**Machine Learning Can Predict Deaths in Patients with Diverticulitis During their Hospital Stay**

Fahad Shabbir Ahmed[1,8], Raza-Ul-Mustafa [2,3,8], Liaqat Ali[4], Imad-ud-Deen[1], Tahir Hameed[5], Asad Ikram[5] , Syed Ahmad Chan Bukhari[7,*]

[1]Department of Pathology, Yale School of Medicine, Yale University, New Haven CT, USA

[2]Department of Computer Engineering and Industrial Automation (DCA) School of Electrical and Computer Engineering (FEEC), Unicamp, Brazil

[3]Department of Computer Science, COMSATS University Islamabad, Sahiwal Campus, Pakistan.

[4]School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China

[5]Girard School of Business Merrimack College North Andover, MA, USA

[6]Department of Neurology, University of New Mexico, Albuquerque NM USA

[7]Division of Computer Science, Mathematics, and Science (Healthcare Informatics), St. John's University, NYC

[8]Co-first

**Corresponding should be addressed to Syed Ahmad Chan Bukhari (bukharis@stjohns.edu)*

## ABSTRACT

**Introduction:** Diverticulitis is the inflammation and/or infection of small pouches known as diverticula that develop along the walls of the intestines. Patients with diverticulitis are at risk of mortality as high as 17% with abscess formation and 45% with secondary perforation, especially patients that get admitted to the inpatient services are at risk of complications including mortality. We developed a deep neural networks (DNN) based machine learning framework that could predict premature death in patients that are admitted with diverticulitis using electronic health records (EHR) to calculate the statistically significant risk factors first and then to apply deep neural network.

**Methods:** Our proposed framework (Deep FLAIM) is a two-phase hybrid works framework. In the first phase, we used National In-patient Sample 2014 dataset to extract patients with diverticulitis patients with and without hemorrhage with the ICD-9 codes 562.11 and 562.13 respectively and analyzed these patients for different risk factors for statistical significance with univariate and multivariate analyses to generate hazard ratios, to rank the diverticulitis associated risk factors. In the second phase, we applied deep neural network model to predict death. Additionally, we have compared the performance of our proposed system by using the popular machine learning models such as DNN and Logistic Regression (LR).

**Results:** A total of 128,258 patients were used, we tested 64 different variables for using univariate and multivariate (age, gender and ethnicity) cox-regression for significance only 16 factors were statistically significant for both univariate and multivariate analysis. The mortality prediction for our DNN out-performed the conventional machine learning (logistic regression) in terms of AUC (0.977 vs 0.904), training accuracy (0.931 vs 0.900), testing accuracy (0.930 vs 0.910) , sensitivity (90% vs 88%) and specificity (95% vs 93%).

**Conclusion:** Our Deep FLAIM Framework can predict mortality in patients admitted to the hospital with diverticulitis with high accuracy. The proposed framework can be expanded to predict premature death for other disease.

**Keywords:** Disease Prediction, Machine Learning, Diverticulitis

## Introduction

Diverticulitis of the colon is the inflammation of the colon segment that has developed outpouching that most commonly occurs in the sigmoid colon. this occurs secondary to the mucosal and submucosal layers of that segment have herniated in the muscular layer of the bowel wall. This disrupts the normal bowel function of the affected bowel segment and cases stasis of fecal matter in that bowel segment leading to accumulation of fecal matter and subsequent bacterial growth and inflammation. in the United States diverticulitis is the eight most common gastrointestinal diagnosis in the outpatient setting with an estimated 2.7 million annual health care visits. This carries an approximate burden of 2.2 billion dollars (1)

Mortality predication using machine learning has now gained momentum in the past few years particularly using deep neural network (2, 3). The ability of DNN to recognize complex patterns helps it to learn about biological risk factors and predict outcomes based on that learning (4, 5). Simply, machine learning is a branch of computer science that deals with the estimation of an outcome based on past experiences using a computer-generated algorithm (6)**.** We set out to create a prediction system that can predict mortality with high accuracy in patients that were admitted to hospital patients with the primary diagnosis of diverticulitis We used FLAIM (machine learning framework) previously to predict mortality in ICU patients (7) and a comparative Logistic regression (LR) model (8). We hypothesized that the DNN would outperform the LR model.

## METHODS

### Dataset and patient selection

We obtained a publicly accessible data-set National Inpatient Sample (NIS) for the year 2014. This database is part of Health-care Cost and Utilization Project (HCUP). NIS is the largest database available with all-payer inpatient health care in the United States. Contains about 7 million admissions per year approximately. All this data is de-identified. All the database tables were aggregated using “key-nis” variable.  Case selection was done using the ICD-9 codes for diverticulitis without hemorrhage (562.11) and diverticulitis with hemorrhage (562.13).

### Data analysis

Patient demographics were calculated and are presented in table 1. A Uni- variate and Multivariate Cox regression analysis was done using SPSS 24.0 was done to evaluate the role of commodities and other parameters to assess the mortality risk associated with this risk factors. The data-set was sorted and binarized to be used for machine learning.

### Machine Learning Methods

Logistic Regression (LR) is the appropriate regression analysis to conduct when the dependent binary variable. To all the regression analyses, logistic regression is a predictive analysis. Moreover, it is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio level independent variables. Later, we used NN for the reason it does not impose any restrictions on the input variables (like

how they should be distributed). Additionally, many studies have shown that NN can better model i.e.  data with high volatility and non-constant variance, given its ability to learn hidden relationships in the data without imposing any fixed relationships in the data. We used Keras a python library that has been used for efficient numerical computations It is high level Application Programming Interface (API) for deep learning models to build and train. Moreover, it wraps Theano and TensorFlow. For fast numerical computational purposes Theano is an efficient library. The parameters we have used for the training and testing is listed in Table 1. On hidden layers we used rectifier and for the output layer we used sigmoid activation function. Finally, the network uses the efficient Adam gradient descent optimization algorithm with a logarithmic loss function, which is called binary-cross entropy in Keras. The performance of the model on the train and test sets recorded during training graphed using a line plot, one for each of the loss and the classification accuracy as shown in Figure 1.

To further investigate the performance of the proposed method, another evaluation metric i.e., Receiver Operating Characteristic (ROC) chart is used. The ROC chart is a plot of true positive rate (TPR) versus false positive rate (FPR) for various thresholds. An ROC chart with more area under the curve (AUC) is considered best. An ideal ROC chart has AUC=100, such chart means that the model is capable of performing with 100 % sensitivity and 100 % specificity.

## RESULTS

From a total of 3.8 million patients 22504 adult patients (age more than 18 years) were extracted with the primary diagnosis of diverticulitis that were with the median age of 60 years, with most patients being women (58.6%) with mostly being Caucasian (75.7%). the median stay in the hospital was 4 / 4.4 days. These patients had a median stay of 14.5 16.2 days of hospital stay and 4.2 10.7 days of intensive care stay. Emergency admission comprised of most admission making 76.7% of all patients that were admitted to the ICU while most admission came through referrals around 490 patients (48%). Mortality among these patients was 16.8% (173), while 441 (43.2%) were transferred out to different health facilities and 407 (40%) patients were discharged home. All of this can be seen in Table 1.

We conducted a univariate and multivariate analysis of a number of different risk factors and the risk factors that were statistically significant for mortality in the multivariate analysis were: higher number chronic conditions, higher number of working diagnosis on admission, higher number of procedures in the current admission. APRDRG high risk and major risk, non-elective admission, Chronic Heart Failure, Chronic Lung Disease, Coagulopathy, Fluid and electrolyte disorder, Renal Failure, Valvular disease and Weight loss. Other statistically significant risk factors included median house hold income (Upper 50th vs Lower 50th). More detailed statistics can be found on in Tables 2. And 3.

The FLAIM-DNN-4RS showed superior performance compared to the logistic regression model in terms of Training set accuracy (93.1% vs 90.0%), Test set accuracy (93.0% vs 0.91%), Sensitivity specificity and area under the receiver-operator curve which are represented by Table 4 and Figure 1. Figure 1 also shows the loss and accuracy during training for FLAIM-DNN-4RS. Table 5 shows the details of the FLAIM-DNN-4RS model.

## DISCUSSION

The incidence of diverticular disease has dramatically increased over the years. Current data suggest that up to 50% of individuals older than 60 years of age suffer from diverticular disease and its complications resulting in a sharp rise of hospitalizations (9).

The current study confirmed high mortality risk in patients of more than 60 years of age ($p<0.001$). It was also seen that the number of chronic conditions also affected the overall mortality of these patients($p<0.05$) admitted with acute diverticulitis significantly, with chronic liver disease, fluid and electrolyte imbalance, Renal failure, weight loss and valvular heart disease ($p<0.001$). We also looked at the clinical risk factors such as chronic NSAID use, Bisphosphonate use, tobacco use, steroid use and antibiotic use which did not affect mortality.

We compared the median household income of the patients and found an increased risk of mortality in the lower 50% of the population (3rd and 4th quartiles ($p<0.005$)). This might be due to better healthcare access, insurance; and possibly education, dietary habits and a healthier lifestyle in patients having household income with in the 3rd and 4th quartile range. This was in contrast to a study conducted in Sweden which showed patients with middle to high household income had a higher risk of diverticular diseases when compared with low household income (10).

Through the application of machine learning methods to the various data present at the time of the new diagnosis of diverticulitis, disease mortality can be moderately predicted (11). The Machine Learning algorithms used in this analysis can provide mortality prediction based upon the values of blood biomarkers at the time of admission and at different time intervals. Disease prediction though this model can help in the prognosis of the disease leading to better outcomes. It can also be used for risk stratification or help providers in clinical decision-making. For diverticulitis, this type of model can also be used to aid in decisions made by physicians and patients. Though future reiteration of this model must be validated using independent samples for a different dataset. (12).

## CONCLUSION

FLAIM-DNN-4RS is a robust model that can predict mortality in diverticulitis patient with very high accuracy we need to further validate this model.

## REFERENCES


1. Reddy VB, Longo WE. The burden of diverticular disease on patients and healthcare systems. Gastroenterol Hepatol (N Y). 2013;9(1):21-7.
2. Cao Y, Fang X, Ottosson J, Naslund E, Stenberg E. A Comparative Study of Machine Learning Algorithms in Predicting Severe Complications after Bariatric Surgery. J Clin Med. 2019;8(5).
3. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface. 2018;15(141).
4. Rabhi S, Jakubowicz J, Metzger MH. Deep Learning versus Conventional Machine Learning for Detection of Healthcare-Associated Infections in French Clinical Narratives. Methods Inf Med. 2019;58(1):31-41.
5. Metsker O, Sikorsky S, Yakovlev A, Kovalchuk S. Dynamic mortality prediction using machine learning techniques for acute cardiovascular cases. Procedia Computer Science. 2018;136:351-8.
6. Deo RC. Machine Learning in Medicine. Circulation. 2015;132(20):1920-30.
7. Ahmed FS, Ali L, Mustafa RU, Khattak HA, Bukhari SAC. A Hybrid Machine Learning Framework to Predict Early Risk of Mortality in Paralytic Ileus Patients using Electronic Health Records. medRxiv. 2019:19006254.
8. Koletsi D, Pandis N. Conditional logistic regression. American journal of orthodontics and dentofacial orthopedics : official publication of the American Association of Orthodontists, its constituent societies, and the American Board of Orthodontics. 2017;151(6):1191-2.
9. Warner E, Crighton EJ, Moineddin R, Mamdani M, Upshur R. Fourteen-year study of hospital admissions for diverticular disease in Ontario. Can J Gastroenterol. 2007;21(2):97-9.
10. Nikberg M, Ji J, Leppert J, Sundquist K, Chabok A. Socioeconomic characteristics and comorbidities of diverticular disease in Sweden 1997-2012. Int J Colorectal Dis. 2017;32(11):1591-6.
11. Thornblade LW, Flum DR, Flaxman AD. Predicting Future Elective Colon Resection for Diverticulitis Using Patterns of Health Care Utilization. EGEMS (Washington, DC). 2018;6(1):1.
12. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med. 2016;375(13):1216-9.

Table 1. Demographics

| Demographics | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | Univariate | | Multivariate | |
| Variables | Sub-categories | Median ± SD | Range | *p*-value | HR (95%CI) | *p*-value | HR (95%CI) |
| Age | | 60 ± 15.2 years | 18-90 years | *<0.001 | 1.07-1.10 | n/a | n/a |
| Length of hospital stay | | 4 ± 4.4 days | 0-121 days | - | - | - | - |
| Number chronic conditions | | 4 ± 2.9 | 1-19 | *<0.001 | 1.16-1.27 | *<0.001 | 1.07-1.21 |
| Number of working diagnosis on current admission | | 8 ± 5.4 | 1-29 | *<0.001 | 1.10-1.17 | *<0.001 | 1.08-1.15 |
| Number of procedures on current admission | | 0 ± 1.9 | 0-15 | *0.036 | 1.01-1.21 | *0.005 | 1.02-1.14 |
| | | n (total = 25504) | Percentage | | | | |
| Primary diagnosis Diverticulitis | | | | *0.005 | 0.23-0.717 | 0.765 | 0.41-1.45 |
| | without hemorrhage | 24348 | 95.5 | - | - | - | - |
| | with hemorrhage | 1156 | 4.5 | - | - | - | - |
| Gender | | | | 0.058 | 0.98-2.19 | n/a | n/a |
| | Male | 10533 | 41.3 | - | - | - | - |
| | Female | 14958 | 58.6 | - | - | - | - |
| Race | | | | 0.106 | 0.62-1.07 | n/a | n/a |
| | Caucasian | 19311 | 75.7 | - | - | - | - |
| | African American | 2131 | 8.4 | - | - | - | - |
| | Asian / Pacific | 129 | 0.5 | - | - | - | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Islanders | | | | | | |
| | Native American | 75 | 0.3 | - | - | - | - |
| | Others | 591 | 2.3 | - | - | - | - |
| | Not known | 1435 | 5.6 | - | - | - | - |
| Death during hospitalization | | 109 | 0.4 | n/a | n/a | n/a | n/a |
| APRDRG Mortality risk | | | | *<0.001 | 4.48-8.10 | *<0.001 | 4.06-7.95 |
| | Low risk | 17522 | 68.7 | *<0.001 | 0.01-0.50 | *<0.001 | 0.02-0.99 |
| | Moderate risk | 5287 | 20.7 | *<0.001 | 0.03-0.22 | *<0.001 | 0.04-0.28 |
| | Major risk | 2179 | 8.5 | 0.587 | 0.58-1.36 | - | - |
| | Extreme risk | 516 | 2.0 | *<0.001 | 6.68-16.96 | *<0.001 | 5.12-13.58 |
| | Major/Extreme risk | 2695 | 11.5 | *<0.001 | 8.91-56.23 | *<0.001 | 5.34-34.19 |
| Admission type | | | | *0.027 | 0.29-0.98 | *<0.001 | 0.45-1.54 |
| | Elective | 20620 | 80.9 | - | - | - | - |
| | Non-Elective | 4804 | 18.8 | - | - | - | - |
| Hospital Service line | | | | *0.007 | 1.17-2.72 | *<0.001 | 0.68-1.68 |
| | Surgical | 6606 | 25.9 | - | - | - | - |
| | Medical | 18898 | 74.1 | - | - | - | - |
| Major operating procedures | | 6604 | 25.9 | *0.009 | 0.37-0.86 | 0.786 | 0.57-1.48 |

Table 2. Comorbidities

| | Comorbidities | | | | | |
|---|---|---|---|---|---|---|
| | | | Univariate | | Multivariate | |
| | n (total = 25504) | % | *p*-value | HR (95%CI) | *p*-value | HR (95%CI) |
| AIDS | 17 | 0.1 | 0.605 | 0.00-7.8e8 | - | - |
| Alcohol Abuse | 583 | 2.3 | 0.276 | 0.12-2.03 | - | - |
| Deficiency Anemia | 2981 | 11.7 | *0.035 | 1.10-2.43 | 0.316 | 0.81-1.92 |
| RA/CVD | 864 | 3.4 | 0.678 | 0.52-2.73 | - | - |
| Chronic blood loss anemia | 169 | 0.7 | 0.571 | 0.08-4.26 | - | - |
| CHF | 1276 | 5.0 | *<0.001 | 2.28-6.28 | *<0.001 | 1.50-3.60 |
| CLD | 4248 | 16.7 | *0.007 | 1.18-2.61 | *<0.001 | 1.01-2.31 |
| Coagulopathy | 545 | 2.1 | *0.007 | 1.35-4.30 | *0.009 | 1.22-4.03 |
| Depression | 2914 | 11.4 | 0.434 | 0.73-2.12 | - | - |
| Uncomplicated DM | 4063 | 15.9 | 0.885 | 0.63-1.72 | - | - |
| Complicated DM | 411 | 1.6 | 0.058 | *1.09-5.66 | 0.538 | 0.50-3.75 |
| Drug Abuse | 397 | 1.6 | 0.630 | 0.36-5.84 | - | - |
| HTN | 13781 | 54.0 | 0.415 | 0.80-1.73 | - | - |
| Hypothyroid | 3386 | 13.3 | 0.895 | 0.61-1.77 | - | - |
| Liver Disease | 807 | 3.2 | 0.853 | 0.41-3.00 | - | - |
| Lymphoma | 149 | 0.6 | 0.881 | 0.16-8.38 | - | - |
| Fluid and electrolyte disorder | 5997 | 23.5 | *<0.001 | 1.62-3.72 | *<0.001 | 1.19-2.88 |
| Metastatic cancer | 200 | 0.8 | 0.250 | 0.72-4.45 | - | - |
| Neurological disorders | 979 | 3.8 | 0.114 | 0.92-3.39 | - | - |
| Obesity | 3892 | 15.3 | 0.091 | 0.36-1.11 | - | - |
| Paralysis | 76 | 0.3 | 0.913 | 0.16-8.04 | - | - |
| PVD | 920 | 3.6 | *0.015 | 1.22-3.81 | 0.292 | 0.76-2.55 |
| Psychosis | 628 | 2.5 | 0.951 | 0.31-3.04 | - | - |
| Pulmonary circulation disorder | 273 | 1.1 | 0.397 | 0.64-3.36 | - | - |
| Renal failure | 1802 | 7.1 | *<0.001 | 2.43-5.49 | *<0.001 | 1.37-3.31 |
| Solid tumors no metastasis | 328 | 1.3 | 0.990 | 0.32-3.21 | - | - |
| PUD without bleeding | 11 | 0.0 | 0.691 | 0.00-9.1e11 | - | - |
| Valvular disease | 838 | 3.3 | *<0.001 | 2.55-6.55 | *0.012 | 1.16-3.29 |
| Weight loss | 916 | 3.6 | *0.021 | 1.10-2.75 | *0.036 | 1.03-2.60 |

Table 3.

| | | Financial | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Median ± SD | Range | Grand total | Univariate | | Multivariate | |
| | | | | | *p*-value | 95%CI | *p*-value | 95%CI |
| Total charges in US Dollars | | 22,234 ± 56,302 (USD) | 207-4,435,000 (USD) | 911,369,152 (911 million USD) | - | - | - | - |
| | | n (total = 25504) | % | Distribution of total cost | - | - | - | - |
| Payers | Medicare | 10746 | 42.1 | 383,686,413 (384 million USD) | - | - | - | - |
| | Medicaid | 2271 | 8.9 | 81,111,854 (81 million USD) | - | - | - | - |
| | Private insurance | 10780 | 42.3 | 385,509,151 (386 million USD) | - | - | - | - |
| | Self-pay | 987 | 3.9 | 35,543,397 (35.5 million USD) | - | - | - | - |
| | No charge | 155 | 0.6 | 5,468,215 (5.5 million USD) | - | - | - | - |
| | Others | 532 | 2.1 | 19,138,752 (19 million USD) | - | - | - | - |
| Median house hold income | | | | | *0.006 | 1.07-1.52 | *0.041 | 1.01-1.46 |
| | 1st Quartile | 5293 | 20.8 | - | *0.006 | 0.28-0.85 | 0.058 | 0.32-1.02 |
| | 2nd Quartile | 6957 | 27.3 | - | 0.516 | 0.56-1.35 | - | - |
| | 3rd Quartile | 6343 | 24.9 | - | 0.069 | 0.98-2.20 | - | - |
| | 4th Quartile | 6523 | 25.6 | - | 0.196 | 0.88-1.99 | - | - |
| | 1st & 2nd vs 3rd & 4$^{th}$ Quartiles | - | - | - | *0.005 | 1.17-2.55 | *0.045 | 1.01-2.31 |

(Upper 50% vs
Lower 50%)

Table 4. Machine Learning Results for FLAIM-DNN-4RS and the Logistic regression model

| Models | Training set accuracy | Test set accuracy | Sensitivity | Specificity | Are under the Receiver operator curve | Ranks |
|---|---|---|---|---|---|---|
| FLAIM-DNN-4RS | 0.931 | 0.930 | 0.90 | 0.95 | 0.977 | 1 |
| Logistic Regression | 0.90 | 0.91 | 0.88 | 0.93 | 0.904 | 2 |

Table 5: Parameters used during training and testing of proposed method

| **Parameters** | **Values** |
|---|---|
| Input layers | Number of features (16) |
| Hidden layers | (2) |
| Activation function on hidden layers | ReLu -Rectified linear units |
| Activation function on output layer | Sigmoid |
| Training epochs | (500,1000, 5000) |
| Batch size | (128,256) |
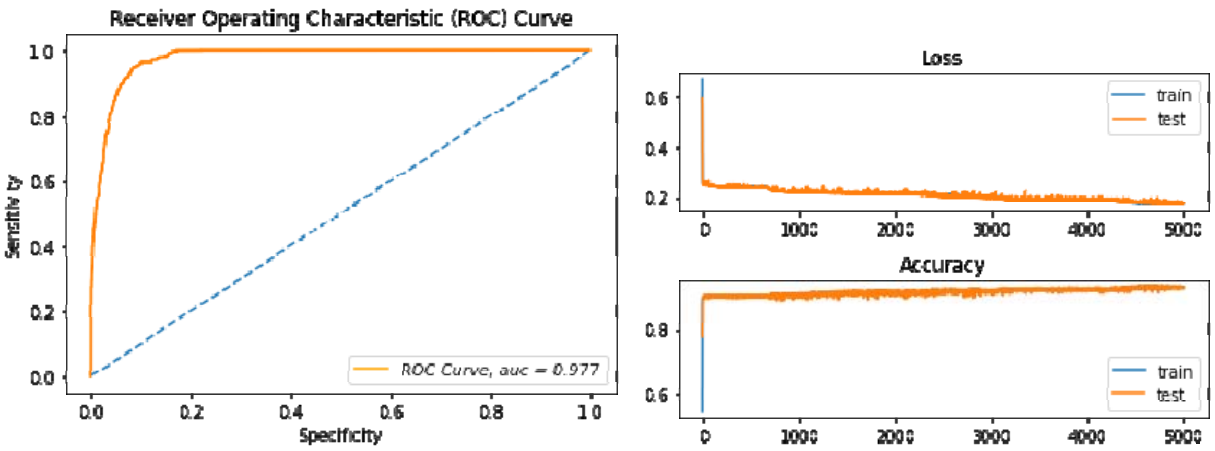
Figure 1. Receiver -Operator Curve for FLAIM-DNN-4RS and the loss and accuracy curve during training