Estimating When to Switch Treatment for Hypotensive ICU Patients Using Longitudinal Targeted Maximum Likelihood Estimation

Andre Waschka and Romain Pirracchio

Department of Statistics, University of California, Berkeley
Department of Anesthesia, University of California, San Francisco

February 23, 2023

Background

Large observational databases are routinely collected in ICU should be used to

- guide clinicians' decision-making process in real time,
- aid in developing more precise (individualized) interventions for patients,
- answer questions of comparative effectiveness of different clinical protocols.

The longitudinal structure of these observational data sets allows researchers to estimate the potential impacts of different time-dependent treatment strategies.

Motivating Question

How can data driven statistical analysis be used to gain a better understanding of the treatment for hypotensive patients, susceptible to septic shock, in intensive care units?

Current Protocol

Hour-1 Surviving Sepsis Campaign Bundle of Care recommends the following steps for patients in septic shock:

1. Measure lactate level. Remeasure if initial lactate is > 2 mmol/L
2. Obtain blood cultures prior to administration of antibiotics
3. Administer broad spectrum antibiotics
4. Begin rapid administration of 30ml/kg crystalloid to be completed within 3 hrs.
5. Apply vasopressors if patient is hypotensive during or after fluid resuscitation

Research Questions

1. **At what time from onset of hypotension should one switch from IV fluids to vasopressors (ignoring the rate the IV fluids)?**

2. **What is the amount of IV fluids a patient should receive before switching to vasopressors?**

Treatment Regimes

Our research questions formulate the following two treatment regimes:

1. **Static Treatment** - considering time to switch from IV fluids to vasopressors
2. **Dynamic Treatment** - accounting for both the time and the amount of IV fluid to switch to vasopressors

MIMIC dataset

MIMIC is an openly available dataset developed by the MIT Lab for Computational Physiology, comprising deidentified health data associated with 40,000 critical care patients. It includes demographics, vital signs, laboratory tests, medications, and more.

Cleaned Dataset

- Data Structure after cleaning.
- Showing the first 32 patients over baseline and treatment at hour 1.

	gender	age	hr0_elixhauser	hr0_Weight_kg	hr0_GCS	hr0_SOFA	PreHypoVolperkg	hr0_HR	hr0_SysBP	hr0_MeanBP	hr0_DiaBP	hr0_RR	hr0_SpO2
1	0	48.29521	0	77.50000	15	13	18.064516	119.00000	91.00000	58.00000	49.00000	35.00000	97.00000
2	1	74.52205	2	109.68333	7	8	63.820088	71.00000	155.00000	99.00000	63.00000	10.00000	100.00000
3	1	23.63611	0	62.50000	15	7	85.247200	92.00000	96.00000	62.66670	46.00000	20.00000	99.00000
4	0	31.11230	1	65.50000	6	8	114.503817	106.00000	113.00000	69.00000	47.00000	28.00000	100.00000
5	0	79.65106	2	85.00000	12	8	102.941176	69.77778	131.00000	79.77778	53.00000	19.44444	98.00000
6	0	71.31736	3	73.40000	15	2	89.008630	108.00000	86.00000	89.66667	50.00000	23.66667	95.66667
7	0	74.31739	2	104.70000	3	12	42.979943	81.00000	168.00000	57.00000	60.00000	13.00000	99.00000
8	0	27.17326	0	120.00000	15	1	136.526188	109.00000	94.75000	72.75000	58.75000	16.00000	100.00000
9	1	79.46297	6	49.00000	3	12	123.571429	103.00000	139.00000	85.00000	49.00000	12.00000	84.00000
10	1	46.76223	3	66.60000	13	7	135.065120	110.00000	111.00000	79.00000	63.00000	22.00000	98.00000
11	0	87.75353	5	74.70000	15	8	66.934404	79.00000	116.00000	64.00000	38.00000	14.00000	97.00000
12	0	57.80676	7	99.80000	9	10	50.100199	130.00000	110.00000	78.00000	63.00000	30.00000	99.00000
13	1	43.52962	0	99.70000	15	0	127.933805	100.00000	121.50000	76.50000	57.00000	22.41379	96.50000
14	0	52.41503	2	78.00000	15	6	42.307692	72.00000	141.00000	84.00000	55.50000	22.00000	100.00000
15	1	79.50059	7	56.70000	15	5	25.793651	124.00000	123.00000	77.00000	54.00000	28.00000	94.00000
16	0	56.23661	4	82.50000	15	6	39.757576	68.00000	124.00000	81.00000	64.00000	16.00000	100.00000
17	1	69.14396	7	72.10000	15	8	6.934813	75.00000	105.00000	70.00000	52.50000	14.00000	98.00000
18	0	32.97108	3	79.60000	7	4	6.281407	86.00000	100.00000	73.00000	57.00000	18.00000	100.00000
19	1	77.04315	3	61.30000	15	15	5.152257	88.80000	111.60000	64.20000	40.50000	23.20000	94.40000
20	1	21.15799	0	77.10000	14	13	79.092088	89.00000	104.00000	68.66670	51.00000	19.00000	100.00000
21	0	79.56675	9	62.10000	4	9	18.148149	61.00000	97.00000	68.00000	58.00000	20.00000	100.00000
22	1	36.77735	4	70.00000	13	8	93.609024	98.83333	81.00000	86.33333	37.50000	14.00000	100.00000
23	0	82.92912	4	73.50000	10	0	33.061224	82.00000	102.00000	70.00000	61.00000	13.00000	94.00000
24	1	79.43806	6	75.20000	15	6	81.313833	74.80000	123.40000	78.40000	53.60000	20.80000	94.60000
25	0	57.65890	4	100.10000	3	9	33.756244	66.00000	121.00000	82.00000	62.50000	36.00000	99.00000
26	1	60.08117	5	50.70000	11	8	64.833708	80.00000	90.00000	63.00000	49.50000	14.00000	100.00000
27	0	79.79948	3	95.90000	15	4	26.120959	84.00000	125.00000	91.00000	61.00000	10.00000	100.00000
28	0	58.21939	3	115.50000	3	10	41.125541	67.00000	99.00000	77.00000	61.00000	14.00000	100.00000
29	0	69.63093	6	76.60000	13	8	2.154047	78.66667	109.33333	76.33333	59.83333	14.66667	97.00000
30	1	78.50149	3	54.90000	15	13	102.914390	109.60000	104.80000	65.20000	62.80000	23.80000	90.80000
31	1	57.77882	1	75.00000	15	7	9.133333	49.00000	115.00000	76.00000	56.50000	14.00000	94.00000
32	1	59.36925	2	102.30000	15	4	25.513225	57.00000	111.00000	94.00000	49.50000	14.00000	100.00000

	hr0_Temp_C	hr0_Arterial_lactate	hr0_Arterial_BE	hr0_mechvent	Hr0Fluidsperkg	hr0_PaO2_FiO2	C1	hr1_vasoflagupdate	hr1_HR
1	39.55555	5.100000	-5.00000000	0	14.193548390	88.00000	uncensored	0	118.00000
2	36.61111	3.400000	-1.00000000	0	0.000000000	380.00000	uncensored	0	71.00000
3	36.22222	2.200000	-6.00000000	0	0.000000000	461.90476	uncensored	0	90.00000
4	37.66667	2.001961	0.00000000	1	0.000000000	1197.50000	uncensored	1	106.00000
5	36.83333	2.300000	0.00000000	1	1.058241176	262.22222	uncensored	0	65.00000
6	36.11111	3.091852	0.00000000	1	0.040871935	292.50000	uncensored	0	102.00000
7	37.16667	2.400000	1.00000000	1	0.000000000	399.00000	uncensored	0	81.00000
8	36.72222	3.100000	-5.00000000	1	18.592188060	354.50000	uncensored	0	112.33333
9	36.33333	5.800000	-9.00000000	1	0.000000000	402.00000	uncensored	0	103.00000
10	37.88889	7.321637	9.32291667	0	1.769760761	192.65625	uncensored	0	111.00000
11	36.38889	4.000000	-1.25773196	0	0.000000000	1209.13108	uncensored	0	79.00000
12	37.94445	7.400000	-12.00000000	1	0.000000000	170.00000	uncensored	0	130.00000
13	37.11111	7.003306	2.50000000	0	1.504513587	423.21429	uncensored	0	98.50000
14	37.83333	3.200000	5.21052632	0	0.000000000	510.30393	uncensored	0	61.00000
15	37.22222	2.400000	0.00000000	0	2.645502646	257.42025	uncensored	0	107.00000
16	36.22222	2.200000	2.00000000	1	0.000000000	365.00000	uncensored	0	68.00000
17	35.94444	2.500000	0.00000000	0	0.000000000	68.47361	uncensored	0	75.00000
18	36.22222	3.700000	-2.00000000	1	0.000000000	1232.49998	uncensored	0	91.00000
19	37.27778	2.100000	-4.00000000	0	8.771506253	99.40000	uncensored	0	80.50000
20	37.72222	2.386486	0.02702703	0	0.000000000	108.75676	uncensored	0	89.00000
21	34.55555	2.100000	-6.00000000	1	2.045088617	866.24999	uncensored	0	62.00000
22	36.37037	3.500000	-10.00000000	1	0.182632619	256.00000	uncensored	0	100.50000
23	38.38889	2.200000	7.00000000	1	2.721088435	597.00000	uncensored	0	80.00000
24	37.11111	5.500000	6.00000000	0	1.555851127	254.00000	uncensored	0	78.66667
25	37.83333	2.200000	0.00000000	1	1.498501499	364.00000	uncensored	0	65.00000
26	36.55556	2.400000	-1.00000000	1	1.359739645	366.00000	uncensored	0	91.40000
27	35.88889	2.058333	0.50000000	0	0.000000000	527.85715	uncensored	0	84.00000
28	33.88889	2.100000	-1.00000000	1	0.000000000	406.00000	uncensored	0	67.25000
29	36.66667	4.331818	-2.03636364	0	0.000000000	154.71875	uncensored	0	80.33333
30	37.66667	3.000000	-1.00000000	0	58.894960530	96.00000	uncensored	1	102.00000
31	36.72222	2.410526	-0.62500000	0	0.000000000	188.59756	uncensored	0	51.00000
32	36.11111	2.951613	1.38461539	0	1.486325187	186.90598	uncensored	0	76.00000

Summary of Data at Baseline

Covariate	Summary
# of Patients	933
% Mortality	11.46%
Avg Age	60.97
Avg SOFA	7.451
Avg Pre-Study Volperkg	62.63
Avg Weight (kg)	84.32
Arterial Lactate	3.347
Mechanical Ventilation	52.84%
Avg GCS	11.22
Avg HR	89.59
Avg SysBP	113.4
Avg DiaBP	52.37

Causal Model and LTMLE

Our analysis with be based on the causal inference model and the LTMLE model.

- The Causal Model will be used to ”translate” the problem that we outlined into a statistical problem.
- We will then have a targeted statistical parameter, the estimand, that the ***LTMLE*** model will estimate.
- The LTMLE incorporates and relies on the ***Super Learner*** - which is a powerful machine learning cross-validation-based estimator selection approach.

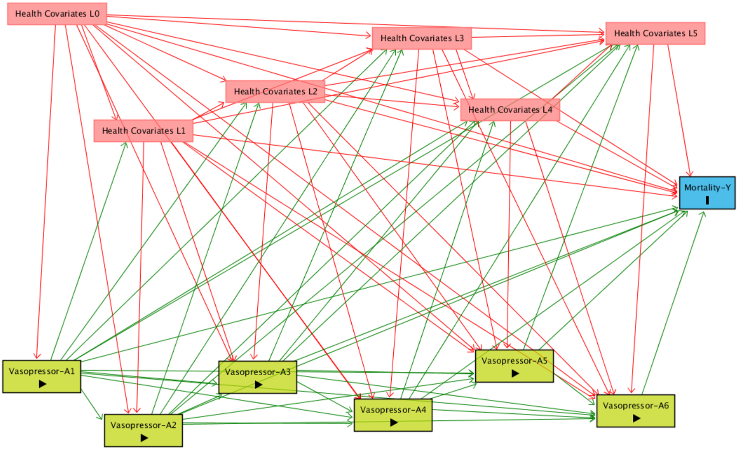
Structural Causal Model

The Structural Causal Model (SCM) reflecting our belief about the time-ordering and relationships between the treatment, covariates, and outcome of interest was:

$$
\begin{aligned}
L_0 &= f_{L_0}(U_{L_0}) \\
A_1 &= f_{A_1}(L_0, U_{A_1}) \\
&\vdots \\
A_6 &= f_{A_6}(L_0, A_1, L_1, A_2, L_2, A_3, L_3, A_4, L_4, A_5, L_5, U_{A_6}) \\
Y &= f_Y(L_0, A_1, L_1, A_2, L_2, A_3, L_3, A_4, L_4, A_5, L_5, A_6, U_Y)
\end{aligned}
$$

where $U_t = (U_Y, U_{L_t}$ and $U_{A_t})$, t = 0, ..., 6 are unmeasured exogenous random variables from some underlying probability distribution P_U . This causal model specifies how each of the variables in the data are generated, with randomness arising only from from the exogenous variables U.

Causal Diagram



Longitudinal Observed Data

- Discrete time scale: Starting with the hour before hypotension and continuing for the first 6 hours of hypotension.
  - $t = 0, .., 6$
- **Covariates** L_t:
  - **Baseline**: gender, age, elixhauser, patient weight in kg, GCS, SOFA, Fluids administered prior to time 0 (Vol per kg)
  - **Time-varying**: HR, SysBP, MeanBP, DiaBP, RR, SpO2, Temp C, Arterial lactate, Arterial BE, mechvent, Fluids administered (Vol per kg), PaO2 FiO2
- **Outcome** Y: Indicator of Mortality in hospital

Longitudinal Observed Data

  - **Treatment** E_t: Indicator of receiving vasopressors at time t. Once a patient goes on treatment, they remain on treatment for the remainder of the observed study.
  - **Right Censoring** C_t: Indicator patient wasn't followed during time t and on.
- **"Intervention"** Nodes: $A_t = (E_t, C_t)$
  - $\bar{A}_t = A_1, ..., A_t$
  - Censoring treated as an additional "Intervention" node: evaluate effect of vasopressors in the absence of censoring

Observed Data and Statistical Model

- Our observed data structure on a randomly sampled subject :

$$\begin{aligned} O &= (L_0, A_1, L_1, A_2, L_2, A_3, L_3, A_4, L_4, A_5, L_5, A_6, Y = L_6) \\ &= (\bar{L}_6, \bar{A}_6) \sim P_0 \end{aligned}$$

  where P_0 is the true underlying distribution from which the data are drawn.
  - The ordering of the elements of O represents their assumed causal ordering.
- Statistical Model $\mathcal{M}$: $P_0 \in \mathcal{M}$
  - Model should reflect real knowledge: large enough to contain the true P_0
  - In practice this generally implies a semi-parametric or non-parametric statistical model.

Returning to Research Questions

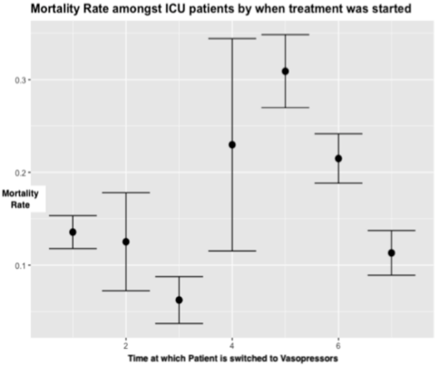
1. **At what time should we switch from IV fluids to vasopressors?**
2. What is the amount of IV fluids a patient should receive before switching to vasopressors?

Time to Switch Results

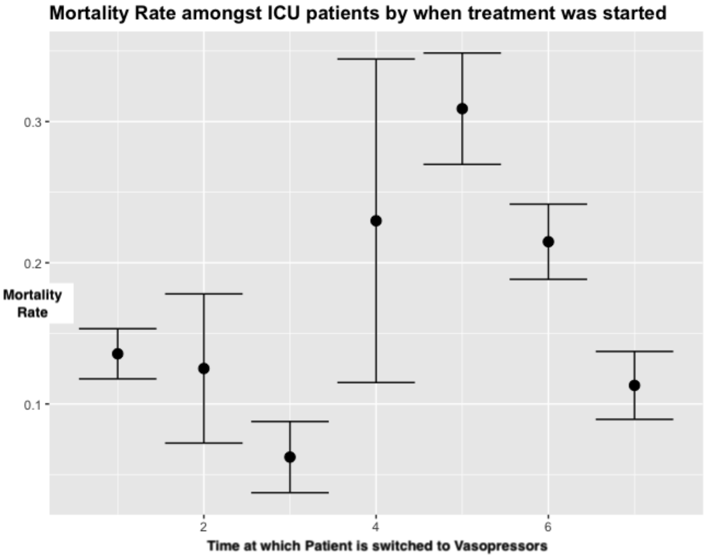
- Our statistical parameter (estimand):

  $$E[Y_{\bar{a}}] = E[...[E(Y|\bar{A}_6 = \bar{a}_6, \bar{L}_5)|\bar{A}_5 = \bar{a}_5, \bar{L}_4)|...|L_0]$$

- Our results show the 7 different estimates we obtained from each of the possible static treatment regimes we defined. (ex: Always treat, Treat starting at hour 2, Never Treat)
- From these results we can see that there is a significant difference in mortality rate between patients treated with vasopressors within the first three hours and patients who were treated after the first three hours.
- The final observation primarily shows patients who were never treated and as a result they also have a low mortality rate.



Time to Switch Results



Research Question 2

1. At what time should we switch from IV fluids to vasopressors?
2. **What is the amount of IV fluids a patient should receive before switching to vasopressors?**

Results for Dynamic Treatment Regimes

- Let V_t be Fluids administered (vol/kg) from the covariates L_t
- For dynamic regimes, we define $d_{t,\theta}$ as a function that incorporates fluid administered to a subject up to time t-1

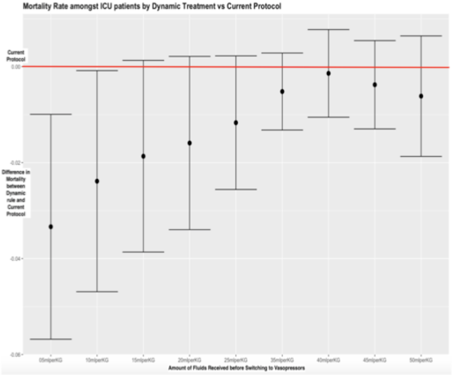
$$d_{t,\theta}(V_t, A_{t-1}) = I(A_{t-1} = 1) + I(A_{t-1} = 0)I(V_t > \theta)$$

  - for θ = (5,10,15,20,25,35,40,45,50)

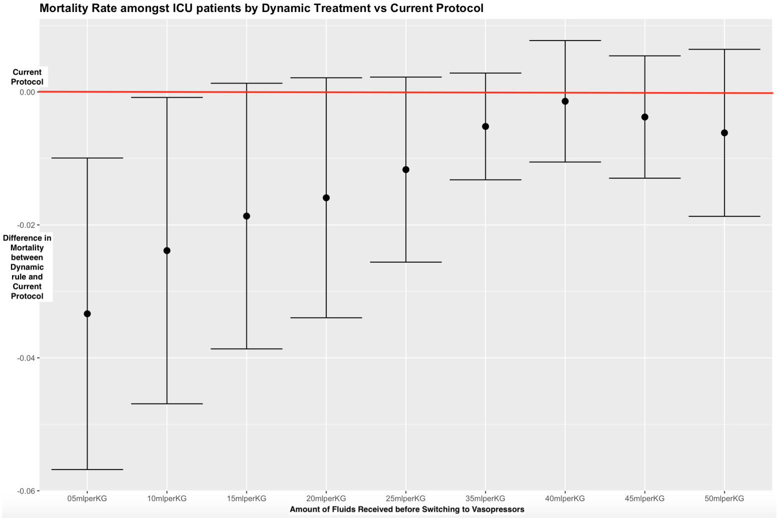
- Our statistical parameter (estimand):

$$\begin{aligned}\Psi_{d_\theta}(P_0) = E[E[...[E(E(Y|\bar{A}_6 = d_\theta(\bar{V}_6), \bar{L}_5)|\\ \bar{A}_5 = d_\theta(\bar{V}_5), \bar{L}_4)|...]|L_0]\\ -[E[...[E(E(Y|\bar{A}_6 = d_{30}(\bar{V}_6), \bar{L}_5)|\\ \bar{A}_5 = d_{30}(\bar{V}_5), \bar{L}_4)|...]|L_0]]\end{aligned}$$

- These results show us the difference in mortality rates between each θ and the currently recommended 30ml/kg threshold before switching to vasopressors. We can see that its beneficial for patients to be switched to vasopressors from fluids after a much smaller amount than the guidelines recommend.



Results for Dynamic Treatment Regimes



Current Work

- Unmeasured Confounding:
  - Unmeasured Confounding is a potential source of concern for our results
  - Unmeasured Confounding is untestable
- Simulations
  - Generate the data
  - Generate potential confounders
  - Estimate their effect

Simulation

3 main simulation types

- Simple Simulation:
  - Uses no information from the data
  - Used to see how a method may work on a generic dataset
- Advanced Simulation:
  - Uses some information about data (variable names, means, etc)
  - Makes strong (and incorrect) assumptions about data distribution
  - Builds a dataset that roughly approximates the true data for a specific problem
- Nonparametric Simulation:
  - Uses the current dataset to build a simulated dataset
  - Assumptions are justifiable
  - Uses machine learning to estimate the probability of landing in a specific ”bin”
  - Covariate value is randomly generated according to the estimated nonparametric probability distribution

Potential Issues and Limitations, Future Developments

- Issues and Limitations:
  - Data transformation loses information
  - Restricted target population
  - Positivity
- Future developments:
  - Moving to a interval based definition of treatments
  - Understanding positivity issues for extreme treatments and critically assessing the causal assumptions
  - Moving to Optimal Dynamic Treatments to be able to identify which of our better rules is the best.

References

Bang H, and Robins JM. Doubly-robust estimation in missing data and causal inference models. Biometrics, 61:962 - 972, 2005.

Decker AL, Hubbard A, Crespi CM, et al. Semiparametric estimation of the impacts of longitudinal interventions on adolescent obesity using targeted maximum-likelihood: accessible estimation with the ltmle package. J Causal Inference. 2014; 2(1):95-108

Pearl J. Causal diagrams for empirical research. Biometrika, 82(4): 669-688, 1995.

Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. Epidemiology. 2014;25(3):418-426.

Robins JM. The control of confounding by intermediate variables. Statistics in medicine, 8(6):679-701, 1989.

Robins JM. A new approach to causal inference in mortality studies with sustained exposure periods - application to control of the healthy worker survivor effect. Mathematical Modelling, 7:1393 - 1512, 1986.

van der Laan MJ, and Gruber S. Targeted Minimum Loss Based Estimation of Causal Effects of Multiple Time Point Interventions. The International Journal of Biostatistics, 8(1), Jan. 2012. ISSN 1557-4679. doi: 10.1515/1557-4679.1370.

References

van der Laan MJ, Rubin D. Targeted maximum likelihood learning. Int J Biostat. 2006;2(1):1-38.

van der Laan MJ, Rose S. Targeted Learning: Causal Inference for Observational and Experimental Data. New York, NY: Springer-Verlag; 2011.

van der Laan MJ, Rose S. Targeted Learning: Causal Inference for Complex Longitudinal Studies a. New York, NY: Springer-Verlag; 2018.

Thank you!