

Applying Machine Learning to Predict Esophageal Cancer Recurrence after Esophagectomy

Hanjia Lyu*

*Department of Computer Science
University of Rochester
Rochester, USA
hlyu5@ur.rochester.edu*

Kevin Kapcio*

*Division of Thoracic and Foregut Surgery
University of Rochester Medical Center
Rochester, USA
kevin_kapcio@urmc.rochester.edu*

Kyle Purrmann

*Division of Thoracic and Foregut Surgery
University of Rochester Medical Center
Rochester, USA
kyle_purrmann@urmc.rochester.edu*

Christian Peyre

*Division of Thoracic and Foregut Surgery
University of Rochester Medical Center
Rochester, USA
christian_peyre@urmc.rochester.edu*

Carolyn Jones

*Division of Thoracic and Foregut Surgery
University of Rochester Medical Center
Rochester, USA
carolyn_jones@urmc.rochester.edu*

Michal Lada

*Division of Thoracic and Foregut Surgery
University of Rochester Medical Center
Rochester, USA
michal_lada@urmc.rochester.edu*

Jiebo Luo

*Department of Computer Science
University of Rochester
Rochester, USA
jlue@cs.rochester.edu*

Abstract—Artificial intelligence and machine learning (ML) models have recently been adapted to healthcare applications with promising results. The objective of this proof-of-concept study was to develop an ML model designed to predict esophageal cancer recurrence after esophagectomy. We conducted a retrospective study of 260 consecutive patients who underwent esophagectomy for esophageal cancer from 2009 through 2018. Patient-specific characteristics were collected. Risk prediction models for different prediction windows were constructed via a sequential forward selection process. To enhance the robustness of this framework, five traditional machine learning algorithms including Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and Naive Bayes (NB) were implemented in our analysis. Model performance was assessed by calculating sensitivity, specificity, positive predictive value (PPV), F1 score, area under the receiver operating characteristic curve (AUC), and overall accuracy using five-fold cross-validation. Feature importance analysis was conducted to provide insights into important risk factors associated with esophageal cancer recurrence.

Index Terms—machine learning, esophageal cancer, esophagectomy, recurrence, prediction

I. INTRODUCTION

Esophageal cancer is the sixth leading cause of cancer-related mortality worldwide and represents an immense societal burden [16], [19], [26]. Esophagectomy remains an important component of the treatment regimen for localized esophageal cancers. However, the cancer recurrence rate following such an operation is significant and is dependent on

patient- and disease-specific characteristics. Recurrence rates following esophagectomy range from 38-52% in the literature [1], [15], [17], [18]. After undergoing the biopsychosocial strain of such an operation, cancer recurrence is an immensely complex problem for both patient and healthcare provider to manage. The utility of an objective tool, personalized to both the patient and clinician at hand, to predict such an outcome and guide clinical decision-making cannot be overestimated.

Patient risk factors and statistical methods have been used in the past to develop risk prediction models for esophageal cancer recurrence following esophagectomy by numerous authors. Patient factors that have been shown to be associated with cancer recurrence include sex, tumor size, tumor location, depth of tumor invasion, tumor differentiation, advanced clinical stage, cervical or celiac lymph node metastatic disease, and adjuvant chemoradiotherapy [1], [3], [9], [18], [24], [29], [30]. These characteristics have been used by many authors to develop statistical risk prediction models via cox multivariate regression analysis. Despite the undisputed utility of these models, they are limited in their applicability due to their inability to continually adapt to the clinical situation at hand when new information becomes available. This is essential in the current fast-paced clinical environment.

Where current risk prediction models fall short, artificial intelligence and machine learning exceed expectations. With the recent application of these systems to medicine, we have the possibility of using large sets of synthetic annotated data to train an algorithm to reliably predict a given patient outcome. In addition, these models can continually adapt to new data

* Equal Contribution.

as it becomes available [2], [10], [12], [20]. Several proof-of-concept studies have recently been performed that have validated the ability of these algorithms to predict risk and augment clinical decision-making [4], [5], [7], [8], [11], [13], [14], [21], [22], [25].

The goal of this study is to demonstrate a proof-of-concept framework for using modern machine learning techniques to develop a prognostic model capable of predicting the likelihood of cancer recurrence in those who have undergone esophagectomy for esophageal cancer. This study represents the pioneering effort in the literature to achieve this specific objective. Such an algorithm would be of immense clinical value in improving our prognostic ability for the purposes of clinical decision-making, patient counseling, and medical or surgical intervention.

II. METHOD

This study, IRB # 5407, was approved by the Institutional Review Board of the University of Rochester. The relevant patient data were obtained via a retrospective chart review. All patients who underwent esophagectomy for esophageal cancer between 01/2009 - 12/2018 were included in this study. Patient characteristics collected for the purpose of this study included details on patient age, sex, past medical and surgical history, clinical tumor staging, adjuvant chemoradiotherapy, esophagectomy procedure type, postoperative pathologic tumor staging, tumor histology, adjuvant and neoadjuvant chemoradiotherapy, postoperative complications, cancer recurrence, mortality, and numerous other characteristics.

A. Data Preprocessing

Nominal variables were converted using one-hot encoding. Subsequently, we imputed missing values (*e.g.*, race) using Multiple Imputation by Chained Equations (MICE) [27]. MICE generates imputations for variables one at a time based on this variable's multivariate regression model [28].

B. Modeling

We constructed recurrence prediction models for different prediction windows, including any-day, 180-day, 360-day, 720-day, and 1080-day windows. In particular, the any-day model is designed to forecast the likelihood of cancer recurrence after esophagectomy. All subjects that had recurrence after esophagectomy were considered positive when building the any-day model. To construct the x -day model, where $x \in \{180, 360, 720, 1080\}$, all subjects that had recurrence within x days after esophagectomy were considered positive.

Inspired by Su et al. [23], we selected features via a sequential forward selection procedure [6]. More specifically, during the feature selection of each model, we intended to maximize the F1 score by adding one feature at a time. During the initial round, models were constructed using each variable and evaluated through five-fold cross-validation. The feature with the highest F1 score was selected and added to the feature set for building the final model. In the second round, models were constructed using the first-added variable

TABLE I: Demographics of 260 patients of the study cohort.

Variable	n (%)	n (%)
	Recurrence-positive subjects (n = 121)	Recurrence-negative subjects (n = 139)
Male sex	102 (84.3)	111 (79.9)
Age		
[18 – 25)	1 (0.8)	0 (0.0)
[25 – 35)	1 (0.8)	2 (1.4)
[35 – 45)	2 (1.7)	3 (2.2)
[45 – 55)	19 (15.7)	23 (16.5)
[55 – 65)	51 (42.1)	51 (36.7)
[65 – 75)	38 (31.4)	51 (36.7)
[75 – 85)	8 (6.6)	8 (5.8)
[85 – 95)	1 (0.8)	1 (0.7)
Race		
American Indian / Alaska Native	0 (0.0)	1 (0.7)
Asian	0 (0.0)	0 (0.0)
Native Hawaiian or other Pacific Islander	0 (0.0)	0 (0.0)
Black or African American	2 (1.7)	2 (1.4)
White	115 (95.0)	131 (94.2)
Unknown / Not reported	1 (0.8)	0 (0.0)
Other	3 (2.5)	5 (3.6)

in combination with one of the remaining variables. The same evaluation was applied. The feature with the highest F1 score was included in the feature set. We stopped adding new features when the F1 score was optimized. However, unlike Su et al. [23] who only constructed models using Logistic Regression, we applied this feature selection method to five traditional machine learning algorithms including Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and Naïve Bayes (NB) to enhance the robustness of the framework. After constructing these five models, we counted the frequency of each feature that was selected. The feature selection frequency was then used to estimate feature importance. The higher the frequency of selection of an individual feature, the more important that feature is considered to be in the recurrence prediction process.

C. Measurements

To assess the prediction performance of the models, we calculated sensitivity, specificity, positive predictive value (PPV), F1 score, accuracy, and area under the receiver operating characteristic curve (AUC) for each of the prediction models' time windows: any-day, 180-day, 360-day, 720-day, and 1080-day.

III. RESULTS

The demographic characteristics of our patient population are summarized in Table I. Out of the 260 patients who underwent esophagectomy, 121 of them (46.5%) experienced a recurrence of esophageal cancer.

The efficacy of the Machine Learning (ML) model in predicting cancer recurrence following esophagectomy was evaluated by sensitivity, specificity, positive predictive value (PPV), F1 score, accuracy, and AUC, and is summarized in Table II. For the any-day prediction window, the average F1 score of five models ranged from 0.78 to 0.82, indicating a

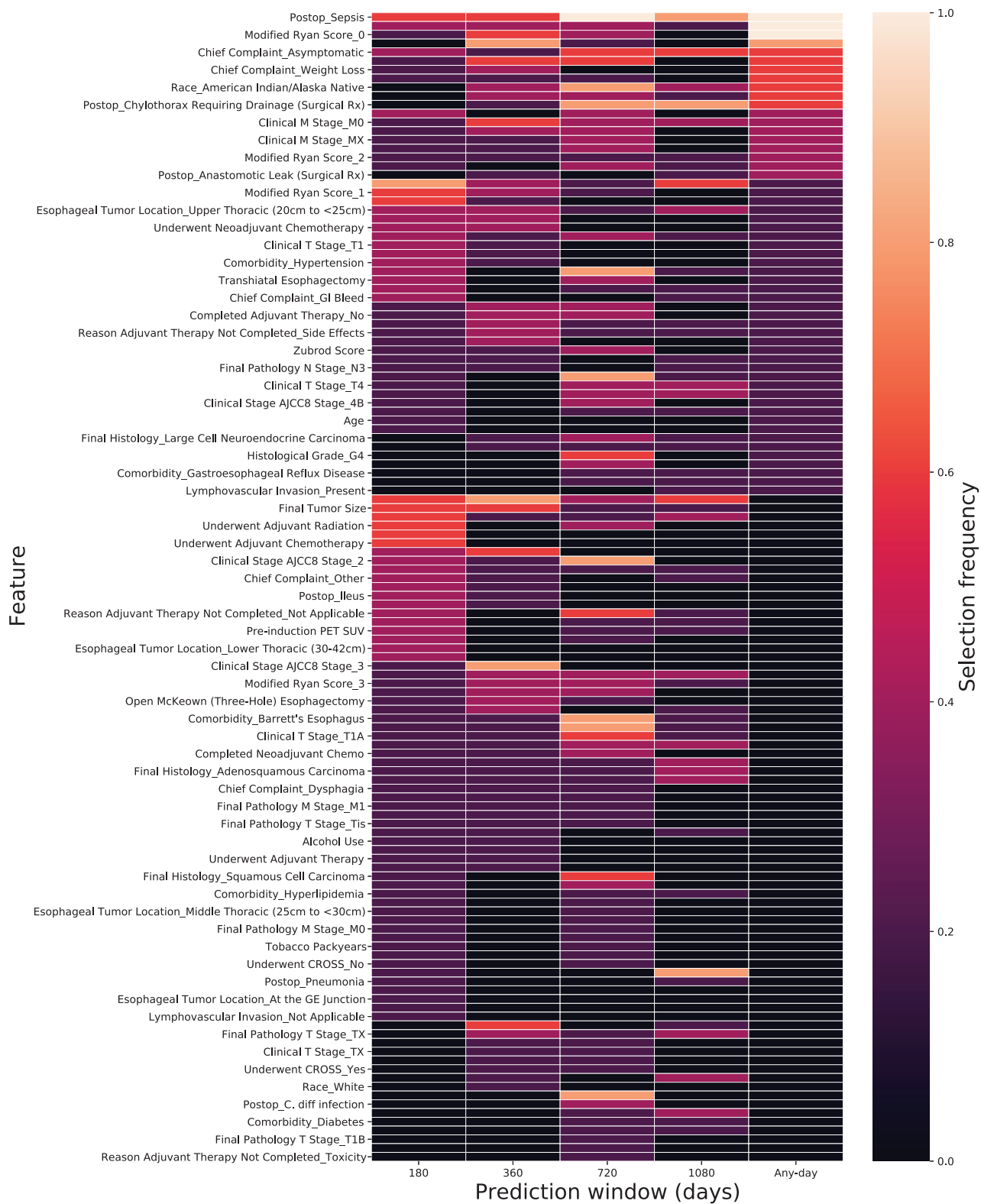


Fig. 1: Feature importance over any-day, 180-day, 360-day, 720-day and 1080-day prediction windows.

TABLE II: Overall performance of the predictive models.

Prediction windows (days)	# positive subjects (%) / # negative subjects (%)	Model	# selected features	Accuracy (SD)	AUC (SD)	Sensitivity (SD)	Specificity (SD)	PPV (SD)	F1 (SD)
Any-day	121 (46.5) / 139 (53.5)	LR	34	0.80 (0.05)	0.81 (0.04)	0.92 (0.05)	0.69 (0.08)	0.72 (0.05)	0.81 (0.04)
		SVM	19	0.77 (0.05)	0.79 (0.06)	0.88 (0.03)	0.69 (0.10)	0.71 (0.06)	0.78 (0.04)
		RF	14	0.82 (0.03)	0.82 (0.05)	0.88 (0.05)	0.77 (0.02)	0.77 (0.02)	0.82 (0.03)
		DT	12	0.79 (0.03)	0.82 (0.02)	0.87 (0.02)	0.73 (0.05)	0.73 (0.03)	0.79 (0.02)
		NB	17	0.80 (0.04)	0.78 (0.06)	0.86 (0.08)	0.75 (0.07)	0.75 (0.05)	0.80 (0.05)
180	37 (14.2) / 223 (85.8)	LR	22	0.92 (0.01)	0.78 (0.05)	0.37 (0.07)	1.00 (0.00)	1.00 (0.00)	0.54 (0.08)
		SVM	48	0.92 (0.02)	0.81 (0.07)	0.51 (0.07)	0.99 (0.02)	0.87 (0.17)	0.65 (0.10)
		RF	9	0.90 (0.03)	0.73 (0.10)	0.49 (0.07)	0.97 (0.03)	0.76 (0.18)	0.58 (0.09)
		DT	9	0.90 (0.02)	0.70 (0.04)	0.51 (0.07)	0.96 (0.02)	0.66 (0.11)	0.57 (0.06)
		NB	19	0.88 (0.02)	0.77 (0.06)	0.60 (0.06)	0.93 (0.02)	0.58 (0.07)	0.58 (0.06)
360	73 (28.1) / 187 (71.9)	LR	19	0.87 (0.02)	0.82 (0.05)	0.67 (0.07)	0.95 (0.02)	0.85 (0.05)	0.74 (0.05)
		SVM	48	0.80 (0.03)	0.79 (0.08)	0.56 (0.12)	0.90 (0.02)	0.70 (0.04)	0.62 (0.08)
		RF	8	0.84 (0.04)	0.83 (0.05)	0.63 (0.11)	0.93 (0.04)	0.79 (0.09)	0.69 (0.09)
		DT	7	0.82 (0.04)	0.79 (0.07)	0.61 (0.13)	0.91 (0.03)	0.73 (0.07)	0.66 (0.10)
		NB	5	0.70 (0.07)	0.71 (0.09)	0.76 (0.13)	0.68 (0.06)	0.49 (0.08)	0.60 (0.10)
720	101 (38.8) / 159 (61.2)	LR	14	0.79 (0.04)	0.80 (0.05)	0.76 (0.09)	0.81 (0.03)	0.72 (0.04)	0.74 (0.05)
		SVM	27	0.80 (0.02)	0.81 (0.02)	0.74 (0.04)	0.83 (0.04)	0.74 (0.05)	0.74 (0.01)
		RF	12	0.83 (0.03)	0.82 (0.04)	0.79 (0.11)	0.86 (0.03)	0.78 (0.02)	0.78 (0.05)
		DT	35	0.77 (0.03)	0.79 (0.02)	0.71 (0.08)	0.81 (0.03)	0.70 (0.03)	0.70 (0.05)
		NB	13	0.78 (0.04)	0.81 (0.05)	0.89 (0.08)	0.71 (0.06)	0.66 (0.05)	0.76 (0.05)
1080	110 (42.3) / 150 (57.7)	LR	12	0.78 (0.01)	0.79 (0.04)	0.85 (0.08)	0.74 (0.05)	0.71 (0.02)	0.77 (0.03)
		SVM	8	0.75 (0.03)	0.79 (0.03)	0.84 (0.08)	0.68 (0.02)	0.66 (0.02)	0.73 (0.04)
		RF	16	0.79 (0.03)	0.79 (0.04)	0.79 (0.08)	0.79 (0.05)	0.73 (0.04)	0.76 (0.04)
		DT	15	0.80 (0.07)	0.80 (0.06)	0.80 (0.09)	0.79 (0.08)	0.74 (0.08)	0.77 (0.07)
		NB	8	0.75 (0.05)	0.77 (0.05)	0.85 (0.07)	0.68 (0.05)	0.66 (0.05)	0.75 (0.06)

good prediction performance. In particular, Random Forest (RF) consistently scored the highest of the four metrics. For AUC and sensitivity, Random Forest (RF) scored the second highest. Decision Tree (DT) performed the best in terms of AUC (0.82). Logistic Regression (LR) had the highest sensitivity (0.92). However, the average F1 score for the five models in the 180-day prediction window decreased compared to the any-day model. The 180-day models suffered from a class imbalance issue with a ratio of 6:1 between recurrence-negative and recurrence-positive subjects. For this prediction window, we did not observe a consistently better model. Furthermore, as the prediction window increased, the issue of imbalance was alleviated. As a result, the prediction performance of the models for 360-day, 720-day, and 1080-day rose. For the 360-day prediction window, Logistic Regression (LR) scored the highest in four metrics (*i.e.*, accuracy, specificity, PPV, and F1) and second highest in AUC and sensitivity. For the 720-day prediction window, Random Forest (RF) was almost consistently superior to the other models, except that it scored second highest in sensitivity. For the 1080-day prediction window, Decision Tree was the best model, with the exception of a relatively low sensitivity score.

Feature importance of each patient characteristic in predicting esophageal cancer recurrence following esophagectomy was estimated, and is summarized in Figure 1. In each prediction window, a lighter color represents a higher degree of importance of a particular risk factor. Feature importance was measured by an average of the frequency that an individual feature was included by the sequential forward selection procedure by each of the five separate machine learning models. Except for binary features, all categorical features were displayed with their associated results in quotations. For instance, “comorbidity_Atrial Fibrillation” measures whether

the patient had atrial fibrillation preoperatively.

For the any-day prediction, final pathology T stage (T0), Stage AJCC8 (stage 1), comorbidity of atrial fibrillation, chief complaints (asymptomatic, weight loss, and other), final pathology T stage (T1A), rare postoperative event (Chylothorax requiring drainage), and clinical N stage (N2), were most frequently selected. It is noteworthy that the importance of esophageal tumor location (cervical esophagus) and the presence of sepsis as a postoperative complication was *consistent* across multiple prediction windows.

IV. DISCUSSION

Esophageal cancers are an aggressive type of malignancy that has a high rate of recurrence even after esophagectomy and adjuvant/neoadjuvant chemo and/or radiation therapy. Currently, there are no effective and accurate means for predicting esophageal cancer recurrence. Esophageal cancer recurrences that are detected based on clinical presentation are associated with considerably worse survival in comparison to those detected by surveillance screening [15]. An accurate method to predict the recurrence of esophageal cancers would allow for more rigorous surveillance during the anticipated time interval of recurrence, facilitating earlier detection and potentially improving survival rates and quality of life.

There are several limitations to our study. The esophageal cancer recurrence details as well as the patient-specific characteristics utilized to construct our machine learning models were obtained from chart review and from a retrospective database and thus are subject to potential bias. Additionally, the subject population used in this study was from a single institution and therefore our results may not be generalizable to other geographical locations or patient populations. Furthermore, the limitations are compounded by the relatively small

sample size as well as the necessity to impute certain missing variables.

V. CONCLUSIONS

This proof-of-concept study demonstrates the utility of using Machine Learning (ML) models to predict esophageal cancer recurrence in patients who have undergone esophagectomy. Overall, our model was able to achieve 82% accuracy in the any-day prediction window in predicting esophageal cancer recurrence, with an accuracy of predicting recurrence reaching as high as 92% in the 180-day prediction window.

Feature importance was estimated and it was found that final pathology T0 and proximal esophageal tumor location were among the most important characteristics for predicting esophageal cancer recurrence in our patient population.

The outcome of our study is promising for predicting esophageal cancer recurrence, however, additional data and multi-center collaboration are required to expand the scope and applicability of the model.

ACKNOWLEDGMENTS

We are grateful to Dr. Wei Zhu for his constructive suggestions. This work was partially supported by the Goergen Institute for Data Science at the University of Rochester.

REFERENCES

- [1] Emmanuele Abate, Steven R DeMeester, Joerg Zehetner, Arzu Oezcelik, Shahin Ayazi, Jesse Costales, Farzaneh Banki, John C Lipham, Jeffrey A Hagen, and Tom R DeMeester. Recurrence after esophagectomy for adenocarcinoma: defining optimal follow-up intervals and testing. *Journal of the American College of Surgeons*, 210(4):428–435, 2010.
- [2] Hamdan O Alanazi, Abdul Hanan Abdullah, and Kashif Naseer Qureshi. A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care. *Journal of medical systems*, 41:1–10, 2017.
- [3] Arianna Barbetta, Smita Sihag, Tamar Nobel, Meier Hsu, Kay See Tan, Manjit Bains, David R Jones, and Daniela Molena. Patterns and risk of recurrence in patients with esophageal cancer with a pathologic complete response after chemoradiotherapy followed by surgery. *The Journal of thoracic and cardiovascular surgery*, 157(3):1249–1259, 2019.
- [4] Yang Cao, Xin Fang, Johan Ottosson, Erik Näslund, and Erik Stenberg. A comparative study of machine learning algorithms in predicting severe complications after bariatric surgery. *Journal of Clinical Medicine*, 8(5):668, 2019.
- [5] Davide Chicco and Giuseppe Jurman. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC medical informatics and decision making*, 20(1):1–16, 2020.
- [6] Francesc J Ferri, Pavel Pudil, Mohamad Hatef, and Josef Kittler. Comparative study of techniques for large-scale feature selection. In *Machine intelligence and pattern recognition*, volume 16, pages 403–413. Elsevier, 1994.
- [7] Benjamin A Goldstein, Ann Marie Navar, and Rickey E Carter. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *European heart journal*, 38(23):1805–1814, 2017.
- [8] Daniel A Hashimoto, Elan Witkowski, Lei Gao, Ozanan Meireles, and Guy Rosman. Artificial intelligence in anesthesiology: current techniques, clinical applications, and limitations. *Anesthesiology*, 132(2):379–394, 2020.
- [9] Dong Young Jeong, Kyung Soo Lee, Joon Young Choi, Myung Jin Chung, Yang Won Min, Hong Kwan Kim, Jae Ill Zo, Young Mog Shim, and Jong-Mu Sun. Surgically resected esophageal squamous cell carcinoma: patient survival and clinicopathological prognostic factors. *Scientific Reports*, 10(1):1–12, 2020.
- [10] Michael I Jordan and Tom M Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015.
- [11] Arman Kilic. Artificial intelligence and machine learning in cardiovascular health care. *The Annals of thoracic surgery*, 109(5):1323–1329, 2020.
- [12] The Lancet. Artificial intelligence in health care: Within touching distance, 2017.
- [13] Qi Li, Qiu-Ling Fan, Qiu-Xia Han, Wen-Jia Geng, Huan-Huan Zhao, Xiao-Nan Ding, Jing-Yao Yan, and Han-Yu Zhu. Machine learning in nephrology: scratching the surface. *Chinese Medical Journal*, 133(06):687–698, 2020.
- [14] Ke Lin, Yonghua Hu, and Guilan Kong. Predicting in-hospital mortality of patients with acute kidney injury in the icu using random forest model. *International journal of medical informatics*, 125:55–61, 2019.
- [15] Feiran Lou, Camelia S Sima, Prasad S Adusumilli, Manjit S Bains, Inderpal S Sarkaria, Valerie W Rusch, and Nabil P Rizk. Esophageal cancer recurrence patterns and implications for surveillance. *Journal of thoracic oncology*, 8(12):1558–1562, 2013.
- [16] Wei-Min Mao, Wei-Hui Zheng, and Zhi-Qiang Ling. Epidemiologic risk factors for esophageal cancer development. *Asian Pac J Cancer Prev*, 12(10):2461–2466, 2011.
- [17] Christophe Mariette, Jean-Michel Balon, Guillaume Piessen, Sylvain Fabre, Isabelle Van Seuningen, and Jean-Pierre Triboulet. Pattern of recurrence following complete resection of esophageal carcinoma and factors predictive of recurrent disease. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 97(7):1616–1623, 2003.
- [18] Satoru Nakagawa, Tatsuo Kanda, Shin-ichi Kosugi, Manabu Ohashi, Tsutomu Suzuki, and Katsuyoshi Hatakeyama. Recurrence pattern of squamous cell carcinoma of the thoracic esophagus after extended radical esophagectomy with three-field lymphadenectomy. *Journal of the American College of Surgeons*, 198(2):205–211, 2004.
- [19] Arjun Pennathur, Michael K Gibson, Blair A Jobe, and James D Luketich. Esophageal carcinoma. *The Lancet*, 381(9864):400–412, 2013.
- [20] Alvin Rajkomar, Jeffrey Dean, and Isaac Kohane. Machine learning in medicine. *New England Journal of Medicine*, 380(14):1347–1358, 2019.
- [21] Cheng-Shyuan Rau, Shao-Chun Wu, Jung-Fang Chuang, Chun-Ying Huang, Hang-Tsung Liu, Peng-Chen Chien, and Ching-Hua Hsieh. Machine learning models of survival prediction in trauma patients. *Journal of clinical medicine*, 8(6):799, 2019.
- [22] Joeky T Senders, Patrick C Staples, Aditya V Karhade, Mark M Zaki, William B Gormley, Marika LD Broekman, Timothy R Smith, and Omar Arnaout. Machine learning and neurosurgical outcome prediction: a systematic review. *World neurosurgery*, 109:476–486, 2018.
- [23] Chang Su, Robert Aseltine, Riddhi Doshi, Kun Chen, Steven C Rogers, and Fei Wang. Machine learning for suicide risk prediction in children and adolescents with electronic health records. *Translational psychiatry*, 10(1):413, 2020.
- [24] Xiao-Dong Su, Dong-Kun Zhang, Xu Zhang, Peng Lin, Hao Long, and Tie-Hua Rong. Prognostic factors in patients with recurrence after complete resection of esophageal squamous cell carcinoma. *Journal of thoracic disease*, 6(7):949, 2014.
- [25] Charat Thongprayoon, Panupong Hansrivijit, Tarun Bathini, Saraschandra Vallabhajosyula, Poemlarp Mekraksakit, Wisit Kaewput, and Wisit Cheungpasitporn. Predicting acute kidney injury after cardiac surgery by machine learning approaches, 2020.
- [26] Sarah B Umar and David E Fleischer. Esophageal cancer: epidemiology, pathogenesis and prevention. *Nature clinical practice Gastroenterology & hepatology*, 5(9):517–526, 2008.
- [27] Stef Van Buuren and Karin Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in r. *Journal of statistical software*, 45:1–67, 2011.
- [28] Ian R White, Patrick Royston, and Angela M Wood. Multiple imputation using chained equations: issues and guidance for practice. *Statistics in medicine*, 30(4):377–399, 2011.
- [29] Tianpeng Xie, Run Xiang, Xiaojun Yang, Yue Cui, and Qiang Li. Analysis of characteristics and factors affecting the recurrence of esophageal cancer within the first year after surgery. *Chinese Journal of Oncology*, 38(2):146–149, 2016.
- [30] Zi-Jiang Zhu, Yang Hu, Yong-Fan Zhao, Xue-Zhong Chen, Long-Qi Chen, and Ying-Tai Chen. Early recurrence and death after esophagectomy in patients with esophageal squamous cell carcinoma. *The Annals of thoracic surgery*, 91(5):1502–1508, 2011.