The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care

Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med*. 2018 ;24(11):1716-1720.

Hyeonhoon Lee

RI paper study (8th)

2022.07.18.

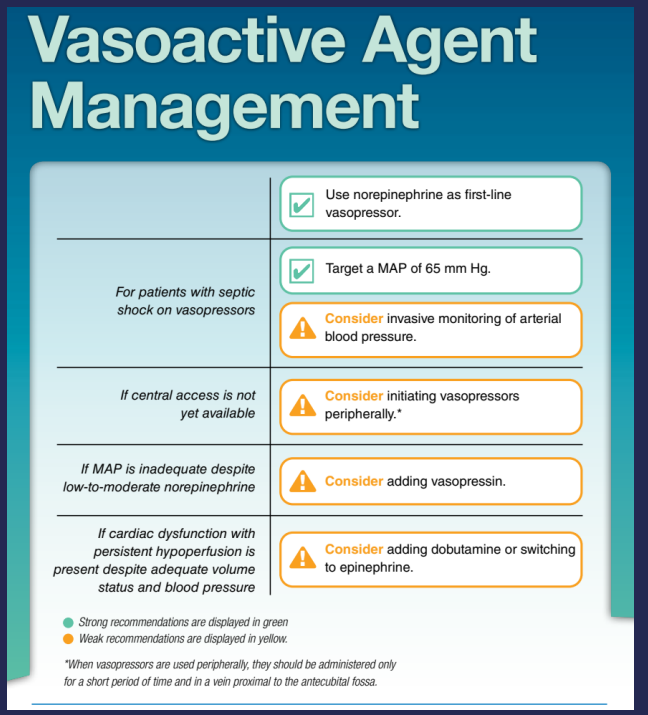
Contents

Introduction

Sepsis in ICU

- A severe *infection* leading to life-threatening *acute organ dysfunction.*
  - *Suboptimal* decisions lead to poorer outcomes.
  - In the last 10–15 years, attempts to develop new treatments to reduce sepsis mortality have uniformly been unsuccessful.

- Challenging in management of *intravenous fluids* and *vasopressors*
  - General guideline: Surviving Sepsis Campaign 2021
  - Real-time, personalized treatment decision?



Komorowski M et al. Nat Med. 2018.

RL for clinical setting

	(Offline) RL	Clinicians
Policy objective	Maximizing an expected return	Maximizing a patient's probability of good outcome
Model	Handling sparse reward signals	Heterogeneous responses to interventions Delayed indications of the efficacy of treatments
Develop	Fixed data from behavior policy	Data (EMR) from suboptimal decisions by clinicians
Deploy	Real-time inference	Real-time decision-making

Komorowski M et al. Nat Med. 2018.

Methods

Patient cohort

- Two large *nonoverlapping* ICU (Intensive care unit) databases (United states)
  - Medical Information Mart for Intensive Care version III (MIMIC-III) : Model development and internal validation
  - eICU Research Institute Database (eRI): external validation

Table 1 | Description of the datasets

	MIMIC-III	eRI
Unique ICUs (*n*)	5	128
Unique ICU admissions (*n*)	17,083	79,073
Characteristics of hospitals, per number of ICU admissions	Teaching tertiary hospital	Nonteaching: 37,146 (47.0%) Teaching: 29,388 (37.2%) Unknown: 12,539 (15.9%)
Age, years (mean (s.d.))	64.4 (16.9)	65.0 (16.7)
Male gender (*n* (%))	9,604 (56.2%)	40,949 (51.8%)
Premorbid status (*n* (%))		
Hypertension	9,384 (54.9%)	43,365 (54.8%)
Diabetes	4,902 (28.7%)	25,290 (32.0%)
CHF	5,206 (30.5%)	15,023 (19.0%)
Cancer	1,803 (10.5%)	11,807 (14.9%)
COPD or RLD	4,248 (28.7%)	18,406 (23.3%)
CKD	3,087(18.1%)	14,553 (18.4%)
Primary ICD-9 diagnosis (*n* (%))		
Sepsis, including pneumonia	5,824 (34.1%)	41,396 (52.3%)
Cardiovascular	5,270 (30.8%)	11,221 (14.2%)
Respiratory	1,798 (10.5%)	9,127 (11.5%)
Neurological	1,590 (9.3%)	7,127 (9.0%)
Renal	429 (2.5%)	1,454 (1.8%)
Others	2,172 (12.7%)	8,747 (11.1%)
Initial OASIS (mean (s.d.))	33.5 (8.8)	34.8 (12.4)
Initial SOFA (mean (s.d.))	7.2 (3.2)	6.4 (3.5)
Procedures during the 72 h of data collection:		
Mechanical ventilation (*n* (%))	9,362 (54.8%)	39,115 (49.5%)
Vasopressors (*n* (%))	6,023 (35.3%)	23,877 (30.2%)
Renal replacement therapy (*n* (%))	1,488 (8.7%)	6,071 (7.7%)
Length of stay, days (median, (IQR))	3.1 (1.8–7)	2.9 (1.7–5.6)
ICU mortality	7.4%	9.8%
Hospital mortality	11.3%	16.4%
90-d mortality	18.9%	Not available

CHF, congestive heart failure; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; ICD-9, International Classification of Diseases version 9; IQR, interquartile range; OASIS, Oxford Acute Severity of Illness Score; RLD, restrictive lung disease; SOFA, sequential organ failure assessment.

Komorowski M et al. Nat Med. 2018.

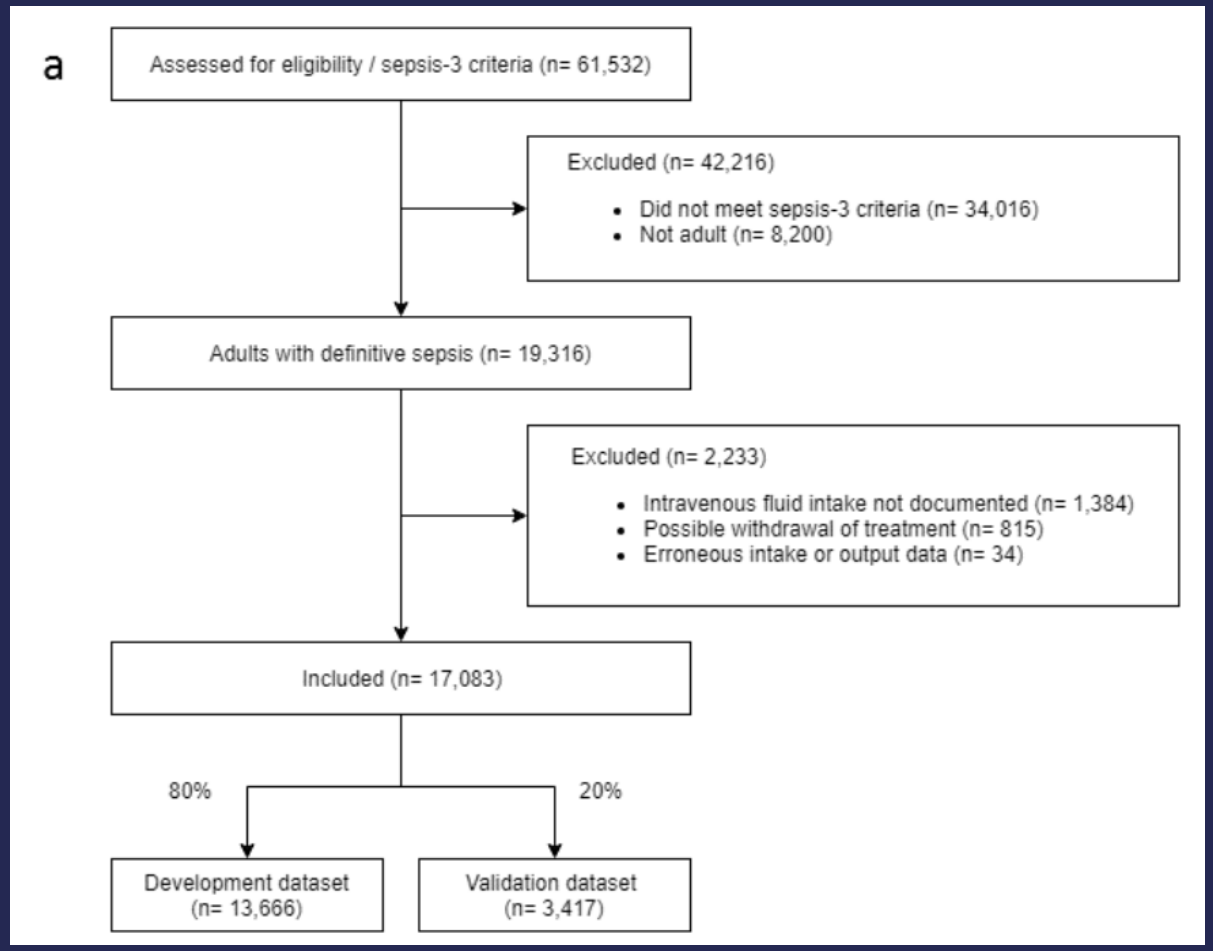
Patient cohort

Comparison of Key ICU Database Characteristics

Characteristics	Amsterdam UMCdb	eICU-CRD	HiRID	MIMIC-IV
Number of centers	1	208	1	1
Center location	Amsterdam, the Netherlands	United States	Bern, Switzerland	Boston, United States
Time period	2003–2016	2014–2015	2005–2016	2008–2019
ICU unique patient count[a]	20,109	139,367	33,905[b]	50,048
ICU admissions/unique patient	100%	120.7%	100.00%	139.1%
ICU patient age, median (IQR)	60–69 (50–59, 70–79)[c]	65 (53–76)	65 (55–75)	64 (51–70)
Gender: male	63.6%	54.0%	64.2%	56.1%
Ethnicity: White	Not reported	77.2%	Not reported	65.6%
Ethnicity: African–American	Not reported	10.6%	Not reported	9.0%
Ethnicity: Other/unknown	Not reported	12.1%	Not reported	24.4%
Mortality: ICU	9.9%	5.5%	6.1%	8.6%
Mortality: ICU elective patients	6.8%	1.9%	Not reported	4.3%
Mortality: ICU urgent patients	18.3%	6.2%	Not reported	8.9%
Mortality: hospital	13.3%	9.0%	Not reported	14.8%
Mortality: 28 d	13.6%[d]	Not reported	Not reported	11.6%[d]
Length of stay ICU, median (SD)	1.0 (0.8–3.1)	1.6 (0.8–3.0)	1.0 (0.8–2.2)	2.0 (1.1–3.9)
Patients with ≥ 1 comorbidity	Not reported	87.8%	Not reported	83.8%
Severity of illness scores	APACHE II, SOFA	APACHE IV, APACHE IVa	APACHE II	APACHE III, Oxford Acute Severity of Illness Score, SOFA
APACHE admission, median (IQR)	APACHE II: 17 (13–22)	APACHE IVa: 51 (37–68)	APACHE II: 16 (12–22)	APACHE III: 41 (30–57)
SOFA first 24 hr, median (IQR)	7 (4–9)	Not reported	Not reported	2 (1–5)
Availability of radiology images	Not reported	Not reported	Not reported	Yes[e]
Availability of clinical notes	Not reported	Deconstructed notes available	Not reported	Not publicly available

Sauer C et al. Crit Care Med. 2022

Patient cohort



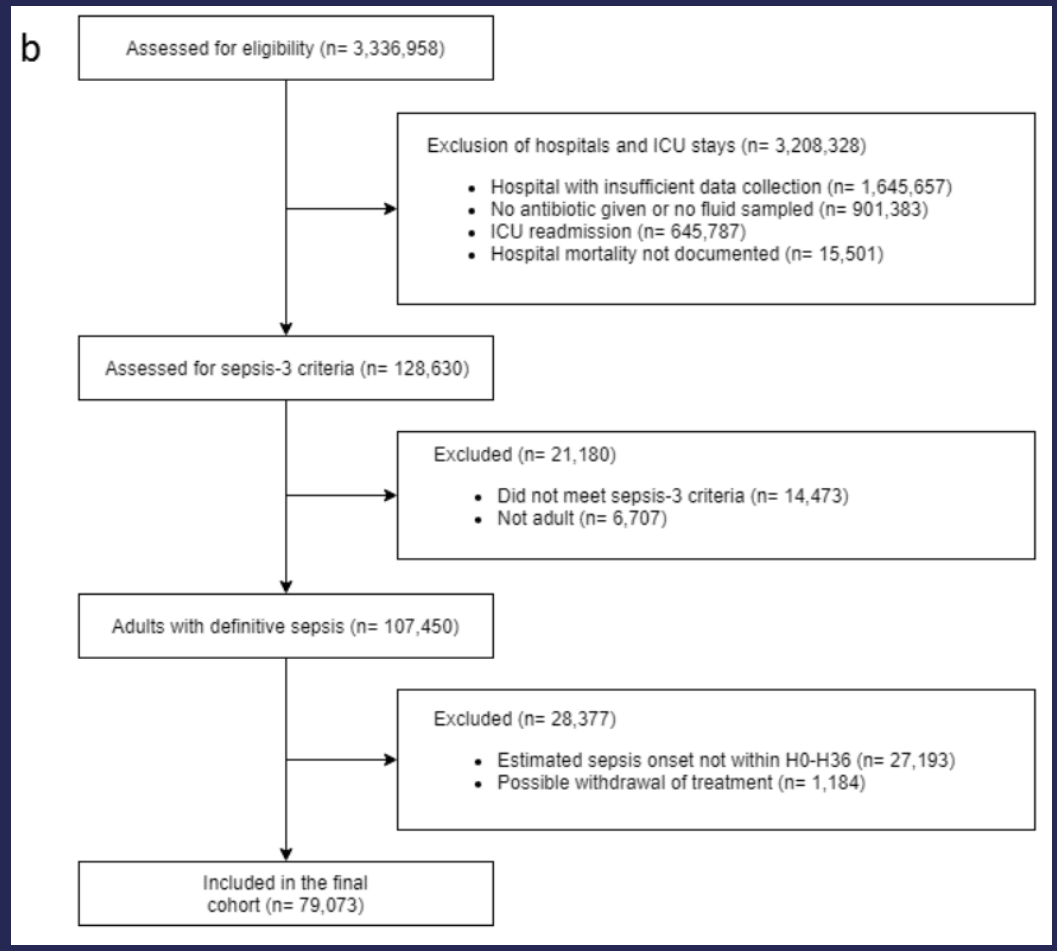


Figure 1.Patient inclusion diagrams in MIMIC-III (a) and eRI (b)

Komorowski M et al. Nat Med. 2018.

Data features

- 48 variables
  - Demographics
  - Elixhauser premorbid status
  - Vital signs
  - Laboratory values
  - Fluids and vasopressors received

Table 2: List of model features

Category	Items	Type	Available in MIMIC-III	Available in eRI
Demographics	Age	Cont.	+	+
	Gender	Binary	+	+
	Weight	Cont.	+	+
	Readmission to intensive care	Binary	+	+
	Elixhauser score (premorbid status)	Cont.	+	-
Vital signs	Modified SOFA*	Cont.	+	+
	SIRS	Cont.	+	+
	Glasgow coma scale	Cont.	+	+
	Heart rate, systolic, mean and diastolic	Cont.	+	+
	blood pressure, shock index	Cont.	+	+
	Respiratory rate, SpO_2	Cont.	+	+
	Temperature			
Lab values	Potassium, sodium, chloride	Cont.	+	+
	Glucose, BUN, creatinine	Cont.	+	+
	Magnesium, calcium, ionized calcium,	Cont.	+	+
	carbon dioxide	Cont.	+	+
	SGOT, SGPT, total bilirubin, albumin	Cont.	+	+
	Hemoglobin	Cont.	+	+
	White blood cells count, platelets	Cont.	+	+
	count, PTT, PT, INR			
	pH, PaO_2, $PaCO_2$, base excess,			
	bicarbonate, lactate, PaO_2/FiO_2 ratio			
Ventilation parameters	Mechanical ventilation	Binary	+	+
	FiO_2	Cont.	+	+
Medications and fluid balance	Current IV fluid intake over 4h	Cont.	+	+
	Maximum dose of vasopressor over 4h	Cont.	+	+
	Urine output over 4h	Cont.	+	+
	Cumulated fluid balance since	Cont.	+	+
	admission (includes preadmission data			
	when available)			
Outcome	Hospital mortality	Binary	+	+
	90-day mortality	Binary	+	-

Supplementary Table 2. Description of the variables included in the datasets. Cont.: continuous; INR: International Normalized Ratio; * Modified SOFA: SOFA based on values in the current 4h time step; PEEP: Positive End Expiratory Pressure; PT: Prothrombin Time; PTT: Partial Thromboplastin Time; SIRS: Systemic Inflammatory Response Syndrome; Shock index: systolic blood pressure/heart rate.

Komorowski M et al. Nat Med. 2018.

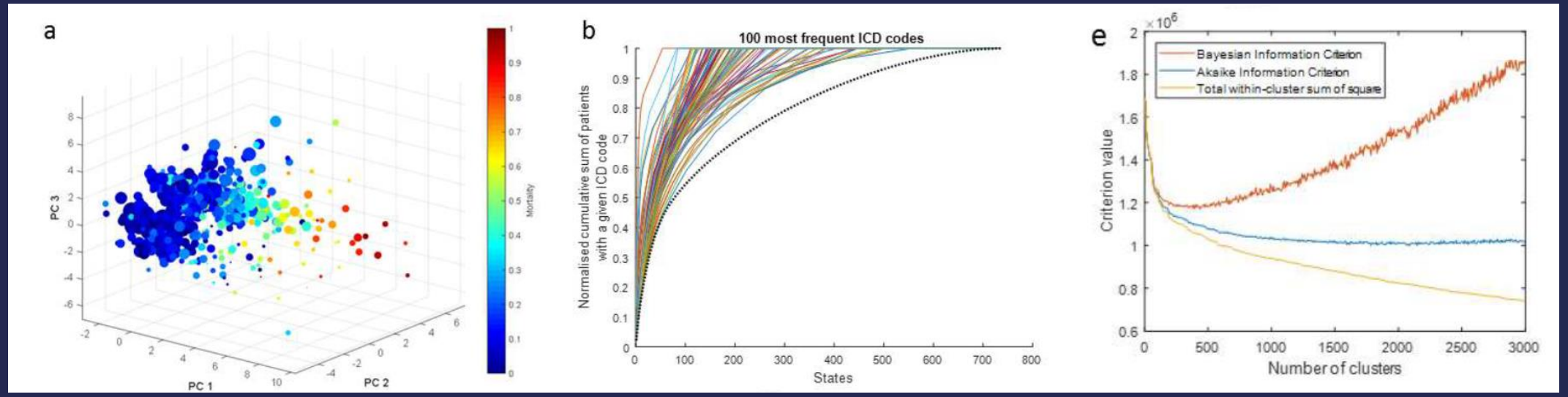
Data preprocessing

- Outlier removal
  - Distribution check using a frequency histogram and univariate statistical analysis.
- Missing value imputations
  - Parameter-specific sample-and-hold approach
  - KNN imputation

Komorowski M et al. Nat Med. 2018.

MDP formulation

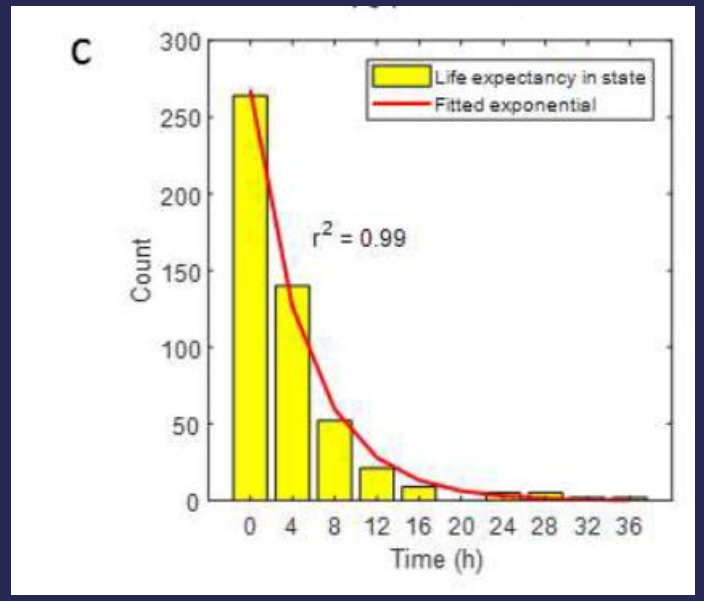
- State space
  - Multidimensional *discrete* time series with 4-h time steps.
  - Up to 72 h of measurements taken around the estimated time of onset of sepsis.
  - K-means++ clustering used to make 750 discrete mutually exclusive patient states.
    - + Two absorbing states (death and discharge of the patient)



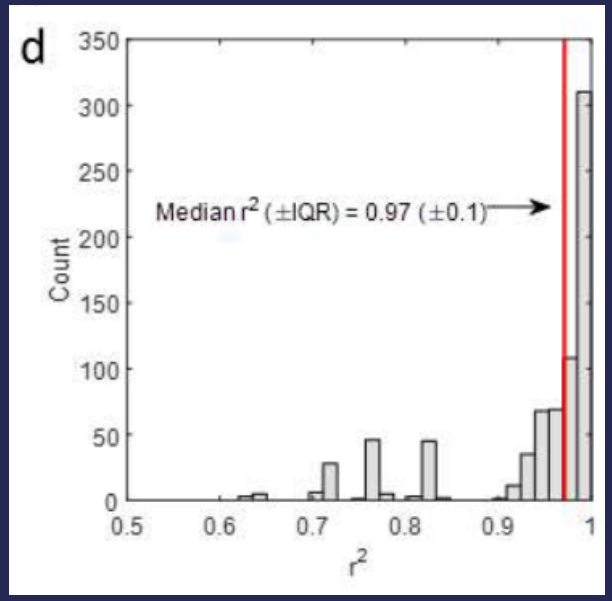
Komorowski M et al. Nat Med. 2018.

MDP formulation

- State space (cont'd)
  - Makov property
    - Transitions are memorylys.
    - The probability to remain in a state follows an exponential decay.



The correlation coefficient r^2 between the data and the fitted function is 0.99 in this example.



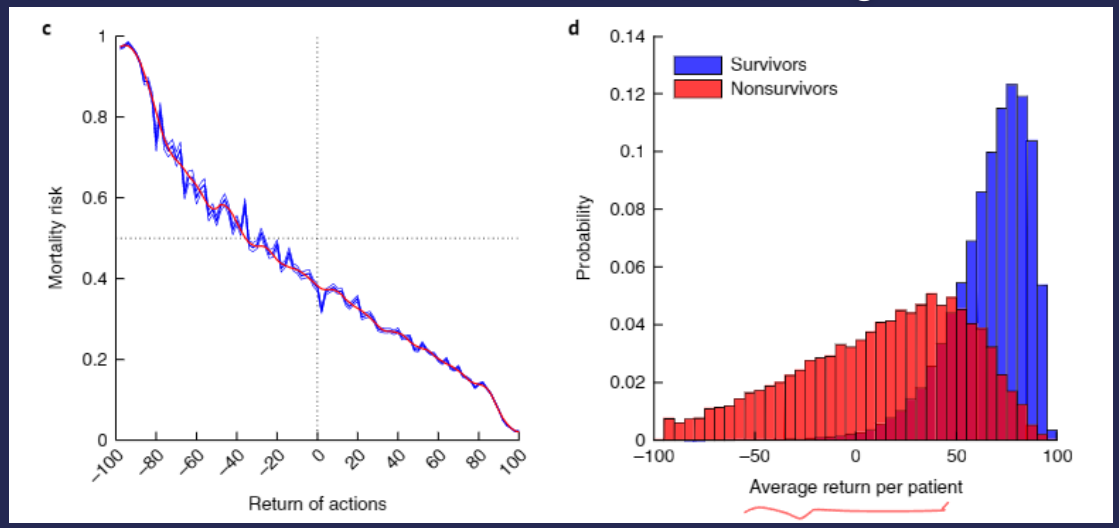
Distribution of the correlation coefficients between life expectancy and exponential decayfunctions in the 750 states of the model.

Komorowski M et al. Nat Med. 2018.

MDP formulation

- Action space (discrete; 5 possible choices)
  - Total volume of intravenous fluids over each 4-h time steps.
  - Maximum dose of vasopressors over each 4-h time steps.

- Reward space
  - Evaluating each treatment options (-100 to + 100)
  - At the end of episodes, +100 for survival and -100 for death, according to either hospital mortality or 90-d mortality



Komorowski M et al. Nat Med. 2018.

Offline RL

- TD learning

$$Q^{\pi}(s,a) \leftarrow Q^{\pi}(s,a) + \alpha \cdot (r + \gamma \cdot Q^{\pi}(s',a') - Q^{\pi}(s,a))$$

$$\pi^{\star}(s) \leftarrow \underset{a}{\operatorname{argmax}}\, Q^{\pi^{\star}}(s,a) \quad \forall s$$

$$V^{\pi}(s) = \sum \pi(s,a) \sum_{s'} T(s',s,a)\,[R(s') + \gamma V^{\pi}(s')]$$

Komorowski M et al. Nat Med. 2018.

Off-policy evaluation

- It was crucial to obtain reliable estimates of the performance of this new policy *without deploying* it, as executing a bad policy would be *dangerous* for patients
- High-confidence off-policy evaluation (HCOPE) method
  - weighted importance sampling (WIS) with bootstrapping (B=2000)
    - Per-step importance ratio
      - Pi1 = New policy (0.99 greedy)
      - Pi2 = Behavior policy (clinician)

$$\rho_t := \pi_1(a_t|s_t)/\pi_0(a_t|s_t)$$

    - Cumulative importance ratio

$$\rho_{1:t} := \prod_{t'=1}^{t} \rho_{t'}$$

    - Average cumulative importance ratio

$$w_t = \sum_{i=1}^{|D|} \rho_{1:t}^{(i)} / |D|$$

      - D, Dataset
    - Trajectory-wise WIS estimator

$$V_{WIS} = \frac{\rho 1:H}{w_H}\left(\sum_{t=1}^{H} \gamma^{t-1} r_t\right)$$

    - Average WIS estimate over all trajectories.

Algorithm 1 Bootstrap Confidence Interval

Input is an evaluation policy π_e, a data set of trajectories, $\mathcal{D}$, a confidence level, $\delta \in [0, 1]$, and the required number of bootstrap estimates, B.

input π_e, $\mathcal{D}$, π_b, δ, B
output $1 - \delta$ confidence lower bound on $V(\pi_e)$.

1: **for all** $i \in [1, B]$ **do**
2: $\tilde{\mathcal{D}}_i \leftarrow \{H_1^i, \ldots, H_n^i\}$ where $H_j^i \sim \mathcal{U}(\mathcal{D})$ // where $\mathcal{U}$ is the uniform distribution
3: $\widehat{V}_i \leftarrow$ **Off-PolicyEstimate**$(\pi_e, \tilde{\mathcal{D}}_i, \pi_b)$
4: **end for**
5: sort$(\{\widehat{V}_i | i \in [1, B]\})$ // Sort ascending
6: $l \leftarrow \lfloor \delta B \rfloor$
7: **Return** $\widehat{V}_l$

Komorowski M et al. Nat Med. 2018, Hanna et al, AAMAS 2017

Model development

- Split dataset : 80 / 20
- Bulid 500 different models using different clustering solutions.
- Select the final model with maximize the 95% confidence lower bound of the AI policy
  - Comparing to 95% confidence upper bound of the clinicians' policy (behavior policy)

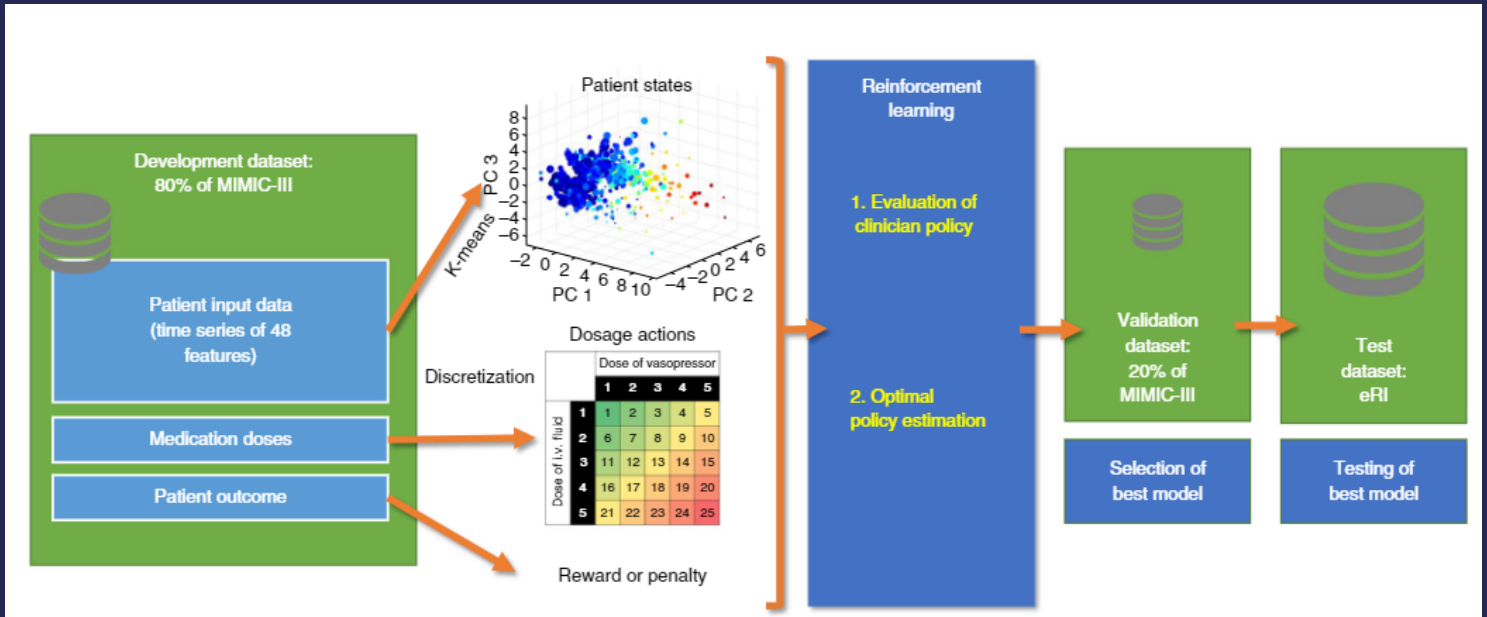
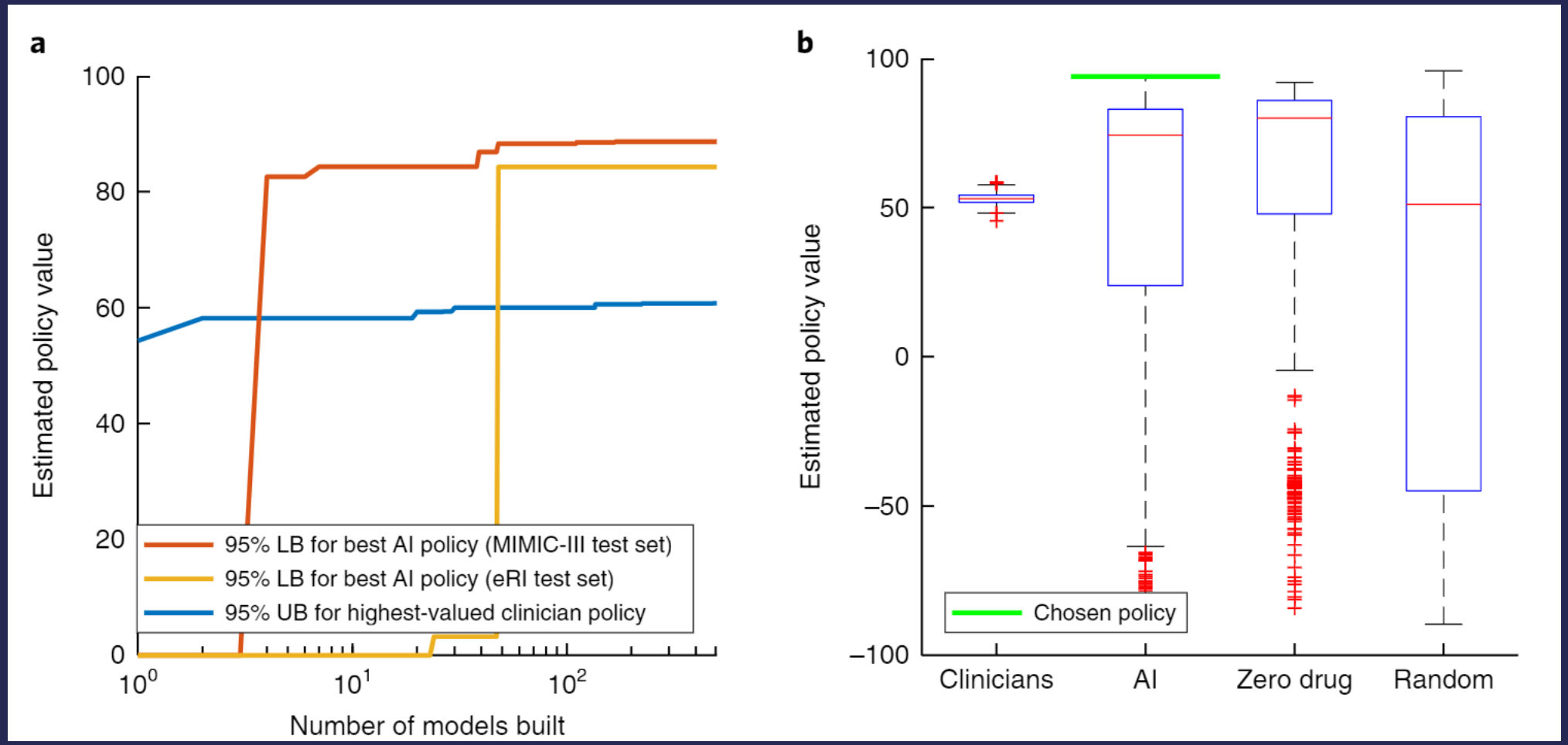


Fig. 1 | Data flow of the AI Clinician. Eighty percent of the MIMIC-III dataset was used to define the elements of the MDP. Time series of patient data were clustered into finite states. The dose of intravenous (i.v.) fluids and vasopressors were discretized into 25 possible actions. Patient survival at 90 d after ICU admission defined reward. Reinforcement learning was used to estimate optimal treatment strategies—the AI policy. The remaining 20% of MIMIC-III data was used to identify the best model among 500 candidates, which was then tested on an independent dataset from the eRI database.

Komorowski M et al. Nat Med. 2018.

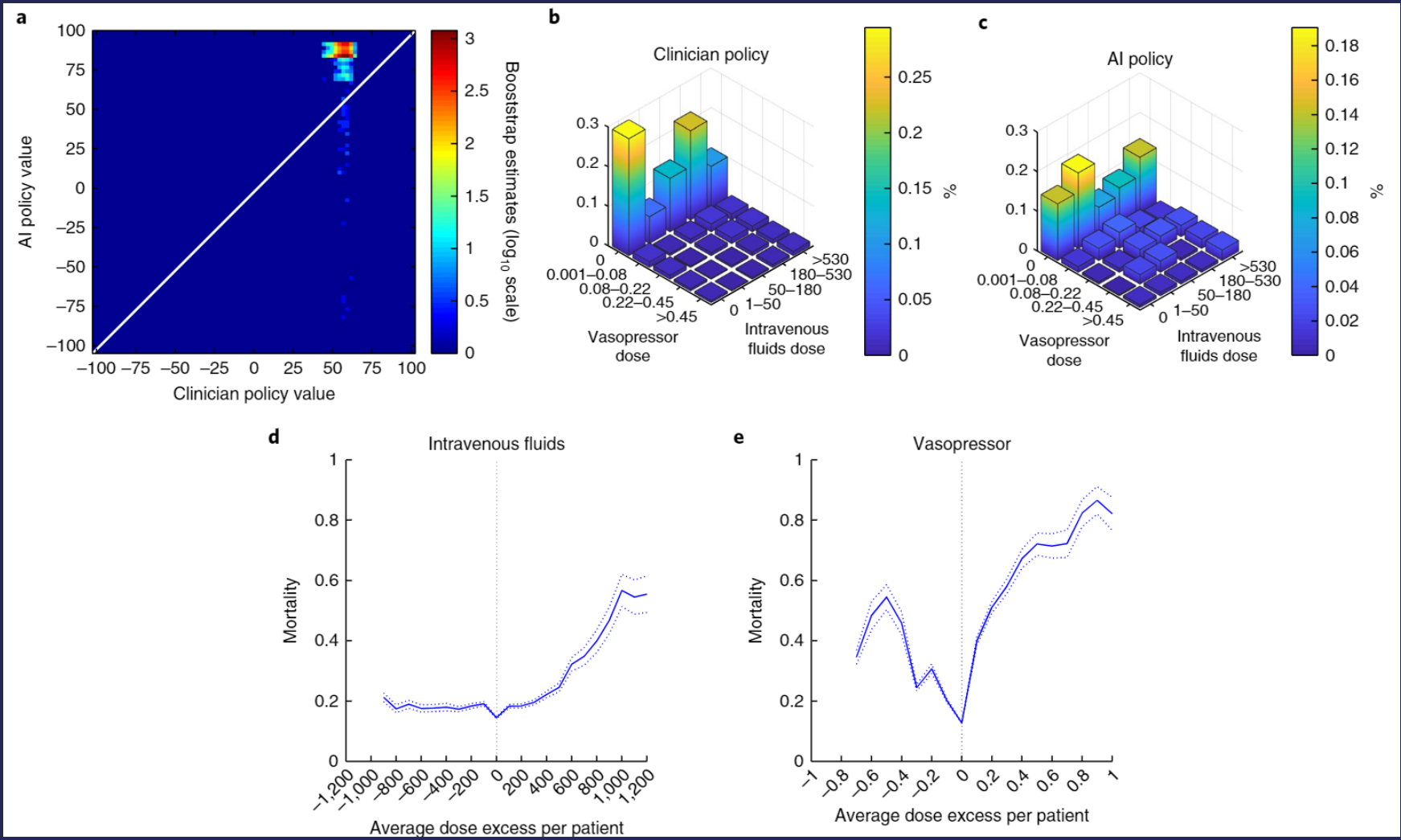
Results

Final model selection



Komorowski M et al. Nat Med. 2018.

Comparison of policies

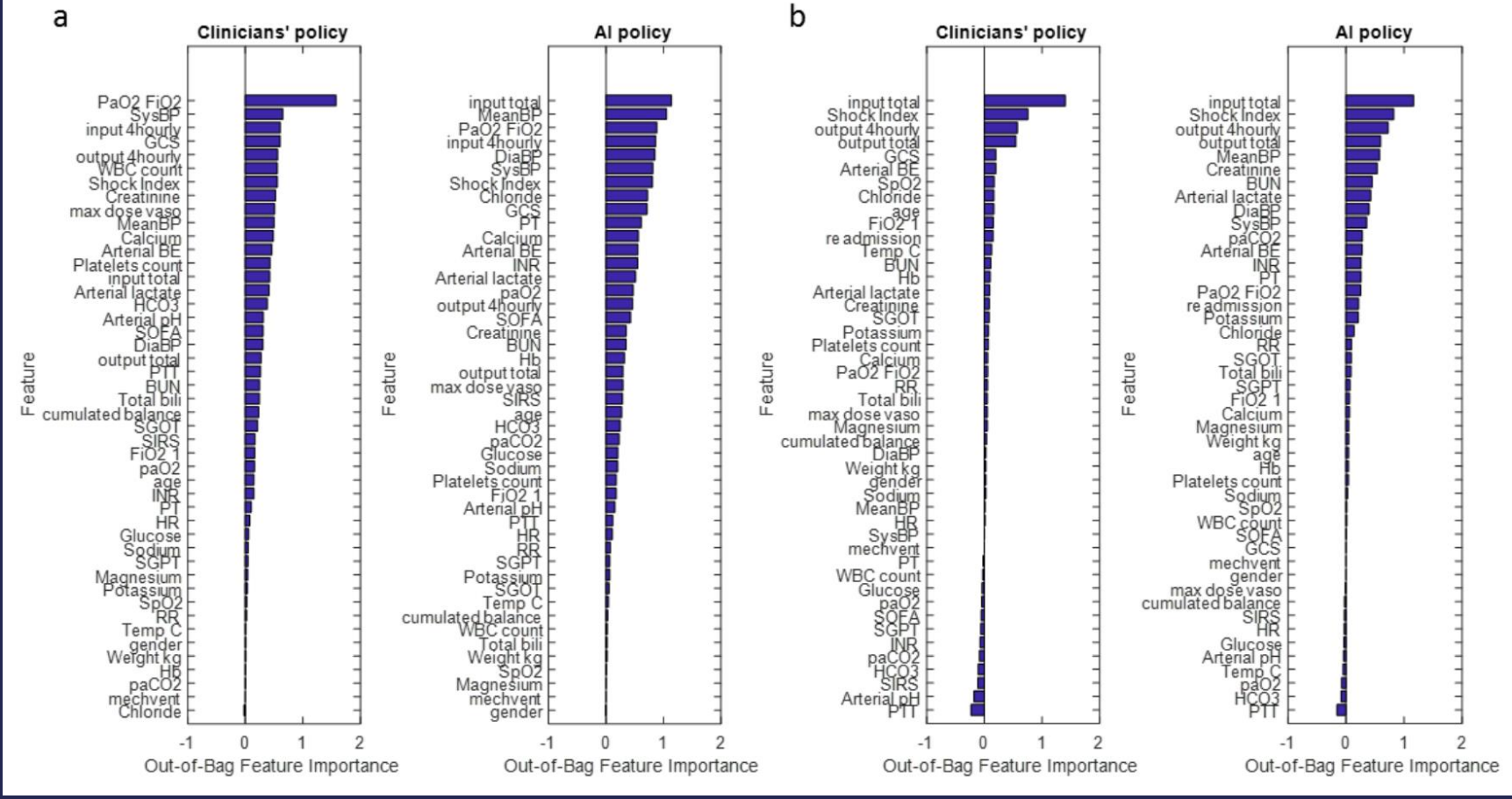


Komorowski M et al. Nat Med. 2018.

Feature importance

- Using Random forest classification model

Figure 2: Feature importance in the clinicians' policy and the AI policy



Komorowski M et al. Nat Med. 2018.

Discussion

Discussion

- Clinical aspect,
  - Early use of low-dose vasopressor has been suggested to play a role in sepsis
    - This may avoid administration of an excessive amounts of fluids, which has been linked with a poorer outcome.
  - Our results support this strategy but importantly allow the treatment to be *individualized for each patient.*
  - Avoiding targeting short-term resuscitation goals and instead following trajectories *toward longer-term survival*
  - A reduction in mortality from sepsis by only a small percentage would represent *several tens of thousands of lives saved annually worldwide.*

Komorowski M et al. Nat Med. 2018.

Discussion

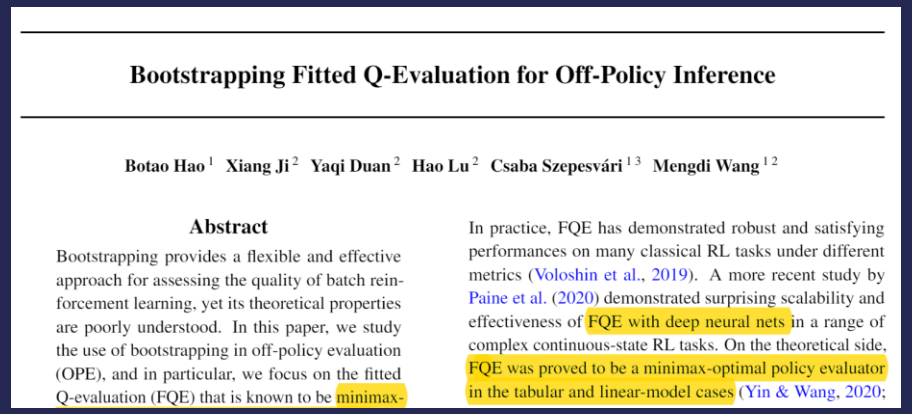
- RL aspect,
  - Neural network? : behavior policy, Fitted Q evaluation

Published as a conference paper at ICLR 2021

BENCHMARKS FOR DEEP OFF-POLICY EVALUATION

Justin Fu*[1] Mohammad Norouzi*[2] Ofir Nachum*[2] George Tucker*[2]
Ziyu Wang[2] Alexander Novikov[3] Mengjiao Yang[2] Michael R. Zhang[2]
Yutian Chen[3] Aviral Kumar[1] Cosmin Paduraru[3] Sergey Levine[1] Tom Le Paine*[3]

[1]UC Berkeley [2]Google Brain [3]DeepMind
justinfu@berkeley.edu,{mnorouzi,ofirnachum,gjt,tpaine}@google.com



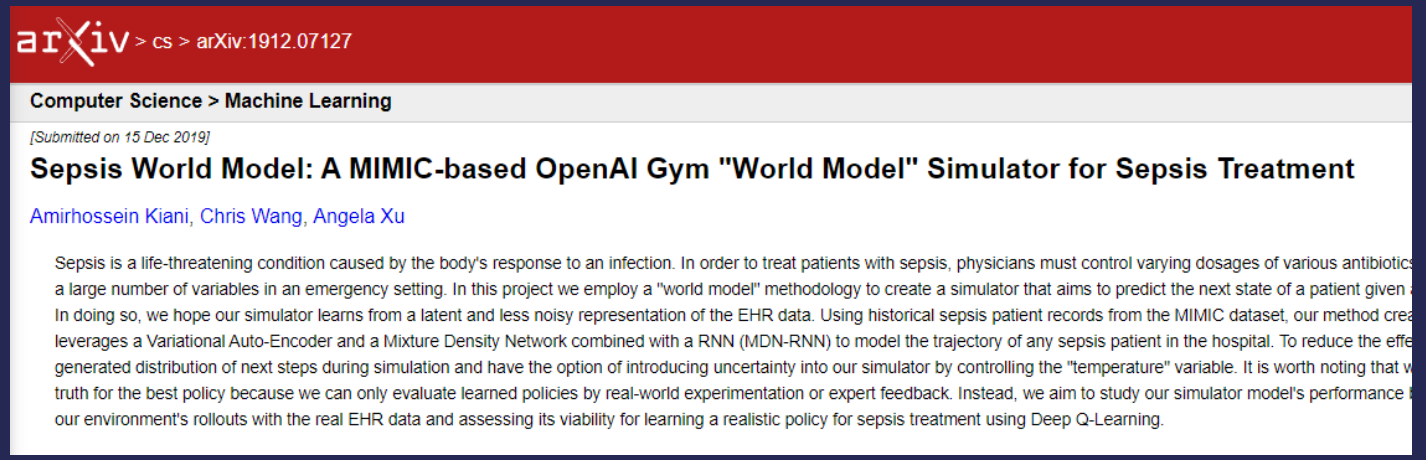
Bootstrapping Fitted Q-Evaluation for Off-Policy Inference

Botao Hao [1] Xiang Ji [2] Yaqi Duan [2] Hao Lu [2] Csaba Szepesvári [1 3] Mengdi Wang [1 2]

Abstract

Bootstrapping provides a flexible and effective approach for assessing the quality of batch reinforcement learning, yet its theoretical properties are poorly understood. In this paper, we study the use of bootstrapping in off-policy evaluation (OPE), and in particular, we focus on the fitted Q-evaluation (FQE) that is known to be minimax-

In practice, FQE has demonstrated robust and satisfying performances on many classical RL tasks under different metrics (Voloshin et al., 2019). A more recent study by Paine et al. (2020) demonstrated surprising scalability and effectiveness of FQE with deep neural nets in a range of complex continuous-state RL tasks. On the theoretical side, FQE was proved to be a minimax-optimal policy evaluator in the tabular and linear-model cases (Yin & Wang, 2020;



arXiv > cs > arXiv:1912.07127

Computer Science > Machine Learning

[Submitted on 15 Dec 2019]

Sepsis World Model: A MIMIC-based OpenAI Gym "World Model" Simulator for Sepsis Treatment

Amirhossein Kiani, Chris Wang, Angela Xu

Sepsis is a life-threatening condition caused by the body's response to an infection. In order to treat patients with sepsis, physicians must control varying dosages of various antibiotics a large number of variables in an emergency setting. In this project we employ a "world model" methodology to create a simulator that aims to predict the next state of a patient given In doing so, we hope our simulator learns from a latent and less noisy representation of the EHR data. Using historical sepsis patient records from the MIMIC dataset, our method crea leverages a Variational Auto-Encoder and a Mixture Density Network combined with a RNN (MDN-RNN) to model the trajectory of any sepsis patient in the hospital. To reduce the effe generated distribution of next steps during simulation and have the option of introducing uncertainty into our simulator by controlling the "temperature" variable. It is worth noting that w truth for the best policy because we can only evaluate learned policies by real-world experimentation or expert feedback. Instead, we aim to study our simulator model's performance t our environment's rollouts with the real EHR data and assessing its viability for learning a realistic policy for sepsis treatment using Deep Q-Learning.

Discussion

- Other RL application in medicine 1

npj | Digital Medicine

www.nature.com/npjdigitalmed

ARTICLE OPEN

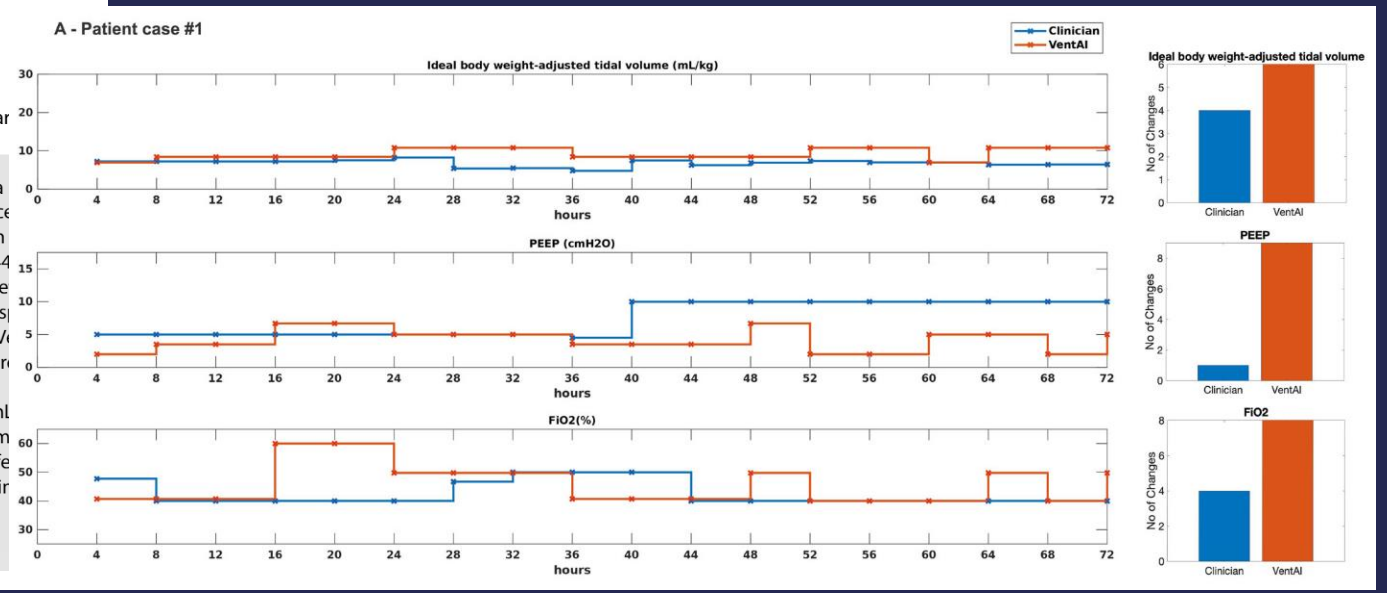
Check for updates

Development and validation of a reinforcement learning algorithm to dynamically optimize mechanical ventilation in critical care

Arne Peine[1,10], Ahmed Hallawa[1,2,10], Johannes Bickenbach[1], Guido Dartmann[3], Lejla Begic Fazlic[3], Anke Schmeink[4], Gerd Ascheid[2], Christoph Thiemermann[5], Andreas Schuppert[6], Ryan Kindle[7,8], Leo Celi[7,8,9], Gernot Marx[1] and Lukas Mar

The aim of this work was to develop and evaluate the reinforcement learning algorithm VentAI, which is able to suggest a dynamically optimized mechanical ventilation regime for critically-ill patients. We built, validated and tested its performance 11,943 events of volume-controlled mechanical ventilation derived from 61,532 distinct ICU admissions and tested it on an independent, secondary dataset (200,859 ICU stays; 25,086 mechanical ventilation events). A patient "data fingerprint" of 44 features was extracted as multidimensional time series in 4-hour time steps. We used a Markov decision process, including a re system and a Q-learning approach, to find the optimized settings for positive end-expiratory pressure (PEEP), fraction of insp oxygen (FiO_2) and ideal body weight-adjusted tidal volume (Vt). The observed outcome was in-hospital or 90-day mortality. V reached a significantly increased estimated performance return of 83.3 (primary dataset) and 84.1 (secondary dataset) compar physicians' standard clinical care (51.1). The number of recommended action changes per mechanically ventilated patient constantly exceeded those of the clinicians. VentAI chose 202.9% more frequently ventilation regimes with lower Vt (5–7.5 mL but 50.8% less for regimes with higher Vt (7.5–10 mL/kg). VentAI recommended 29.3% more frequently PEEP levels of 5–7 cm and 53.6% more frequently PEEP levels of 7–9 cmH_2O. VentAI avoided high (>55%) FiO_2 values (59.8% decrease), while prefe the range of 50–55% (140.3% increase). In conclusion, VentAI provides reproducible high performance by dynamically choosir optimized, individualized ventilation strategy and thus might be of benefit for critically ill patients.

npj Digital Medicine (2021)4:32 ; https://doi.org/10.1038/s41746-021-00388-6



Discussion

- Other RL application in medicine 2

Journal of *Personalized Medicine*

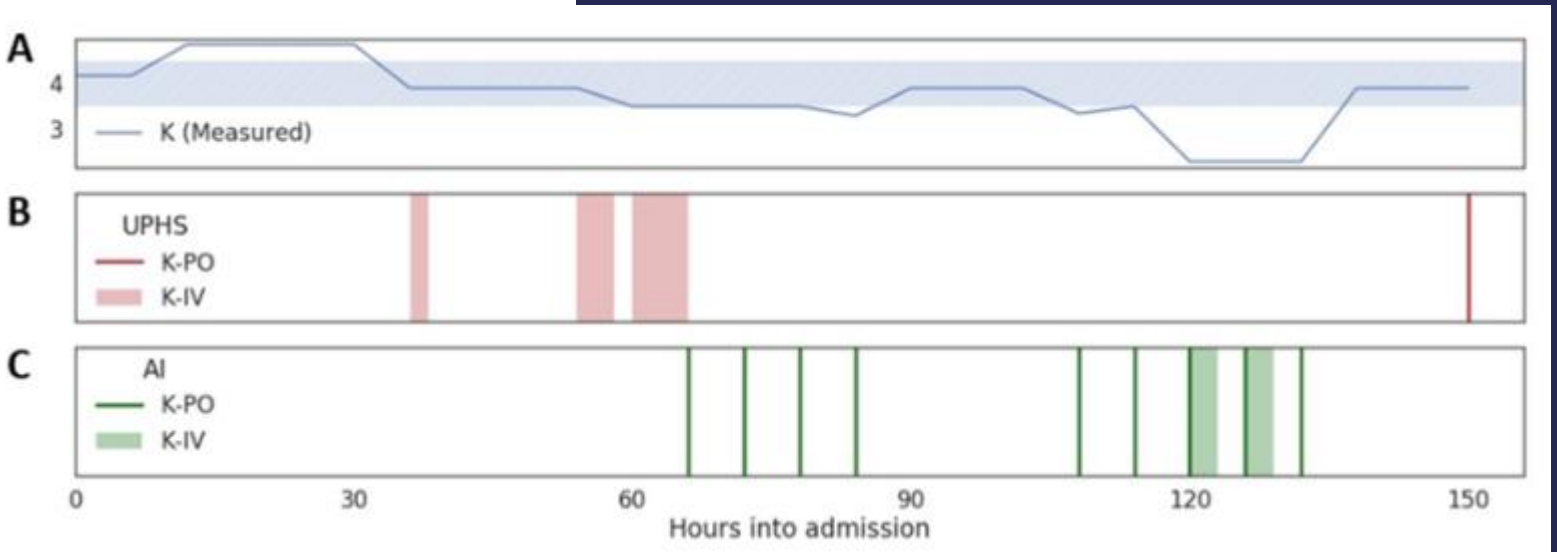
MDPI

Article

Guiding Efficient, Effective, and Patient-Oriented Electrolyte Replacement in Critical Care: An Artificial Intelligence Reinforcement Learning Approach

Niranjani Prasad [1,†], Aishwarya Mandyam [1,2,†], Corey Chivers [3], C. William Hanson III [3,4], Barbara E. Engelhardt [1,2] and Krzysztof La

1 Department of Computer Science, Princeton Un
prasad.niranjani@gmail.com (N.P.); amandyam(
2 Gladstone Institutes, San Francisco, CA 94158, U
3 University of Pennsylvania Health System, Phil
corey.chivers@pennmedicine.upenn.edu (C.C.);
hansonb@uphs.upenn.edu (C.W.H.III)
4 Department of Anesthesiology and Critical Car
5 Penn Medicine Predictive Healthcare, Universit
Philadelphia, PA 19104, USA
6 Leonard Davis Institute of Healthcare Economi
* Correspondence: klaudanski@gmail.com; Tel.: -
† These authors contributed equally to this work.




Thank you