

Summary: Venous thromboembolisms (VTEs) - blood clots that form within patients' veins and can subsequently travel through the circulatory system to critical organs like the heart, lungs, and brain - are one of the most common causes of hospital mortality in America. In this work we demonstrate the use of contextual multi-armed bandits, a machine learning algorithm, to learn from a dataset of 7,433 patients and identify treatment strategies which can increase the effectiveness of prophylactic anticoagulation therapy for patients in the intensive care unit (ICU).

Context: More than half of hospitalized patients in the US are at risk for VTEs such as deep vein thromboses (DVTs) or pulmonary embolisms (PEs) [1]. PEs alone are estimated to cause between 5% and 10% of all hospitalized patient deaths in the US and UK [2,3]. Fortunately, prophylactic anticoagulation therapy with pharmacologics like UFH has been shown to effectively reduce patients' risk of VTEs [4]. Prior research suggests that two laboratory tests, the activated partial thromboplastin time (aPTT) and anti-factor Xa chromogenic assay (anti-Xa), provide useful metrics for titrating patient heparin dosages [5-7]. Specifically, patients who fall outside of "therapeutic ranges" for these laboratory values are at higher risk of bleeding or developing a VTE. Finding the right heparin dose to achieve these therapeutic ranges can be difficult, however, when the patient's state is complex, multidimensional, and dynamic, and when the patient responses to medical interventions are heterogeneous. This is frequently the case for hospitalized patients [8,9]. While previous work has employed reinforcement learning algorithms to find strategies which optimize time spent in the therapeutic range for aPTT [10], no work to date has developed such a strategy for anti-Xa therapeutic range targeting. Additionally, as aPTT and anti-Xa test results are often discordant, it is unclear the degree to which targeting the therapeutic range for aPTT also achieves therapeutic range targeting for anti-Xa values and vice versa [11,12]; our work aims to shed light on this question.

Objective: To (1) learn treatment strategies that effectively increase the time a patient spends in the therapeutic range for aPTT and anti-Xa values by analyzing historical patient data, and (2) evaluate the degree to which treatment strategies designed to target the aPTT therapeutic range agree or disagree with treatment strategies designed to target the anti-Xa therapeutic range.

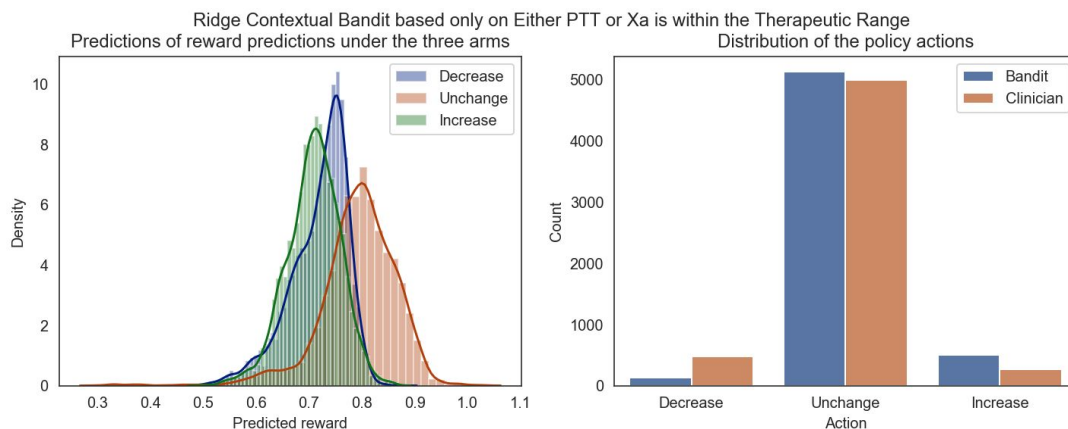
Data: We extracted and analyzed the Electronic Health Records (EHRs) of 9,026 patients collected between 2012 and 2018 (5,428 (60.1%) were male and the mean (SD) age was 60.2 (17.8) years) corresponding to 42,360 simultaneous observations of aPTT and anti-Xa test results in the Stanford Translational Research Integrated Database Environment (STRIDE) [13]. In addition to aPTT and anti-Xa test results, demographic information and other available lab tests relevant to the heparin dosing problem were included in our analysis.

Methods: We used contextual multi-armed bandits [14] to learn treatment strategies that increase the overall expected time a patient spends in the therapeutic range for the aPTT and anti-Xa tests. In the training phase, the contextual multi-armed bandit algorithm learns to map from the patient's context vector (e.g. demographics and relevant lab test results) to a predicted reward for each action. We assigned a reward of +1 to actions that led to the patient's anti-Xa and/or aPTT lab test results being within the therapeutic range in the subsequent observation and discretized actions to be one of increasing, decreasing, or maintaining the current heparin dose. We used ridge regression [15] to estimate the function mapping from patient context to predicted reward. After training, the learned treatment strategy, or policy, was extracted by taking the action that had the highest predicted reward according to our bandit model for the given patient context vector. We evaluated our policy on a held-out test set of 1,048 patients using off-policy policy evaluation (OPPE) with a weighted doubly robust estimator [16].

Results: We learned three treatment strategies from the STRIDE data. The first strategy, the aPTT-optimizing treatment policy, learned to predict and optimize actions which increase the expected fraction of time a patient spends in the therapeutic range for aPTT. The second, anti-Xa-optimizing strategy learned to predict and optimize actions which increase expected time spent in the therapeutic range for anti-Xa. The third learned policy was rewarded for and subsequently optimized time spent in the therapeutic range for either anti-Xa *or* aPTT (both counted equally toward the model’s rewards). These policies were compared against the treatment strategy employed by clinicians in the ICU. Under the aPTT-optimizing policy and with common assumptions of nonconfounding and coverage, our OPPE results show that, on average (SD), 59% (34%) of the patient’s observations would be in the therapeutic range for aPTT, and 32% (40%) would be in the therapeutic range for anti-Xa, and 75% (33%) would be in the therapeutic range for either. Under our anti-Xa-optimizing policy, 55% (55%) of patient observations would be in the therapeutic range for aPTT, 36% (53%) would be in the therapeutic range for anti-Xa, and 75% (47%) would be in the therapeutic range for either. By comparison, under the observed clinician strategy, 55% (50%) of a patient’s observations would be in the therapeutic range for aPTT, 37% (48%) would be in the therapeutic range for anti-Xa, and 76% (43%) would be in the therapeutic range for either. Of the 5,750 observations in the test set, our aPTT-optimizing policy’s suggested action aligned with clinician on 3,829 (66%); our anti-Xa-optimizing policy actions aligned with the clinician on 1,246 (22%); and our aPTT-optimizing policy and anti-Xa-optimizing policy agreed with each other 1,795 (31%) observations.

Conclusions and Relevance: To our knowledge, our work is the first demonstration of using artificial intelligence methods to learn a heparin dosing strategy that optimizes for the proportion of time a patient spends in the therapeutic range of the anti-Xa laboratory test. Our results demonstrate that our learned policy performs comparably to clinicians in terms of expected proportion of time a patient spends in the anti-Xa therapeutic range. Our findings also suggest that targeting the anti-Xa therapeutic range is a more challenging task than targeting the therapeutic range for aPTT. Furthermore, the relatively small overlap between suggested actions for our aPTT-optimizing and anti-Xa-optimizing strategies suggest that optimizing for the aPTT therapeutic range does not necessarily lead to optimal performance for anti-Xa therapeutic range compliance. Further work is needed to evaluate which of these strategies lead to reduced risk of mortality and thrombotic events in a sequential decision making context.

Figure: Illustration of our model’s predicted rewards associated with each action for all patient observations in the test set.



References:

- [1] Anderson Jr, Frederick A., et al. "Estimated annual numbers of US acute-care hospital patients at risk for venous thromboembolism." *American journal of hematology* 82.9 (2007): 777-782.
- [2] Alikhan, R., et al. "Fatal pulmonary embolism in hospitalised patients: a necropsy review." *Journal of clinical pathology* 57.12 (2004): 1254-1257.
- [3] Sandler, D. A., and J. F. Martin. "Autopsy proven pulmonary embolism in hospital patients: are we detecting enough deep vein thrombosis?." *Journal of the Royal Society of Medicine* 82.4 (1989): 203-205.
- [4] Geerts, William H., et al. "Prevention of venous thromboembolism: American College of Chest Physicians evidence-based clinical practice guidelines." *Chest* 133.6 (2008): 381S-453S.
- [5] Basu, Dilip, et al. "A prospective study of the value of monitoring heparin treatment with the activated partial thromboplastin time." *New England Journal of Medicine* 287.7 (1972): 324-327.
- [6] Hull, Russell D., et al. "Optimal therapeutic level of heparin therapy in patients with venous thrombosis." *Archives of Internal Medicine* 152.8 (1992): 1589-1595.
- [7] Levine, Mark N., et al. "A randomized trial comparing activated thromboplastin time with heparin assay in patients with acute venous thromboembolism requiring large daily doses of heparin." *Archives of internal medicine* 154.1 (1994): 49-56.
- [8] Hylek, Elaine M., et al. "Challenges to the effective use of unfractionated heparin in the hospitalized management of acute thrombosis." *Archives of internal medicine* 163.5 (2003): 621-627.
- [9] Wheeler, Arthur P., Robert DB Jaquiss, and John H. Newman. "Physician practices in the treatment of pulmonary embolism and deep venous thrombosis." *Archives of internal medicine* 148.6 (1988): 1321-1325.
- [10] Nemati, Shamim, Mohammad M. Ghassemi, and Gari D. Clifford. "Optimal medication dosing from suboptimal clinical examples: A deep reinforcement learning approach." 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2016.
- [11] Zehnder, James, Elizabeth Price, and Jing Jin. "Controversies in heparin monitoring." *American journal of hematology* 87.S1 (2012): S137-S140.
- [12] 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.
- [13] Lowe, Henry J., et al. "STRIDE—An integrated standards-based translational research informatics platform." *AMIA Annual Symposium Proceedings*. Vol. 2009. American Medical Informatics Association, 2009.
- [14] Lattimore, Tor, and Csaba Szepesvári. "Bandit algorithms." preprint (2018).
- [15] Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*. Vol. 1. No. 10. New York: Springer series in statistics, 2001.
- [16] Dudik, Miroslav, John Langford, and Lihong Li. "Doubly robust policy evaluation and learning." arXiv preprint arXiv:1103.4601 (2011).