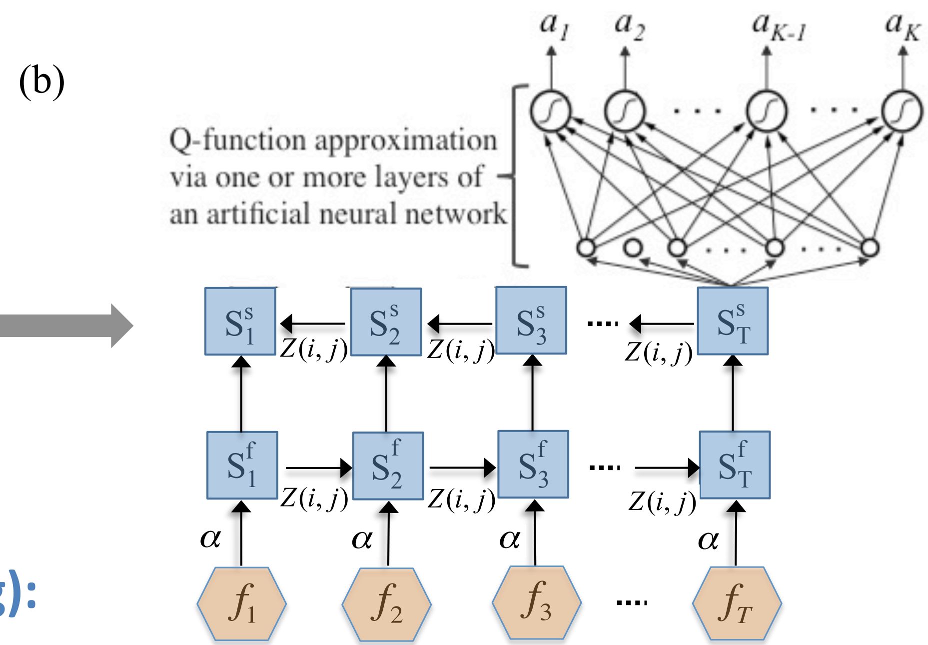
- Dynamic Programming (recursive estimation)

$$Q^*(s,a) = E_{s'}[\ \underbrace{r(s')}_{\text{immediate reward}} \ + \ \gamma \underbrace{\max_{a'} Q^*(s',a')}_{\text{subsequent rewards (discounted)}} \ \mid s,a]$$

- Two Problems to address:
  - State estimation
  - Learning of Optimal Policy (or the Q-function)

# End-to-end State Estimation and Reinforcement Learning




## Discrete Action-Spaces (Q-Learning):

- Optimal state-action value function:

$$Q^*(s,a) = \max_{\pi} E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots \mid s_t = s, a_t = a, \pi]$$

- Dynamic Programming (recursive estimation):

$$Q^*(s,a) = E_{s'}[\ r(s')\ +\ \gamma \max_{a'} Q^*(s',a')\ \mid s,a]$$

- (Neurally) fitted Q-iteration:

$$Loss(\theta_i) = E_{s,a,r,s'}[\ (Q(s,a;\theta_i)\ -\ y)^2],\ \text{where}\ \ y = r(s') + \gamma \max_{a'} Q(s',a';\theta_{i-1})$$

| | **Clinician**<br>**Mean [5%, 95%]** | **RL Agent** |
|---|---|---|
| Time to 1st Pressor | **6** [1, 40] | **11** [1, 49] * |
| Time to 1st Fluid | **4** [1, 22] | **1** [1, 21] * |
| Volume of 1st Pressor | **0.12** [0.03, 0.068] | **0.02** [0.003, 0.88] * |
| Volume of 1st Fluid | **34** [5, 1000] | **28** [4.5, 1085] |
| Mean Pressor | **0.02** [0.001, 0.27] | **0.0001** [0, 0.17] * |
| Mean Fluid | **23** [1 138] | **89** [1, 765] * |
| Maximum Pressor | **0.22** [0.04, 0.9] | **0.05** [0 0.99] * |
| Maximum Fluid | **411** [19.9, 1600] | **964** [25.5 1505] * |