

Assessing Machine Learning and Deep Learning Models for Suggested Dosing of Anesthetic Induction Medications

Presenting Author: Samir Kendale, MD; NYU Langone Health Department of Anesthesiology, Perioperative Care, and Pain Medicine

Introduction: As we continue to learn to apply machine learning techniques to the perioperative space, we can begin to envision automation of more complex aspects of anesthesiology. Prediction of hypotension combined with closed-loop vasopressor administration, and processed EEG-targeted medication infusions are examples of how processing large amounts of well labeled data can potentially be used for intraoperative management. Induction of general anesthesia requires careful and deliberate decision-making regarding medication choice and dosage. The training of the anesthesiologist in the complex interactions between physiology and pharmacology allows for selection of appropriate induction medications. We hypothesized that we would be able to train a suite of machine learning models to predict induction doses of commonly used medications for induction of general anesthesia, in this pilot study focusing only on propofol and fentanyl.

Methods: For this pilot study, we used 6 months of data (January to June 2019) extracted from the electronic health record (Epic Systems, Verona, WI) and initially investigated two drugs commonly used during induction of general anesthesia: propofol and fentanyl. Split-set validation was used in which 70% of data were used for training and 30% for testing. Models included age, sex, weight, body mass index, preoperative comorbidities, preoperative medications, and natural language processed procedure text as features. Three algorithms were trained on each target, using 5-fold cross validation: extreme gradient boosting model, deep learning, and elastic net, using mean absolute error as the primary performance metric. The best performer was further tuned and applied to the test set for validation.

Results: After tuning, the extreme gradient boosting model for prediction of propofol dosing performed best, with a mean absolute error of 37 mg and a root mean squared error of 54 mg on the training set, and a mean absolute error of 37 mg and root mean squared error of 53 mg on the test set. The extreme gradient boosting model for prediction of fentanyl dosing performed best, with a mean absolute error of 44 mcg and a root mean squared error of 63 mcg on the training set, and a mean absolute error of 44 mcg and root mean squared error of 66 mcg on the test set. Plotting of the errors suggested a normal distribution of error for both models.

Conclusions: We were able to develop two models for prediction of induction doses of propofol and fentanyl. These models could be deployed preoperatively for aid in tailoring an anesthetic plan by providing suggested induction doses, and may ultimately be daisy-chained to other models that can predict postinduction hypotension, resulting in a feedback loop for more precise induction medication selection. Performance of these models would likely improve with the addition of more data, as well as with the addition of more potentially relevant features, and should then be validated externally. Just as with any machine learning model, accurate representation is highly dependent on trusting of the labeled data. Future development of these models may include training of a multi target regression neural network model to incorporate the interaction between induction medications, though this will require significantly more data than used in this pilot study. Long-term possibilities include automated induction of general anesthesia.