# Review for OPRE-2023-04-0206

This paper develops a model-based algorithm for personalized and adaptive dosing for heparin. The paper establishes a patient-level model of how heparin metabolizes based on aPTT measurements, and this model has several patient-specific parameters. Based on the historical observations for a patient, the algorithm uses a scenario-generation approach to compute the next dose, which takes the uncertainty of the parameters into account. The paper proves that this algorithm is asymptotically consistent, and then validates the performance of this approach via numerical simulations.

## Comments

The paper studies an important topic, and I commend the authors for using deep medical knowledge on heparin to derive the model in this work. The direction of using data-driven, machine learning algorithms in the medical setting has a ton of potential for real-world impact. This paper provides a comprehensive model and framework for personalized heparin dosing via a data-driven approach, backed by both theoretical and empirical evaluations. However, I have several major concerns with this work, which I elaborate on below.

1. **Contribution as a prescriptive algorithm.** At a high level, a major concern is in regards to the nature of the main contribution. The paper is framed as a prescriptive guide for personalized heparin dosing. The paper presents a model for heparin dosing (backed by domain knowledge), develops an algorithm that is tailored for this model, shows theoretical guarantees, and numerically evaluates the algorithm with respect to this model. However, since all models are wrong, I find this approach to be insufficient to show that this algorithm would actually work well on real patients. In general, I find such a claim to be a difficult one to make without a randomized control study, or perhaps a very careful and rigorous off-policy evaluation (e.g. if there is a dataset where one can evaluate exact counterfactuals). These are often very difficult or infeasible, and hence if the authors cannot perform such a rigorous evaluation, then it is difficult to claim that this paper has a meaningful contribution to heparin dosing.

   One alternative is to frame the main contribution as the insights that arise from this model (which in my opinion, is one of the main benefits of a model-based approach). The authors list a couple of insights at the end, but I find these too vague and they were not rigorously validated. An example of an insight that could arise from a model-based approach such as

this would be: "A one-size-fits-all policy works well for most patients, but patients with characteristic X benefits drastically from personalization, so physicians should be extra vigilant with such patients." This statement is completely made up, but this is an example of a much more specific and actionable insight, compared to the ones currently listed.

2. **Comparison to model-free methods.** As a methodological framework, another concern is similar in spirit to the previous concern, which is the lack of comparison of this particular model against other approaches that do not use the same model. The authors claim that the main advantage of their method relative to existing approaches is that it is model-based — however, the authors do not provide any reasons for why a model-based approach is superior to model-free approaches, and there are no attempts to compare to other any other methods, either theoretically or empirically. Personally, I generally think of model-based approaches as inferior to model-free approaches in terms of providing a prescriptive algorithm, since the former relies on a specific model being correct. The following sentence found in the abstract seems contradictory: "While many proposed solutions are model free, they require complex models".

   The literature review mentions some model-free methods, and the authors write *"they can be thought of as providing a one size fits all dosing policy for patients. Therefore, even though the overall empirical performance of these methods seems strong, the resulting policy might not be suitable for a specific patient"*.

   (a) The authors do not specify what "might not be suitable for a specific patient" means: if the authors are implying that the performance will be worse, then this should be shown either theoretically and/or empirically.

   (b) The authors should expand on what a "one size fits all" policy refers to, because in my mind, the method developed in this paper is also one size fits all. That is, the method in this paper effectively produces a function from "history of actions and observations for a patient" to a "dose" and this same function is used to determine the next dosage for every patient. A model-free approach would also be able to output such a mapping, so it is not the case that a model-free approach is inherently different than a model-based approach. The only way (that I see) in which a dosing policy would be fundamentally inferior is if it was non-adaptive to new data.

3. **No learning across patients.** Practically, a major disadvantage of the proposed method is that it does not leverage learning across patients: it uses the data for each patient separately. I find it hard to believe that treating each patient separately is advantageous compared to learning across patients, since it is not leveraging the tons of offline data that exists. In fact, the ability to learn across patients is more natural in a model free-approach: one could learn a mapping from (patient covariates, history of observations) to next dose.

4. **Algorithm.** Several parts of the algorithm seemed ad-hoc or unspecified.

(a) The prediction part of the algorithm keeps some parameters fixed ($\alpha$, $\beta$), while computing the MLE over the other parameters. The parameters that are fixed are the ones where there are $m$ scenarios that are sampled. I found this approach to be quite confusing without much discussion. Why are the parameters separated in this way? Is it because the parameter spaces $\mathcal{A}$ and $\mathcal{B}$ can be discretized? (If so, the other parameter spaces can also be discretized?) Proving a negative result about the naive MLE would be helpful (see comment c).

(b) How are the $m$ scenarios chosen? I could not find where this is specified. Choosing these scenarios is likely the most important part of the algorithm from a practical perspective, but there seems to be no discussion on this (as well as choosing $m$).

(c) The main result of Theorem 1 says that the optimal sequence of doses is asymptotically optimal. Is it the case that the "naive MLE estimation procedure" (described in the middle of page 22) would not satisfy the same theorem? I have never seem this type of statement before, so I'm not sure how to think about it. Could the authors use a more "standard" metric like regret, or alternatively show that a sublinear regret is impossible, and hence one needs to loosen the benchmark? Otherwise, since the authors are introducing a new metric, the authors should show some type of comparison (e.g. show that the same result doesn't hold on a naive algorithm), to demonstrate the significance of this result.

5. "Safety" is something that is greatly emphasized, but this notion is not formalized, and it's not clear which part of the algorithm guarantees safety. The algorithm samples different parameters over an uncertainty set, but it still computes an expectation over the sampled parameters (rather than a worst-case mechanism, which is my interpretation of what safety should represent). It doesn't seem like there are any theoretical results related to safety.

In summary, I am not convinced of the contribution of this paper as a prescriptive algorithm for heparin (comment 1), and I also have major concerns about the methodology (comments 2-5). This paper could potentially be reframed as insights for heparin dosing (as in comment 1), or perhaps as a more general methodological contribution to model-based RL (i.e. not specific to heparin). However, as such would require a substantially different paper, I unfortunately recommend rejection.