Machine Learning Approach to Optimize Sedation Use in Endoscopic Procedures

Shorabuddin SYED[a,1], Mahanazuddin SYED[a], Fred PRIOR[a,b], Meredith ZOZUS[c], Hafsa Bareen SYEDA[a], Melody L. GREER[a], Sudeepa BHATTACHARYYA[a,d] and Shashank GARG[e]

[a]*Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR, USA*

[b]*Department of Radiology, University of Arkansas for Medical Sciences, Little Rock, AR, USA*

[c]*Department of Population Health Sciences, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA*

[d]*Department of Biological Sciences and Arkansas Biosciences Institute, Arkansas State University, Jonesboro*

[e]*Division of Gastroenterology and Hepatology, University of Arkansas for Medical Sciences, Little Rock, AR, USA*

Abstract. Endoscopy procedures are often performed with either moderate or deep sedation. While deep sedation is costly, procedures with moderate sedation are not always well tolerated resulting in patient discomfort, and are often aborted. Due to lack of clear guidelines, the decision to utilize moderate sedation or anesthesia for a procedure is made by the providers, leading to high variability in clinical practice. The objective of this study was to build a Machine Learning (ML) model that predicts if a colonoscopy can be successfully completed with moderate sedation based on patients' demographics, comorbidities, and prescribed medications. XGBoost model was trained and tested on 10,025 colonoscopies (70% - 30%) performed at University of Arkansas for Medical Sciences (UAMS). XGBoost achieved average area under receiver operating characteristic curve (AUC) of 0.762, F1-score to predict procedures that need moderate sedation was 0.85, and precision and recall were 0.81 and 0.89 respectively. The proposed model can be employed as a decision support tool for physicians to bolster their confidence while choosing between moderate sedation and anesthesia for a colonoscopy procedure.

Keywords. Colonoscopy, machine learning, electronic health records, moderate sedation, quality improvement.

1. Introduction

About 17-20 million endoscopy procedures are performed in the United States annually [1]. Traditionally, moderate sedation has been used to perform these procedures. However, in 10-33% of cases, patients do not tolerate the sedation well leading to an aborted or incomplete procedure that is repeated with anesthesia [2]. Deep sedation provided by an anesthesiology provider is an alternate to moderate sedation. It is highly

---

[1] Corresponding Author, Shorabuddin Syed, University of Arkansas for Medical Sciences, Little Rock, AR, USA; E-mail: ssyed@uams.edu.

successful but expensive and not as widely available [3]. At present, the decision to utilize moderate sedation or anesthesia for a procedure is made by the providers leading to a high variability in clinical practice. Prediction scores to determine tolerance of moderate sedation for endoscopy procedures have been developed. However, these scores are modest in predicting the success of moderate sedation (AUC ≤70%) [4].

Artificial Intelligence (AI) techniques including Machine Learning (ML) and Deep Learning (DL) are being used in a variety of clinical applications ranging from diagnosis to outcome prediction [5-7]. ML is an iterative process in which the algorithm tries to find the optimal combination of both model variables and variable weights with the goal of minimizing error in the predicted outcome [8]. With ease of data extraction from Electronic Health Record (EHR) systems, prediction accuracy of the AI models is surpassing clinical experts in certain domains [9]. A limited number of studies have applied AI techniques to improve quality of endoscopy procedures, specifically on the optimal method of sedation. To address this knowledge gap, we built a ML model to determine if a colonoscopy can be successfully completed with moderate sedation.

2. Methods

2.1. Data Source and Study Population

This is an Institutional Review Board (IRB) approved retrospective study that included all patients undergoing colonoscopy at the University of Arkansas for Medical Sciences (UAMS) from May 1st, 2014 to August 31st, 2020. UAMS is an integrated health care delivery organization that utilizes the EPIC platform (Epic Systems Corp, Verona, WI) for electronic health records (EMR). All adults (>=18 years of age) who had a colonoscopy with moderate sedation at UAMS during the study period were identified in the EMR. Patients who had colonoscopy: 1) with anesthesia, 2) without sedation, 3) with another procedure on same day, or 4) with documented poor bowel preparation were excluded from the study.

2.2. Feature Selection

The features included in this study were based on literature and feedback from a panel of board-certified gastroenterologists [4, 10]. The selected features included Body mass index (BMI) recorded at the time of the procedure, demographics (age, gender, race, ethnicity, zip code), history of smoking or alcohol use, illicit substances use, visit type (inpatient, outpatient, emergency), comorbidities, current medications that can affect tolerance of moderate sedation, provider performing the procedure, and total procedure time. Co-morbid illness assessment was done by calculating the Elixhauser Comorbidity Index (ECI) score using the Van Walraven algorithm [11, 12]. Patients were divided into 4 groups based on BMI: BMI<30, class I obesity (BMI 30-34.99), class II obesity (BMI 35-39.99) and class III obesity (BMI 40 and above). Medications that can alter the tolerance or efficacy of moderate sedation were grouped into seven categories: 1) opiates, 2) antihistamines, 3) antiepileptic, 4) non-benzodiazepine hypnotics, 5) anti-depressants or anti-anxiety, 6) benzodiazepines, and 7) antipsychotics. Total number of these medicines were calculated for each patient from their medication list at the time of the procedure. Most of the colonoscopies were performed by gastroenterologists and each

provider was treated as a separate category. Hepatologists and surgeons performed significantly fewer procedures. Therefore, they were kept as two groups.

2.3. Data Labelling, analysis and Model Design

The primary outcome of this study was to determine if a procedure could be successfully performed with ≤ 10 units of moderate sedation. Midazolam and Fentanyl are used for moderate sedation at UAMS. Based on literature and feedback from a panel of board-certified gastroenterologists [4, 10], a sedation unit was defined as: 1) Midazolam 1 mg = 1 unit, and 2) Fentanyl 25 mcg = 1 unit. Each procedure in the dataset was labelled as low or high sedation based on the cutoff of 10 sedation units.

The features were normalized by subtracting the mean and dividing by the standard deviation. The dataset was randomly split into a training and a test set (70% - 30%). We developed the model to predict required sedation level for colonoscopy procedures using XGBoost [13]. The model was tuned and trained with the following hyperparameter settings: max_depth = 3, eta = 0.1, gamma = 0.2, min_child_weight = 4, subsample = 1, and nrounds = 145, the other hyperparameters were set to their default values. As the dataset was imbalanced, we used 10-fold cross validation. Performances of the model on the test set was evaluated by the following metrics: AUC, prediction accuracy, precision, and F1 scores.

3. Results

After applying the exclusion criteria, the final dataset included 10,025 colonoscopies performed on 8,681 distinct patients during the study period. Procedures done with low and high sedation were 7,404 and 2,621 respectively. Demographics and clinical characteristics of patients undergoing colonoscopy procedure for both training and validation set is shown in Table 1. The top three patient comorbidities in the dataset were hypertension (73%), obese (43%), and depression (36%). Before the colonoscopy schedule date, 25.1% of patients were on opioids (any of the aforementioned seven opioid categories).

The 10 fold cross-validated AUC on the testing set was 0.762 [0.753-0.777], F1-score on predicting procedures that need moderate sedation was 0.85, precision and recall were 0.81 and 0.89 respectively. However, the F1 score for predicting high sedation procedure was 0.46 .The AUC-ROC of the prediction model is shown in Figure 1.
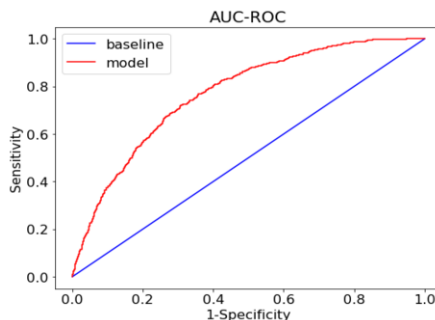
Figure 1. Receiver Operating Characteristics (ROC) curve describing XGBoost model's accuracy of predicting colonoscopies that can be successfully completed with moderate sedation.

We used a secondary model, Random Forest (RF) [14] to identify which discrete features were most influential in predicting sedation levels for the colonoscopy procedures. The top three features were: 1) physician performing the procedure, 2) total procedure time, and 3) patient's age.

Table 1. Clinical characteristics of patients undergoing colonoscopy procedure for both training and validation dataset.

Variables	Training dataset (n = 7,017) Patient count = 6,301	Test dataset (n = 3,008) Patient Count = 2,842
Age		
Mean ± SD	57.57 ± 12.59	58.20 ± 12.52
Range	18-84	18-83
Female No. (%)	4,070 (64.59%)	1,761 (61.96%)
Race		
White	60.31%	59.96%
Black	33.74%	33.92%
Other	5.95%	6.08%
Visit Type		
Outpatient	92.36%	92.22%
Inpatient	7.58%	7.75%
Procedure Time		
Mean ± SD	26.51 ± 12.47	26.73 ± 12.62
Range	5-90	5-87
Procedure Type		
Moderate Sedation (%)	5,158 (73.51%)	2,246 (74.67%)
High Sedation (%)	1,859 (26.49%)	762 (25.33%)

4. Discussion and Conclusion

To our knowledge, this is the first study to evaluate the use of ML to predict if a patient will require low or high dose of moderate sedation to have a successful colonoscopy. The XGBoost model achieved reliable accuracy (over 80%) in predicting procedures that could successfully be completed with moderate sedation. This is clinically meaningful as these patients can be scheduled for a colonoscopy with moderate sedation with a low rate of failure. Physician performing the procedure, total procedure time and age of the patient were the most important features and are consistent with observations in clinical practice. In addition, most of these features are included in the high conscious sedation requirements (HCSR) score built by McCain et al, 2020 [4]. According to the HCSR score, patient age has significant association to the level of sedation needed for successful completion of the procedure [4]. Physicians tend to have a significant variation in their sedation administration during the procedure that depends on their prior training, experience and patient population. Higher sedation doses are needed for longer procedures and for younger patients, and sedation requirements are expected to decrease with increasing age or shorter duration of a procedure.

There are multiple limitations to this study that we wish to acknowledge. We only included medication prescribed at UAMS and not active medications that the patients were on before the procedure. The active medication list is documented as free text and needs extensive chart review. In addition, the model's accuracy to predict procedures needing high sedation was less when compared to moderate sedation, further analysis is needed to understand features corresponding to deep sedation procedures. To overcome the mentioned limitations and improve the model's prediction accuracy, we intend to add more data to this model, to increase its accuracy over time so that it can be implemented in daily clinical practice to guide the choice of sedation for a colonoscopy.

References

[1] Peery AF, Dellon ES, Lund J, Crockett SD, McGowan CE, Bulsiewicz WJ, et al. Burden of gastrointestinal disease in the United States: 2012 update. Gastroenterology. 2012;143(5):1179-87.

[2] Hreinsson JP, Kalaitzakis E, Gudmundsson S, Björnsson ES. Upper gastrointestinal bleeding: incidence, etiology and outcomes in a population-based setting. Scandinavian journal of gastroenterology. 2013;48(4):439-47.

[3] Krigel A, Chen L, Wright JD, Lebwohl B. Substantial Increase in Anesthesia Assistance for Outpatient Colonoscopy and Associated Cost Nationwide. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2019;17(12):2489-96.

[4] McCain JD, Stancampiano FF, Bouras EP, DeVault KR, Gilbert EL, Ryan T, et al. Creation of a score to predict risk of high conscious sedation requirements in patients undergoing endoscopy. Gastrointestinal endoscopy. 2020;91(3):595-605.

[5] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology. 2017;2(4):230-43.

[6] Syeda HB, Syed M, Sexton KW, Syed S, Begum S, Syed F, et al. Role of Machine Learning Techniques to Tackle the COVID-19 Crisis: Systematic Review. JMIR medical informatics. 2021;9(1):e23811.

[7] Syed M, Syed S, Sexton K, Syeda HB, Garza M, Zozus M, et al. Application of Machine Learning in Intensive Care Unit (ICU) Settings Using MIMIC Dataset: Systematic Review. 2021;8(1):16.

[8] Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC medical informatics and decision making. 2019;19(1):281.

[9] Richens JG, Lee CM, Johri S. Improving the accuracy of medical diagnosis with causal machine learning. Nature Communications. 2020;11(1):3923.

[10] Benson M, Hubers J, Caldis M, Gopal D, Pfau P. Safety and Efficacy of Moderate Sedation in Super Obese Patients Undergoing Lower and Upper GI Endoscopy: a Case-Control Study. Obesity surgery. 2020;30(9):3466-71.

[11] van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Medical care. 2009;47(6):626-33.

[12] Syed S, Baghal A, Prior F, Zozus M, Al-Shukri S, Syeda HB, et al. Toolkit to Compute Time-Based Elixhauser Comorbidity Indices and Extension to Common Data Models. Healthcare informatics research. 2020;26(3):193-200.

[13] Chen T, Guestrin CJPotnASICoKD, Mining D. XGBoost. A Scalable Tree Boosting System. 2016.

[14] Breiman L. Random Forests. 2001;45(1 %J Mach. Learn.):5–32.