A User Study on Data-Driven Dosage of Heparin

Joe Rowley

A thesis presented for the degree of

Bachelor of Science

Advised by:

Sri Kurniawan

Mohammad M. Ghassemi

University California, Santa Cruz

December 2015

Maybe include keywords section here:

Retrospective data driven dosing, HCI, Web Application Development

Abstract

The intention of this study was to determine the feasibility of using statistical drug dosing tools in a mock clinical setting. We wanted to learn about both the technical implementation and user experience challenges involved with creating and using computer aided dosing tools. First a web application was developed using existing models for dosing heparin in ICU patients. This application was next developed into an interactive user survey. We found that Doctors were generally more confidant and faster when dosing using the tools. [[add other results here too. It's good to avoid numbers in abstracts from what I recall]] In closing we recommend some best practices for future development of statistical drug dosing tools.

Acknowledgements

The models used in this tool were introduced Mohammad M. Ghassemi, Stefan E. Richter, Ifeoma M. Eche, Tszyi W. Chen, John Danziger, and Leo A. Celi in their paper *A data-driven approach to optimized medication dosing: a focus on heparin*. This project would not have been possible without Mohammad's support.

Additionally, data used in these models was sourced from the MIMIC-II and MIMIC-III (Saeed et al. (2011)) medical databases, a project supported by the MIT Lab for Computational Physiology.

I also have many advisors who helped me with this project:

- Mohammad M. Ghassemi
- Sri Kurniawan
- Matthew Guthaus

Additionally the following people were very influential in preparing me for this project:

- James Davis
- Phokion Kolaitis
- Philip Strong

Finally, this project would not have been possible without the unwavering moral and technical support of my parents Howard Rowley and Carol Diamond.

Table of Contents

1 Introduction

The primary goal of this study was to determine the technical implementation challenges and user experience challenges involved with a statistical dosing tool. Specifically, we implemented a heparin dosing tool that was first introduced in the paper *A data-driven approach to optimized medication dosing: a focus on heparin* (Ghassemi et al. 2014). In the paper a statistical model is generated for predicting an optimal infusion rate for heparin for patients in intensive care units (ICUs). In this study, we reimplemented their existing models into an interactive web application, which allows the user to generate dosage guidelines based on a patient's features. Next a user study application was created, which presents the doctors with a patient case. The doctor initially provides their own suggestion for the infusion rate and is then presented with the model's prediction for an optimal infusion rate. After the doctor views the model, they have the opportunity to adjust their initial dose

and provide a reasoning for their dose adjustments. Their interaction with the tool yields many metrics, including number of dose adjustments, speed of dosing, number of times guidelines were viewed.

2 Background

2.1 Traditional Dosing of Heparin

Heparin is an anticoagulant medication administered intravenously to patients with blood clots. Specifically it is often used for strokes, pulmonary embolism (blockage in an artery in the lung), acute coronary syndrome (e.g. heart attack), and many other conditions. The general protocol for dosing heparin is to initially order baseline lab results for the patient. These values combined with context of the patients condition allow a doctor to determine the infusion rate and if a bolus (a single large initial dose) is required. The primary metric used for monitoring patient progress is activated partial thromboplastin time (APTT). After the infusion is started the patient is monitored and the dose is adjust periodically. Most often new labs are ordered every 6 hours [citation to UW and Beth Israel].

2.2 Predicting Therapeutic Outcome

The objective of the dosing tool is determine an ideal initial infusion rate, such that patient has the highest probability of reaching a therapeutic state in 6 hours time. A therapeutic state is defined as an APTT value of between 60 seconds and 100 second. To predict an optimal dose Ghassemi et al.

employed the use of multi-variant logistic regression.

2.2.1 Logistic Regression Approach

The objective of the model was to determine the probability that patient would end in one of three categories: sub-therapeutic, therapeutic, or supra-therapeutic. A patient can by definition only be in one category at a time. This lead Ghassemi et al. to the following equation:

(1) $1 = P(sub{-}therapeutic) + P(therapeutic) + P(supra{-}therapeutic)$

The challenge of using logistic regression to predict therapeutic outcomes is that the probability of a therapeutic outcome is not a strictly monotonic function. So using logistic regression directly to predict P(therapeutic) will not work. That being said, the probability of sub-therapeutic and supra-therapeutic outcomes are monotonically decreasing and increasing functions respectively. Given these properties, we can use multinomial logistic regression to determine functions for sub-therapeutic and supra-therapeutic aPTT. Using equation 1 we can then determine the probability of a therapeutic outcome.

Although at this point a case been has made that multinomial logistic regression may be an appropriate tool, there are a few considerations the Ghassemi et al. took with regards to the form of regression. There are many forms of multinomial logistic regression with subtle but important differences. There is a natural order to the classifications (sub-therapeutic, therapeutic, supra-therapeutic), so ordinal logistic regression could be considered. The challenge with ordinal logistic regression, however, is that it makes the assumption of proportional odds, so regression intercept terms vary across

classes while regression coefficients are shared across classes. In the words of the original authors *"it assumes that explanatory features maintain identical effects across varying ranges of the outcome"*[citation appendix]. Instead the ordinal regression they instead used a modified version of standard multinomial logistic regression, which utilized a flexible reference category. That is instead of using a single class to predict the outcome that is not being modeled, they used all classes. So when modeling supra-therapeutic aPTT, standard multinomial logistic regression would reference either therapeutic or sub-therapeutic aPTT classes, while their approach referenced both classes. This choice was demonstrated to have significant effects on the validity of their models [citation to appendix, section 2].

2.2.2 Feature Selection

There are innumerable potential features one could consider as predictors of patient outcome. In their paper, Ghassemi et al. determined the salient features as follows: age, SOFA score, Elixhauser, Heparin dose, Measurement time, Creatinine, Ethnicity, Gender, ICU type, presence of pulmonary embolism, obesity, presence of end stage renal failure.

2.3 Summary of Their Results

The original paper and associated appendix document the efforts taken to validate the models. To summarize, they partitioned their dataset, dedicating 70% to training and 30% for validation. They then compared the calculated the predicted classification for each patient using both a full fea-

tured model and a weight-only model. The full featured model was more accurate than the weight-only approach as measured by Volume Under the Receiver Operating Characteristic Surface (the multiple class version of Area Under Receiver Operating Characteristic Curve), with values of 0.48 vs 0.42 [citation].

In closing, they noted that a randomized controlled trial would be necessary to determine if this mechanism is more effective for dosing. This thesis project represents the initial steps towards this trial.

3 Calculator Explanation

This project is composed of two key components. The first is a "dose calculator" which contains all the needed components to determine the optimal dose for a given patient. The second component is a survey that incorporates the calculator and is used to study how doctors interact with the dosing tool. We will now describe the calculator.

3.1 Functional Requirements

The purpose of the calculator is to determine the functions for the probability of sub-therapeutic, therapeutic, supra-therapeutic for a range of infusion amounts for a given patient. This functionality required some key steps: 1. Access Data and filter it. 2. Calculate Static Models 3. Make a prediction based on a given patient's features.

3.2 Architecture

3.3 Implementation

4 Survey Explanation

4.1 Goals of Survey

what the goals of the survey were. - what did we want to learn? hypothesis how the survey was conducted. how we decided what information to ask about. what information that was logged.

4.2 Architecture and Implementation

talk about the stack, tools used etc. how it was implemented. user experience.

5 Survey Results

5.1 Cumulative Results

how many difference people took survey, specialties, time since last dosed heparin.

5.2 Dose differences

how did the doses differ in part 1 and part 2 of the survey.

5.3 Time differences

time difference in part 1 and part 2.

maybe add some graphs here.

5.4 Understanding/Confidence in Results

....

5.5 Other results

things that weren't expected.

other things of note

6 Retrospective / Next Steps

6.1 Shortcomings in the Survey

- errors

- lack of data
- better explanation of how the model/logistic regression works. explain

what features are being used in the prediction and what aren't. suggest a potential user interface that gives weights to things e.g. The patients suggested dose deviates from a baseline dose curve due to the patients ethnicity and creatinine and mean SOFA scores.

6.2 Places for Future Research

how did the doses differ in part 1 and part 2 of the survey.

6.3 Implementation Differences in future version

- transfer model to device so it can do queries more quickly
- instead of static data source have real medical database
- easily interchangeable models
- better analytics

7 Conclusion

Summary of work that was done, then summary of results. Then summary of places for future work.

Appendix 1: Some extra stuff

- link to test software
- link to source code git repos

- screenshots and links to alternative design for testing aPTT over time
- links to data sources/notebook of documentation
- extended results from survey.

  .. tables with stats like ave, std deviation etc for each patient 1-10 ..-

1. First ordered list item
2. Another item

   Unordered sub-list.
3. Actual numbers don't matter, just that it's a number –1. Ordered sub-list

   –1. Ordered sub-list –1. Ordered sub-list

4. And another item.
5. first item in the list
6. second item in the list

- subitem
- subitem

1. third item in the list

- an entry
- another entry

- some sub entry without leading bullet
- – some sub entry with leading bullet
- another entry for another entry
- – blablabla
- – blublublu
- – * dfdf
- – * · also some way

Appendix 2: Some extra stuff

This could be a list of papers by the author for example Also tutorials/people/stack overflow that was helpful. tom's markdown -> latex

References

Ghassemi, M.M. et al., 2014. A data-driven approach to optimized medication dosing: A focus on heparin. *Intensive care medicine*, 40(9), pp.1332–1339.

Saeed, M. et al., 2011. Multiparameter intelligent monitoring in intensive care iI (mIMIC-iI): A public-access intensive care unit database. *Critical Care Medicine*, 39, pp.952–960.