Use of Predictive Analytics to Identify Risk of Postoperative Complications among the Women Undergoing the Hysterectomy Surgery for Endometrial Cancer

By

Behnam Alimohammadisagvand

Supervised By

Prof. Pavel Andreev and Prof. Wojtek Michalowski

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of

Master of Science in Health Systems

uOttawa

The University of Ottawa

Abstract

The fourth most frequent cancer among women in Canada is endometrial cancer (EC) [1] with about 6500 new cases diagnosed in 2015. The standard of care for the treatment of EC is a hysterectomy surgery [2] that involves the surgical removal of the uterus. There are two types of hysterectomy surgery [8]: open or laparotomy and minimally invasive surgery (MIS). Regardless of a type, the EC surgery may entail postoperative complications [3]. These complications may vary from simple ones such as fever to severe complications that may result in postoperative morbidity or mortality. Identification of patients who are at risk of developing postoperative complications is a clinically important problem. It allows managing these patients differently before the surgery to prevent postoperative complications. Hence, the primary goal of this research is to develop a data-driven predictive model for postoperative complications to be used by a surgeon at the time of consult, prior to the hysterectomy surgery. The developed model aims to predict patients at minimal or elevated risk of developing postoperative complications. For this study, we employed data of EC patients who had hysterectomy surgery at the Cancer Center of The Ottawa Hospital (TOH). The dataset including 81 attributes associated with 644 patients who underwent a hysterectomy surgery between January 1, 2012, and March 31, 2015, were analyzed in this study. Data were collected in 3 different periods of time: 1) within 4-5 weeks prior to surgery, 2) within 24 hours after surgery, and 3) within 30 days following discharge. Ten predictive models were developed to be used prior to surgery using a set of 40 attributes. The performance of the models was assessed by F-measure, G-mean, and sensitivity as an alternative measure through a 10-fold cross-validation. A model employing PART technique with undersampling gave the best overall performance (F-measure=21.28, G-mean=55.4, and Sensitivity= 60.9) among other models prior to surgery. However, to improve the performance of the developed models, we expanded the study objectives and also developed postoperative prior to discharge. Ten predictive models were developed to be used prior to discharge using a set of 40 attributes. The PART technique with undersampling gave the best overall performance (F-measure=21.8, G-mean=56.5, and Sensitivity= 53.6) among other models prior to discharge.

As a result, it was not found a statistically significant difference between the best-performing predictive model prior to surgery and prior to discharge with regards to performance measures. Hence, shifting the time of prediction from prior to surgery to prior to discharge did not improve the performance of the predictive models. Accordingly, the PART technique with undersampling prior to surgery was selected as the best-performing predictive model to predict which patients are at a given risk category (elevated or minimal risk) of developing postoperative complications following the hysterectomy surgery.

To my late father, who died from cancer and, unfortunately, didn't stay in this world long enough to see his son’s achievement

Acknowledgement

Foremost, I would like to express my sincere gratitude to my supervisors Professor Pavel Andreev and Professor Wojtek Michalowski for the continuous support of my Master studies and research, for their patience, motivation, enthusiasm, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor for my Master's studies.

Apart from my supervisors, I would like to thank the rest of my thesis committee: Professor Herna Viktor from the Faculty of Engineering and Dr. Michael Fung-Kee-Fung from The Ottawa Hospital, for their encouragement, and insightful comments. Moreover, many thanks to Professor Tony Quon from the University of Ottawa and Professor Szymon Wilk and Professor Krystyna Napierała from the Poznan University of Technology for all their help and support regarding data mining methods and approaches.

I would also offer my grateful acknowledgment to Dr. Tien Le from the Department of Obstetrics and Gynecology at the Ottawa Hospital for contributing his experience and expertise to the development of this research. I am also grateful to NSERC for its financial support of this research. In addition, I would like to thank my friend, Farzad Firouzbakht, for providing valuable feedback throughout this research.

\Last but not least, I would like to thank my family: my mom, my sisters, and my brother for supporting me spiritually throughout carrying out this thesis and my life in general.

Table of Contents

Abstract II

Acknowledgement V

Table of Contents VI

List of Figures IX

List of Tables X

Chapter 1: Introduction 1

- 1.1 Treatment of EC 1
- 1.2 Problem Statement and Motivation 2
- 1.3 Research question and research objectives 3

Chapter 2: Literature Review 4

- 2.1 Risk factors related to postoperative complications of hysterectomy surgery for EC 4
- 2.2 Data mining 6
- 2.3 Predictive analytics using data mining methods 7
  - 2.3.1 Classification methods 7
- 2.4 Applications of predictive analytics in healthcare 7

Chapter 3: Research Design 9

- 3.1 Pre-processing of the raw data 9
- 3.2 Model development 9
  - 3.2.1. Data mining techniques 9
    - C4.5 decision trees 9
    - PART decision rules (PART**)** 10
    - Naïve Bayes (NB) 10
  - 3.2.2 Solutions for imbalanced data 10
    - Resampling methods 11
    - Algorithmic methods 12
    - BRACID algorithm 12
- 3.3 Models Evaluation 13
  - *Confusion Matrix* 13
  - 3.3.1 Sensitivity, Specificity 14
  - 3.3.2 Geometric Mean (G-mean) 14

3.3.3 F-measure 14
Chapter 4: Data Analysis and Model Development 16
4.1 Pre-processing of the raw data 16
4.1.1 Dataset description 16
4.1.2 Removing the redundant attributes 17
4.1.3 Identifying the target attributes 17
4.1.3 Analyzing the attributes for possible merging 17
4.1.4 Dealing with missing values 18
4.1.5 Identifying imbalance ratio of the target (class) attribute 18
4.2 Model Development 19
4.2.1 Analysis of the target attribute 19
4.2.1 Development of predictive models that are used prior to surgery 19
Step 1: Developing three predictive models considered as baseline models 20
Step 2: Developing three predictive models by use of undersampling 20
Step 3: Developing three predictive models by use of oversampling 20
Step 4: Developing one predictive model by use BRACID 20
4.2.2 Development of predictive models to be used before discharge 20
4.3 Testing and evaluating the predictive models 21
4.3.1 Performances of the predictive models prior to surgery 21
4.3.2 Performances of the predictive models prior to discharge 24
4.3.3 Comparison of the best predictive models prior to surgery and prior to discharge 26
Chapter 5: Discussion and Conclusion 29
5.1 Key findings 29
5.2 Limitations and future work 30
5.3 Contributions of the thesis 31
References 33
Appendixes 36
Appendix 1. 36
Appendix 2. 40
Appendix 3. 43
Appendix 4 43
Appendix 5 44

Appendix 6 44
Appendix 7 45
Appendix 8 48
Appendix 9 49

List of Figures

Figure 1: Confusion Matrix for performance evaluation 14

List of Tables

Table 1: *Summary of studies that identified risk factors associated with postoperative complications of EC surgery* 5
Table 2: *Administrative attributes* 17
Table 3: *Redundant attributes* 17
Table 4: *Summary of the attributes with missing values in both categories of the target attribute* 18
Table 5*: Number of patients with/without postoperative complications over the four years of surgery* .. 19
Table 6*: Results averaged over five runs of 10-fold cross-validation*........ 22
Table 7: *Ranking of the ten predictive models prior to surgery (based on F-measure)* 23
Table 8: *Ranking of the ten predictive models prior to surgery (based on G-mean)* 23
Table 9: *Ranking of predictive models with undersampling Oversampling prior to surgery (based on F-measure)* 24
Table 10: *Ranking of predictive models with undersampling and oversampling prior to surgery (based on G-mean)* 24
Table 11: *Results averaged over five runs of 10-fold cross-validation*........ 24
Table 12: *Ranking of the 10 predictive models prior to discharge (based on F-measure)*........ 25
Table 13: *Ranking of the 10 predictive models prior to discharge (based on G-mean)* 25
Table 14: *Ranking of predictive models with undersampling and oversampling prior to discharge (based on F-measure)* 26
Table 15: *Ranking of predictive models with undersampling and oversampling prior to discharge (based on G-mean)* 26
Table 16: *Comparison of the best performing predictive model prior to surgery with the best performing predictive model prior to discharge with regards to performance measures*........ 26
Table 17: *Ranks of the best performing model prior to surgery and prior to discharge (based on F-measure)* 27
Table 18: *Ranks of the best performing model prior to surgery and prior to discharge (based on G-mean)* 27

Chapter 1: Introduction

Cancer is one of the principal causes of death worldwide [4]. Although research on cancer has resulted in better understanding of the disease and methods of its treatment, cancer is still the leading cause of death in Canada, and it is responsible for about 30% of all deaths [5]. Statistics [6] show that there is 40% probability for an average Canadian to develop cancer in his or her lifetime. In Canada in 2015, 196,900 new cases of cancer were diagnosed, and 78,000 cancer-related deaths occurred. Cancer affects both sexes with almost the same rate of mortality. The four most common types of cancer for women in Canada are breast, lung, colorectal, and endometrial cancer[1]. This research is about the issues around surgical management of endometrial cancer (EC) that is the fourth most frequent cancer among women in developed countries [1]. Worldwide it affects an estimated 142,000 women each year and results in an about 42,000 deaths [4].

1.1 Treatment of EC

The standard of care for the treatment of EC is a hysterectomy surgery [2] that involves the surgical removal of the uterus. There are two types of hysterectomy surgery [8]:

1. Open or laparotomy surgery involving large abdominal incision [7],

2. Minimally Invasive Surgery (MIS) that includes two methods [8]:

    - standard laparoscopy, and
    - robotic-assisted surgery.

Unlike laparotomy, standard laparoscopy is done through small incisions. During this procedure, the surgeon makes the incisions of up to half an inch on patients' abdomens, and through these incisions, plastic tubes called ports are placed. Then, a camera and laparoscopic instruments are protruded through the ports which enable the surgeon to see the surgical field and to operate on uterus without a need for a large abdominal incision [9].

Several studies have demonstrated that standard laparoscopy surgery compared to laparotomy has been associated with less postoperative pain, quicker recovery, shorter hospital stays, and also resulting in lower hospitalization costs [10]. Despite these potential benefits, standard laparoscopy has not been widely utilized for the management of EC [10] because of a slow learning curve for the surgeons and the challenge of performing operations on obese patients [11]. An alternative to standard laparoscopy is robotic-assisted surgery.

Robotic-assisted surgery is performed significantly more often comparing to standard laparoscopy as reported in [12]. This procedure is similar to standard laparoscopy with a difference being that a computer-guided robot is used by surgeons to perform the procedure. The

---

[1] Canadian Cancer Statistics 2015

computer station is located in the operating room and surgeon controls the robot's movements. In fact, robotic-assisted surgery allows a surgeon to access difficult areas more easily and provides a better view of the operating theater compared to standard laparoscopy [11].

Use of robotic-assisted surgery is associated with some limiting factors such as patient’s age, obesity measured by the BMI, uterus size, cancer grade, and accessibility to the uterus [13]. So, if patient’s characteristics do not meet relevant inclusion criteria for robotic-assisted surgery, then laparotomy or standard laparoscopy are the alternatives.

1.2 Problem Statement and Motivation

Each type of EC surgery may involve postoperative complications [3]. These complications may vary from simple ones such as fever to severe complications that may contribute to postoperative morbidity or mortality. Different studies (e.g. [14]) investigated postoperative complications of hysterectomy surgery. Some of these complications are urinary tract infection, surgical site infection, bleeding, organ damage, heart attack or stroke, and venous thromboembolism.

In the study we assume that the following four types of hospital visits that occurred within first 30 days after discharge are proxy for postoperative complications associated with the EC surgery:

- readmission to hospital,
- a visit to the Emergency Room (ER),
- re-operation, and
- admission to the Intensive Care Unit (ICU).

In the charts, these visits were directly associated with prior EC surgery.

Identification of patients who are at risk of developing postoperative complications is an interesting and clinically important issue since it might allow to decrease postoperative complications rate by more targeted patients’ management prior to the surgery. Ability to do so should not only decrease the cost of EC management but also should contribute to the improved patient outcomes.

This research is an attempt to address the issue of possible postoperative complications by using a data-driven approach for developing a predictive model that should help surgeons, prior to surgery, with the identification of patients who are at elevated or minimal risk of developing postoperative complications. Because in general there are only very few patients who experience complications and are readmitted to the hospital (including ER or ICU), we decided to categorize patients into more general groups; specifically, those who had any type of re-admission and those who had none. Therefore, we assumed that patients at elevated risk of postoperative complications are those who were readmitted, and patients at minimal risk of postoperative complications are those who were not readmitted within 30 days after the EC surgery.

Identification of patients who are at risk of developing postoperative complications can also be done after the surgery and before discharge. In fact, if a predictive model prior to surgery doesn't show satisfactory performance, developing a predictive model for use prior to patient's discharge may result in better performance. Although developing such a model seems to be more likely to produce better results because of more complete data, it also has its drawbacks. In principle, a predictive model should be most helpful to a surgeon if used before surgery because results produced by such a model can be taken into account while preparing the patient for the surgery as well as during postoperative management. Moreover, such a model can help to identify those patients at elevated risk who are most likely to benefit from intensive secondary strategies and closer follow-up right after the surgery. So, shifting the time of prediction from "prior to surgery" to "prior to discharge" might postpone the introduction of secondary strategies and closer follow-up for patients at elevated risk to the time just prior to the discharge.

1.3 Research question and research objectives

The research question for this study is as follows:

For the EC patients, is it possible to predict, prior to the surgery or prior to discharge from the hospital, which ones are at a given risk category (elevated or minimal risk) of developing postoperative complications within 30 days after discharge window?

To answer this question, we consider risk prediction as a binary classification problem where two risk categories define class labels. To address the research question, we established the following objectives:

- To build predictive models prior to surgery and prior to discharge to classify patients into risk categories.
- To assess the accuracy of the resultant models.
- To identify the best-performing model.

Chapter 2: Literature Review

The purpose of this review is to summarize the existing research on predicting postoperative complications that patients may face after hysterectomy surgery, and on use of data mining for development of predictive models of postoperative complications. The review was carried out as follows:

Search process: query with the following keywords and operators were employed: (“endometrial cancer,” or “hysterectomy surgery,”) and (“postoperative complication” or “adverse event”). Another search was carried out by using the query: “prediction,” and “postoperative complication,” and “data mining.”

Databases: The following databases were used: PubMed and Scopus.

Year limit: Search results before the year 2000 were excluded.

Language: Only studies in English were included.

The search resulted in identifying 453 articles. Abstracts of these articles were screened by the researcher, and 85 relevant articles were identified as relevant. The selected 85 articles deal with at least on one of the following:

1) Risk factors associated with postoperative complications after hysterectomy surgery.

2) Prediction of postoperative complications after hysterectomy surgery.

3) Use of predictive modeling for identification of postoperative complications.

2.1 Risk factors related to postoperative complications of hysterectomy surgery for EC

One of the main streams of identified research is concerned with the identification of the risk factors associated with postoperative complications of this surgery (please see Table 1 for details).

Risk Factor	Study (year, country)
Age	[15] (2003, United States) [16] (2011, United States) [17] (2016, United States) [18] (2012, United States)
History of myocardial infarction	[18] (2015, United States)
Blood transfusion	[18] (2015, United States)
Type of Surgery	[18] (2015, United States) [19] (2013, United States) [20] (2015, United States) [21] (2004, Germany) [22] (2012, Australia)
Abdominal wall thickness	[23] (2011, United States)
Grade of tumor	[22] (2012, Australia)
Morbidity score	[22] (2012, Australia)
Obesity	[22] (2012, Australia) [21] (2004, Germany) [3] (2015, United States)
Previous radiation	[15] (2003, United States)
Previous abdominal surgery	[15] (2003, United States)
Malignancy	[15] (2003, United States)
Smoking	[16] (2013, United States)
History of vascular disease	[16] (2013, United States)
Uterine size	[23] (2011, United States)

Table 1: *Summary of studies that identified risk factors associated with postoperative complications of EC surgery*

Finan *et al.* [23] reviewed 119 EC patients who underwent robotic-assisted surgery with the aim to establish if measurements of uterine volume and a wall thickness of abdomen are important modalities to predict "difficult patients." A difficult patient in their study is the one who started with robotic-assisted surgery and subsequently procedure needed to be converted to the laparotomy. By using a regression model, the authors identified that measurements of uterine volume greater than 793 cm^3 as well as abdominal wall thickness are significant risk factors allowing to identify a difficult patient. Liang *et al.* [18] retrospectively studied 395 EC patients who underwent robotic-assisted surgery to investigate the risk of complications after the surgery. The authors used an extended length of stay (LOS) as a proxy for these complications, specifically, if standard LOS of 1 day was extended. Using multivariate regression model, the authors identified the two risk factors: age of a patient and history of myocardial infarction. Tozzi *et al.* [21] retrospectively reviewed 383 patients among whom 191 underwent laparotomy and 192 MIS (65% robotic, 35% laparoscopy). Using multivariate regression model, the authors identified the following features as the risk factors for postoperative complications: age of a patient and type of surgery. In fact, patients who underwent laparotomy were at a higher risk of developing postoperative complications. Bakkum-Gamez *et al.* [16] investigated risk factors associated with postoperative complication after laparoscopy surgery. The authors used surgical site infection (SSI) within 30 days after surgery as a proxy for the postoperative complications. Using multivariate regression model, the authors identified the following risk factors: age, BMI, American Society of Anesthesiologists (ASA) score[1] greater than 2, and smoking. Kondalsamy-Chennakesavan *et al.* [22] aimed to identify significant risk factors to predict the occurrence of postoperative complications for patients following open or laparoscopic surgery. In their study, postoperative complications are defined as any event which resulted in death, or which was life-threatening, or which resulted in significant disability within six weeks after surgery. Using multivariate regression model, the authors identified the following risk factors: co-morbidity score, BMI, grade of cancer, and hemoglobin level.

All the studies summarized in Table 1 were about hypothesis testing, and therefore they relied on regression models for identification of the risk factors. So, they were looking at the factors that contribute to the risk of postoperative complications instead of focusing on prediction of risk of postoperative complications.

In essence, statistical analysis of data usually amounts to individually testing priori hypotheses about relationships of a particular type (e.g., linear relationships) between predictors and outcomes in a data sample. If the specific statistical models and hypotheses are misspecified, then a predictor that could have significant diagnostic value for the accurate prediction of an outcome might be missed.

The focus of this research is not to construct and test a hypothesis but instead to build a data-driven model that will predict patient's membership in one of the risk categories (elevated or minimal). To do so, we will use data mining techniques to develop a predictive model without putting forward any a priori hypothesis. Thus, although finding significant risk factors using a hypothesis test is important, our study employs different data mining techniques without defining these factors. However, later the resulted attributes in the predictive model can be compared with the risk factors in the previous studies to validate their relevance. With attention to the focus of this research, data mining is shortly explained herein.

2.2 Data mining

"Data mining involves the use of data analysis tools to discover previously unknown patterns and relationships in large dataset" [24]. It is a part of knowledge discovery process, and it contributes to this process by analyzing the data and capturing patterns if any [24]. Knowledge discovery requires an iterative process of cleaning, integrating, selecting, transforming, and mining of data for identifying possible patterns [25].

---

[1] ASA score (ranging from 1 to 5) is used for assessing the fitness of patient before surgery. Score 1 refers to a normal healthy patient while score 5 refers to moribund patients who are not expected to survive without the operation.

2.3 Predictive analytics using data mining methods

Predictive analytics, a subset of data mining, is an area that has been growing in popularity in recent years. Predictive analytics, which includes data mining methods, are used to construct predictive models [26]. In the context of predictive analytics, data mining is the process of building a representative model that fits the observational data. This model serves two purposes: on the one hand, it predicts the target attribute based on the input attributes. On the other side, we can use it to understand the relationship between the target attribute and all the input attributes.

At a very fundamental level, most predictive analytics problems can be categorized into either numeric prediction or classification problems. In the case of numeric prediction, we try to predict the numeric value of a target attribute using the values assumed by the input attributes, as is done in a traditional regression modelling.

In classification or class prediction, we try to use the information from the input attributes to sort the data samples into two or more distinct classes.

2.3.1 Classification methods

The most straightforward type of a classification problem with two classes is called binary classification. In binary classification, the target attribute has only two possible values, and each object needs to be assigned one of these two categories.

Since we use data mining in this study and want to be consistent with commonly-used terms in data mining, we should clarify that instead of using the term "item," term "object," mainly used in literature to refer to items in data sets, will be used in this text. Bearing in mind that this study analyzes the data of EC patients, we can just state that each patient in this study is an object according to the above definition.

As the aim of this study is to predict membership of an object in one of two risk categories that are known, this implies that our problem belongs to the supervised learning problems where class labels (values of target attribute – in this case risk categories) are known. To this end, we will develop a predictive model that allows classifying unseen objects (new patients) into one of these two risk categories.

2.4 Applications of predictive analytics in healthcare

The goal of predictive analytics in healthcare is to derive models that can use the patient-specific information to predict the outcome of interest and to support medical decision-making [27]. Predictive analytics may be used for the development of predictive models for patient's prognosis, diagnosis, length of hospital stay, and treatment planning. For the sake of this study, we focus on the models predicting postoperative complications after surgery. Hsieh et al. [28] attempted to predict postoperative complications after endovascular aneurysm repair (EVAR). EVAR is an advanced minimally invasive surgical technology for the aortic surgery. The postoperative complication in their study is morbidity after the surgery. The authors retrospectively examined 140 charts of the patients who underwent EVAR surgery from 2000 to

2009. Patients in the study were described by 173 attributes including demographics and preoperative data, while postoperative complications were used as the target attribute. To predict postoperative morbidity, the authors developed three models: Bayesian network (BN), artificial neural network (ANN) and support vector machine (SVM). A receiver operating characteristic (ROC) curve was used to evaluate the accuracy of the models. In this series, the SVM model was selected as the best predictive model for EVAR.

Thottakkara et al. [29] retrospectively examined charts of 50,318 adult patients who underwent any inpatient surgery between 2000 and 2010. Patients in the study were described by 285 demographics, socio-economic, administrative, pharmacy, and laboratory attributes. In their study, the postoperative complication was identified as acute kidney injury (AKI) in the first seven days after surgery. To make a prediction before the surgery, they compared four predictive models: Naïve Bayes, generalized additive model (GAM), logistic regression, and SVM. After testing, they concluded that GAM and SVM showed superior performance with regards to ROC compared to the other models.

Kluska et al. [30] aimed to predict postoperative complications after radical hysterectomy surgery for treatment of cervical cancer. They retrospectively examined charts of 107 patients each described by 10 attributes including patients' preoperative features. The postoperative complication in their study was defined as injuries to large blood vessels and related organs. The authors developed 11 predictive models to predict postoperative complication after the radical hysterectomy. SVM showed the best model satisfying the quality criteria such as the area under ROC curve, sensitivity, and specificity.

Conducted literature review shows that there are no studies specifically dedicated to predicting postoperative complications of EC patients. Therefore, to the best of our knowledge, this research is the first attempt to address this gap.

Chapter 3: Research Design

The study follows a research design typical for data mining:

1.Pre-processing of the raw data.

2. Model Development.

3. Model Evaluation.

3.1 Pre-processing of the raw data

This step focused on exploring the dataset and its attributes, and it involves:

- removing of the redundant attributes,
- identifying the target (class) attribute,
- analyzing the attributes and their domains for possible merging,
- dealing with missing values, and
- defining the imbalanced ratio, if any.

3.2 Model development

To build a predictive model that can generate understandable rules that can be interpreted by physicians, we employ three techniques that are frequently considered in the clinical domain [27]: PART decision rules (PART), C4.5 decision trees, and naïve Bayes (NB). It's worth mentioning that although some other techniques (e.g., SVM and ANN) have shown superior performance in clinical problems, we do not use them for the comprehensibility reasons. Because techniques such as SVM are not able to generate rules that are untestable for physician and they can't easily interpret the results of such techniques, unlike our proposed techniques that create rules that can be easily understood by physicians.

3.2.1. Data mining techniques

C4.5 decision trees

Decision trees are one of the most intuitive and frequently used data mining techniques. C4.5 is one of the decision trees that have frequently been used in clinical problems[31]. A C4.5 decision tree [32] is easy to set up and also easy to interpret. A C4.5 decision tree is used to separate a dataset into classes belonging to the target attribute. And, C4.5 decisions trees are used when the target attribute is categorical in nature.

A C4.5 decision tree approaches the classification problem by dividing the data into subsets based on the values of the input attributes. The attributes that help to achieve the cleanest levels of such separation are considered significant in their influence on the target attribute and end up at the root and closer-to-root levels of the tree.

PART decision rules (PART)

Decision rules induction is a data mining technique of deducing if-then rules from a data set. These symbolic decision rules explain an inherent relationship between the attributes and class labels in the dataset. PART is the decision rules that have frequently been used in clinical problems [31]. PART uses the sequential covering technique to build a rule set. The sequential covering is an iterative procedure of extracting rules from the data set. The sequential covering approach attempts to find all the rules in the dataset class by class. A decision rule has two parts, an antecedent (if) and a consequent (then). An antecedent is an object found in the data. A consequent is an object that is found in combination with the antecedent. Decision rules are created by analyzing data for frequent if/then patterns and using the criteria support and confidence to identify the most important relationships. Support is an indication of how frequently the objects appear in the data. Confidence indicates the number of times the if/then statements have been found to be true.

More importantly, classification using rules provides a simple framework to identify a relationship between attributes and the class label that is not only used as a predictive model but a detailed model. Single rule learners are the simplest form of data mining model and indicate the most powerful predictor in the given set of attributes. Given the wide reachability of rules, rule induction is commonly used as a tool to express the results of data mining, even if other data mining techniques are used.

Naïve Bayes (NB)

Naive Bayes is a simple probabilistic data mining technique [33] The naïve Bayes algorithm finds its roots in statistics and probability theory. In general, this data mining technique tries to predict class labels based on attributes by best approximating the relationship between input and output variables. NB has frequently been used in clinical problems [31] The naïve Bayesian algorithm is built on Bayes' theorem and provides a probabilistic way of building a model. This technique calculates the probability for each value of the class attribute for given values of input attributes. Despite their naive design and apparently simple assumptions, naive Bayes techniques have worked quite well in many complex real-world situations. In fact, naive Bayes techniques are simple, efficient and robust to noise and irrelevant attributes [33]

3.2.2 Solutions for imbalanced data

One of the potential challenges in developing predictive models is dealing with the imbalanced dataset. Class imbalance complicates the development of predictive models because it limits the ability of a model to make accurate predictions [34]. Indeed, it has been shown that most of the data mining techniques suffer when dealing with the prediction of class-imbalanced data [34]. Proposed solutions for class imbalance can be divided into *resampling methods* and *algorithmic methods* [35]. The first group contains approaches to adjust the data by changing the distribution of objects between the classes and the second concerns revisions to the model development process to accommodate the imbalance. In other words, the algorithms methods can be applied directly to imbalanced datasets.

Resampling methods

Resampling methods are used to balance the class distribution by either increasing (oversampling) the objects in the minority class, or by decreasing (undersampling) the objects in the majority class [36, 37]. Random under and oversampling methods have received significant attention to counter the effect of imbalanced datasets and were used in various studies. In most situations, objects are removed or added until the balance between the classes is achieved.

Another resampling method is the synthetic minority oversampling technique (SMOTE) [38]. The essence of SMOTE technique is oversampling the minority class by creating synthetic objects instead of just duplicating objects from a minority class. Although SOMTE has shown better results than random oversampling and undersampling [38], it will not be used in this study. In fact, using SMOTE to create synthetic objects (in our case these are non-existing patients) will result in developing predictive models based on synthetic (artificial) patients, and it can be difficult for the clinicians to understand and interpret such models. In light of that, we use random undersampling and oversampling methods in this study.

Another resampling approach to deal with imbalanced problem is a hybrid approach represented as a combination of undersampling with oversampling. These hybrid from of resampling was not considered in this study for the following reasons:

1. Combining random undersampling with random oversampling may require defining a trade-off between undersampling and oversampling. Thus, choosing a trade-off to which the majority and minority class should be under-sampled and oversampled may be very challenging and lead to generating samples that are too strongly biased towards the minority class and leads to deteriorated specificity and overall accuracy [39] .

2. Hybrid resampling has not demonstrated a better performance in terms of sensitivity. Wilk and Stefanowski [39] evaluated the performance of three hybrid resampling approaches and compared with random oversampling in experiments on nine imbalanced datasets and demonstrated that such approaches may outperform models developed by random undersampling or oversampling in terms of overall accuracy but do not outperform them in terms of sensitivity.

Stratified random sampling, as an alternative to random resampling techniques, has not been considered in this study. Using a stratified random sampling requires meeting several conditions which are challenging to meet for the dataset employed in this study. First, every object of the majority class must me identified and classified as one, and only one subclass of the majority class. It is challenging to sorting each object of the majority class into a single stratum. This challenge is not possible to overcome in this study.

Algorithmic methods

We also consider algorithmic methods in the model development. Different algorithms have addressed the issue of imbalanced data by applying different approaches [40-43]. One of the algorithms that showed good results is called Bottom-up induction of Rules and Cases for Imbalanced Data (BRACID) [44].

BRACID algorithm

BRACID is a bottom-up rule-based data mining technique that starts from the most specific rule that can cover a single example and then generalizes this rule until it cannot be further generalized. As a result, BRACID can concentrate on the minority class and does not disregard examples of a minority class and in such sense is addressing the problem of class imbalance [45]. BRACID has shown superior performance in dealing with several imbalanced datasets [44]. This algorithm produces decision rules that are readable and comprehensible and able to better recognize the minority class than other compared data mining techniques [44]. More importantly, the classification performance of BRACID (concerning all measures) has been better than different techniques specialized for class imbalance[44]. Napierala and Stefanowski [44] evaluated the performance of BRACID in experiments on several imbalanced datasets and demonstrated that BRACID outperforms other techniques such as C4.5, PART, and K-NN. Therefore, we use BRACID to develop predictive models.

Altogether, in this study we construct 10 predictive models as follows: 3 baseline models by applying the PART, C4.5, and NB techniques to the imbalance dataset, 3 predictive models by applying the PART, C4.5, and NB techniques to the balanced dataset after resampling by random undersampling, 3 predictive models by applying the PART, C4.5, and NB techniques to the balanced dataset after resampling by random oversampling, and 1 model by applying the BRACID technique to the imbalanced dataset.

However, since the nature of this study follows a data-driven approach, there might be a possibility that predictive models derived from available (even expanded) data do not have desired predictive performance. In such case, we also developed predictive models that can be used after the surgery but before discharge. Such models were also considered in the literature [46], and their use can be considered as an attempt to flag potentially risky patients for more thorough discharge teaching and follow-ups. By extending the time that we tend to use our model from “prior to surgery” to “prior to discharge,” we hypothesized that more attributes are identified as significant attributes to use for the prediction problem. Thus, it might improve the performance of the predictive models.

3.3 Models Evaluation

Model evaluation is one of the most important steps in the procedure of developing predictive models using data mining because it gives us a clear view regarding the strength of each model and enables to compare different models based on standard performance measures.

There is a number of ways to evaluate the quality of prediction of the models. Most often it is done using 10-fold cross-validation with the measures such as F-Measure, G-mean, and sensitivity.

Cross-validation [47] is a statistical method in which data set is split into two segments: one used to train a model and the other used to validate it. The primary form of cross-validation is k-fold cross-validation in which the data is first partitioned into k equally (or nearly equally) sized segments or folds. Then, k repeats of training and validation are performed such that each time a different fold of the data is held-out for validation while the remaining k-1 folds are used for learning. Cross-validation is used to evaluate or compare classifiers as follows: in each repeat, one or more classifier use k-1 folds of data to learn one or more models, and subsequently, the learned models are used to make predictions for data retained in the validation fold. In data mining 10-fold cross-validation (k = 10) is the most common approach [48].

The models' performance should be analyzed based on a series of measures that are recommended in the literature. To this end, we use F-measure and G-mean as two primary performance measures to evaluate our models' performance. We also used sensitivity as an alternative measure. To calculate the two primary performance measures for each model, we first need to define sensitivity and specificity of a predictive model using the confusion matrix.

Confusion Matrix

A confusion matrix (Figure 1) is often built to test a predictive model, which shows the experimental results of that model. In fact, from the confusion matrix, apart from other more elaborated measures one can construct more straightforward measures concerning the recognition of the positive and negative classes as follows:

- The predicted class is positive, and the actual class is also positive → this is a "True Positive" or TP.
- The predicted class is positive, and the actual class is negative → this is a "False Positive" or FP.
- The predicted class is negative, and the actual class is positive → this is a "False Negative" or FN.
- The predicted class is negative, and the actual class is also negative → this is a "True Negative" or TN.

	Predicted positive	**Predicted negative**
True positive	TP	FN
True negative	FP	TN

Figure 1: *Confusion Matrix for performance evaluation*

3.3.1 Sensitivity, Specificity

The sensitivity of a prediction model is defined as its ability to identify actual cases, i.e., for our study correctly; it measures the proportion of patients with elevated risk of postsurgical complications which are correctly identified as such. The specificity of a prediction model is defined as its ability to identify negative cases, i.e., for our study these are patients with minimal risk of postoperative complications which are correctly identified as such. These measures are calculated as follows:

Sensitivity = $\frac{TP}{TP+FN}$ Specificity = $\frac{TN}{TN+FP}$

However, sensitivity and specificity may not be sufficient to measure the performance of a predictive model. Generally, analysts would want to balance between both sensitivity and specificity. Performance measures that try to balance between sensitivity and specificity are Geometric mean (G-mean) and F-measure.

3.3.2 Geometric Mean (G-mean)

The G-mean measure was used by Kelement et al. [49], and it is the geometric mean of the sensitivity and specificity and is used to evaluate how balanced the performance of the predictive model is. The fundamental idea of G-mean is to maximize the recognition of both minority and majority classes while keeping these accuracies balanced [50].

G-mean = $\sqrt{S_n . S_P}$

where $S_n\ (sensitivity) or (recall) = \frac{TP}{TP+FN}$ and $S_P\ (specificity) = \frac{TN}{TN+FP}$

A high value of G-mean indicates that the performance of both sensitivity and specificity are high simultaneously. In other words, the higher the G-mean, the higher the accuracy is in both positive and negative objects simultaneously.

3.3.3 F-measure

The F-measure as it was also proposed in [51] and it aggregates precision and recall:

Precision = $\frac{TP}{TP+FP}$ Recall = $\frac{TP}{TP+FN}$

Where precision is the ratio of correctly predicted positive objects to the total predicted positive objects.

$$F_\beta = (1+\beta^2) \cdot \frac{Precision\ .\ recall}{\beta^2\ .\ Precision+recall}$$

A high value of F-Measure indicates that the model performs better on the positive class and also ensures that both recall and precision are reasonably high. F-measure also depends on the β factor, which is a parameter that takes values from 0 to infinity and is used to control the influence of recall and precision separately. It can be shown that when $\beta=0$ then F-measure reduces to precision and conversely when $\beta\to\infty$ then F-measure approaches recall. To adjust the relative importance of precision versus recall, β is considered to be 1 in the studies [44, 51]

As a conclusion, because in this study we aim not only to correctly predict patients at the Minimal class but also to correctly predict patients at the Elevated class, we use G-mean and F-measure as the two main performance measures of this study. These two performance measures try to balance between the sensitivity and the specificity and are not just biased to one of the classes per se [52]. Also, using these two measures is important in the avoidance of overfitting of the Minimal class and helpful in providing more weight to patients correctly classified in the Elevated class. Hence, the model that has the highest F-measure score and G-mean compared to other models will be selected as the best performing model.

However, in the case when one model does not perform well on both F-measure and G-mean we will use sensitivity as an alternative measure to choose the best performing model. The reason why we prefer sensitivity over specificity is that sensitive aids us to determine how well the model can classify patients of the Elevated class. Finding these patents in this study is of interest to us.

Chapter 4: Data Analysis and Model Development

The study was approved by TOH Research Ethics Board. Data were collected through a retrospective chart study of EC patients who had hysterectomy surgery at the Cancer Center of the Ottawa Hospital (TOH). Specifically, we included patients who underwent the laparotomy surgery and the robotic-assisted surgery between January 1, 2012, and March 31, 2015. It should be mentioned that during this period only 34 patients underwent the laparotomy surgery with only one reported postoperative complication. So, the laparotomy patients are excluded from this study.

4.1 Pre-processing of the raw data

4.1.1 Dataset description

In total records of 644 patients (objects) were identified for the study. These patients underwent the laparotomy surgery and the robotic-assisted surgery from January 1, 2012, to March 31, 2015. Each object was described by 81 attributes. These attributes are divided into the categories of patient history including preoperative comorbidities and laboratory test results, operative attributes, and 30-day after discharge attributes for identification of postoperative complications. Values of the attributes were collected in 3 different periods of time: values of 38 attributes were collected within 4-5 weeks prior to surgery, values of 39 attributes were collected within 24 hours after surgery, and values of 4 attributes describing different types of re-admissions were collected within 30 days following discharge. A full list of all attributes and their domains is presented in Appendix 1.

Since in this study we aim to develop a predictive model that can predict which patients are at a given risk category (elevated or minimal risk) of developing postoperative complications within 30 days after discharge, we don’t use an extended LOS as a proxy of complications. The extended LOS can be used as a proxy for complications before discharge. Hence, a proxy for postoperative complications within 30 days after discharge is (1) readmission to hospital, (2) a visit to the Emergency Room (ER), (3) re-operation, or (4) Intensive Care Unit (ICU) admission.

As shown in Table 1, some studies identified risk factors that are associated with postoperative complications that EC patients may experience after the hysterectomy surgery. Finding the most significant risk factors identified in previous studies helped us to compare these risk factors with our dataset’s attributes and see if our research also includes the same attributes that were already proven to be significant risk factors associated with postoperative complications of the hysterectomy surgery. Hence, we can eventually find out those significant attributes that can be appropriately used for prediction of postoperative complications of hysterectomy surgery but are missing in the initially collected dataset of our study.

4.1.2 Removing the redundant attributes

By using descriptive statistics, we were able to find out the frequencies of the binary attributes, mean values of the numerical attributes, and missing values of all the attributes. IBM SPSS Statistics 23 was used for descriptive analysis. Results of this analysis are presented in Appendix 2. Accordingly, we explored the redundant attributes and missing values of the study.

The exploratory analysis identified 13 administrative attributes (listed in Table 2, description of attributes in Appendix 1) that were removed from the dataset.

Administrative Attributes				
Study number	MRDx, Reoperate_30day	Campus	Proc_Date	Proc_Start_Tim
OutRoom_Time	PACU_leave_Avail	ORout2PACU_Disch_Min	First_PAR2PAR13_Min	Disch_Loc
Adm_Type	MRDx_Text	Procedure_Desc		

Table 2: *Administrative attributes*

Redundant Attributes			
Deficiency Anemia	Decubitus Ulcer	Traumatic Injury Arising In Hospital	Obstetrical Complication(fetus)(Endocri Complic)
Depression	Drug Abuse	Peri Vase Dis	AIDS/HIV
Liver Disease	Lymphoma	Other Neurological Disorders	Peptic Ulcer
Paralysis	Psychoses	Lyte Problem	VTE
Weight Loss	Anesthesia Complication	Cardiac Complication	Acute_LOS

Table 3: *Redundant attributes*

We were able to identify 20 redundant attributes (listed in Table 3, description of attributes in Appendix 1). The redundant attribute is such that has a single value for all objects in the entire dataset.

4.1.3 Identifying the target attributes

The target attribute is considered “risk of postoperative complications” with two values: elevated, and minimal. The “elevated” class is assigned to those patients who had recorded value for at least one of the four proxy attributes in the charts, and the “minimal” class label was assigned to all remaining patients (i.e., those patients that did not return to a hospital during 30 days after discharge).

4.1.3 Analyzing the attributes for possible merging

The exploratory analysis helped us to identify those attributes with binary value domain that can be considered as values from a domain of a higher level attribute. For example, two attributes “Diabetes Complicated” and “Diabetes Uncomplicated” with binary values yes or no can be treated as values from the domain [complicated, uncomplicated, none] of new attribute “Diabetes.” Also, the same approach was applied to the attributes Hypertension_Complc and

Hypertension_NON_Complc which resulted in a single attribute called Hypertension with values from the domain [complicated, uncomplicated, none].

Furthermore, the binary attributes “Procedure,” “Hysterectomy approach,” and “Robotic use“ were combined into a single attribute called “Surgery type” with values from the domain [Robotic, Laparotomy].

4.1.4 Dealing with missing values

The highest number of missing values was observed for the attribute “First pain score” and “Max pain score.” And the lowest number of missing values was observed for the attributes “ASA,” “Nausea,” and “PAR score” The following table is a summary of information for the attributes with missing values in both classes.

Attribute	Overall (%)	Class: Elevated (%)	Class: Minimal (%)
ASA	5 (0.77)	0	5 (0.87)
Nausea	5 (0.77)	0	5 (0.87)
Vomit	8 (1.2)	0	8 (1.2)
First pain score	26 (4)	2 (3.1)	24 (4.1)
Max pain score	26 (4)	2 (3.1)	24 (4.1)
PAR score	5 (0.77)	0	5 (0.77)

Table 4: *Summary of the attributes with missing values in both categories of the target attribute*

Many factors can cause missing values. For example, value is not collected, value was not recorded, or patients may refuse to answer a question for whatever reason. All of these types of missing data can cause problems for the development of predictive models such as loss of efficiency, complications in handling and analyzing data, or bias resulting from differences between missing and complete data. In fact, missing values may limit the ability to draw definitive conclusions from data or could lead to incorrect inferences [53]. However, considering that prevalence of the missing values in a dataset is very low (0.3%), we decided not to impute the missing values.

Normalization was not performed as a preprocessing step since the final list of attributed included binary or ordinal data and does not require normalization [55].

4.1.5 Identifying imbalance ratio of the target (class) attribute

We can notice that there is a significant class imbalance between the “elevated” and “minimal” classes where 89.6% of patients belong to the minimal risk class. In fact, in imbalanced data, there is a minority class or a few classes which include much smaller numbers of objects than the majority class [54]. There is not a specific recommendation about the degree of an imbalance between class cardinalities. Some studies have suggested that severe imbalance ratios are 1:10, 1:100 or even greater [44]. In our data, the imbalance ratio is 1:10.

The exploratory analysis produced a dataset with the number of attributes decreased to 40 (see Appendix 4), with 21 of these attributes (see Appendix 3) associated with data items collected prior surgery. The set of all 40 attributes will be used when developing predictive models prior to discharge (after the surgery), and the set of 21 attributes will be used when developing predictive models prior to surgery.

4.2 Model Development

In this section, we first present data analysis of the target attribute, and then we describe stages of predictive models development. We developed models at two different prediction times (before surgery and before discharge).

4.2.1 Analysis of the target attribute

The data analysis revealed different frequencies of postoperative complications over the four years period when data was collected (see Table 5). To test if there is a dependence between patients' risk class allocation and years we conducted Independence chi-squared [56] test.

	Years of Surgery				
	2012	**2013**	**2014**	**2015**	**Total**
Patients at the elevated class	24	22	19	4	**69**
Patients at the minimal class	172	162	187	54	**575**
Total	**196**	**184**	**206**	**58**	**644**

Table 5: *Number of patients with/without postoperative complications over the four years of surgery*

The results (see Appendix 5) were not statistically significant (P-Value = 0.554) showing that we cannot reject null hypothesis which states that proportion of patients with minimal and elevated risk are similar (independent) across four years. Therefore we decided to ignore the longitudinal character of data and use the entire dataset for developing prediction models.

4.2.1 Development of predictive models that are used prior to surgery

Modeling implementation started with building predictive models including C4.5 decision trees, PART decision rules (PART), naïve Bayes (NB), and the BRACID technique. These methods are frequently considered in clinical problems. To implement 10 predictive models before surgery, we used the set of 21 attributes that are collected before surgery and were explored earlier (Listed in Appendix 3). We should also note that we didn't use any feature selection methods in this study because 1) The PART, C4.5, and BRACID techniques generating rules and threes have an internal mechanism which selects the most relevant attributes for the prediction problem; and 2) Both sets of attributes that we used for the prediction problems prior to surgery and prior to discharge have already a relatively small number of attributes.

To develop 10 predictive models prior to surgery, we performed the following steps:

Step 1: Developing three predictive models considered as baseline models

We developed three baseline models by applying PART, C4.5, and NB techniques to the imbalanced dataset, resulting in three predictive models that help us to assess the impact of resampling methods on the results. To do so, we used the PART and NB techniques of WEKA 3.7 where default values of learning algorithms' parameters. Also, we used the J48 algorithm of WEKA 3.7. J48 is an open source Java implementation of the C4.5 technique in the Weka 3.7. Applying these techniques to the dataset developed three predictive models considered as baseline models of this study. As an example of these three predictive models, the predictive model using C4.5 technique is presented in Appendix 6.

Step 2: Developing three predictive models by use of undersampling

We developed three predictive models by applying the PART, C4.5, and NB techniques to the balanced dataset. To balance the data we used the random undersampling. In such case, we used the filter "SpreadSubSample" in WEKA 3.7 and also selected the value of 1 for the "distributionSpread." This step resulted in three predictive models that were developed on a balanced dataset created by random undersampling. To illustrate an example of these three models, the PART technique with undersampling is presented in Appendix 7.

Step 3: Developing three predictive models by use of oversampling

We developed three predictive models by applying PART, C4.5, and NB techniques to the balanced dataset created by random oversampling. In such case, we used the filter "Resample" in WEKA and also selected the value of 1 for "BiasToUniformClass." We also put the value of 179 for the "sampleSizePercent" that oversamples the minority class (in our case it refers to the patients with the elevated class and makes it equal to the majority class. This step resulted in three predictive models that were developed on our balanced dataset resampled by random oversampling. To present an example of these three models, the C4.5 technique with oversampling is shown in Appendix 8.

Step 4: Developing one predictive model by use BRACID

The model including the BRACID techniques was implemented using Windows Powershell from Microsoft; it is a configuration management framework with its associated scripting language. Some of the rules generated with the BRACID technique to the imbalanced dataset prior to surgery are given in Appendix 9.

4.2.2 Development of predictive models to be used before discharge

To develop predictive models that are used before discharge, we followed the same steps as we used to develop predictive models prior to surgery. However, all of the predictive models were developed using the set of 40 attributes that are collected before discharge and were explored earlier (listed in Appendix4)

4.3 Testing and evaluating the predictive models

The performance of all predictive models is assessed by averaging the values of the G-mean, F-measure, and sensitivity using a stratified 10-fold cross-validation repeated 5 times to reduce the variance of results. It should be noted that the process of performing undersampling and oversampling was carried out as a part of the cross-validation. In essence, resampling can result in overfitting if it is done before cross-validation.

Finally, to be consistent with [31], we applied a non-parametric Friedman test (with α=0.05) that globally compared the performance of multiple combinations of preprocessing methods (random undersampling and random oversampling) and classification techniques over multiple data sets. So, the differences between the values of the measures were evaluated for statistical significance by Friedman test. Eventually, this process aided us to 1) Find out if there is an advantage of using preprocessing methods (random undersampling and random oversampling); and 2) to select the best performing predictive model to be used at each perdition time.

4.3.1 Performances of the predictive models prior to surgery

A summary of performance measures for the 10 predictive models that were developed using the set of 21 attributes with values collected prior to surgery is given in Table 6. As shown in this table, the values of sensitivity for the three baselines that were developed by applying the C4.5, PART, and NB technique to the imbalanced dataset are significantly low, meaning that none of these three predictive models can correctly predict patients at the elevated class with respect to postoperative complications. Having a large value of specificity, these models can only correctly predict patients at the Minimal class. Also, the values of F-measure and G-mean that are also resulted from sensitivity are considerably low for the baselines. In fact, Low values of G-mean and F-measure indicate that the balance between classification performances on the Minimal and Elevated class is poor. Indeed, it indicates the overfitting to the Minimal class and the poor degree to which the Elevated class is marginalized.

Applying the C4.5, PART, NB, and BRACID technique to the balanced dataset with undersampling and oversampling resulted in 6 predicative models with higher values of sensitivity, F-measure, and G-mean compared to those of three baselines. Although it decreased the value of specificity, it makes the avoidance of overfitting the Minimal class and underfitting the Elevated class, resulting in performing much better on the Elevated class compared to the three baselines.

Applying the BRACID technique to the imbalanced dataset resulted in a predictive model that also increased the values of sensitivity, F-measure, and G-mean compared to those of three baselines.

	Imbalanced dataset			Balanced dataset by undersampling			Balanced dataset by oversampling			Imbalanced dataset
	C4.5	**PART**	**NB**	**C4.5**	**PART**	**NB**	**C4.5**	**PART**	**NB**	**BRACID**
Sensitivity (%)	0	3.2	4.6	59.4	60.9	52.18	33.04	31.58	54.8	38.2
Specificity (%)	100	98.9	98.4	51.8	50.6	57.72	73.18	72.18	57.5	72
F-Measure (%)	0	5.6	7.88	20.9	21.28	20.64	18.54	17.38	21.5	20.1
G-mean (%)	0	17.4	21.2	55.4	55.4	54.7	49.2	47.7	56.1	49.5

Table 6: *Results averaged over five runs of 10-fold cross-validation*

To choose the best performing predictive model prior to surgery, we were consistent with our two main performance measures: G-mean and F-measure. We also used sensitivity as an alternative where we could not choose the best performing model only based on G-mean and F-measure.

To this end, Table 7 reports the results of the Friedman test based on F-measure for all the 10 predictive models. In this test, all the 10 predictive models were asked to record how well they perform with regards to F-measure on a scale of 1 to 10, with 1 being poor and 10 extremely good. A Friedman test was then carried out to see if there were differences in 10 predictive models on F-measure. The null hypothesis (i.e., that all predictive models performed equally) was rejected (P-value=0.000). According to the critical difference between ranks, the PART technique with undersampling and the NB technique with oversampling were the best performing predictive models.

Observations for G-mean are similar to those for F-measure. Obtained results presented in Table 8 emphasized that the PART technique with undersampling and the NB technique with oversampling were the best performing predictive models (P-value=0.000).

To select the best performing predictive model between the PART technique with undersampling and the NB technique with oversampling we used sensitivity an alternative measure. Since the sensitivity of the PART technique with undersampling is higher than that of the NB technique with oversampling, we considered the PART technique with undersampling as the best performing predictive model prior to surgery.

Predictive Model		Mean Rank
Imbalanced dataset	C4.5	1.00
	PART	2.20
	NB	2.80
Balanced dataset by Undersampling	C4.5	7.70
	PART	**8.30**
	NB	7.20
Balanced dataset by Oversampling	C4.5	5.50
	PART	4.40
	NB	**9.10**
Imbalanced dataset	BRACID	6.80

Table 7: *Ranking of the ten predictive models prior to surgery (based on F-measure)*

Predictive Model		Mean Rank
Imbalanced dataset	C4.5	1.00
	PART	2.20
	NB	2.80
Balanced dataset by Undersampling	C4.5	8.20
	PART	**8.60**
	NB	8.00
Balanced dataset by Oversampling	C4.5	5.60
	PART	4.20
	NB	**8.80**
Imbalanced dataset	BRACID	5.60

Table 8: *Ranking of the ten predictive models prior to surgery (based on G-mean)*

Based on the results of Table 7 for F-measure and according to the critical difference between ranks, both random undersampling and oversampling improved the F-measure of the baseline models. Observations for G-mean as presented in Table 8 also confirm that both random undersampling and oversampling improved the G-mean of the baseline models.

To compare undersampling with oversampling in regards to F-measure and G-mean respectively, we conducted the Friedman test where the null hypothesis (i.e., that all predictive models developed by undersampling performed equally to the predictive models developed by oversampling) was rejected (P-value=0.021 and 0.006 respectively) As shown in Table 9 and Table 10, The C4.5 and PART techniques with undersampling achieved higher rank compared to the C4.5 and PART techniques with oversampling. But, the NB technique with undersampling achieved lower rank compared to NB technique with oversampling.

Predictive Model		Mean Rank
Balanced dataset by Undersampling	C4.5	3.90
	PART	4.50
	NB	3.60
Balanced dataset by Oversampling	C4.5	2.50
	PART	1.40
	NB	5.10

Table 9: *Ranking of predictive models with undersampling Oversampling prior to surgery (based on F-measure)*

Predictive Model		Mean Rank
Balanced dataset by Undersampling	C4.5	4.20
	PART	4.60
	NB	4.20
Balanced dataset by Oversampling	C4.5	2.00
	PART	1.20
	NB	4.80

Table 10: *Ranking of predictive models with undersampling and oversampling prior to surgery (based on G-mean)*

4.3.2 Performances of the predictive models prior to discharge

A summary of performance measures for the 10 predictive models that were developed using the set of 40 attributes with values collected prior to discharge is given in Table 11. As shown in this table, the results are very similar to what we already noticed in the predictive models prior to surgery.

	Imbalanced dataset			**Balanced dataset by undersampling**			**Balanced dataset by oversampling**			**Imbalanced dataset**
	C4.5	**PART**	**NB**	**Un + C4.5**	**Un + PART**	**Un + NB**	**Ov + C4.5**	**Ov + PART**	**Ov + NB**	**BRACID**
Sensitivity (%)	0.84	11.86	17.4	52.74	53.6	38.84	24.62	11.86	17.8	24.5
Specificity (%)	98.76	93.5	80.5	61.46	59.6	78.5	81.2	93.5	94.8	87.8
F-Measure (%)	1.52	14.4	21.6	22.24	21.8	24.42	17.5	14.4	21.62	20.7
G-mean (%)	7.1	33.1	36.4	56.8	56.5	55.2	44.7	33.1	40.6	41.5

Table 11: *Results averaged over five runs of 10-fold cross-validation*

Table 12 reports the results of the Friedman test based on F-measure for all the10 predictive models. The null hypothesis (i.e., that all predictive models performed equally) was rejected (P-value=0.000). According to the critical difference between ranks, the NB technique with undersampling was the best performing predictive model.

Table 13 reports the results of the Friedman test based on G-mean for all the 10 predictive models. The null hypothesis (i.e., that all predictive models performed equally) was rejected (P-value=0.000). According to the critical difference between ranks presented in Table 13 for G-mean the PART and C4.5 techniques with undersampling were the best performing predictive models.

To select the best predictive model among the NB technique with undersampling, the PART technique with undersampling, and the C4.5 technique with undersampling we used sensitivity as an alternative measure. Since the PART technique with undersampling has a higher sensitivity compared to those of the NB and C4.5 techniques with undersampling, the PART technique with undersampling is considered as the best performing predictive model prior to discharge.

Predictive Model		Mean Rank
Imbalanced dataset	C4.5	1.00
	PART	2.50
	NB	6.90
Balanced dataset by Undersampling	C4.5	8.00
	PART	8.00
	NB	**9.40**
Balanced dataset by Oversampling	C4.5	4.00
	PART	2.50
	NB	6.90
Imbalanced dataset	BRACID	5.80

Table 12: *Ranking of the 10 predictive models prior to discharge (based on F-measure)*

Predictive Model		Mean Rank
Imbalanced dataset	C4.5	1.00
	PART	2.70
	NB	4.20
Balanced dataset by Undersampling	C4.5	**9.20**
	PART	**9.20**
	NB	8.60
Balanced dataset by Oversampling	C4.5	7.00
	PART	2.70
	NB	5.00
Imbalanced dataset	BRACID	5.40

Table 13: *Ranking of the 10 predictive models prior to discharge (based on G-mean)*

Based on the results of Table 12 for F-measure and according to the critical difference between ranks, both random undersampling and oversampling improved the F-measure of the baseline models significantly except where the NB technique with oversampling achieved the same rank of the NB baseline model with regard to F-measure. Observations for G-mean as presented in Table 13 also confirm that both random undersampling and oversampling improved the G-mean of the baseline models significantly.

To compare undersampling with oversampling in regards to F-measure and G-mean respectively, we conducted the Friedman test where the null hypothesis (i.e., that all predictive models developed by undersampling performed equally to the predictive models developed by oversampling) was rejected (P-value=0.001 and 0.000 respectively) As shown in Table 14 and Table 15, all the predictive models developed by undersampling achieved higher rank compared to the predictive models that were developed by oversampling.

Predictive Model		Mean Rank
Balanced dataset by Undersampling	C4.5	4.40
	PART	4.20
	NB	5.60
Balanced dataset by Oversampling	C4.5	2.00
	PART	1.00
	NB	3.80

Table 14: *Ranking of predictive models with undersampling and oversampling prior to discharge (based on F-measure)*

Predictive Model		Mean Rank
Balanced dataset by Undersampling	C4.5	5.20
	PART	5.20
	NB	4.60
Balanced dataset by Oversampling	C4.5	3.00
	PART	1.00
	NB	2.00

Table 15: *Ranking of predictive models with undersampling and oversampling prior to discharge (based on G-mean)*

4.3.3 Comparison of the best predictive models prior to surgery and prior to discharge

Extending the time that we tend to use our predictive model from “prior to surgery” to “prior to discharge” resulted in more attributes that are used to develop predictive models. However, to study if such a time extension improved the predictive model based on performance measures we need to compare performance measures of the best performing predictive model prior to surgery with those of the best performing model prior to discharge as shown in Table 16.

	Time of Prediction	
	Prior to surgery	**Prior to discharge**
The best-performing predictive model	The PART technique with undersampling	The PART technique with undersampling
Sensitivity (%)	**60.9**	53.6
Specificity (%)	50.6	59.6
F-Measure (%)	21.28	**21.8**
G-mean (%)	55.4	**56.5**

Table 16: *Comparison of the best performing predictive model prior to surgery with the best performing predictive model prior to discharge with regards to performance measures*

To this end, we conducted the Friedman test (Ranks are presented in Table 17 and 18) to compare G-mean and F-measure of the PART technique with undersampling prior to surgery with the G-mean and F-measure of the PART technique with undersampling prior to discharge.

Predictive Model	Mean Rank
The PART technique with Undersampling prior to surgery	1.40
The PART technique with Undersampling prior to surgery	1.60

Table 17: *Ranks of the best performing model prior to surgery and prior to discharge (based on F-measure)*

Predictive Model	Mean Rank
The PART technique with Undersampling prior to surgery	1.20
The PART technique with Undersampling prior to surgery	1.80

Table 18: *Ranks of the best performing model prior to surgery and prior to discharge (based on G-mean)*

The result of the Friedman test based on G-mean (p-value=0.180) and F-measure (p-value= 0.655) proved that the null hypothesis (i.e., that the PART technique with undersampling prior to surgery performed equally to the NB technique with undersampling prior to discharge) was not rejected. That means the best performing predictive model prior to discharge did not perform statistically better than the best performing predictive model prior to surgery. In other words, extending the time of prediction from “prior to surgery” to “prior to discharge” does not improve classification of patients into risk categories regarding possible postoperative complications. Hence, more attributes that were collected after surgery and used in our analysis to develop predictive models prior to discharge are not statistically significant predictors with regards to classification of patients into a risk category.

While analyzing the results, we also noticed that the performance measures of all predictive models are not significantly high. For example, the BRACID technique, unlike what we found out in the literature, did not perform well in this study. It could shed light on the fact that nature of imbalanced data sets (characteristics of the minority class distribution) influences on the classification performance of an algorithm.

Current studies [57] on imbalanced data difficulty factors have shown that objects of an imbalanced dataset are safe or unsafe objects – borderline, rare and outliers. So, there are four types of objects in the minority class of an imbalance dataset; safe, borderline, rare, and outlier. Safe objects are those located in the homogenous regions populated by the objects from one class only. These objects are considered to be easier, compared to other three types, to learn by a predictive model. The borderline objects refer to those objects located in the regions around decision boundary between classes. The rare objects are assigned to those that are isolated pairs or triples of minority class objects, located in the majority class region, which are far from the decision boundary, so they are not borderline objects. Outlier objects refer to those singular objects of the minority class that are located in the majority class region.

It is worth mentioning that collecting information about local characteristics of the minority class and distinguishing between safe and unsafe- borderline, rare and outlier objectives are useful to differentiate the performance of predictive models. Napierala and Stefanowski [57] showed that safe datasets are easy to learn for all the predictive models. However, Datasets with a lot of borderline objects are difficult. And datasets with rare and especially outlier objects are extremely difficult to recognize for all predictive models.

To understand the characteristics of the minority class of the data, we carried out an analysis in Weka 3.7 to explore the distribution of the minority class of our imbalanced dataset. This analysis revealed that out of all the objects in the minority class: 1 (1.5%) is safe, 9 (13.0%) are borderline, 17 (24.6%) are rare and 42 (60.9%) are outlier.

Because 59 (85.5%) objects of the minority class are just rare and outlier objects, we can confirm that this dataset is extremely difficult for all the predictive models including the BRACID technique to learn. This result is consistent with what Napierala and Stefanowski showed in [57] where they applied seven different predictive models to 26 datasets. They noticed that when there is a lot of rare and outlier objects in the minority class, it is impossible to recognize more than 30% of the minority objects (for some data even no objects are correctly classified).

Also, they showed that when the dataset is imbalanced with mostly safe objects, a PART technique with oversampling or undersampling performs satisfactorily with regards to performance measures. However, for all the imbalanced datasets with so many rare and outlier objects, the PART technique with oversampling or undersampling didn’t achieve a sensitivity higher than 52% which is also consistent with our results.

Chapter 5: Discussion and Conclusion

In this chapter, we summarize the major findings of our research. Then the limitations of this research are presented future works concerning the constraints of the study. Afterwards, we will discuss the contributions of this research.

5.1 Key findings

The results of this study successfully provide support for its two objectives which include: 1. Developing predictive models prior to surgery and prior to discharge to classify patients into risk categories.2. Evaluating the accuracy of the resultant models to select a best-performing model between both prediction times used in the analysis (prior to surgery and prior to discharge).

It is worthy to note that this research also showed that, for both prior to surgery and prior to discharge models, both resampling methods including random undersampling and oversampling improved the performance of the baseline models significantly. So, using any type of resampling methods will assist us to more accurately classify which patients are at a given risk category (elevated or minimal risk) of developing postoperative complications within 30 days after discharge

Comparing random oversampling with undersampling among all predictive models developed prior to surgery and prior to discharge indicates that the predictive models developed by random undersampling (except one) performs better compared to the predictive models developed by random oversampling.

Also, the findings of this study indicate that the BRACID technique, for both prior to surgery and prior to discharge models, did not outperform the predictive models developed by random undersampling. That reveals the fact that the BRACID technique performance also depends on the difficulty of the minority class distribution and such a technique does not outperform predictive models developed by undersampling, where there are a lot of rare or outlier objects in the minority class.

Finally, considering the possible lack of data that was required for developing predictive models prior to surgery, we also reconsidered the time of prediction. However, the results of this study reveal that shifting the time of surgery from “prior to surgery” to “prior to discharge” did not significantly improve the performance of the best performing predictive model prior to surgery compared to the best predictive model prior to discharge. In other words, attributes that are collected after surgery are not significant predictors with regards to a risk category (elevated or minimal risk) of postoperative complications. The sensitivity value of the best performing model of the study is 60.9 % which is relatively low, resulting in relatively low values of F-measure and G-mean. Hence, future study should address the issue of the relatively low-performance measures by developing predictive models that include risk factors that were not included in this study and can increase the performance measures.

In conclusion, the PART technique with undersampling prior to surgery is selected as the best performing predictive model in this study to classify patients into two risk categories with respect to postoperative complications. The rules generated by this predictive model are provided in Appendix 7.

5.2 Limitations and future work

To identify the specific type of the complication, one needs detailed postoperative admission data, and it was not available for this research. In fact, the data included only information about different types of return visits to the hospital that occurred within 30 days of the surgery. Although this information did not point out to a particular type of postoperative complication, it was sufficient to identify some problems that patient experienced. So, future research can address this limitation by collecting more postoperative data regarding a particular type of complication to develop a predictive model that can predict what certain type of complication a patient my encounter within 30 days of hysterectomy surgery.

In our research, to develop predictive models prior to surgery and prior to discharge, we used two sets of attributes that did not include all the significant attributes that were previously identified as risk factors associated with postoperative complications of EC surgery in the literature (given in Table 2). In essence, the two sets of attributes used in the study are "known knowns," referring to only some of those attributes that are relevant for the prediction problem in literature and we have access to them. However, there are known unknowns, referring to those four attributes that are also relevant for the prediction problem in literature but we could not have access to them and use them for developing models.

Lack of a number of relevant attributes for the prediction problem can lead to predictive models with poor performance. In our case, we believe that the relatively small rate of performance measures of our selected model may be resulted due to known unknowns and also unknown unknowns.

To address the issue of not having enough relevant attributes for the prediction problem further research can be done to develop predictive models with considering more attributes that reported in the literature as risk factors but not available in the initial dataset and therefore we were not able to use them in the analysis. Some of the known risk factors that were not in the initial dataset are *Tumor Grade*, *Prior Abdominal Surgery, Uterine size (>793*cm^3*), and Blood Transfused*.

Additional limitations are that this study considered only two types of hysterectomy surgery-Robotic-assisted surgery, and laparotomy. As a whole, the result of this study will not support laparoscopy patients undergoing laparoscopy for the management of her EC. In fact, our predictive model is not able to predict if laparoscopy patients are a risky patient with respect to postoperative complications.

To deal with this limitation future research should also study laparoscopy patients undergoing a hysterectomy surgery and develop predictive models that can predict if laparoscopy patients are a risky patient with respect to postoperative complications.

Lack of information about surgeon's experience might be another limitation of our study. Surgeon's experience may be hypothesized as a significant predictor with regards to the risk of postoperative complications; hence, future research can investigate if such an attribute is statistically significant for prediction of the risk categories. If so, such attribute can improve the performance of the predictive models.

Last but not least, the target attribute (risk of postoperative complications) in this study has two classes- Elevated and minimal. Although the predictive model can determine which patient belongs to the "elevated" class or "minimal," such a model is not able to identify to which proxy attributes (readmission to hospital, a visit to the ER, re-operation, admission to the ICU) a patient exactly belongs to. As a result, future research may consider all these proxies as classes of the target attribute and attempt to develop a predictive model to classify patients into five risk categories with the following classes:

- minimal,
- readmission to hospital,
- a visit to the Emergency Room (ER),
- re-operation, and
- admission to the Intensive Care Unit (ICU).

Such a model will provide more information about the type of postoperative complication that a patient may encounter after hysterectomy surgery.

5.3 Contributions of the thesis

In this research, we have developed a PART technique with undersampling to be used prior to the hysterectomy surgery. Such a model can predict which patients are at a given risk category (elevated or minimal risk) with respect to postoperative complications following the hysterectomy surgery.

While the predictive model of this study was developed based on the data from TOH (Ottawa, ON), the rules generated by the model can be valid for other Canadian hospitals. The patient cohorts of other Canadian hospitals are not significantly different from the population attending TOH, they all have relatively same demographic, preoperative, and postoperative attributes, and the selection criteria for choosing the surgery type are similar to those of TOH, the surgical procedure and the quality of it are similar across hospitals hence the results of the study have external validity.

Such a model can be most helpful to physicians as results produced by such a model can be taken into account while preparing the patient for the surgery and during postoperative management.

Indeed, identification of the elevated/minimal-risk patients for the hysterectomy surgery can be performed by using the PART technique with undersampling prior to hysterectomy surgery. As a result, it allows for the follow-up and treatment strategies as the additional management methods in addition to the hysterectomy surgery per se.

In practice, such a model can be integrated into hospital information systems, enabling integration of a risk prediction as a clinical decision support tool for outpatient care. Thus, it can practically provide an assessment tool for therapeutic interventions based on the risk of postoperative complications.

Furthermore, identifying elevated/minimal-risk patients assists in allowing better–informed counselling for patients and families and a more thorough approach to earlier versus later follow-up after hospital discharge.

Moreover, this study fills the gap related to lack of predictive abilities regarding the risk stratification of the EC patients. Finally, it contributes to the predictive analytics field by exploring ways of dealing with clinical data that are sparse and represent some computational challenges.

References

1. Jiang, T., L. Huang, and S. Zhang, *Preoperative serum CA125: a useful marker for surgical management of endometrial cancer.* BMC cancer, 2015. **15**(1): p. 1.
2. Seamon, L.G., et al., *Minimally invasive comprehensive surgical staging for endometrial cancer: robotics or laparoscopy?* Gynecologic oncology, 2009. **113**(1): p. 36-41.
3. Mahdi, H., et al., *The impact of obesity on the 30-day morbidity and mortality after surgery for endometrial cancer.* Journal of minimally invasive gynecology, 2015. **22**(1): p. 94-102.
4. Zolbanin, H.M., D. Delen, and A.H. Zadeh, *Predicting overall survivability in comorbidity of cancers: A data mining approach.* Decision Support Systems, 2015. **74**: p. 150-161.
5. Organization, W.H., *International statistical classification of diseases and health related problems (The) ICD-10*. 2004, World Health Organization.
6. Canadian Cancer Society, S.C., P.T.C.R. Public Health Agency of Canada, and cancer.ca/statistics, *Special topic: Predictions of the future burden*

of cancer in Canada. Caanadian Cancer Society, 2015.

7. Magrina, J.F., et al., *Robotic radical hysterectomy: comparison with laparoscopy and laparotomy.* Gynecologic oncology, 2008. **109**(1): p. 86-91.
8. Ramos, A. and A.N. Fader, *Minimally Invasive Surgery in Gynecology: Underutilized?* Current Obstetrics and Gynecology Reports, 2015. **4**(3): p. 152-158.
9. Reich, H., J. DeCAPRIO, and F. McGLYNN, *Laparoscopic hysterectomy.* Journal of Gynecologic Surgery, 1989. **5**(2): p. 213-216.
10. Gehrig, P.A., et al., *What is the optimal minimally invasive surgical procedure for endometrial cancer staging in the obese and morbidly obese woman?* Gynecologic oncology, 2008. **111**(1): p. 41-45.
11. Bandera, C.A. and J.F. Magrina, *Robotic surgery in gynecologic oncology.* Current Opinion in Obstetrics and Gynecology, 2009. **21**(1): p. 25-30.
12. Magrina, J.F. and V.L. Zanagnolo, *Robotic surgery for cervical cancer.* Yonsei medical journal, 2008. **49**(6): p. 879-885.
13. Obstetricians, A.C.o. and Gynecologists, *Choosing the route of hysterectomy for benign disease.* ACOG Committee Opinion, 2009(444): p. 1156-8.
14. Clarke-Pearson, D.L. and E.J. Geller, *Complications of hysterectomy.* Obstetrics & Gynecology, 2013. **121**(3): p. 654-673.
15. Chi, D.S., et al., *Ten-year experience with laparoscopy on a gynecologic oncology service: analysis of risk factors for complications and conversion to laparotomy.* American journal of obstetrics and gynecology, 2004. **191**(4): p. 1138-1145.
16. Bakkum-Gamez, J.N., et al., *Predictors and costs of surgical site infections in patients with endometrial cancer.* Gynecologic oncology, 2013. **130**(1): p. 100-106.
17. Mahdi, H., D. Lockhart, and K.A. Maurer, *Impact of age on 30-day mortality and morbidity in patients undergoing surgery for endometrial cancer.* Gynecologic oncology, 2015. **137**(1): p. 106-111.
18. Liang, M.I., et al., *Predicting Inpatient Stay Lasting 2 Midnights or Longer After Robotic Surgery for Endometrial Cancer.* Journal of minimally invasive gynecology, 2015. **22**(4): p. 583-589.
19. Kumar, S., et al., *Risk of postoperative venous thromboembolism after minimally invasive surgery for endometrial and cervical cancer is low: a multi-institutional study.* Gynecologic oncology, 2013. **130**(1): p. 207-212.

20. Dowdy, S.C., et al., *Factors predictive of postoperative morbidity and cost in patients with endometrial cancer.* Obstetrics & Gynecology, 2012. **120**(6): p. 1419-1427.
21. Tozzi, R., et al., *Analysis of morbidity in patients with endometrial cancer: is there a commitment to offer laparoscopy?* Gynecologic oncology, 2005. **97**(1): p. 4-9.
22. Kondalsamy-Chennakesavan, S., et al., *Risk factors to predict the incidence of surgical adverse events following open or laparoscopic surgery for apparent early stage endometrial cancer: results from a randomised controlled trial.* European Journal of Cancer, 2012. **48**(14): p. 2155-2162.
23. Finan, M.A., et al., *Magnetic resonance or computerized tomography imaging to predict difficulty of robotic surgery for endometrial cancer.* Journal of Robotic Surgery, 2012. **6**(2): p. 131-137.
24. Phyu, T.N. *Survey of classification techniques in data mining*. in *Proceedings of the International MultiConference of Engineers and Computer Scientists*. 2009.
25. Paolo, G., *Applied data mining: statistical methods for business and industry.* 2003.
26. Siegel, E., *Predictive analytics: The power to predict who will click, buy, lie, or die*. 2013: John Wiley & Sons.
27. Bellazzi, R. and B. Zupan, *Predictive data mining in clinical medicine: current issues and guidelines.* International journal of medical informatics, 2008. **77**(2): p. 81-97.
28. Hsieh, N.-C., et al. *Using intelligence techniques to predict postoperative morbidity of endovascular aneurysm repair*. in *Asian Conference on Intelligent Information and Database Systems*. 2011. Springer.
29. Thottakkara, P., et al., *Application of Machine Learning Techniques to High-Dimensional Clinical Data to Forecast Postoperative Complications.* PloS one, 2016. **11**(5): p. e0155705.
30. Kluska, J., M. Kusy, and B. Obrzut. *Prediction of radical hysterectomy complications for cervical cancer using computational intelligence methods*. in *International Conference on Artificial Intelligence and Soft Computing*. 2012. Springer.
31. Wilk, S., et al., *Application of Preprocessing Methods to Imbalanced Clinical Data: An Experimental Study*, in *Information Technologies in Medicine*. 2016, Springer. p. 503-515.
32. Quinlan, J.R., *C4. 5: Programming for machine learning.* Morgan Kauffmann, 1993. **38**.
33. Rish, I. *An empirical study of the naive Bayes classifier*. in *IJCAI 2001 workshop on empirical methods in artificial intelligence*. 2001. IBM.
34. Sun, Y., A.K. Wong, and M.S. Kamel, *Classification of imbalanced data: a review.* International Journal of Pattern Recognition and Artificial Intelligence, 2009. **23**(04): p. 687-719.
35. Garcı, S., et al., *Evolutionary-based selection of generalized instances for imbalanced classification.* Knowledge-Based Systems, 2012. **25**(1): p. 3-12.
36. Lin, S.-J., C. Chang, and M.-F. Hsu, *Multiple extreme learning machines for a two-class imbalance corporate life cycle prediction.* Knowledge-Based Systems, 2013. **39**: p. 214-223.
37. Nunes, C., et al. *Class Imbalance in the Prediction of Dementia from Neuropsychological Data*. in *Portuguese Conference on Artificial Intelligence*. 2013. Springer.
38. Chawla, N.V., et al., *SMOTE: synthetic minority over-sampling technique.* Journal of artificial intelligence research, 2002. **16**: p. 321-357.
39. Stefanowski, J. and S. Wilk. *Selective pre-processing of imbalanced data for improving classification performance*. in *International Conference on Data Warehousing and Knowledge Discovery*. 2008. Springer.
40. García, S., A. Fernández, and F. Herrera, *Enhancing the effectiveness and interpretability of decision tree and rule induction classifiers with evolutionary training set selection over imbalanced problems.* Applied Soft Computing, 2009. **9**(4): p. 1304-1314.

41. Kubat, M. and S. Matwin. *Addressing the curse of imbalanced training sets: one-sided selection*. in *ICML*. 1997. Nashville, USA.
42. Milaré, C.R., G.E. Batista, and A.C. Carvalho, *A hybrid approach to learn with imbalanced classes using evolutionary algorithms.* Logic Journal of IGPL, 2010: p. jzq027.
43. Nguyen, C.H. and T.B. Ho. *An imbalanced data rule learner*. in *European Conference on Principles of Data Mining and Knowledge Discovery*. 2005. Springer.
44. Napierala, K. and J. Stefanowski, *BRACID: a comprehensive approach to learning rules from imbalanced data.* Journal of Intelligent Information Systems, 2012. **39**(2): p. 335-373.
45. Berthold, M. and D.J. Hand, *Intelligent data analysis: an introduction*. 2003: Springer Science & Business Media.
46. Zhang, W., et al., *A Clinical Prediction Model for Postcardiac Surgery Atrial Fibrillation in an Asian Population.* Anesthesia and analgesia, 2016.
47. Witten, I.H. and E. Frank, *Data Mining: Practical machine learning tools and techniques*. 2005: Morgan Kaufmann.
48. Flach, P., *Machine learning: the art and science of algorithms that make sense of data*. 2012: Cambridge University Press.
49. Klement, W., et al., *Predicting the need for CT imaging in children with minor head injury using an ensemble of Naive Bayes classifiers.* Artificial intelligence in medicine, 2012. **54**(3): p. 163-170.
50. Kubat, M., R. Holte, and S. Matwin. *Learning when negative examples abound*. in *European Conference on Machine Learning*. 1997. Springer.
51. He, H. and E.A. Garcia, *Learning from imbalanced data.* IEEE Transactions on knowledge and data engineering, 2009. **21**(9): p. 1263-1284.
52. Gu, Q., L. Zhu, and Z. Cai, *Evaluation measures of the classification performance of imbalanced data sets.* Computational intelligence and intelligent systems, 2009: p. 461-471.
53. Kaiser, J., *Dealing with missing values in data.* Journal of systems integration, 2014. **5**(1): p. 42.
54. Longadge, R. and S. Dongre, *Class imbalance problem in data mining review.* arXiv preprint arXiv:1305.1707, 2013.
55. Witten, I.H., et al., *Data Mining: Practical machine learning tools and techniques*. 2016: Morgan Kaufmann.
56. Cohen, L., L. Manion, and K. Morrison, *Research methods in education*. 2013: Routledge.
57. Napierala, K. and J. Stefanowski, *Types of minority class examples and their influence on learning classifiers from imbalanced data.* Journal of Intelligent Information Systems, 2016. **46**(3): p. 563-597.
58. Freeman, E.A. and G.G. Moisen, *A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa.* Ecological Modelling, 2008. **217**(1-2): p. 48-58.
59. Kazem, M.A., *Predictive models in cancer management: A guide for clinicians.* The Surgeon, 2017. **15**(2): p. 93-97.

Appendixes

Appendix 1.

A full list of all 81 attributes and their domains initially collected for this study

Name	Description	Time of Collection	Value domain
Age at admission	Age of patients	Before surgery	Numerical (years)
(Body Mass Index) BMI > 40	Weight in kilograms divided by height in meters	Before Surgery	Binary (Yes, No)
ASA score	A score for assessing the fitness of patient before surgery. Score 1 refers to an average healthy patient while score 5 relates to moribund patients who are not expected to survive without the operation	Before Surgery	Integer, ranging from 1 to 5
Elixhauser Comorbidity Score	A score for measuring patient's comorbidity. The lower the score, the less comorbid a patient	Before Surgery	Integer, ranging from 1 to 20
Alcohol abuse	Addicted to alcohol or heavy drinker	Before Surgery	Binary (Yes, No)
Study number	A number assigned to each patient	Before Surgery	Numerical
Campus	Campus of surgery	Before Surgery	Binary (Yes, No)
Deficiency Anemia	Deficiency of red blood cells or hemoglobin in the blood	Before Surgery	Binary (Yes, No)
Cardiac Arrhythmia	Irregular heartbeat	Before Surgery	Binary (Yes, No)
Blood Loss Anemia	Decrease in some red blood cells	Before Surgery	Binary (Yes, No)
MRDx	Most responsible diagnosis	Before Surgery	Integer, ranging from 1 to 4
Congestive Heart Failure	Inability of the heart to keep up with the demands on it	Before Surgery	Binary (Yes, No)
Chronic Pulmonary Disease	A lung disease characterized by chronic obstruction of lung airflow	Before Surgery	Binary (Yes, No)
Rheumatoid Arthritis/collagen	Inflammation in the joints	Before Surgery	Binary (Yes, No)
Metastatic Cancer	Cancer that has spread from the place where it first started to another location in the body	Before Surgery	Binary (Yes, No)
Solid Tumor without Metastasis	Cancer cells have not spread in body	Before Surgery	Binary (Yes, No)
Adm_Type	Type of admission	Before Surgery	Binary (Yes, No)
Coagulopathy	A condition in which the blood's ability to coagulate (form clots) is impaired	Before Surgery	Binary (Yes, No)
Depression		Before Surgery	Binary (Yes, No)
Diabetes, Complicated	Diabetes associated with end organ damage such as peripheral neuropathy	Before Surgery	Binary (Yes, No)

Diabetes, Uncomplicated	Diabetes without any of the complicated conditions	Before Surgery	Binary (Yes, No)
Drug Abuse	The habitual taking of addictive or illegal drugs	Before Surgery	Binary (Yes, No)
Fluid and Electrolyte Disorders	An imbalance of minerals in the body. For the body to function properly	Before Surgery	Binary (Yes, No)
AIDS/HIV	If patients are diagnosed with AIDS	Before Surgery	Binary (Yes, No)
Hypertension, Complicated	Hypertension associated with severe conditions	Before Surgery	Binary (Yes, No)
Hypertension, Uncomplicated	Hypertension associated with uncomplicated conditions	Before Surgery	Binary (Yes, No)
Hypothyroidism	Abnormally low activity of the thyroid gland	Before Surgery	Binary (Yes, No)
Liver Disease	If patients are diagnosed with any liver disease	Before Surgery	Binary (Yes, No)
Lymphoma	Cancer of the lymph nodes	Before Surgery	Binary (Yes, No)
Other Neurological Disorders	Such as confusion, pain and altered levels of consciousness	Before Surgery	Binary (Yes, No)
Obesity	If patient is diagnosed with obesity	Before Surgery	Binary (Yes, No)
Peptic Ulcer Disease excluding bleeding	A break in the lining of the stomach	Before Surgery	Binary (Yes, No)
Peripheral Vascular Disorders	A narrowing of the arteries other than those that supply the heart or the brain	Before Surgery	Binary (Yes, No)
Paralysis	The loss of the ability to move (and sometimes to feel anything) in part or most of the body	Before Surgery	Binary (Yes, No)
Psychoses	If patient is diagnosed with any psychoses	Before Surgery	Binary (Yes, No)
Pulmonary Circulation Disorders	Affecting the blood vessels leading to or from the lungs	Before Surgery	Binary (Yes, No)
Renal Failure	A medical condition in which the kidneys fail to filter waste products from the blood adequately	Before Surgery	Binary (Yes, No)
Valvular Disease	Any damage to or a defect in one of the four heart valves	Before Surgery	Binary (Yes, No)
Weight Loss	How much weight a patient has lost	Before Surgery	Numerical (lb.)
Robot-Assisted	If the surgery is Robotic or not	During Surgery	Binary (Yes, No)
Hysterectomy Type	Robotic, Standard laparoscopy, Laparotomy	During Surgery	Integer ranging from 1 to 3
Procedure	How If it is an MIS	During Surgery	Binary (Yes, No)
Proc_Date	Date of Surgery	During Surgery	Numerical
Procedure Time	Time of Surgery	During Surgery	Numerical
Procedure Length in Minutes	Duration of surgery	During Surgery	Numerical (Minutes)
OutRoom_Time	Time when surgery ends	During Surgery	Numerical
PACU Ready to Leave	If a patient is ready to leave the post-anesthesia care unit	During Surgery	Binary (Yes, No)
OR Out Room to PACU	Time in minutes from when the patient	During Surgery	Numerical (Minutes)

Ready to Leave	leaves the theater to when the patient is clinically ready to leave the PACU		
Nauseated During Clinical Session	If a patient is extremely nauseated within 24 hours of the surgery	After Surgery (in 24 hours)	Binary (Yes, No)
Vomited During Clinical Session	Indicator of an assessment of "Vomited" during the clinical session	After Surgery (in 24 hours)	Binary (Yes, No)
First Pain Score	First pain score assessment after the surgery. “0” refers to no pain and “10” refers to the worst pain.	After Surgery (in 24 hours)	Integer ranging from 0 to 10
Maximum Pain Score	Maximum pain score assessment after the surgery. “0” refers to no pain and “10” refers to the worst pain.	After Surgery (in 24 hours)	Integer ranging from 0 to 10
First_PAR2PAR13_Min	Time in minutes from the first pain score to the max pain score	After Surgery (in 24 hours)	Numerical (Minutes)
Post Anesthetic Recovery (PAR) Score on Arrival	A cumulative score measured by activity, respiration, circulation, and consciousness level of patients. Patients with a score of 10 or greater are eligible for discharge	After Surgery (in 24 hours)	Numerical
Acute_LOS	Only acute lengths of stay	After Surgery (in 24 hours)	Numerical
RIW (Resource Intensity Weight)	The intensity of resource use (relative cost) associated with different diagnostic, surgical procedure and demographic characteristics of a patient	After Surgery (in 24 hours)	Numerical
Total LOS in Days	The entire length of stay, in days (includes both acute and ALC (alternate level of care) days)	After Surgery (in 24 hours)	Numerical
Any PSI	Post-admission diagnoses indicative of an adverse event that may be linked to suboptimal safety and quality of care during the encounter	After Surgery (in 24 hours)	Binary (Yes, No)
Disch_Loc	Location of discharge	After Surgery (in 24 hours)	Binary (Yes, No)
Hemorrhagic Event	Any events that accompanied by or produced by hemorrhage	After Surgery (in 24 hours)	Binary (Yes, No)
Anesthesia Complication	Complications associated with anesthesia	After Surgery (in 24 hours)	Binary (Yes, No)
Cardiac Complication	If patient has encountered a cardiac complication	After Surgery (in 24 hours)	Binary (Yes, No)
Central Nervous System Complication	If patient has encountered a central nervous system complication	After Surgery (in 24 hours)	Binary (Yes, No)
Lung_Complc	If patient has encountered a lung complication	After Surgery (in 24 hours)	Binary (Yes, No)
Fluid_Adv_Event	If patient has encountered a Fluid adverse event	After Surgery (in 24 hours)	Binary (Yes, No)
Delirium	An acutely disturbed state of mind that occurs in fever, intoxication, and other	After Surgery (in 24 hours)	Binary (Yes, No)
Endocri_Complc	If patient has encountered a Endocrine	After Surgery	Binary (Yes, No)

	Complication	(in 24 hours)	
Drug Related Complication	If patient has encountered a drug-related complication	After Surgery (in 24 hours)	Binary (Yes, No)
Obstetrical Complication (fetus)	If patient has encountered an obstetrical complication	After Surgery (in 24 hours)	Binary (Yes, No)
Gastrointestinal Complication	Complications such as constipation, impaction, bowel obstruction, diarrhea, and radiation enteritis	After Surgery (in 24 hours)	Binary (Yes, No)
Hospital-Acquired Infection	If a Hospital-Acquired Infection is caused	After Surgery (in 24 hours)	Binary (Yes, No)
Severe Event	If a critical surgical event occurred	After Surgery (in 24 hours)	Binary (Yes, No)
Surgical Complication	If a surgical complication occurred	After Surgery (in 24 hours)	Binary (Yes, No)
Traumatic Injury Arising In Hospital	Physical injuries of sudden onset and severity which require immediate medical attention	After Surgery (in 24 hours)	Binary (Yes, No)
Decubitus Ulcer	A sore developed by an invalid because of pressure caused by lying in bed in one position	After Surgery (in 24 hours)	Binary (Yes, No)
VTE	Venous thromboembolism (VTE) is a disease that includes both deep vein thrombosis (DVT) and pulmonary embolism (PE)	After Surgery (in 24 hours)	Binary (Yes, No)
Readmission to hospital in 30 days	***Postoperative complications within 30 days of surgery***	After Surgery (in 30 days)	Binary (Yes, No)
Emergency room in 30 days		After Surgery (in 30 days)	Binary (Yes, No)
Re-Operation in 30 days		After Surgery (in 30 days)	Binary (Yes, No)
ICI Admission in 30 days		After Surgery (in 30 days)	Binary (Yes, No)

Appendix 2.

Summary of the frequency for the binary attributes and Mean (± Standard Deviation) for the numerical attributes. Missing values of all the attributes are also provided used in the study

Attributes	Frequency in Robotic Surgery (%) (n = 312)	Frequency in Laparotomy Surgery (%) (n = 332)	Total Frequency (%) (n = 644)	Total missing values (%)
Age at Admission	61.01 ± 8.68	63.42 ± 11.27	61.92 ± 10.4	0
BMI > 40	21.8	16	18.9	0
ASA Score	2.54 ± 0.54	2.62 ± 0.56	2.57 ± 0.55	5 (0.87)
Elixhauser Comorbidity Score	4.01 ± 2.33	5.06 ± 4.27	4.40 ± 3.51	0
Alcohol Abuse	0	0.6	0.3	0
Deficiency Anemia	0	0	0	0
Cardiac Arrhythmia	1	1.8	1.5	0
Blood Loss Anemia	0	0.6	0.3	0
Congestive Heart Failure	0.3	0.6	0.4	0
Chronic Pulmonary Disease	0.3	3.3	1.8	0
Rheumatoid Arthritis/collagen	0	0.3	0.1	0
Metastatic Cancer	2.6	10.5	6.4	0
Solid Tumor without Metastasis	95.5	93.4	91.4	0
Coagulopathy	0	0.6	0.3	0
Depression	0.3	0.3	0.3	0
Diabetes, Complicated	0.6	5.1	2.8	0
Diabetes, Uncomplicated	8.3	13	10.5	0
Drug Abuse	0	0	0	0
Fluid and Ecletrolyte Disorders	0	0.3	0.1	0
AIDS/HIV	0	0	0	0
Hypertension, Comlicated	0	0	0	0
Hypertension, Uncomlicated	5.8	13.6	9.3	0
Hypothyroidism	1.9	1.5	1.8	0
Liver Disease	0	0	0	0
Lymphoma	0	0	0	0

Other Neurological Disorders	0	0.3	0.1	0
Obesity	5.1	7.5	6.1	0
Peptic Ulcer	0	0	0	0
Paralysis	0	0.3	0.1	0
Psychoses	0	0	0	0
Pulmonary Circulation Disorders	0	2.1	1	0
Renal Failure	0.3	0	0.1	0
Valvular Disease	0	0.3	0.1	0
Weight Loss	0	0	0	0
Procedure Length in Minutes	255.7 ± 48 (Minutes)	162.7 ± 44.5 (Minutes)	207.6 ± 65 (Minutes)	0
Nauseated During Clinical Session	37.9	24.8	32	5 (0.87)
Vomited During Clinical Session	9.2	3	6.4	8 (1.2)
First Pain Score	2.34 ± 2.56	3.84 ± 3	3.05 ± 2.93	24 (4.1)
Maximum Pain Score	4.41 ± 1.9	5.42 ± 2.3	4.91 ± 2.17	24 (4.1)
Post Anesthetic Recovery (PAR) Score on Arrival	11.33 ± 1.24	11.2 ± 1.5	11.29 ± 1.36	5 (0.77)
RIW (Resource Intensity Weight)	1.25 ± 0.17	1.5 ± 1.3	1.354 ± 0.9	0
Total LOS in Days	1.07 ± 0.44	4.5 ± 6.4	2.78 ± 4.84	0
Acute LOS	1.07 ± 0.44	4.4 ± 6	2.71 ± 4.52	0
Any PSI	5.1	14.2	9.5	0
Hemorrhagic Event	0.3	4.2	2.2	0
Anesthesia Complication	0	0	0	0
Cardiac Complication	0	0.6	0.3	0
Central Nervous System Complication	0	0.6	0.3	0
Delirium	0	0.6	0.3	0
Drug Related Adverse Event	0.3	0	0.1	0
Obstetrical Complication(fetus)	0	0	0	0
Gastrointestinal Complication	0	4.8	2.4	0

Hospital-Acquired Infection	0	2.1	1	0
Severe Event	0	0.9	0.4	0
Surgical Complication	4.2	3.6	3.8	0
Traumatic Injury Arising In Hospital	0	0	0	0
Decubitus Ulcer	0	0	0	0
VTE	0	0	0	0
Readmission in 30 days	5.4	13.3	9.2	0
Emergency room in 30 days	2.2	6.6	4.4	0
Re-Operation in 30 days	1.3	1.8	1.5	0
ICI Admission in 30 days	0.6	0.9	0.7	0

Appendix 3.

Set of 21 attributes that is used for the prediction problem prior to surgery

ASA	Age	BMI40	Elix_Co_score
Alcohol Abuse	Arrhyth	BloodLoss_Ane	Heart_Fail
COPD	Rheum_Arthr	Met_Cancer	cancer_no__Met
Coag_dis	Diabete	Hypertension	Valve_Dis
Hypothyroid	Obese	Pulm_dis	Renal_Fail
surgery Type			

Appendix 4

Set of 40 attributes (19 attributes in the following table as well as the set of 20 attributes in Appendix 5) used for the prediction problem prior to discharge

Proc_Lent_Min	Nausea	Vomit	Firs_Pain_score
Max_Pain_score	PAR_Scor_Init	RIW	LOS
PSI	Hemorrh	CNS_complc	Delirium
Drug_Adv_Event	Fluid_Adv_Event	GI_Complc	Infection
Lung_Complc	Severe_Advers	Surgic_complc	

Appendix 5

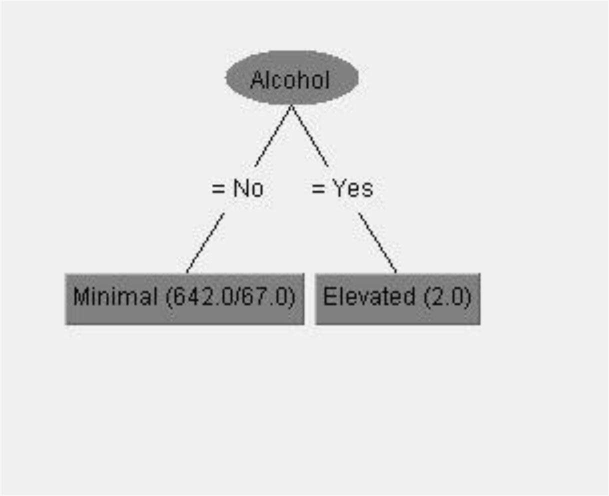
Results of the Independence chi-squared to test if there is a dependence between patients' risk class allocation and years

Tabulated statistics: Year of Surgery, Risk of Postoperative Complications

```
Rows: Year of Surgery   Columns: Postoperative Complication


              2012          2013          2014          2015          All

Elevated      24            22            19            4             69
              21.00         19.71         22.07         6.21          69.00

Minimal       172           162           187           54            575
              175.00        164.29        183.93        51.79         575.00

All           196           184           206           58            644
              196.00        184.00        206.00        58.00         644.00

Cell Contents:      Count
                    Expected count


Pearson Chi-Square = 2.139, DF = 3, P-Value = 0.544
```

Appendix 6

The predictive model using the C4.5 technique applied to the balanced dataset prior to surgery



Appendix 7

Rules generated by the PART technique with undersampling for the prediction problem prior to surgery

Rule	Class (Elevated or Minimal)	Support	Confidence (%)
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND take Robotic Surgery AND without a Metastatic Cancer AND has an ASA score of 2 AND younger than 63	Minimal	10	100
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND take Robotic Surgery AND not diagnosed with diabetes AND has an ASA score of 2	Minimal	9	77.7
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND take Laparotomy Surgery AND not diagnosed with diabetes AND without a Metastatic Cancer AND has an ASA score of 1 or 2	Minimal	4	100
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND take Laparotomy Surgery AND has a Body Mass Index greater than 41 AND not diagnosed with diabetes	Elevated	10	90
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND take Laparotomy Surgery AND has a non-complicated Hypertension AND has an Elixhauser Comorbidity Score greater than 1	Elevated	4	100
IF patient is not diagnosed with Chronic Pulmonary Disease AND take Robotic Surgery AND without a Metastatic Cancer AND not diagnosed with diabetes AND has a Body Mass Index greater than 41 AND her an ASA score greater than 2 AND older than 43	Minimal	15	80
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND	Elevated	4	100

take Laparotomy Surgery AND without a Metastatic Cancer AND has Non-Complicated diabetes AND has a Body Mass Index greater than 41			
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND is obese AND has no hypertension	Minimal	3	100
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND younger than 55 AND has an ASA score greater than 2	Elevated	7	100
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND take Robotic surgery AND without a Metastatic Cancer AND Younger than 63 AND has an ASA score greater than 2	Minimal	13	92.3
IF patient is not Alcohol abuse AND not diagnosed with Chronic Pulmonary Disease AND has a Body Mass Index greater than 41	Minimal	4	75
has a Body Mass Index greater than 41 AND has Non-Complicated diabetes AND younger than 67	Minimal	4	75
without a Metastatic Cancer AND has a Body Mass Index less than 41 AND without diabetes AND take a Robotic surgery AND has an ASA score greater than 2 AND younger than 68	Elevated	5	100
IF patient has a Body Mass Index less than 41 AND without diabetes AND take a Robotic surgery AND	Elevated	6	66.6
IF patient has a Body Mass Index less than 41 AND without diabetes AND take a Laparotomy surgery AND has an ASA score less than 3 AND has an Elixhauser Comorbidity Score less than 11	Elevated	35	71.5

IF patient has a Body Mass Index less than 41 AND without diabetes AND take a Laparotomy surgery AND Older than 61 AND has an Elixhauser Comorbidity Score greater than 1 and less than 11	Elevated	13	77
IF patient has a Body Mass Index less than 41 AND without diabetes AND with a Metastatic Cancer AND take a Laparotomy surgery AND Older than 61	Elevated	3	100
IF patient is with a Metastatic Cancer	Minimal	4	100
IF patient has a Body Mass Index less than 41 AND without diabetes AND	Minimal	3	100

Appendix 8

The predictive model using the C4.5 technique with oversampling prior to surgery

Appendix 9

Some of the rules generated with the BRACID technique to the imbalanced dataset prior to surgery

(ASA =2.0) ***And*** *(Age in 56.0:69.0)* ***And*** *(BMI41 = No)* ***And*** *(Elix_Co_score =9)* ***And*** *(Alcohol = No)* ***And*** *(BloodLoss_Ane = No)* ***And*** *(Heart_Fail = No)* ***And*** *(COPD = No)* ***And*** *(Rheum_Arthr = No)* ***And*** *(Met_Cancer = No)* ***And*** *(cancer_no__Met = Yes)* ***And*** *(Coag_dis = No)* ***And*** *(Diabete = Non)* ***And*** *(Hypothyroid = No)* ***And*** *(Obese = No)* ***And*** *(Pulm_dis = No)* ***And*** *(Valve_Dis = No) ->* ***Minimal***

(ASA = 2.0) ***And*** *(Age in 54.0:55.0)* ***And*** *(BMI41 = No)* ***And*** *(Elix_Co_score =4)* ***And*** *(Alcohol = No)* ***And*** *(Arrhyth = No)* ***And*** *(BloodLoss_Ane = No)* ***And*** *(Heart_Fail = No)* ***And*** *(COPD = No)* ***And*** *(Rheum_Arthr = No)* ***And*** *(Met_Cancer = No)* ***And*** *(cancer_no__Met = Yes)* ***And*** *(Coag_dis = No)* ***And*** *(Diabete = Non)* ***And*** *(Hypertension = Non)* ***And*** *(Hypothyroid = No) (Obese = No)* ***And*** *(Pulm_dis = No)* ***And*** *(Renal_Fail = No)* ***And*** *(Valve_Dis = No)* ***And*** *(surgery Type = Laparotomy) ->* ***Minimal***

(ASA = 2.0) ***And*** *(Age in 79.0:88.0)* ***And*** *(BMI41 = No)* ***And*** *(Elix_Co_score in 4.0:10.0)* ***And*** *(Alcohol = No)* ***And*** *(Arrhyth = No)* ***And*** *(BloodLoss_Ane = No)* ***And*** *(Heart_Fail = No)* ***And*** *(COPD = No)* ***And*** *Rheum_Arthr = No)* ***And*** *(Met_Cancer = No)* ***And*** *(cancer_no__Met = Yes)* ***And*** *(Coag_dis = No)* ***And*** *(Hypothyroid = No)* ***And*** *(Obese = No)* ***And*** *(Pulm_dis = No)* ***And*** *(Renal_Fail = No)* ***And*** *(Valve_Dis = No)* ***And*** *(surgery Type = Laparotomy) ->* ***Minimal***

(ASA = 2.0) ***And*** *(Age in 49.0:71.0)* ***And*** *(Elix_Co_score in 9.0:16.0)* ***And*** *(Alcohol = No)* ***And*** *(BloodLoss_Ane = No)* ***And*** *(Heart_Fail = No)* ***And*** *(COPD = No)* ***And*** *(Rheum_Arthr = No)* ***And*** *(cancer_no__Met = Yes)* ***And*** *(Coag_dis = No)* ***And*** *(Diabete = Non)* ***And*** *(Pulm_dis = No)* ***And*** *(Renal_Fail = No)* ***And*** *(Valve_Dis = No) ->* ***Minimal***

(ASA =3.0) ***And*** *(Age in 53.1:54.9)* ***And*** *(BMI41 = No)* ***And*** *(Elix_Co_score=4)* ***And*** *(Alcohol = No)* ***And*** *(Arrhyth = No)* ***And*** *(BloodLoss_Ane = No)* ***And*** *(Heart_Fail = No)* ***And*** *(COPD = No)* ***And*** *(Rheum_Arthr = No)* ***And*** *(Met_Cancer = No)* ***And*** *(cancer_no__Met = Yes)* ***And*** *(Coag_dis = No) (Diabete = Non)* ***And*** *(Hypertension = Non)* ***And*** *(Hypothyroid = No)* ***And*** *(Obese = No)* ***And*** *(Pulm_dis = No)* ***And*** *(Renal_Fail = No)* ***And*** *(Valve_Dis = No) (surgery Type = Robat) ->* ***Elevated***

(ASA=2) ***And*** *(Age in 73.1:77.8)* ***And*** *(BMI41 = No)* ***And*** *(Elix_Co_score =4)* ***And*** *(Alcohol = No)* ***And*** *(Arrhyth = No)* ***And*** *(BloodLoss_Ane = No)* ***And*** *(Heart_Fail = No)* ***And*** *(COPD = No)* ***And*** *(Rheum_Arthr = No)* ***And*** *(Met_Cancer = No)* ***And*** *(cancer_no__Met = Yes)* ***And*** *(Coag_dis = No)* ***And*** *(Diabete = Non)* ***And*** *(Hypertension = Non)* ***And*** *(Obese = No)* ***And*** *(Pulm_dis = No)* ***And*** *(Renal_Fail = No)* ***And*** *(Valve_Dis = No) ->* ***Elevated***