DR IVAN DUDURYCH (Orcid ID : 0000-0002-7747-156X)

Article type : Original Article

Predicting Outcomes of Pelvic Exenteration Using Machine Learning

PelvEx Collaborative

Submitting Author:

Ivan Dudurych M.B., B.A.O., B.Ch., M.Sc.

Department of Surgery

St Vincent's University Hospital

Dublin, Ireland

Email: duduryci@tcd.ie

Keywords:

Colorectal Surgery,

Artificial Intelligence (AI),

Machine Learning (ML),

Pelvic Exenteration,

Artificial Neural Network (ANN)

All authors declare no conflict of interest

Abstract

Aim: We aim to compare machine learning (ML) with neural network performance in predicting R0 resection (R0), length of stay >14 days (LOS), major complication rates at 30 days post-operatively (COMP) and survival greater than one year (SURV) for patients having pelvic exenteration for locally advanced and recurrent rectal cancer.

Method: A deep learning computer was built, and programming environment established. The *PelvEx Collaborative* database was used which contains anonymized data on patients who underwent pelvic exenteration for locally advanced or locally recurrent colorectal cancer between 2004 and 2014. Logistic Regression (LR), Support Vector Machine (SVM) and Artificial Neural Network (ANN) were trained. 20% of the data was used as a test set for calculating prediction accuracy for R0, LOS, COMP and SURV. Model performance was measured by plotting Receiver Operating Characteristic (ROC) Curves and calculating the Area Under ROC (AUROC).

Results: ML models and ANNs were trained on 1,147 cases. The AUROC for all outcome predictions ranged from 0.608 to 0.793 indicating modest to moderate predictive ability. The models performed best at predicting length of stay >14 days with an AUROC of 0.793 using preoperative and operative data. Visualised LR Model weights indicate varying impact of variables on the outcome in question.

Conclusion: This paper highlights the potential for predictive modelling of large international databases. Current data allow moderate predictive ability of both complex ANNs and more classic methods.

What does this paper add to the literature?

Providing predictions of outcomes for pelvic exenteration surgery is difficult. The application of machine learning to patient, peri-surgical and surgical variables results in moderate ability to predict outcomes of pelvic exenteration surgery. Based on this dataset, there is no benefit to using more complex artificial neural networks.

Introduction:

Machine Learning (ML) can be utilised as a statistical technique which allows the creation of models that learn from examples without any prior conditional programming. Supervised ML determines how to assign importance to variables (e.g. Body Mass Index, Age) and train them to determine the probability of an event/risk (e.g. risk of 30-day complication). A trained model can process new information and perform accurate analysis on new data. Supervised ML provides the opportunity for predictions in both diagnosis and prognosis. However, it does not guarantee the identification of pathological pathways or modifiable risk factors [1]. Artificial Intelligence (AI) is a subset of ML which focuses on the application of "Artificial Neural Networks" to obtain a prediction.

Data relating to patients are ever increasing, with more demographic, laboratory, radiological and outcome data (survival, quality of life etc.) being generated, resulting in gigabytes of data per visit. It quickly becomes difficult for clinicians to process all these data points using traditional statistical analysis. ML may facilitate prompt assessment of the data, providing summaries and useful predictions to aid clinicians in their decision-making [2–4]. Previously published papers have used ML to provide accurate predictions of cancer staging (based on clinical, histopathological and genetic data), hospital length of stay, risk of complications, and need for intervention or re-intervention [5–12]. This enhanced predictive power may improve our ability to counsel patients with our perioperative planning [13–16].

The aim of this study is to assess the predictive accuracy of several machine learning models including Logistic Regression (LR), Support Vector Machine (SVM) and Artificial Neural Networks (ANN), with regards to clear resection margins, length of hospital stay, perioperative complication rate and survival for advanced rectal cancer requiring exenteration.

Method:

Development Environment Setup

A deep learning computer was assembled with the following specifications:

- Threadripper 2 2950x CPU
- NVidia RTX 2080ti GPU
- 48GB of RAM

Once built, a development environment was set up. The programming was done in Python (Version 3.6, Python Software Foundation). The LR and SVM models were built using the Sci-Kit Library [17]. The ANN was built using Keras on a TensorFlow® backend and trained on a

Nvidia 2080ti GPU using NVidia cuDNN [18]. Data were visualized using Matplotlib [19]. These tools are open source and free to use.

Data pre-processing

The PelvEx dataset was used in this study [20–22]. This is a retrospective international observational cohort study assessing patient outcomes following pelvic exenteration for colorectal cancer over a ten-year period between 2004 and 2014. Ethical approval was granted at the individual institutional level, and data were collected and submitted centrally by principal investigators in their respective centres. A colorectal cancer diagnosis was based on histological and/or radiological imaging.

Missing BMI and Duration of Surgery was imputed using R statistical software following techniques described by Su *et al* [23,24]. The rest of the cases containing missing data were excluded from the study. In total 1,147 cases were eligible for model training and included in the study. Categorical variables were one-hot encoded, i.e. categorical variables with no natural ordering were converted into a binary form for easier machine learning processing. The dataset was standardised by removing the mean and scaling to unit variance: $z = x - \mu / \delta$ where μ is the mean of the training samples and δ is the standard deviation of the training samples. The resulting dataset was randomised, stratified and split once to ensure that all models were trained and tested on the same data (Figure).

The dataset was split into:

- 60% Training data – used for training the models.
- 20% Validation data – used for internally validating the models as they were training.
- 20% Test data – used for internally testing the resulting models.

Model Selection

The Sci-kit learn library was used for the Machine Learning models, while the Tensorflow library was used for the Neural Network model. 2 SVM kernels were tested (linear and sigmoid) and the best performing (sigmoid) used. The best ANN model was selected by using the Talos [25] toolkit to create and test multiple fully connected ANN trained to predict 1-year survival and retrained for other outcomes using the same model. The highest performing one was used in this study.

Assessing Model Performance

The goal of the analysis was to assess AI prediction of pelvic exenteration outcomes and compare ML to ANN methods. The target outcomes were:

- R0 Resection
- Length of Stay >14 days
- Complications within 30 days
- Survival >1 year

Due to the imbalance in the outcome data for events, e.g. death, weighing was applied to give a correct positive prediction more importance and reduce false negatives. The test set of cases was used to assess model performance. Test cases were fed into the trained models in order to output predictions. The output prediction was compared against the known true outcome for each case. This comparison was used to plot a Receiver Operating Characteristic (ROC) Curve and the Area Under ROC (AUROC), which are used to assess the model performance. An AUROC of 1 indicates perfect predictive ability, whereas an AUROC of 0.5 indicates no better than random guessing.

Results

1,147 patients underwent pelvic exenteration for advanced or recurrent rectal cancer (Table 1). There were 692 males representing 60% of the cases. 610 cases had locally advanced rectal cancer (53%), and 537 cases (47%) had locally recurrent colorectal cancer. The median age was 63 years and the median BMI was 26. The median LOS was 15 days. 424 patients (37%) had a clinically significant complication within 30 days. 803 exenterations (70%) achieved an R0 resection, 290 had an R1 resection (25%) and 54 had an R2 resection (5%).

Predictive Performance

The results indicated a moderate predictive ability of all three AI methods. For the prediction of R0 resection the models achieved an AUROC between 0.608 and 0.707 (Figure 2A). LR had the best AUROC of 0.707, followed by SVM with 0.682 and ANN performed the worst with an AUROC of 0.608.

All three models were better at predicting length of stay >14 days, achieving an AUROC between 0.757 and 0.793 (Figure 2 B). LR performed best with 0.793 followed by ANN with 0.789 and finally SVM with 0.757.

Complications at 30 days had a predicted AUROC between 0.719 and 0.756 (Figure 3A). In this instance ANN performed best, with an AUROC of 0.756 followed by LR with 0.742 and finally SVM with 0.719.

Predicting 1-year survival achieved an AUROC between 0.671 and 0.740 (Figure 3B). LR performed best with an AUROC of 0.740, followed by ANN with 0.693 and finally SVM with 0.671. Overall, LR performed best in predicting 1-year survival, length of stay and R0 resection, while ANN performed best in predicting complication within 30 days post-operatively. SVM was the worst performing model overall.

Logistic Regression Weights

The weights from trained LR models were extracted to assess the positive/negative influence of the training variables on prediction (Figure 4). To predict outcomes using logistic regression, all variables were used at the same time. Due to this, odds ratios are unavailable, as they were not individually analysed by the logistic regression model. The weights indicate in which direction a variable influences the prediction; positive weights result in a positive prediction; negative weights result in a negative prediction. It is the combination of all the variables along with their weights that results in the overall prediction. The full table of weights can be viewed in the appendix table A1.

R0 Resection

The weights show that the biggest indicator of a successful R0 resection is whether or not the surgery is performed for Recurrent CCA. Surgery performed for Locally Advanced CCA has a negative impact on a successful R0 resection. This is followed by the type of surgery (Total Exenteration) and then gender. The second largest negative influence was BMI, followed by No Neoadjuvant Therapy.

Length of Stay >14 days

The weights indicate that in the context of all the other variables, Duration of Surgery has the largest impact on predicting a LOS >14 days, followed by the number of lymph nodes excised and whether it was a redo operation. In this case, a positive prediction means that the patient has a prolonged stay in hospital; whereas a shorter duration of surgery predicts less nodes excised and no surgical reintervention. A high BMI has the biggest influence on a negative prediction, followed by male gender and having neoadjuvant radiotherapy.

Complications within 30 days

Interventional Radiology input has the biggest impact on a positive prediction for Complications within 30 days. This is followed by surgical reintervention and increasing age. The largest influences on negative prediction are No Neo-Adjuvant Therapy, longer duration of surgery and a greater number of units of blood transfused.

Survival >1 year

The survival model weights indicate that an R0 resection has the biggest influence on a positive prediction, followed by neo-adjuvant chemotherapy. The largest negative influence is having an R2 resection, followed by a longer LOS.

Discussion and conclusions

Logistic Regression outperforms ANN and SVM in all predictions but the 30-day complication rates. In every prediction task all three models achieved similar performance based on their AUROCs. This indicates that individual ML models are better at processing the complex combination of patient and surgical factors for predicting a specific outcome.

Overall, LR was the easiest model to implement. It required the least computing power and does not carry the "black box" feature common to many ML models. Due to the complex mathematics and connections in some models, it is not possible to view how the variables contributed to the final prediction. This means that while a tool can give accurate and meaningful predictions, studying its structure will not provide any insights into the underlying pathological or clinical processes. Despite its enhanced ability to process complex relationships between variables, the computationally expensive ANN did not outperform the LR model. SVM was the worst performing model overall. While there are no studies directly comparing surgeons to these predicting tools, proxy clinical calculators based on deterministic sets of rules achieve roughly the same accuracy [15].

Predicting R0 resection from clinical data is an important prognostic indicator, with which the models had moderate success. This is the most important factor in determining long-term survival [13]. Clinical Decision Tools can help bridge the gap between translating big data into actual expected clinical outcomes and help in counselling patients both pre- and post-operatively. The development of a risk calculator could be a useful tool for illustrating risk and benefit in a format that is easy to explain.

The variable weights from the logistic regression model illustrate a complex interplay. Depending on the desired outcome prediction, a variable may have positive or negative effects. For example, an increased BMI influenced R0 prediction negatively, meaning that R0 was less likely with a higher BMI in the context of all the other variables. However, it had a positive influence on survival prediction. Another example is the apparent influence of duration of surgery on length of stay and complications within 30 days. A longer operation correlated with predicting a longer stay in hospital; however it also predicted no complication within 30 days. There is an additional interesting result where locally recurrent CCA yields a higher prediction of achieving an R0 resection, yet predicts shorter survival compared with locally advanced CCA, which conversely predicts a longer survival but less R0 resection. The scale to which a variable influences a prediction also changes based on the question. It is evident that manually calculating a patient's predicted course would be time consuming and prone to error. Using a prediction model may help save time and highlight the most important modifiable factors individualised for the surgeon and the patient. As we progress in the digital age, AI is likely to become a routine part of clinical practice.

We acknowledge that there are significant limitations to this analysis, mainly related to the retrospective nature of the study. There was also a high number of missing data points which reduces the accuracy of the analysis. However, as this is the largest international dataset on pelvic exenteration outcomes for rectal cancer, it is unlikely to be improved upon in the near future. Further, the goal of the study was to compare various models that are as close to each other as possible, as in the future, focus may be directed towards optimising an individual model, such as logistic regression. Regardless of any limitations, this work provides valuable insight into the feasibility of AI and the best models to apply.

This paper demonstrates the application of AI to a large international collaborative database on pelvic exenteration outcomes and shows moderate predictive capabilities for R0 resection, length of stay >14 days, complications within 30 days and survival >1 year. With the current available data there was no benefit to using the more complex ANN when compared with LR and SVM. Variable weights extracted from the LR model showed the influence of various variables on predicted outcomes. This work highlights great potential for further studies combining the AI model with individual patient's data to help decision making and optimisation of treatment.

References

1. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med. 2019 04;380(14):1347–58.

2. Adams K, Papagrigoriadis S. Creation of an effective colorectal anastomotic leak early detection tool using an artificial neural network. Int J Colorectal Dis. 2014 Apr 1;29(4):437–43.

3. Burke HB, Goodman PH, Rosen DB, Henson DE, Weinstein JN, Harrell FE, et al. Artificial neural networks improve the accuracy of cancer survival prediction. Cancer. 1997;79(4):857–62.

4. Anand SS, Smith AE, Hamilton PW, Anand JS, Hughes JG, Bartels PH. An evaluation of intelligent prognostic systems for colorectal cancer. Artificial Intelligence in Medicine. 1999 Feb 1;15(2):193–214.

5. Horta AB, Salgado C, Fernandes M, Vieira S, Sousa JM, Papoila AL, et al. Clinical decision support tool for Co-management signalling. Int J Med Inform. 2018;113:56–62.

6. Thornblade LW, Flum DR, Flaxman AD. Predicting Future Elective Colon Resection for Diverticulitis Using Patterns of Health Care Utilization. EGEMS (Wash DC) [Internet]. [cited 2019 Apr 30];6(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5983027/

7. Gründner J, Prokosch H-U, Stürzl M, Croner R, Christoph J, Toddenroth D. Predicting Clinical Outcomes in Colorectal Cancer Using Machine Learning. Stud Health Technol Inform. 2018;247:101–5.

8. Wang Q, Wei J, Chen Z, Zhang T, Zhong J, Zhong B, et al. Establishment of multiple diagnosis models for colorectal cancer with artificial neural networks. Oncol Lett. 2019 Mar;17(3):3314–22.

9. Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. Sci Rep [Internet]. 2018 Feb 21 [cited 2019 Apr 30];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5821847/

10. Maheshwari K, Cywinski J, Mathur P, Cummings KC, Avitsian R, Crone T, et al. Identify and monitor clinical variation using machine intelligence: a pilot in colorectal surgery. J Clin Monit Comput [Internet]. 2018 Sep 24 [cited 2019 May 3]; Available from: https://doi.org/10.1007/s10877-018-0200-x

11. Amirkhan R, Hoogendoorn M, Numans ME, Moons L. Using recurrent neural networks to predict colorectal cancer among patients. In: 2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings [Internet]. Institute of Electrical and Electronics Engineers Inc.; 2018 [cited 2019 May 3]. p. 1–8. Available from: https://research.vu.nl/en/publications/using-recurrent-neural-networks-to-predict-colorectal-cancer-amon

12. Ponzio F, Macii E, Ficarra E, Di Cataldo S. Colorectal Cancer Classification using Deep Convolutional Networks. An Experimental Study. In SciTePress; 2018 [cited 2019 May 3]. p. 58–66. Available from: https://iris.polito.it/handle/11583/2696438

13. Yang TX, Morris DL, Chua TC. Pelvic exenteration for rectal cancer: a systematic review. Dis Colon Rectum. 2013 Apr;56(4):519–31.

14. Soguero-Ruiz C, Hindberg K, Mora-Jiménez I, Rojo-Álvarez JL, Skrøvseth SO, Godtliebsen F, et al. Predicting colorectal surgical complications using heterogeneous clinical data and kernel methods. Journal of Biomedical Informatics. 2016 Jun 1;61:87–96.

15. Weller GB, Lovely J, Larson DW, Earnshaw BA, Huebner M. Leveraging electronic health records for predictive modeling of post-surgical complications. Stat Methods Med Res. 2018 Nov 1;27(11):3271–85.

16. Sanger PC, van Ramshorst GH, Mercan E, Huang S, Hartzler AL, Armstrong CAL, et al. A Prognostic Model of Surgical Site Infection Using Daily Clinical Wound Assessment. Journal of the American College of Surgeons. 2016 Aug 1;223(2):259-270.e2.

17. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12:2825–2830.

18. Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems [Internet]. 2015. Available from: http://tensorflow.org/

19. Hunter JD. Matplotlib: A 2D graphics environment. Computing in Science & Engineering. 2007;9(3):90–95.

20. PelvEx Collaborative. Surgical and Survival Outcomes Following Pelvic Exenteration for Locally Advanced Primary Rectal Cancer: Results From an International Collaboration. Ann Surg. 2019;269(2):315–21.

21. PelvEx Collaborative. Factors affecting outcomes following pelvic exenteration for locally recurrent rectal cancer. BJS. 2018;105(6):650–7.

22. Srinivasaiah N, Shekleton F, Kelly ME, Harji D, Malietzis G, Askari A, et al. Minimally invasive surgery techniques in pelvic exenteration: a systematic and meta-analysis review. Surg Endosc. 2018 Dec 1;32(12):4707–15.

23. Su Y-S, Gelman A, Hill J, Yajima M. Multiple Imputation with Diagnostics (mi) in R: Opening Windows into the Black Box. Journal of Statistical Software, Articles. 2011;45(2):1–31.

24. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2019. Available from: https://www.R-project.org

25. Hyperparameter Optimization for Keras Models. Contribute to autonomio/talos development by creating an account on GitHub [Internet]. Autonomio; 2019 [cited 2019 Jun 18]. Available from: https://github.com/autonomio/talos

Tables and Figures

Figure 1 Flow chart illustrating the case allocation for this study.

Figure 2 **A** ROC Curve for R0 resection. **B** ROC Curve for Length of stay >14 days. (LR – Logistic Regression, SVM – Support Vector Machine, NN – Artificial Neural Network, ROC – Receiver Operating Characteristic)

Figure 3 **A** ROC Curve for Complications within 30 days. **B** ROC Curve for Survival >1 year. (LR – Logistic Regression, SVM – Support Vector Machine, NN – Artificial Neural Network, ROC – Receiver Operating Characteristic)

Figure 4 Stacked bar chart of variable weights for Logistic Regression prediction models. Positive values promote a “yes” prediction, negative values promote a “no” prediction. (BMI – Body Mass Index, CCA – Colorectal Cancer)

Total cases 1147	**Median (IQR)** **Percent (#)**	**Std**
Age (years)	63 (15)	11.4
Male: Female	60:40	
BMI	26 (5)	4.15
Surgery Duration (min)	450 (190)	176
Blood Units	2 (3)	5
Number of Nodes Excised	6 (15)	12
Number of Nodes Positive	0 (1)	3

Survival (months)	29 (35)	29
Length of Stay (days)	15 (15)	18
Re-Admission w/i 30d	6% (68)	
Complication w/i 30d	37% (424)	
Surgical Reintervention	10% (118)	
Interventional Radiology	7% (83)	
Death w/i 30 days	1% (13)	
No Neoadjuvant Rx	23% (267)	
Chemo-Radiotherapy	61% (700)	
Radiotherapy	11% (122)	
Chemotherapy	5% (58)	
Total Exenteration	38% (432)	
Posterior Exenteration	45% (521)	
Anterior Exenteration	6% (68)	
Modified Exenteration	11% (126)	
Sacral Bone Resection	19% (213)	
R0 resection	70% (803)	
R1 resection	25% (290)	
R2 resection	5% (54)	
Locally Advanced CCA	53% (696)	
Recurrent CCA	47% (451)	

Table 1 Summary of used variables in the database of the 1,147 cases used in this study. (BMI – Body Mass Index, CCA – Colorectal Cancer)

Figure 1

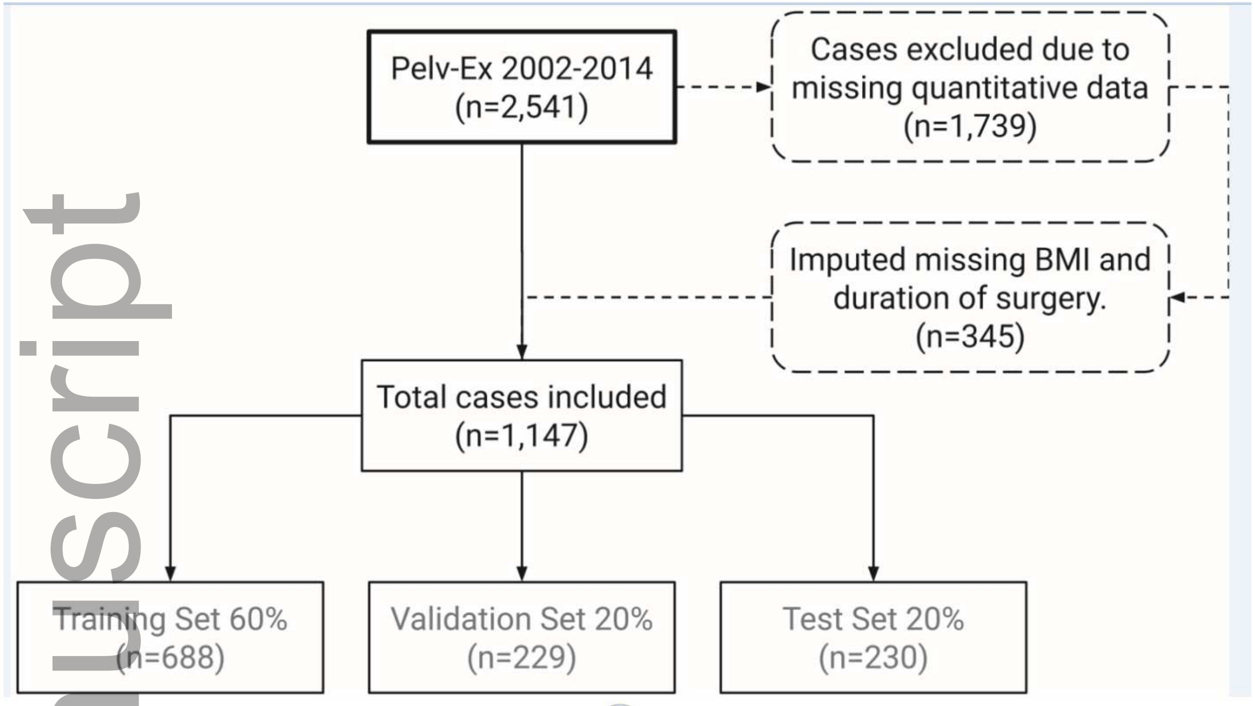


Figure 2

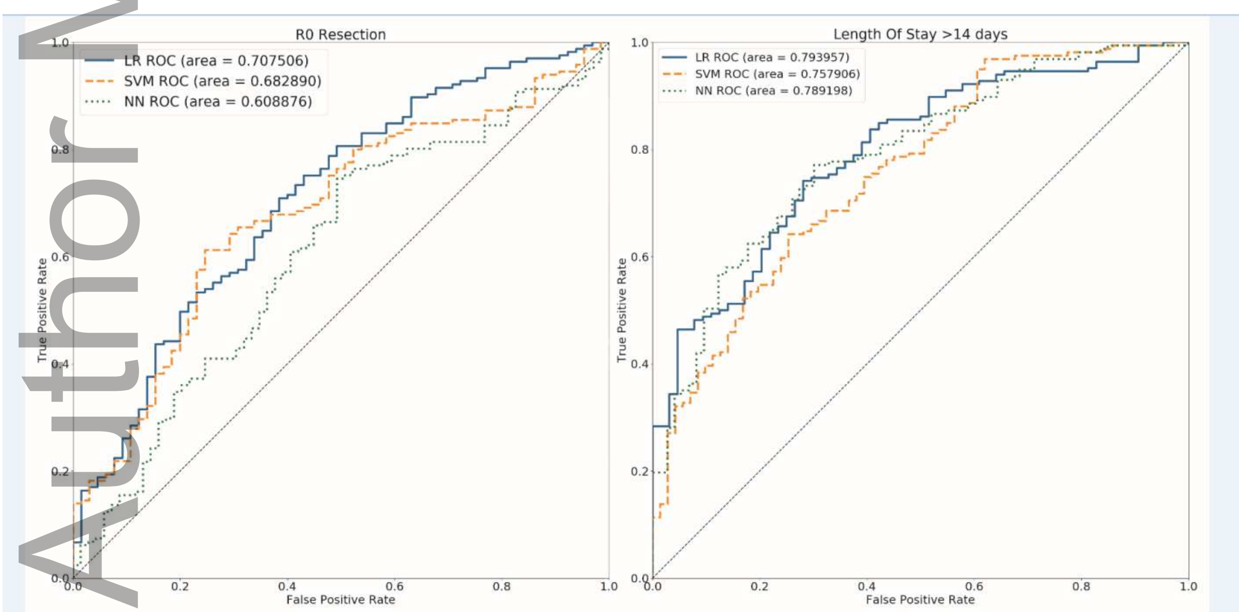


Figure 3

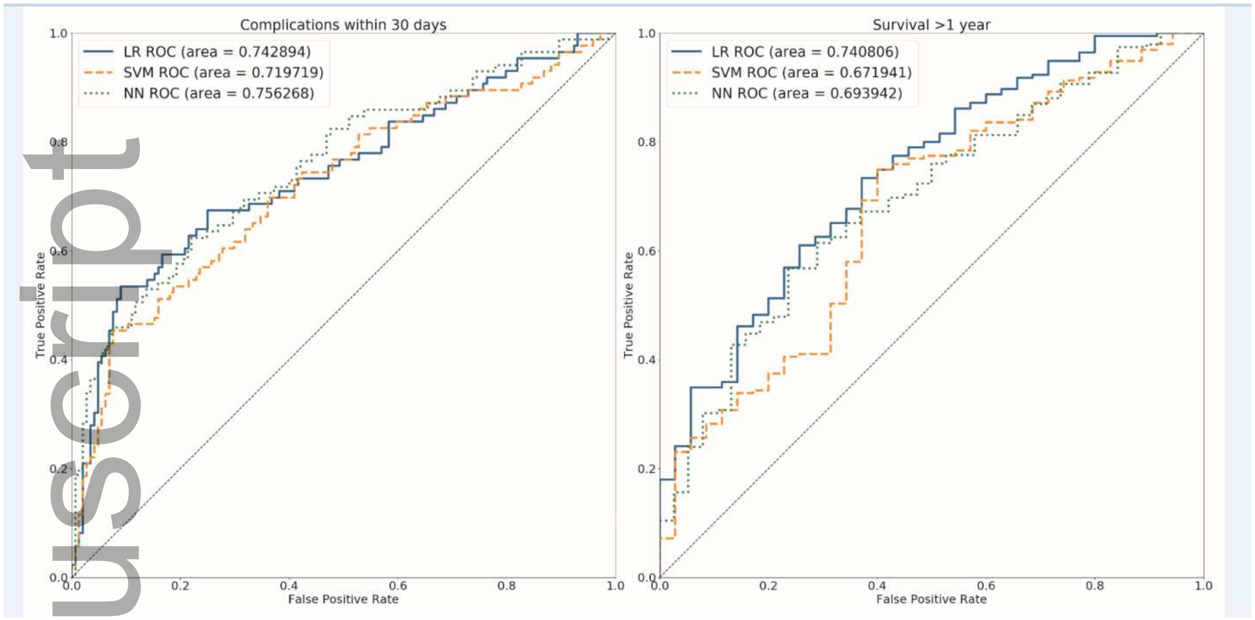
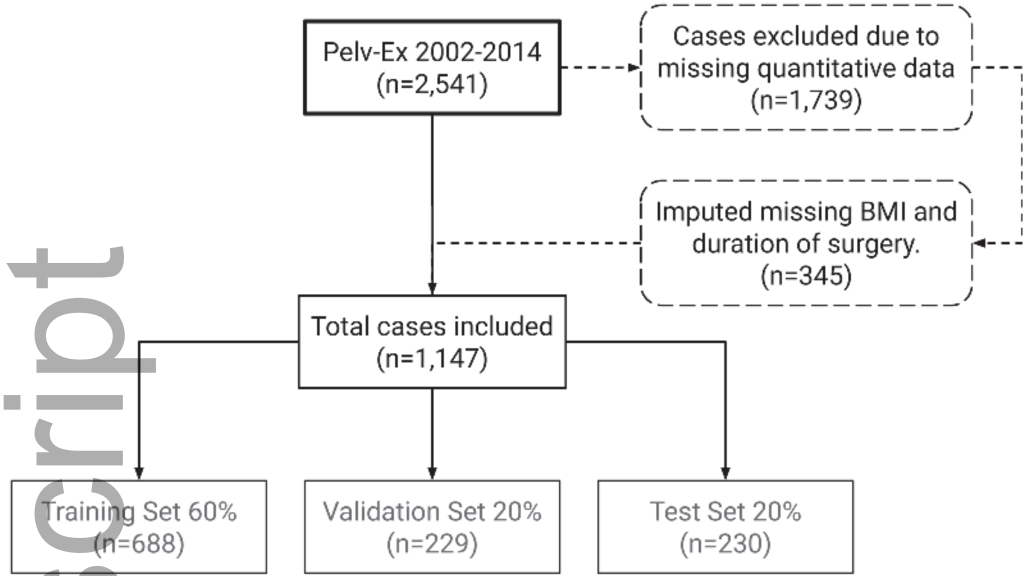
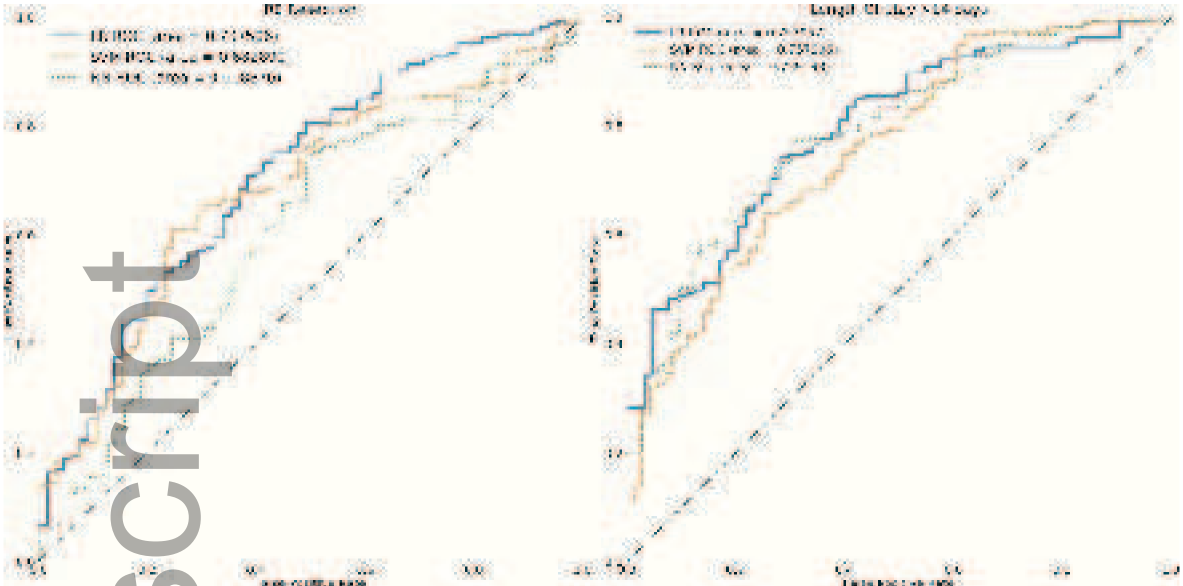

Complications within 30 days
LR ROC (area = 0.742894)
SVM ROC (area = 0.719719)
NN ROC (area = 0.756268)
True Positive Rate
False Positive Rate
Survival >1 year
LR ROC (area = 0.740806)
SVM ROC (area = 0.671941)
NN ROC (area = 0.693942)
True Positive Rate
False Positive Rate

Figure 4

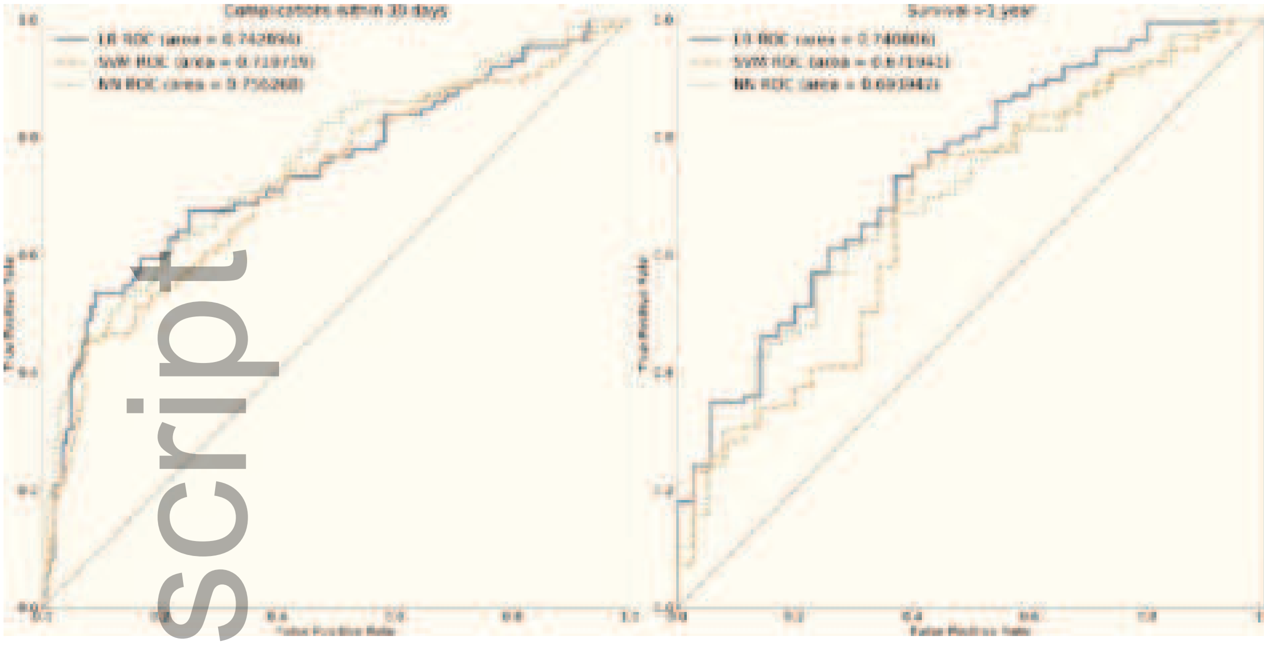

Logistic Regression Weights
Age
Male
Female
BMI
Surgery Duration
Blood Units
Number of Nodes Excised
Number of Positive Nodes
Length of Stay
Re-Admission w/i 30d
Complications w/i 30d
Surgical Reintervention
Interventional Radiology
No NeoAdj Rx
Chemo-Radiotherapy
Total Exenteration
Posterior Exenteration
Anterior Exenteration
Modified Exenteration
Sacral Bone Resection
R0 Resection
R1 Resection
R2 Resection
Locally Advanced CCA
Recurrent CCA
R0 Resection
Length of Stay >14 days
Complications w/i 30d
Survival >1 year



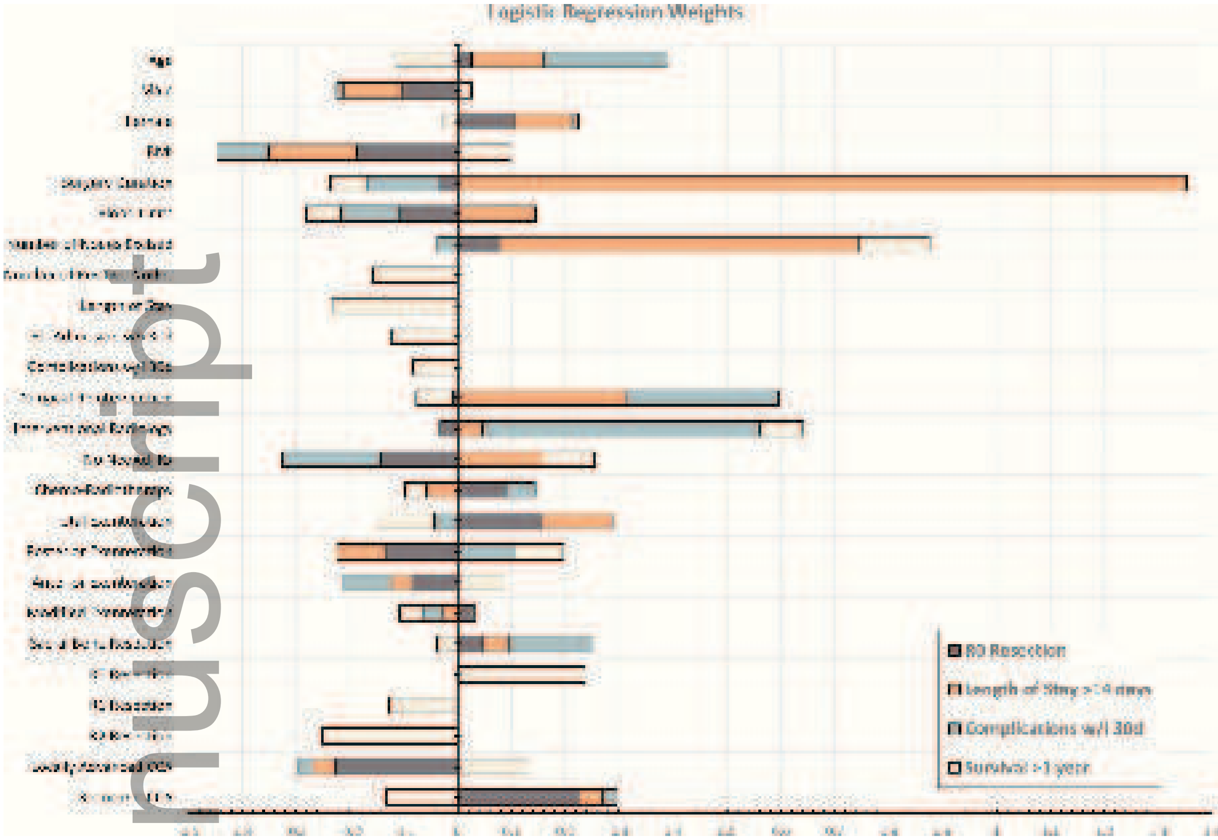
codi_15235_f1.tif



codi_15235_f2.tif



codi_15235_f3.tif



codi_15235_f4.tif