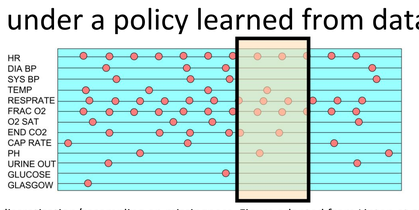
Missingness as Stability: Understanding the Structure of Missingness in Longitudinal EHR data and its Impact on Reinforcement Learning in Healthcare

Scott Fleming[1,2], Kuhan Jesapragasan[2], Tony Duan[2], Daisy Ding[1], Saurabh Gombar[3,4], Nigam Shah[1,4], Emma Brunskill[2]

[1]*Center for Biomedical Informatics Research, Stanford University*
[2]*Department of Computer Science, Stanford University*
[3]*Department of Pathology, School of Medicine, Stanford University*
[4]*Department of Medicine, School of Medicine, Stanford University*

Introduction

- Most reinforcement learning (RL) algorithms assume regularly sampled time series data with no missingness (i.e. ***not*** EHR data).
- Recent RL in healthcare papers tend to preprocess EHR data by
  a. Resampling data to uniformly sized windows,
  b. Mean pooling lab values within the same bin,
  c. Imputing with Last-Observation-Carried-Forward (LOCF) imputation.
- This process strips any missingness information.
- We show that maintaining missingness information could increase estimated expected reward under a policy learned from data.



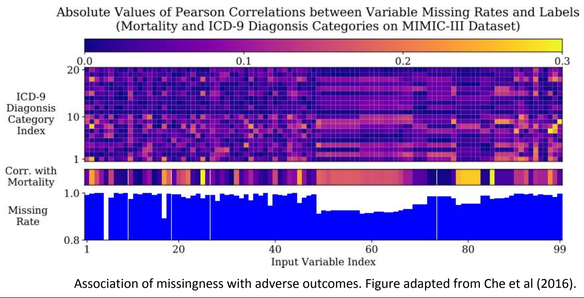
Impact of discretization/resampling on missingness. Figure adapted from Lipton et al (2016).

Clinical Application

- **Venous Thromboembolisms** (VTEs) are blood clots that can form within a patient's veins and travel to critical organs like the heart, lungs, and brain, blocking the flow of blood.
- VTEs are a leading cause of preventable hospital death in America (> 2x the number of deaths from breast cancer each year!).
- Common treatment in ICU is anticoagulation with unfractionated heparin (UFH).
- UFH is typically titrated so that activated partial thromboplastin time (aPTT) and anti-Xa assay values fall within established therapeutic ranges.
- Too much UFH ⟶ uncontrolled bleeding
- Too little UFH ⟶ VTEs + potential strokes/PEs/etc
- We model UFH dosing as RL problem
- We analyzed EHRs of 8,983 patients in Stanford Hospital ICU on UFH.

Related Work

- Agniel et al (2018) showed that information *about* lab tests (e.g. whether the test was performed and when) was a better predictor of diagnosis codes than the *actual result* of the test.
- Lipton et al (2016), Sharafoddini et al (2019), and Agor et al (2019) all found that the simple practice of incorporating a binary missingness indicator into the feature space improved performance of supervised learning algorithms.
- Che et al (2018) found EHR missingness was correlated with mortality and various ICD-9 diagnosis categories; built a model to incorporate missingness, improved predictive performance.
- Nemati and Ghassemi (2016) used RL for optimal heparin dosing. Their work did not employ formal Off-Policy Policy Evaluation methods and did not address/incorporate missingness information.



Association of missingness with adverse outcomes. Figure adapted from Che et al (2016).

Dynamics under missingness

Methods: We assess whether aPTT dynamics are different when an auxiliary lab test is missing from the EHR by comparing the following sets of r.v.s:

$$X_{L,missing} = \{|\text{aPTT}_{t+1} - \text{aPTT}_{t-1}| : L_t \text{ missing}\}$$
$$X_{L,not\ missing} = \{|\text{aPTT}_{t+1} - \text{aPTT}_{t-1}| : L_t \text{ not missing}\}$$

We perform a Mann-Whitney *U* test to test against the null hypothesis that the two sets have the same distribution:

Table 1: Mann-Whitney U Test evaluating whether aPTT dynamics are significantly different under concurrent laboratory test missingness vs. non-missingness. (**bold**: significant after family-wise error rate control). Only significant results are shown here; see Table 3 in Appendix for full results.

Concurrent Lab, L	Median $X_{L,missing}$	Median $X_{L,not\ missing}$	p-value
Prothrombin Time (PT)	27.10	35.44	**1.61e-10**
International Normalized Ratio (INR)	27.10	35.44	**1.61e-10**
Total Bilirubin (TBILI)	32.71	36.21	**2.01e-4**
Aspartate Aminotransferase (AST)	32.67	36.24	**2.12e-4**
Alanine Aminotransferase (ALT)	32.71	36.23	**2.57e-4**
Erythrocyte Sedimentation Rate (ESR)	34.99	39.13	**2.20e-5**

Results: aPTT dynamics appear to be less volatile when the results of certain laboratory tests are missing from the EHR (esp. those related to liver failure)
Does this mean that missingness is a proxy for stability of aPTT?
Is this information useful for learning and evaluating a policy from data?

RL with Missingness Information

Method: We use Reinforcement Learning to learn a heparin dosage management policy and evaluate its estimated performance using EHR data.

Assume we are given a dataset $D = \{\tau^i = (s_0^i, a_0^i, r_0^i, s_1^i, \ldots, s_{|\tau^i|}^i)\}_{i=1}^N$ where x_t^i represents the i^{th} patient visit's trajectory, consisting of observations at time *t*, actions taken a_t^i, and rewards accrued r_t^i. In consultation with Stanford clinicians, we defined the state, action, and rewards as follows:

$$\text{Reward}_t = \begin{cases} -1 & \text{if aPTT test} \notin [40, 80] \text{ seconds at time } t \\ -1 & \text{if anti-Xa test} \notin [0.3, 0.7] \text{ IU/mL at time } t \\ -1 & \text{for every observation regardless} \end{cases} \qquad \text{Action}_t = \begin{cases} \text{Increase heparin dosage} \\ \text{Decrease heparin dosage} \end{cases}$$

We first learn an optimal policy with Fitted Q-Iteration, an Off-Policy Policy Optimization (OPPO) Algorithm (Ernst et al, 2005):

Algorithm 1: Fitted Q-Iteration for Off-Policy Policy Learning

Collect dataset D of N trajectories, $\tau^1, \tau^2, \ldots, \tau^N$
i.e. $D = \{\tau^i = (s_0^i, a_0^i, r_0^i, s_1^i, \ldots)\}_{i=1}^{|D|}$
Initialize $\hat{Q}_0(s, a) \leftarrow 0\ \forall s \in \mathcal{S}, a \in \mathcal{A}$
for $k = 1, 2, \ldots, K$ **do**
  for $i = 1, 2, \ldots, N$ **do**
    for $t = 1, 2, \ldots, |\tau^i| - 1$ **do**
      $y_t^i \leftarrow r_t^i + \gamma \max_a \hat{Q}_{k-1}(s_{t+1}^i, a)$
    end
  end
  $\hat{Q}_k \leftarrow \arg\min_Q \sum_{i=1}^N \sum_{t=1}^{|\tau^i|-1} \left(Q(s_t^i, a_t^i) - y_t^i\right)^2$
end
$\widehat{\pi}^*(s) \leftarrow \arg\max_a \hat{Q}_K(s, a)$
return $\widehat{\pi}^*$

We then estimate the performance of our learned policy using only the available data generated under the clinician policies. We experiment with four popular Off-Policy Policy Evaluation (OPPE) algorithms (Voloshin et al, 2019):

Policy	Standard	Step-Wise
IS	$\hat{V}_{IS}^{\pi_e} = \frac{1}{N}\sum_{i=1}^N \rho_{0:H_i}^i \left(\sum_{t=0}^{H_i} \gamma^t r_t^i\right)$	$\hat{V}_{step-IS}^{\pi_e} = \frac{1}{N}\sum_{i=1}^N \sum_{t=0}^{H_i} \rho_{0:t}^i \gamma^t r_t^i$
WIS	$\hat{V}_{WIS}^{\pi_e} = \sum_{i=1}^N \frac{\rho_{0:H_i}^i}{\sum_{j=1}^N \rho_{0:H_j}^j} \left(\sum_{t=0}^{H_i} \gamma^t r_t^i\right)$	$\hat{V}_{step-WIS}^{\pi_e} = \sum_{i=1}^N \sum_{t=0}^{H_i} \frac{\rho_{0:t}^i}{\sum_{j=1}^N \rho_{0:t}^j} \gamma^t r_t^i$

where $H_i = |\tau_i| - 1$ represents the number of state-action pairs in the patient history and $\rho_{j:j'}^i = \rho_{j:j'}(\tau^i, \pi_e, \pi_b) = \prod_{t=j}^{\min(j', |\tau^i|-1)} \frac{\pi_e(a_t^i|s_t^i)}{\pi_b(a_t^i|s_t^i)}$ represents the importance weights. We experiment with including missingness information during just OPPO, just OPPE, and both OPPO and OPPE (except where confounding is clearly an issue i.e. OPPO with missingness info and OPPE not).

Results:

Table 2: Impact of Including Missingness Information in Patient State Representation on Off-Policy Policy Optimization (OPPO) and Off-Policy Policy Evaluation (OPPE). Numbers reported are the average ± 1 standard deviation of OPPE estimates over the validation folds from 5-fold CV.

Missingness Information Used?		Expected Return (Off-Policy Evaluation)			
OPPO	OPPE	Standard IS	Standard WIS	Stepwise IS	Stepwise WIS
No	No	-0.547 ± 0.074	-2.159 ± 0.236	-1.137 ± 0.097	-3.843 ± 0.156
No	Yes	-0.538 ± 0.076	-2.155 ± 0.246	-1.118 ± 0.100	-3.838 ± 0.178
Yes	Yes	**-0.504 ± 0.086**	**-2.049 ± 0.253**	**-1.046 ± 0.108**	**-3.710 ± 0.232**

Conclusions

- Dynamics of aPTT were significantly less volatile when some of the auxiliary lab values were missing from the EHR, suggesting that missingness may indicate physiological stability.
- Including an indicator for missing lab values consistently (but not significantly) increased the estimated reward of our learned policy.
- This could be because (1) the policies learned with a missingness-augmented state space are actually better, or (2) including missingness information yields better propensity score estimates and/or higher confidence estimators of the expected reward under our learned policy.
- Lack of significance may well be attributable to high variance of IS estimators.
- Binary indicators for missingness are a simple but potentially effective way to increase estimated rewards for Off-Policy RL tasks in healthcare.

Open Questions/Future Work

- Does an increase in estimated reward from OPPE correspond to actual improved performance?
- How do policies trained with missingness information differ from policies trained without?
- Are the propensity scores computed more accurate when missingness info is included?
- Do our findings hold when applied to data from different hospital systems?
- Do our findings hold for different clinical contexts/problems (e.g. sepsis)?

References

Agniel, D., Kohane, I. S., & Weber, G. M. (2018). Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. *Bmj*, *361*, k1479.

Agor, J., Özaltın, O. Y., Ivy, J. S., Capan, M., Arnold, R., & Romero, S. (2019). The value of missing information in severity of illness score development. *Journal of biomedical informatics*, *97*, 103255.

Brill-Edwards, Patrick, et al. "Establishing a therapeutic range for heparin therapy." Annals of internal medicine 119.2 (1993): 104-109.

Bussey, Henry, John L. Francis, and Heparin Consensus Group. "Heparin overview and issues." Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy 24.8P2 (2004): 103S-107S.

Che, Z., Purushotham, S., Cho, K., Sontag, D., & Liu, Y. (2018). Recurrent neural networks for multivariate time series with missing values. *Scientific reports*, *8*(1), 6085.

Ernst, D., Geurts, P., & Wehenkel, L. (2005). Tree-based batch mode reinforcement learning. Journal of Machine Learning Research, 6(Apr), 503-556.

Lipton, Z. C., Kale, D., & Wetzel, R. (2016, December). Directly modeling missing data in sequences with RNNs: Improved classification of clinical time series. In *Machine Learning for Healthcare Conference* (pp. 253-270).

Lowe, H. J., Ferris, T. A., Hernandez, P. M., & Weber, S. C. (2009). STRIDE–An integrated standards-based translational research informatics platform. In AMIA Annual Symposium Proceedings (Vol. 2009, p. 391). American Medical Informatics Association.

Nemati, S., Ghassemi, M. M., & Clifford, G. D. (2016, August). Optimal medication dosing from suboptimal clinical examples: A deep reinforcement learning approach. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2978-2981). IEEE.

Price, Elizabeth A., et al. "Discordant aPTT and anti-Xa values and outcomes in hospitalized patients treated with intravenous unfractionated heparin." Annals of Pharmacotherapy 47.2 (2013): 151-158.

Sharafoddini, A., Dubin, J. A., Maslove, D. M., & Lee, J. (2019). A New Insight Into Missing Data in Intensive Care Unit Patient Profiles: Observational Study. *JMIR medical informatics*, *7*(1), e11605.

Voloshin, C., Le, H. M., Jiang, N., & Yue, Y. (2019). Empirical Study of Off-Policy Policy Evaluation for Reinforcement Learning. arXiv preprint arXiv:1911.06854.

Acknowledgements

Scott Fleming was supported by a Stanford Graduate Fellowship, a National Defense Science & Engineering Graduate Fellowship, and an NIH BD2K grant (5-T32-LM-12409-3). We received support and resources from Stanford Research IT regarding data. Research IT, and the Stanford Clinical Data Warehouse (CDW), are supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through grant UL1 TR001085. The content of studies done using the CDW is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.