

Research Article

A Machine Learning Approach for Prediction of Surgical Outcomes in Elective Surgery

Dennis Muriithi^{1,*} , Virginia Mwangi² 

¹Faculty of Science and Technology, Chuka University, Chuka, Kenya

²Nursing Department, Chuka County Referral Hospital, Chuka, Kenya

Abstract

The aim of this research was to design a Machine Learning (ML) approaches to predict surgical outcome associated with perioperative risks factors among patients undergoing elective surgery. The research employed descriptive cross-sectional survey and a sample size of 292 patients. Only adult patients undergoing elective surgery were considered. Machine Learning (ML) Algorithm such as Logistic regression, Support vector machine, k-nearest neighbors and random forest were used to provide insights into how different factors such as patient related perioperative risk, procedure related perioperative risk and health system related perioperative risk influence the likelihood of successful surgical outcome. The study found that Random Forest model achieved the highest cross validation accuracy of 100%, which means it correctly classified all data points in the test set. It implies that the random Forest model was the most suitable for classifying surgical outcome among elective surgery patient at Chuka County Referral Hospital. It had a Kappa of 1 indicating a perfect agreement between its predictions and the ground truth in comparison with other algorithms. In addition, Random Forest model achieves a perfect score (1.0) for sensitivity, precision, F1-Score, and balanced accuracy. This suggests that the model is doing extremely well at correctly classifying both positive and negative cases. Availability of main surgical supplies (health system related perioperative risk factors) had the highest score indicating that it was more important factor for the models predictions than other perioperative risk factors. In this study, the Machine Learning analysis identified unknown parameters associated with successful surgical outcome. An application of Machine Learning algorithms as a decision support tool could enable the medical health practitioners to predict the surgical outcome of patients undergoing elective surgery and consequently optimize and personalize clinical management of patient.

Keywords

Machine Learning, Algorithms, Elective Surgery, Surgical Outcome, Perioperative Risks

1. Introduction

Every patient undergoing surgery is at risk for morbidity and mortality since surgery is a significant kind of treatment that carries inherent risk for even the healthiest patients. Physical condition of each patient is crucial to the success of surgery [4]. Every surgical patient presents with a different

collection of comorbidities and risk factors. In addition to the risks associated with surgery and anesthesia that apply to all patients, each surgical procedure has its own unique set of management challenges and potential complications [3]. A report by the National Confidential Enquiry into Patient

*Corresponding author: Kamuriithi2011@gmail.com (Dennis Muriithi)

Received: 31 May 2024; **Accepted:** 27 June 2024; **Published:** 20 August 2024



Copyright: © The Author(s), 2024. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Outcome and Death (NCEPOD) recommended that all elective high-risk patients should be evaluated and thoroughly investigated in pre-assessment clinics. The report emphasized on the need of developing a UK-wide approach that permits swift and easy identification of patients who are at high risk of postoperative morbidity and mortality as cited by [10]. The World Health Organization (WHO) reported that over 313 million surgical procedures are carried out year globally, and that surgical mortality and complications in developing nations are expected to be between 0.4% and 0.8% and 3–17%, respectively [6, 11]. It was suggested that the probability of a bad outcome for each individual should be calculated as part of the assessment of high-risk surgical patients. To ensure that patients receive the appropriate level of care, this needs to be explained to them, carefully documented, and used to categorize them. While there is a high risk of severe morbidity or short-term death, it is prudent to proceed cautiously while considering whether to proceed with surgery at all.

A study was conducted on perioperative patient outcomes in the African surgical outcomes, carried out in 25 African countries including Kenya and it was concluded that patients in Africa were twice as likely to die following surgery as compared to the global average for postoperative fatalities, despite having a low risk profile and few complications following surgery [2]. They went on to say that efforts to widen access to surgical care in Africa should be accompanied by better monitoring for patients who experience postoperative problems who are showing signs of physiologic decline. This study focused on patients undergoing elective surgery. Elective surgery refers to surgery that is planned ahead of time because there is no medical emergency involved and surgery can be postponed for up to a year in this case. It is therefore important that patients undergoing elective surgery and health care workers be able to understand the perioperative risks and their prevention in order to reduce length of hospital stays and avoid suffering and loss of life that result from complications of surgical procedures. Patients who have complications are more likely to die even 5 years after surgery. In recent times, attempts to involve patients in decision-making over their care have generally been made. However, there are a lot of gaps in knowledge among health care workers of assessing the risk of complications and communicating that risk to patients as part of the consent process, assisting them in deciding whether or not to proceed with surgery.

Prior to surgery, perioperative risks are evaluated by taking into account the patient's medical history, doing a physical examination, and running several standard tests in addition to any additional that could be considered required. In this study the patient related perioperative risk factors, procedure related perioperative risk factors and health system related perioperative risk factors were crucial in predicting the surgical outcome for a patient undergoing elective surgery. Patient-related perioperative risk factors are influencing and incidental elements that raise the patient's risk of injury or likelihood of suffering loss during the perioperative period. This includes

factors such as comorbidities, anxiety, obesity, cigarette smoking and alcohol consumption. Procedure related perioperative risk factors refer to factors that influence the chance or probability of suffering harm or loss during the intraoperative period, example type and complexity of surgery and duration of surgery. Health system-related perioperative risks factors refer to health system related factors namely; equipment failure, availability of supplies and staffing that influences the chance or probability of suffering harm or loss during the perioperative period. In this study, patients undergoing elective surgery had their surgical outcomes and perioperative hazards evaluated [5]. Following elective surgery, the patient's surgical results were assessed. Subsequently, interventions ought to be focused on perioperative risk mitigation and enhancing health systems to lower rates of morbidity and mortality, re-admission, and shorten hospital stays, which lowers the cost of surgery, as well as to minimize blood loss and pain [10, 7, 1].

Data from the Chuka County Referral Hospital's records department revealed that cancellations and delays were commonplace for elective surgery. This was brought on by inadequate surgical supplies, inadequate staffing, and inadequate patient preparation prior to surgery. 526 individuals were readmitted to CCRH in 2021 after undergoing elective surgery, and 6 patients passed away following the procedure (Chuka County Referral Hospital records department, 2021).

Machine learning (ML) is a subfield of artificial intelligence (AI) that focuses on the development of algorithms capable of learning from and making predictions or decisions based on data [12]. Classification is a supervised machine learning algorithm that involves classifying and predicting the categorical outcome variable. Several supervised machine-learning algorithms are used for classification. These algorithms include multinomial logistic regression, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forests, and Naïve Bayes classifiers. The algorithm involves training computational models to recognize patterns and relationships within large datasets, thereby enabling the automation of analytical tasks and discovering insights that might be difficult for humans to discern. These models learn by example, improving their performance over time as they are exposed to more data. The application of machine learning models to classify surgical outcome among patient undergoing elective surgery provides a comprehensive understanding of the factors contributing to unsuccessful surgical outcome and develop targeted interventions to support health systems. The aim of this study is to design a Machine Learning (ML) approaches to predict surgical outcome associated with perioperative risks among patients undergoing elective surgery.

2. Material and Methods

2.1. Study Area

Chuka County Referral Hospital (CCRH) in Tharaka-Nithi County served as the study's site. It is a public health facility situated in Chuka division, Chuka Igambang'ombe constituency along Meru-Nairobi road. Since the hospital offers specialist surgical procedures in addition to curative, preventative, and promotive treatments, it was taken into consideration for this study. With a strong patient flow, the hospital is the biggest in the county and provides referral services to outlying clinics. It has two operation theatres and performs approximately 2100 operations annually. The study was carried out in the surgical wards, operating theatres and surgical out-patient clinics.

2.2. Study Design

A descriptive cross-sectional survey was used to collect data on perioperative surgical risks among patients at the hospital's surgical units. The selected design was chosen because it facilitated the precise collection of data regarding the current status of the scenario (perioperative risks), which was then utilized to draw conclusions. It also made it possible to collect information in an appropriate manner, both quantitatively and qualitatively.

2.3. Eligibility Criteria

Inclusion criteria: Only adult patients undergoing elective surgery at Chuka County Referral Hospital were considered. Exclusion criteria: Children and patients who were seriously ill at the time of the study were not included.

2.4. Study Population

The study targeted all the surgical patients that were undergoing elective surgery during study period (2nd June to 30th December 2023) at Chuka County Referral Hospital in Tharaka-Nithi County, Kenya. A research administered structured questionnaire for data collection. A sample size of 292 patients was considered for this study.

2.5. Machine Learning Approaches

Machine Learning (ML) Algorithm can play crucial roles in understanding and predicting the surgical outcome. ML models aims to provide insights into how different factors such as patient related perioperative risk, procedure related perioperative risk and health system related perioperative risk influence the likelihood of successful surgical outcome. Logistic regression, Support vector machine, k-nearest neighbors and random forest algorithms complement this approach by offering more robust techniques for classification and

prediction. These machine learning algorithms can capture complex relationships between predictors and surgical outcome, allowing for more accurate predictions and identification of high-risk patient undergoing elective surgery.

Logistic regression is a supervised learning algorithm that classifies new data by calculating the probabilities of the data belonging to each class. Odds ratios are an informative way of interpreting the impact each of our predictors has on the odds of a case belonging to the positive class [8]. The Support Vector Machine (SVM) algorithm finds a hyperplane (a surface with one less dimension than there are predictors) that best separates the classes. It is sensitive to the values of its hyper-parameters, which must be tuned to maximize performance. Kernel-Nearest Neighbors (kNN) is a popular and versatile algorithm used in machine learning for both classification and regression tasks. It works on the principle that data points close together tend to be similar. So, to classify a new data point, kNN identifies the k closest data points (neighbors) from the training data set. The class label that is most frequent among these k neighbors becomes the predicted class for the new data point. Random forest is a powerful machine learning algorithm that utilizes the combined strength of multiple decision trees to make predictions. It is a popular choice for various tasks due to its accuracy, flexibility, and robustness.

The following steps were involved in Machine Learning modeling process:

1) Data Collection and Preparation

This step involves collection of data that is of good quality, accurate and relevant for the study in hand. Also, cleaning the data by handling missing values, outliers, and inconsistencies as well as perform transformations like scaling or encoding categorical features for the algorithms to work effectively.

2) Exploratory Data Analysis (EDA)

This step involves getting familiar with the data by analyzing its statistical properties and visualizing the relationships between features and the target variable. EDA helps identify patterns, trends, and potential issues in the data.

3) Model Selection and Training

1. Choosing the Algorithm such as Random Forest, kNN, SVM, or Logistic Regression.
2. Splitting Data: The data is divided into two sets: training data and testing data. The training data is used to build the model, and the testing data is used to evaluate the model's performance on unseen data.
3. Training the Model: The chosen algorithm is trained on the training data. This involves the algorithm learning the underlying patterns and relationships between features (perioperative risks factors) and the surgical outcome. Hyper-parameter tuning was done during this phase to optimize the model's performance. Hyper-parameters are settings specific to each algorithm that control its behavior.

4) Model Evaluation

1. Performance Metrics: After training, the models

performance was evaluated on the testing data using accuracy, precision, recall.

2. Model Interpretation: Based on chosen algorithm, interpret the models predictions to gain insights into the perioperative risks factors influencing the surgical outcome.

3. Results and Discussion

3.1. Patient Profile

The study sought to establish the patient demographic characteristics such as gender, age and education level as shown in Table 1.

Table 1. Patient Information.

Gender	Frequency	Percent
Female	128	43.84
Male	164	56.16
Age (Years)	Frequency	%

Gender	Frequency	Percent
18-35	100	34.25
36-50	192	65.75

Education	Frequency	%
Primary	64	21.92
Secondary	168	57.53
College	60	20.55

Table 1 shows the patient information of 292 survey respondents across gender, age bracket, and education level. There were slightly more males (164) than females (128) in the survey. This means that males make up 56.16% of the respondents while females make up 43.84%. Also, the majority of the respondents (192 or 65.75%) fall within the 36-50 age bracket. Only 100 respondents (34.25%) are between 18-35 years old. It was noted that over half (168 or 57.53%) of the respondents have a secondary level of education. Primary education comes in second at 21.92% (64 respondents), and 20.55% (60 respondents) have a college education.

3.2. Model Cross Validation Accuracy

Table 2. Model Cross Validation Accuracy.

Models	CV Accuracy	95% CI	P-value	Kappa
Logistic Regression	0.9403	(0.8541, 0.9835)	0.1581	0.7452
K-Nearest Neighbour	0.9104	(0.8152, 0.9664)	0.4414	0.3599
Random Forest	1	(0.9464, 1.0000)	0.0006	1
Support Vector Machine	0.9552	(0.8747, 0.9907)	0.07075	0.7048

Table 2 shows the performance of four machine learning models on a classification task. The accuracy metric represents the proportion of data points that each model classified correctly. In this case, the Random Forest model achieved the highest accuracy of 100%, which means it correctly classified all data points in the test set. It implies that the random Forest model is the most suitable for classifying surgical outcome among elective surgery patient in Chuka County Referral Hospital. A 95% Confidence Interval represents the range of values that we can be 95% confident the true accuracy lies within. For example, the 95% CI for the Random Forest is (0.9464, 1.0000), indicating that we can be 95% confident that the true accuracy falls somewhere between 94.64% and 100%.

Kappa is the agreement between the model's predictions

and the actual labels, while accounting for chance agreement. A Kappa of 1 indicates perfect agreement, while 0 indicates no agreement beyond chance. For instance, Random Forest achieved a Kappa of 1, which suggests perfect agreement between its predictions and the ground truth, while Support Vector Machine has a Kappa of 0.7048, indicating a moderate level of agreement. Random Forest model achieved the highest accuracy (1.0) and Kappa (1.0), suggesting the best performing model in classifying surgical outcome among elective surgery patient in Chuka County Referral Hospital. Support Vector Machine has a high accuracy (0.9552) but a lower Kappa (0.7048) compared to Random Forest. It also has a p-value that is not statistically significant at a 95% confidence level. Other models such as Logistic Regression and

K-Nearest Neighbors models considered in this study have lower accuracy and Kappa scores compared to Random Forest and Support Vector Machine.

3.3. Evaluation of Model Performance

Table 3 shows performance metrics for four Machine Learning models on a binary classification task. Sensitivity (Recall) represents the proportion of positive cases that the model correctly identified. A high sensitivity (close to 1) indicates the model is good at identifying true positives and

avoids many false negatives. Specificity represents the proportion of negative cases that the model correctly identified. A high specificity (close to 1) indicates the model is good at identifying true negatives and avoids many false positives. Precision represents the proportion of positive predictions that were actually correct. A high precision indicates the model is good at making precise predictions, with very few false positives. Balanced Accuracy: This metric is the average of sensitivity and specificity, giving equal weight to both. It can be useful when the classes are imbalanced, meaning there are more examples of one class than the other.

Table 3. Evaluation of Model Performance.

Performance Metrics	Logistic Regression	K-Nearest Neighbour	Random Forest	Support Vector Machine
Sensitivity	0.933	0.983	1	1
Specificity	1	0.286	1	0.571
Precision	1	0.922	1	0.952
F1-Score	0.966	0.952	1	0.976
Balanced Accuracy	0.967	0.635	1	0.786

On the basis of the model Performance Metrics the following highlights were made;

Random Forest model: Achieves perfect scores (1.0) for sensitivity, precision, F1-Score, and balanced accuracy. This suggests that the model is doing extremely well at correctly classifying both positive and negative cases.

Logistic Regression & Support Vector Machine: Both models have high sensitivity (1.0) but lower specificity compared to Random Forest. This implies that the models might be very good at identifying positive cases but struggling to distinguish some negative cases from positive ones, leading to potential false positives. In this case, the models might be better choices despite their lower overall accuracy, because they have higher specificity.

K-Nearest Neighbor: Has the lowest overall performance among the models here. It has a high sensitivity (0.983) but a low specificity (0.286), indicating it might make many false positive predictions.

3.4. Box-and-Whisker Plot of Results

Box-and-whisker chart is used to visualize the distribution of data. It shows the following information: The center of the data (usually the median), the spread of the data (represented by the box) and the presence of any outliers (data points that fall outside a certain range). The results of random forest algorithm demonstrate a higher ability in classifying the unsuccessful surgical outcome and overall classification ability.

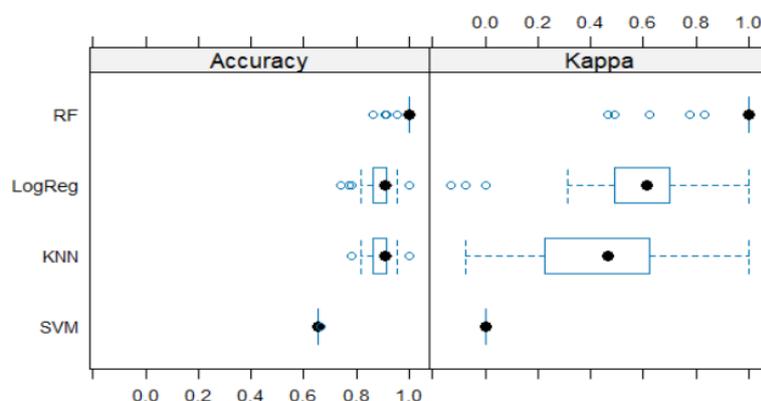


Figure 1. Comparison of the models.

3.5. Importance Predictors in the Models

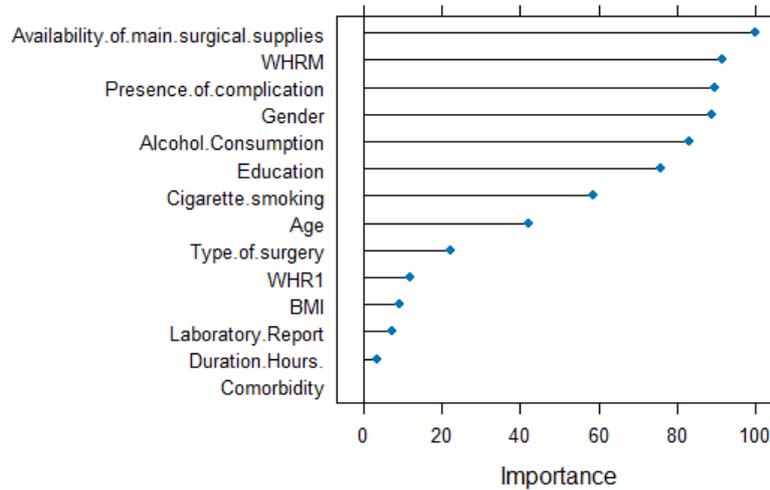
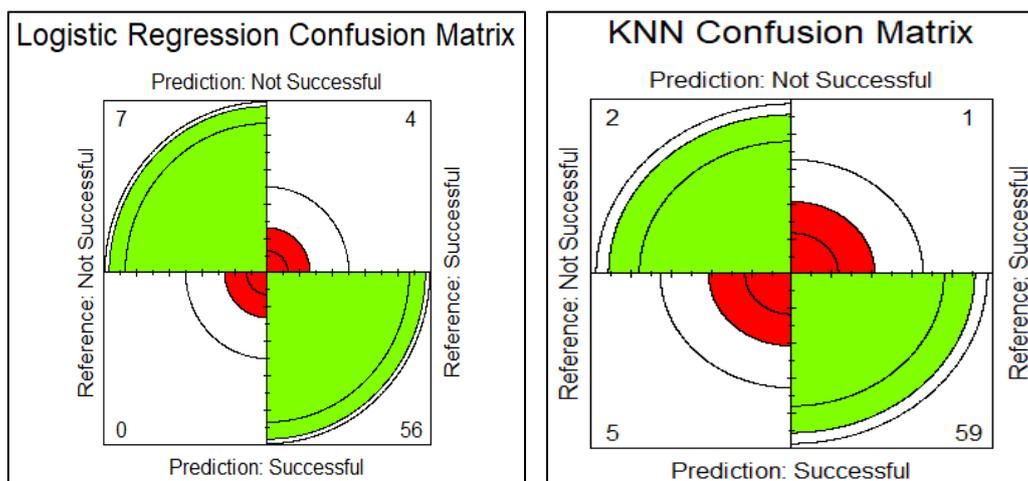


Figure 2. Importance Predictors in the Models.

The relative importance of each feature can be compared by the heights of the lines. For instance, the line for "Availability of main surgical supplies" appears to have a higher score than the line for "Education", indicating that availability of surgical supplies was a more important factor for the model's predictions than the patient's educational background. According to [9], having sufficient surgical equipment is essential for providing safe surgical care. In earlier research from Nigeria, Cameroon, Sierra Leone, Somalia, Ethiopia, and Malawi, shortages of surgical supplies were discovered. A system

supporting the equipment must be in place to guarantee that the right equipment is purchased, utilized, and maintained as intended, following adequate training and a secure supply chain of consumables, [10]. It is important to note that the specific interpretation of these feature importance scores can vary depending on the machine learning algorithm used. In this case, the importance scores were a useful tool for gaining insights into which factors were most influential for the Machine Learning models performance on classification task.

3.6. Model Confusion Matrices



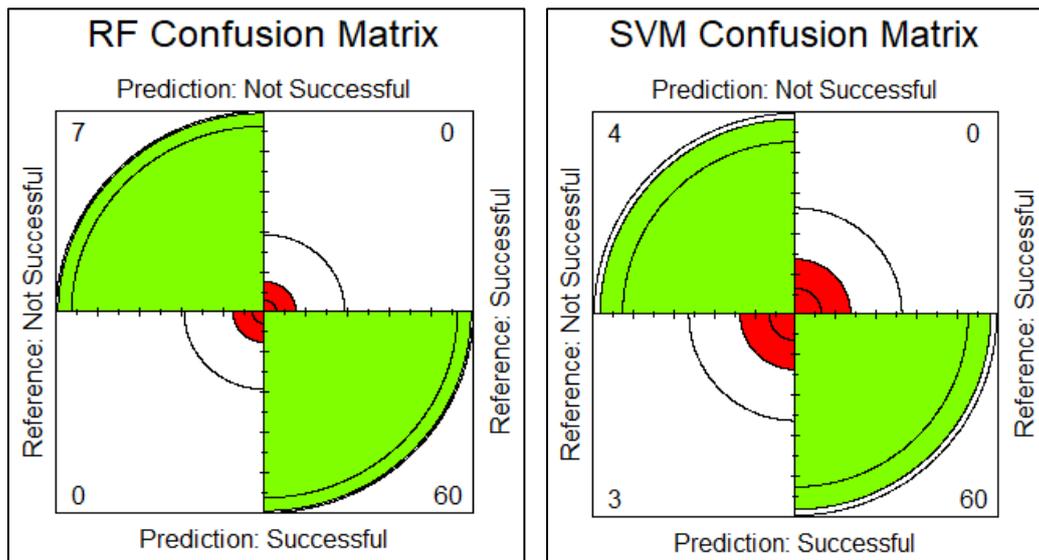


Figure 3. Model Confusion Matrices.

A confusion matrix is a table that allows visualization of the performance of an algorithm, particularly in classification tasks. Figure 2 shows the performance of a binary classification models, where the surgical outcome variable has two classes: successful and not successful. In this case, Rows represent the actual classes (reference), Columns represent the predicted classes and Values in each cell represent the number of instances in a particular category. For instance, RF confusion matrix shows that all cases were correctly classified in the respective categories. There were no misclassifications. Nevertheless, there were misclassification in other models but mean misclassification error was very low.

4. Conclusion

The application of data science to health care, in particular the use of machine learning predictive models, shows great promise. In this study, the Machine Learning analysis identified unknown parameters associated with successful surgical outcome. An application of Machine Learning algorithms as a decision support tool could enable the medical health practitioners to predict the surgical outcome of patients undergoing elective surgery and consequently optimize and personalize clinical management of patient. The study recommends use of large data set to verify the 100% model accuracy of Random forest model.

Abbreviations

BMI	Body Mass Index
CI	Confidence Interval
CV	Cross-Validation
EDA	Exploratory Data Analysis
ML	Machine Learning

WHO	World Health Organization
WHR1	Waist-to-Height Ratio(1=Male)
WHRM	Waist-to-Hip Ratio (Male)

Acknowledgments

The authors acknowledge the staff of Chuka County Referral Hospital who offered great help during data collection. Our special appreciation goes to the surgical patients who willingly contributed their experiences during the study period.

Author Contributions

Dennis Muriithi: Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, writing – original draft, Writing – review & editing

Virginia Mwangi: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Resources, Validation, Visualization, writing – original draft, Writing – review & editing

Funding

This work was supported and Financed by the authors.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Alsaigh H., Airuwaili R., Alsaleh I., Alghadoni A. (2020) Hospital Readmission after Surgery: Rate and predisposing factors. *International Journal of Medical Research & Health Sciences* volume 9, issue 10
- [2] Biccard, B. M., Madiba, T. E., Kluys, H. L., Munlemvo, D. M., Madzimbamuto, F. D., Basenero, A.,... & Arrey, O. (2018). Perioperative patient outcomes in the African Surgical Outcomes Study: a 7-day prospective observational cohort study. *The Lancet*, 391(10130), 1589-1598.
- [3] Fleisher, L. A., Fleischmann, K. E., Auerbach, A. D., Barnason, S. A., Beckman, J. A., Bozkurt, B.,... & Wijeyesundera, D. N. (2015). 2014 ACC/AHA guideline on perioperative cardiovascular evaluation and management of patients undergoing noncardiac surgery: executive summary: a report of the American College of Cardiology/American Heart association Task Force on practice guidelines. *Journal of nuclear cardiology: official publication of the American Society of Nuclear Cardiology*, 22(1), 162-215.
- [4] Gabriel, R. A., Sztain, J. F., Hylton, D. J., Waterman, R. S., & Schmidt, U. (2018). Postoperative mortality and morbidity following non-cardiac surgery in a healthy patient population. *Journal of anesthesia*, 32(1), 112-119.
- [5] Mazmudar, A., Vitello, D., Chapman, M., Tomlinson, J. S., & Bentrem, D. J. (2017). Gender as a risk factor for adverse intraoperative and postoperative outcomes of elective pancreatic resection. *Journal of Surgical Oncology*, 115(2), 131-136.
- [6] Meara JG, Leather AJM, Hagander L, Alkire BC, Alonso N, Ameh EA. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Lancet* 2015; 386: 569–624. Meersch, M., Schmidt, C., & Zarbock, A. (2017). Perioperative acute kidney injury: an under-recognized problem. *Anesthesia & Analgesia*, 125(4), 1223-1232.
- [7] Merkow, Ryan P., et al. (2015) 'Underlying reasons associated with hospital readmission following surgery in the United States.' *JAMA*, Vol. 313, no. 5.
- [8] Muriithi. D. K, Kihoro. J and Waititu. A (2012). Ordinal Logistic Regression Versus Multiple Binary Logistic Regression model for predicting student loan allocation. *Journal of Agriculture Science and Technology* Vol. 14(1), pp 190-203.
- [9] Oosting, R. M., L. Wauben, S. G., Madete, J. K., Groen, R. S., & Dankelman, J. (2020). Availability, procurement, training, usage, maintenance and complications of electrosurgical units and laparoscopic equipment in 12 African countries. *BJS Open*, 4(2), 326-331. <https://doi.org/10.1002/bjs5.50255>
- [10] Virginia. M. M, Gitonga.L Nyamu.H & Kainga.S. (2023). Influence of patient related perioperative risks on surgical outcomes among patients undergoing elective surgery at Chuka County Referral Hospital. *ijhsr (www.ijhsr.org)*. Vol. 13(11), pp 109-121. <https://doi.org/10.52403/ijhsr.20231114>
- [11] World Health Organization (2018). WHO global report on trends in prevalence of tobacco smoking 2000-2025, second edition. Geneva: World Health Organization <http://www.who.int/tobacco/publications/surveillance/trends-tobacco-smoking-second-edition>
- [12] Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science/SN Computer Science*, 2(3). <https://link.springer.com/article/10.1007/s42979-021-00592-x>