

This paper proposes a model based approach for personalized heparin dosing using a piecewise-linear treatment effect model. Their algorithm aims to maintain a patient's aPTT within a safe range, dynamically adjusting patient doses based on their prior history. They evaluate their method on MIMIC data and demonstrate improved safety.

This was a very interesting read, and I commend the authors for tackling an important and impactful problem. I requested two highly-qualified referees to evaluate the paper and I'm grateful for their detailed feedback. Unfortunately the reviews are somewhat negative and aligned on the key issues. I'll summarize the issues (raised by both reviewers) here, but please read their reports for more details.

1. How valid is the model? Both reviewers point out that the piecewise-linear approximation is introduced without justification. Model based approaches are only superior to model free approaches if the model is accurate. There is no characterization of the approximation error introduced by this approximation, no justification for why it was chosen, and no comparison to model free methods that can avoid this. Given that this is a huge motivation for the paper, it needs better justification.
2. Evaluation. This is related to the previous point in that the empirical results do not support the assumed model. Can the authors do off-policy optimization to support the claim that their approach may actually help in practice?
3. Asymptotic optimality vs regret. The authors show that their method is asymptotically optimal, but so are many naïve methods (and model free methods as well). A regret result would better show the learning rate, and how long one can anticipate poor decisions before reaching a safe state.
4. Safety. Related to the previous point, there is no result demonstrating safety, which is a key point of motivation for adopting this model based approach.

The reviewers also raise a number of issues with the assumptions and the design of the algorithm. For instance, why is there no learning based on offline data? Again, this seems important for the stated goal of safety.

Addressing these concerns (both theory and empirical) may require a fundamentally different paper, so I regretfully recommend reject & resubmit. However, if the authors are able to address these issues, this could be a valuable contribution to the journal.