Predicting a Surgical Site Hospital Acquired Infection: A case study

Nicole Bartsch

Research assignment presented in partial fulfilment of the requirements for the degree of
MCom (Quantitative Management)
at Department of Logistics, Stellenbosch University

Supervisor: Dr Isabelle Nieuwoudt

March 2020

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2020

Abstract

In this study a logistic regression model for a private healthcare group, was used to determine the predicted number of Surgical Site Infections (SSIs) of an operative procedure at a healthcare facility. The purpose of this study is to determine the Standard Infection Ratio (SIR) which compares the actual number of SSIs that were contracted by patients at a hospital against the number of SSIs predicted. A SIR of above 1 is regarded as a bad result as the model predicted less infections to occur at a hospital than the actual number of infections that did occur. A SIR of below 1 is an ideal and good result that hospitals should aspire to achieve. The SIR is calculated across three hospitals, across three years (2016, 2017 and 2018) and across five operative procedure groups (HYST, SB, BILI, CARD and KPRO).

Specific significant risk variables were taken into account per operative procedure group. These variables ranged from whether the patient was a diabetic or not, the age of the patient, which hospital the patient was admitted to, the BMI of the patient and the ASA score of the patient. Since the American Society of Anesthesiologists Classification (ASA) score is not captured electronically per patient, a logic was developed to determine the ASA score of a patient based on their clinical coding information and level of care they received.

The logistic regression model was developed per operative procedure group and determines the probability of a patient contracting an SSI. A Hosmer-Lemmeshow goodness of fit test was conducted to compare the actual events against the predicted events across 10 subgroups of the model's population. Finally, the SIR was calculated by dividing the actual number of SSIs by the predicted number of SSIs at a hospital.

There is a clear difference in the SIR results across the three hospitals that were considered, over the three years being analysed. Hospital A needs to focus on the operative procedure group CARD and Hospital B needs to focus on all five operative procedures except for the operative procedure group SB where they scored an SIR of below 1. Hospital C achieved exceptional SIR results with all operative procedure groups across all three years having an SIR result of below 1. Both Hospital A and Hospital B need to improve the infection prevention and control practices at their hospitals as well as schedule interventions to decrease the number of SSIs occurring at their hospitals.

Opsomming

In hierdie studie is 'n logistieke regressiemodel vir die private gesondheidsorgeenheid, gesondheidsorgeenheid gebruik om die voorspelde aantal chirurgiese lokale infeksies (SSIs) na 'n operasie by een van gesondheidsorgeenheid se hospitale, te bepaal. Die doel van hierdie studie is om die Standaard Infeksie Verhouding (SIR) te bepaal wat die werklike aantal SSIs wat deur pasiënte in 'n hospitaal opgedoen is met die aantal voorspelde SSI's te vergelyk. 'n SIR van groter as 1 word as 'n slegte resultaat beskou, aangesien die model voorspel het dat minder infeksies in 'n hospitaal sou voorkom as die werklike aantal infeksies wat wel voorgekom het. 'n SIR van minder as 1 is 'n ideale en goeie resultaat waarna hospitale behoort te mik. Die SIR word bereken oor drie hospitale, oor drie jaar (2016, 2017 en 2018) en oor vyf operatiewe prosedure goepe (HYST, SB, BILI, CARD en KPRO).

Spesifieke beduidende veranderlikes is per operatiewe prosedure groep in ag geneem. Hierdie veranderlikes het gewissel tussen of die pasiënt 'n diabeet was of nie, die ouderdom van die pasiënt, in watter hospitaal die pasiënt opgeneem is, die BMI van die pasiënt en die ASA-telling van die pasiënt. Aangesien die *American Society of Anesthesiologists Classification* (ASA) telling nie elektronies per pasiënt opgeneem word nie, is 'n logika ontwikkel om die ASA-telling van 'n pasiënt te bepaal op grond van hul kliniese koderingsinligting en die versorgingsvlak wat hulle ontvang het.

Die logistieke regressiemodel is per operatiewe prosedure groep ontwikkel en bepaal die waarskynlikheid dat 'n pasiënt 'n SSI kan opdoen. 'n Hosmer-Lemmeshow geskiktheidstoets is uitgevoer. Uiteindelik is die SIR bereken deur die werklike aantal SSI's te deel deur die voorspelde aantal SSI's vir 'n hospitaal.

Daar is 'n duidelike verskil in die SIR-resultate in die drie hospitale wat beskou is gedurende die drie jaar wat geanaliseer was. Hospitaal A moet fokus op die operasionele prosedure groep CARD en Hospital B moet fokus op al vyf operatiewe prosedure groepe, behalwe die operatiewe prosedure groep SB waar hulle 'n SIR van onder 1 behaal het. Hospital C het uitsonderlike SIR-resultate behaal deurdat alle operasionele prosedure groepe gedurende al drie jare 'n SIR-uitslag van minder as 1 behaal het. Beide Hospitaal A en Hospitaal B moet klem lê op die verbetering van infeksievoorkomings en beheerpraktyke by hul hospitale, sowel as intervensies bewerkstellig om die aantal SSI's wat by hul hospitale voorkom, te verminder.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ASA:: American Society of Anesthesiology

BILI:: The operative procedure group made up of open and laparoscopic liver, bile duct and pancreatic surgery, including resections, excisions, ablations, biopsies, repairs and ostomies.

BMI:: Body Mass Index

CARD:: The operative procedure group made up of procedures on the heart valves or septum but does not include coronary artery bypass graft, surgery on vessels, heart transplantation, or pacemaker implantation.

CAUTI:: Catheter Associated Urinary Tract Infections

CCHF:: Crimean Congo Hemorrhage Fever

CLABSI:: Central Line Associated Bloodstream Infections

CPB:: Cardiopulmonary Bypass

CPT:: Current Procedural Terminology

DF:: Dengue Fever

DHF:: Dengue Hemorrhagic Fever

EHR:: Electronic Health Record

HAI:: Hospital Acquired Infections

HL:: Hosmer-Lemmeshow

HYST:: The operative procedure group made up of open and laparoscopic hysterectomies, and/ or removal of ovaries or tubes, with or without lymphadenectomies and exenteration.

ICD:: International Classification of Diseases and Related Health Problems

KPRO:: The operative procedure group made up of arthroplasty of the knee.

NNIS:: American National Nosocomial Infections Surveillance

NNHISA:: National Health Information System of South Africa

NHSN:: American National Health and Safety Network

OR:: Operations Research

QM:: Quantitative Management

PAS:: Patient Admin System

SB:: The operative procedure group made up of open or laparoscopic small bowel surgeries and resections including revisions and repairs, ileostomy and jejunostomy.

SIR:: Standardized Infection Ratio

SSI:: Surgical Site Infections

TB:: Tuberculosis

VAP:: Ventilator Associated Pneumonia

WHO:: World Health Organisation

CHAPTER 1

Introduction

The South African Constitution's Bill of Rights says that everyone has the right to have access to healthcare services within South Africa. *Healthcare* is the continuous improvement of a human's health, either by the treatment, prevention and cure of disease, illness and injury [28]. Healthcare within South Africa may be divided into a public sector and private sector, with the public sector being larger than the private sector. The private sector consists of mainly patients whom are covered by hospital insurance or have medical aid coverage [6]. Tourists or overseas nationals who join a private health insurer or sign up for a local medical aid will receive care, in the private sector, that is on par with healthcare in their home country. South Africa's private sector boasts the highest standard of healthcare throughout Africa and there are more than 200 private hospitals across South Africa [6].

Hospitals generate, on a daily basis, large amounts of data, for example, administration details, test and result details, trials, etc. In general, analysing these type of data may lead to improvements in a system. In the healthcare environment, for example, it may help to rapidly improve the quality of healthcare and support a wide range of current healthcare barriers and issues such as Electronic Healthcare Records (EHRs), population health management and disease surveillance within a patient [6]. Most healthcare systems within Europe, the United States of America and any developed country are making use of EHR systems and by implementing an EHR system within South Africa, it may aid the nation's healthcare system.

1.1 Hospital Acquired Infections (HAIs)

As stated by the American National Health and Safety Network (NHSN): "*Hospital acquired infections* (HAIs) are complications that emanate from a stay in a medical facility [26]." Differently stated, HAIs or *nosocomial infections* are described as an infection, not present on admission, but rather an infection that develops itself or is contracted by the patients during their stay in the hospital or in the period after a patient's hospital stay [26]. HAIs may be divided into *surgical site infection* (SSI), *catheter-associated urinary tract infection* (CAUTI), a *central line-associated bloodstream infection* (CLABSI) and *ventilator-associated pneumonia* (VAP).

HAIs are a significant problem throughout the world and are an even greater burden in developing countries such as South Africa. Within the public healthcare sector of South Africa there is a severe lack of staff and training. Similarly, in the private sector, human resources dedicated to patient surveillance activities are often insufficient [6]. A high HAI rate is experienced throughout South Africa and a leading cause may be the lack of trained and motivated staff throughout South Africa's public and private healthcare systems. A study completed by the Human Resource Department for Health Corporation, states that nurses within the private sector were generally satisfied, whereas nurses within the public sector were generally dissatisfied due to their pay, workload and the resources that were available to treat patients within the public healthcare system in South Africa [25]. Although there are numerous problems in the South African healthcare sector, this research will only focus on the patients surveillance activities within the private sector.

An example of a HAI is Klebsiella Pneumonia which may result in multiple infections that a patient contracted in a healthcare setting. Klebsiella infections are linked to many different types of HAIs. Klebsiella infections are the result of Klebsiella bacteria being present in a hospital. The Klebsiella infection may be found in a person's intestine, however, the infection is not caused there [19]. In order for a person to contract the Klebsiella bacteria a person must be directly exposed to the bacteria. For example, a patient contracting Pneumonia or a bloodstream infection due to the Klebisella bacteria, the patient would have had to be exposed to the Klebsiella bacteria entering their respiratory tract or their blood. Many people and therefore patients who are hospitalised and are receiving treatments are susceptible to contracting the Klebsiella infection due to being exposed to the bacteria in the hospital environment [19].

Two medical professionals namely, Dr ST Hlope and Dr NH Mckerrow conducted a research based study through the University of Kwazulu-Natal and the Department of Health, on hospital acquired Klebsiella Pneumonia infections in pediatric intensive care units [15]. The study emphasised that HAIs are a significant problem in the intensive care units. A nosocomial infection prolongs a patient's stay by 5 to 10 days [15]. By decreasing the number of nosocomical infections within a hospital, the hospital will decrease the patient's costs and decrease mortality rates. The study states that HAIs such as Klebsiella Pnemonia within Neonatal ICUs has increased rapidly within South Africa across the recent years [15]. The study suggests that the cause of these outbreaks may be directly linked to under staffing, overcrowding and breakdown in infection control measures.

McKibben *et al.* [22] performed an analysis to determine and evaluate the influence of HAIs in *neonates* (any infant of less than four weeks old) based on any additional costs and/or any additional hospital stay. The study focused on all neonatal patients from a specific university hospital that were admitted from October 1993, into the neonatal intensive care unit, and were discharged alive before December 1993. This included 515 neonates, 69 (13%) which had one or more HAI, 45 (20 neonates with an HAI, 25 neonates suspected of contracting an HAI were matched to 45 controls.) Many contributing variables were taken into account such as gestational age, surgery, artificial ventilation and the utilisation of a catheter. Central vascular catheter utilisation was the only factor significantly associated with an HAI. The average additional duration of stay in the hospital for neonates with an HAI was 24 days (a range between 30 and 54 days). The average additional charges for patients who had contracted an HAI was 9635 Euros. Accommodation accounted for 72% of the additional charges, fees for 22%, pharmaceuticals for 5% and ancillary items for 1% of these extra charges. Overall the fees and fees billed per day were similar for neonates who contracted one or more HAIs and for neonates with a suspected HAI [21].

In the United States of America, within a hospital, the total number of HAIs and associated deaths were calculated. The three main sources of data included the National Nosocomial Infections Surveillance (NNIS) system of the NHSN, Centers for Disease Control and Prevention as well as data from the National Hospital Discharge Survey. Based off the NHSN data, the total number of patients who contracted a HAI and whose death was due to a HAI was used to determine the number of deaths. In 2002, the total number of HAIs in U.S. hospitals was approximately 1.7 million. The 1.7 million HAIs consisted of 33 269 HAIs that were contracted by high risk newborns, 19 059 HAIs contracted by low risk newborns which were located in general ward nurseries, 417 946 HAIs contracted by high risk adults and children in ICUs and 1 266 851 HAIs that were contracted by adults and children in other locations than ICUs. A total of 98 987 deaths were estimated to have been linked to HAIs in hospitals within the USA. This included 35 967 deaths linked to pneumonia, 30 665 deaths associated from bloodstream infections, 13 088 deaths caused by urinary tract infections, 8 205 deaths from SSIs, and 11 062 deaths were due to other types of infections. In the United States of America, HAIs are regarded as a significant factor which leads to an increased number of deaths. HAIs also place a huge financial impact on the patient, medical insurer and the healthcare provider [18].

The Science of Healthcare Epidemiology of America studied and estimated the number of mortalities, costs of HAIs and the proportion of HAIs within hospitals in the United States. The methodologies made use of the most up to date public data to determine the number of HAIs and any mortalities associated with an HAI per annum. The range of the number of infections that could have been prevented within a year, deaths and annual costs were calculated by multiplying the infection, mortality, and billed fees with the ranges of preventability for each HAI. The lowest and highest risk reductions, over the past 10 years, were used to calculate the range of preventability within the US. In order to calculate the incremental cost of HAIs, a systematic review was performed which made use of costs from studies in the general US patient populations [31]. This study showed that up to 65%-70% of patients whom had contracted either CLABSI or CAUTI may have be preventable and 55% of patients who had contracted VAP or SSI may be have been prevented. CAUTI may be regarded as the most preventable HAI, whereas CLABSI and VAP have the highest number of preventable deaths. CLABSI also has the highest cost impact linked to all patients whose CLABSI could have been prevented. In conclusion, comprehensive implementation of prevention and intervention strategies could prevent a significant amount of HAIs, save lives and save billions of dollars throughout the United States of America [31].

The most common HAI contracted by patients are SSIs. In the United Kingdom (UK), between 5% and 10% of patients who were admitted to theatre and underwent surgery are estimated to develop an SSI which results in an increased length of stay as well as an increase in the patient's risk of mortality. Approximately 1 billion pounds is spent, from the country's financial resources, on SSIs each year within the UK. The majority of SSIs are preventable and certain infection prevention controls can be taken before, during and after any surgical procedures to reduce the risk of a patient contracting an SSI [4]. CAUTI may effect any patient that had a catheter inserted during their hospital stay [11], while CLABSI is a result of an infection which develops after any central line placement (*i.e.* the insertion of a catheter) and are associated with both increased mortality and morbidity [12]. Any patient who is on a ventilator during their stay in hospital is at risk of developing VAP.

HAIs, for example, affect approximately two million patients who are admitted to acute care hospitals within the United States of America. A rising concern is that nosocomial pathogens are becoming more resistant to antimicrobial agents [8]. The increase in the resistance of antimicrobial agents may result in an increased length of stay for a patient in hospital as well as increased heathcare costs [8].

A study conducted on the efficacy of nosocomial infection control estimated that the incident of HAIs contributes approximately 45 million dollars towards infection prevention and intervention expenses, per year in the United States of America. HAIs make up approximately 6% of all mortalities in the USA, exceeding the amount of mortalities associated with breast cancer [20].

1.2 Reporting of Hospital Acquired Infections

Approximately one in ten patients in acute care hospitals at any one time have acquired an infection after admission to a hospital. HAIs are a great concern across all health facilities and have a negative influence on the population's health as well as the demands on scarce resources across hospitals [26]. Over recent years, consumers have indirectly demanded healthcare information, including the clinical performance of healthcare providers [22]. Many healthcare organizations are publicly disclosing information regarding institutional, physician and patient experience performance. The aim of publicly reporting on healthcare performance is intended to keep an open and honest relationship with the organization's past, present and future patients and allow for consumers to make better decisions regarding their healthcare [30]. The reporting of healthcare facilities' performance publicly has taken several forms. The public reporting of healthcare may be distributed via performance reports (static and dynamic reports) which would describe the results of clinical indicators such as medical care in terms of mortality, selected complications, infection rates, or medical errors. Infection rates are becoming increasingly important to consumers and patients, forcing healthcare service providers to publicly report on their infection rates and ways in which they aim to improve and decrease these rates [22].

The reporting of HAIs in a format in which the consumer and patient may understand, is becoming a priority for most healthcare facilities. HAIs have been reported on in many different forms, including total sum of HAIs, total HAIs per 1000 discharged patients, total HAIs per 1000 catheter days, and a risk-adjusted rate of HAIs per 1000 catheter days. Reporting infection rates visually by means of graphical interfaces to illustrate the distribution of these occurrences across all reporting hospitals should be a priority in all healthcare settings [27].

Currently, all HAIs are recorded and calculated as a rate. For example, the SSI rate is the number of SSIs divided by the number of operative procedure PER 1000 patients (a rate). However, rates cannot reflect the differences in the risk between populations, resulting in a loss of comparability.

The *Standardized Infection Ratio* (SIR) is a statistical measure used to monitor HAIs over time, at a national, private, public, or hospital level. The **actual** number of HAIs is compared to the **predicted number of infections** within a hospital. The predicted number of HAIs is an estimate based on healthcare data, patient data, and are also risk adjusted. Risk adjustment takes into account that certain hospitals may see and treat sicker patients and more complex medical cases than other hospitals. By making use of the SIR, hospitals will be able to prevent HAIs in the future.

The NHSN encourages all heathcare providers to convert from a rate methodology to the SIR. The rates are pooled statistical means which is calculated by summing the number of infections and dividing the total by the number of device days (per HAI). *Device days* refer to the number of days a patient may require to use a particular device, for example the number of days a patient is on a ventilator [9]. Calculating HAI rates based on a pooled mean, may result in false calculated rates as well as a rate that reflects differences within populations and this may result in the loss of the comparability of HAIs over time and across healthcare facilities.

Apart from reporting on HAIs, during recent years, hospital organisations have been pressurised to decrease the cost of healthcare whilst increasing the quality of healthcare. A huge focus has been placed on reducing HAIs by improving infection prevention techniques at hospitals as well as the correct monitoring of HAIs within hospitals. At the simplest form, cleanliness and hygiene are two very important factors of quality healthcare within a hospital. Maintaining a clean environment within a hospital results in preventing HAIs and promotes patient and visitor confidence. However, more and more hospitals also want to predict a HAI outbreak in order to improve even more on HAI prevention.

1.3 Problem Description

One of the leading hospital groups, within the private healthcare sector in Southern Africa, consists of 49 hospitals and 2 day clinics in South Africa as well as 3 hospitals in Namibia. This healthcare facility places science at the heart of their care by aiming to provide the highest care, since their five core business values are patient safety, mutual trust and respect, teamwork, client focus and performance driven [32].

The HAI in hospitals that will be the focus in this study is SSI. The NHSN divides all operative procedures that occur in theatres into operative procedure groupings to calculate a SSI ratio per operative procedure grouping based on the SIR methodology. During this study the focus will be on SSIs within hospitals in Southern Africa, for five operative procedure groupings, labeled BILI, SB, CARD, KPRO and HYST.

By developing the SIR for SSIs, hospitals and operational staff will be able to track and aim to reduce the number of SSIs, and thus the SIR for SSIs. This will therefore result in a decrease in costs and expenses for the hospitals as well as the patients and the patients' health insurance. The cost of treating infections which often result in septicemia for the patient are extremely high for all parties involved, not to mention the health risk to the patient.

SIR for SSIs developed by the NHSN allows hospitals to benchmark internationally. The healthcare facilities consists of three main international divisions across Switzerland, the Middle East and Southern Africa. The SIR for SSIs will assist in comparing the three divisions to each other, which will result in decisions being made on where the biggest improvement focus should be aimed.

In this study, the researcher will calculate a SIR per specific operative procedure group per hospital per year for all 49 hospitals over a period of three years. These values will assist the hospitals in decreasing and combating the number of infections throughout the hospitals. After the SIRs of all hospitals were calculated, three of these hospitals, referred to as Hospital A, Hospital B and Hospital C, will be discussed and compared in detail. In general, all three hospitals have unique patient profiles within their hospital. Each hospital is admitting a variety of patients from being at high risk of dying to patients being at low risk of dying. SSIs within all three hospitals are leading causes of mortalities. For specific hospitals, specific operative procedures may result in higher SSIs.

Furthermore, it is a constant battle at hospitals to ensure that accurate and real-time data is recorded and reported on, since a 'negative' stigma may be built around the capturing of infections. This will lead to the situation where not all infections are captured and may result in under reporting. By measuring and incorporating the SSI SIR across hospitals, one would also be able to track under reporting at certain hospitals.

Reducing under reporting together with tracking infections obtained by surgical patients, as well as the process of how to improve and decrease the number of SSIs, will ensure a 'best care, always' approach as described by the healthcare facilities slogan.

1.4 Objectives and layout of the thesis

The objectives of this research project are

I To obtain a literature study required for the project.

II To obtain all the data required for the project.

III To build a logistic regression model, to calculate the predicted SSIs across hospitals per operative procedure grouping.

IV Calculate the SIR per hospital per operative procedure group and compare three of the hospitals.

V To provide the hospitals with the SIR results in order to aid them in preventing any SSIs across their specific hospital and patients.

The thesis commences in Chapter 2 with some literature review on the prediction of HAIs. This is followed in Chapter 3 with the process of data collection, thus addressing Objective I. The methodology followed in this research project is the topic of Chapter 4, with the results given in Chapter 5. Thus, Objectives II to IV are addressed in Chapters 4 and 5. Finally, some concluding remarks are made in Chapter 6.

CHAPTER 2

Literature

The use of Operations Research models and specifically simulation models are becoming an integral part of improving systems in healthcare and hospitals, as stated by the Virginia Polytechnic Institute and State University. In recent years, healthcare costs have increased along with healthcare organisations being put under pressure to provide improved healthcare to their patients. Typically, discrete event simulation is an effective tool for allocating scare resources while minimising healthcare costs and increasing the patient's satisfaction or to optimise the utilisation of theatres [17]. Over the last few years, there has been a rapid growth in the number of software based methods to solve complex optimisation and simulation problems in healthcare facilities [17]. One such application is systems to predict HAI outbreaks.

The focus of the literature in this chapter is to emphasise the importance of measuring and predicting HAIs. The chapter commences in §2.1 with a few real-life HAIs studies and corresponding mathematical models. In §2.2 the use of a Standard Infection Ratio (SIR) to predict SSIs will be discussed. The method to predict SSIs in this study, logistic regression, is described in §2.3.

2.1 Modeling Healthcare Associated Infections (HAIs)

A predicative model for dengue outbreak was developed using multiple rulebase classifiers which included decision trees, rough set classifiers, associate classifiers and naive bayes [5]. Dengue Fever (DF) and Dengue Hemorrhagic Fever (DHF) have become an increased public health related issue in Malaysia and the World Health Organization (WHO) has reported these two diseases as rising pandemics. Thus, if the early detection of a dengue outbreak could be improved through forecasting or prediction methods, strategic outbreak planning and decision making can commence to decrease and limit the repercussions following an outbreak. The aim of the classification model by Baker *et al.* [5], is to predict a dengue outbreak. Several classifiers are investigated to study the performance of different rule based classifiers individually as well as in combination. The results stated that the multiple classifiers produce up to 70% better accuracy with a higher count of quality rules compared to the single classifier [5]. The rule based classifiers are selected as rules and are compared to a model which includes weights in a neural classifier and probability values in a Bayesian classifier.

A study was conducted at a hospital in Mexico City to understand and determine the nosocomial outbreak in an intensive care unit due to an HAI in newborn infants resulting in neonatal septicemia. Forty six newborns presented either one or multiple infections during their stay in the neonatal unit over a period of time. Of those 46 newborns, Sepsis was recorded in 41 of them. Apart from appropriate measures dealing with hygiene and education of personnel, predictive modeling such as decision trees and regression will result in the implementation of various measures to eliminate or decrease the risk of outbreaks occurring. This may also limit the number of deaths per neonatal ICU and improve the overall status of the hospital [3].

A Japanese study to develop and internally cross-check a SSI prediction model, dedicated towards a SIR model, was conducted for this country [13]. The data analyses included all data reported to the Japanese Nosocomial Infections Surveillance system for patients who were admitted to theatre and underwent a surgical procedure. The data obtained was for a period of 2 years, specifically from the beginning of 2010 to the end of 2012. The statical method used to build the model was logistic regression analyses, with the objective being to predict all SSIs. A SSI prediction model was built for each of the operative procedure categories, after the variable selection took place based on the data collected from the Japanese Nosocomial Infections Surveillance system. In the case of over fitting of the model, standard bootstrapping techniques were applied. The study sample comprised of 349 987 cases from 428 participating healthcare facilities throughout Japan. Of all the patients that were admitted for surgery, 7.0% of the patients contracted an SSI. In conclusion, the SSI predictive models developed for Japan resulted in higher accuracy than the average SSI model. The SSI predictive models will be used to help conquer unnecessary SSIs contracted by patients which will improve the healthcare facilities performance and identify patients that are at high risk of obtaining an SSI in specific procedure categories [13].

A study completed at the Second Surgical University Clinic in Vienna, Austria, analysed the risk factors for severe bacterial HAIs after patients had Valve Replacements and Aortocoronary Bypass operations [23]. The analysis consisted of evaluating 246 patients in total of which 84 patients undergoing valve replacements or 162 patients undergoing coronary bypass operations. A multiple logistic regression model was built to determine the ability to predict a HAI across a patient. The variables or risk factors considered were age, sex, diabetes mellitus, duration of cardiopulmonary bypass (CPB) which is also known as the amount of minutes the patient spent in theatre, amount of blood restored on the day of operation, repeat thoracotomy for bleeding, intraaortic balloon pumping, reoperation, emergency operation, and the professional status of the surgeon [23].

The results of the Second Surgical University Clinic showcased that for patients who were admitted for a bypass procedure, the only significant variable associated with a HAI was repeat thoracotomy which scored a p-value of 0.0004. However, the classification analysis conducted revealed that repeat thoracotomy could not be the only variables used because the variable is too unspecific for a reliable prediction of a HAI [23].

A univariate analysis emphasised that a restoration on a patient of more than 2.5 liters of blood with a p-value equal to 0.0001, reoperation with a p-value equal to 0.0821, duration of operation (p-value = 0.0061), duration of CPB (p-value = 0.0318), and intraaortic balloon pumping (p-value = 0.0281) were linked with HAIs following valve replacements. The patients who underwent a valve replacement operation, resulted in a well performing model in predicting HAIs. The classification analysis that was conducted, showed a good fit between the observed HAIs and predicted HAIs. The model accurately and correctly predicted 75% of the patients who correctly contracted an infection and 96% of the patients who correctly did not contract an infection [23].

2.2 HAIs modeling through the Standard Infection Ratio (SIR)

The methodology to calculate the SIR is very similar to calculating the Standard Mortality Index. By making use of HAI data, the SIR methodology compares the actual or observed number of HAIs reported versus the number of HAIs predicted. Therefore, a SIR greater than 1.0 indicates that more HAIs were observed than predicted and a SIR less than 1.0 indicates that more HAIs were predicted than observed [9].

By making use of a SIR, healthcare providers and facilities are able to benchmark results across different hospitals or healthcare facilities, the benchmarking of results can be used to measure progress over a period of time. A SIR aids in comparing the number of actual infections experienced at a facility to the number of infections predicted by the model for the same facility. The number of predicted infections may be calculated by using various methodologies such as a logistic regression model or a binomial logistic regression model.

A study conducted through the biophysics and mathematical engineering department of the University of Turkey developed a mathematical model to predict epidemic diseases by making use of the SIR method [16]. Epidemic diseases and infections in healthcare such as Tuberculosis (TB), AIDS, Crimean-Congo Hemorrhage Fever (CCHF) are all regarded as major health problems and therefore it is necessary that measures are taken to decrease the high numbers of these epidemic diseases throughout the world. Taking appropriate precautions starts by developing mathematical models to determine certain predictions [16].

Patterns of infections from surgery are ongoing and very sensitive to seasonal fluctuations. A vital part of infection prevention control management, within a healthcare facility, is to record the epidemiological features of infections in these patients over time. A study conducted at a 750 bed sized university hospital in Chiang Mai, Thailand, aimed to describe and explore the predictive risk factors or variables of the SSIs [24]. The study focused on patients admitted to theatre who underwent specific operations. The methodology used was based on the guidelines from the NNIS to identify and diagnosed infections and thereafter calculate the SIR. The methodology selected and used to predict infections, by making use of a number of significant risk factors, was the application of a multiple logistic regression model and the study included data from September 1998 to March 2000. The data used included 4193 patients and of these patients a total of 4437 major operations were analysed. The analysis identified 192 SSIs, 76 CAUTIs, 26 CLABSIs, and 39 cases of VAP, resulting in an infection rate of 4.3 SSIs per 100 operations, 11.0 CAUTIs per 1000 urinary catheter days, 6.1 CLABSIs per 1000 central line days, and 11.0 VAPs per 1000 ventilator days [24].

The resulting SIR for the SSIs in the Chiang Mai study was 2.3, with the SIR for CAUTI equal to 2.1, CLABSI equal to 1.1 and for VAP the SIR was 0.8. The factors associated with the prediction for SSIs were the duration of the operation in minutes, American Society of Anesthesiologists (ASA) score, and the degree of wound contamination. In conclusion, all of the SIR results identified, except the SIR for VAP, were above the average NNIS results and above a SIR of 1 [24].

A book, *Disease Control Priorities in Developing Countries*, was published based on analyses of diseases within developing countries. The SIR mathematical model was applied for the prediction of the HIV/AIDS patients population and it was concluded that the numerical results obtained from the model are expressing the trend of the exact data and this confirms a good fit of the model [16]. After a thorough analysis and investigation it may be concluded that the

SIR model developed and applied can predict the number of infected individuals, however, the model is sensitive to fluctuations in real data [16].

The SIR statistics do have certain limitations which are regarded as 'not serious' limitations and the SIR remains the most effective and trustworthy statistic available for all risk-adjustment purposes within infection control practices in hospitals. SIR is the most simplistic risk-adjustment statistic available for comparison over a period of time within hospitals as well as across multiple hospitals [14]. According to an article written by Tracy Gustafson from Cambridge University, the three main reasons to favour the SIR above an alternative statistical method for standardising rates and comparing rates within hospitals over time are that

1. The SIR provides a more reliable estimate of the infection rate when "smaller" denominators or numerators are calculated or reported, therefore instances where there are fewer actual infections at a hospital or fewer infections predicted for a specific hospital.

2. The mix of patients and the type of cases from a specific hospital rarely change over time, apart from random seasonal fluctuations or outliers [14].

The study conducted at Cambridge University emphasises the excellent job the SIR does of risk adjusting for the different procedures over a certain period of time. For example, at Hospital Y and Hospital Z for the months of June and July there were exactly an equal amount of 200 operative procedures completed as well as exactly the same SSI rate of 6.0%. The risk of the procedures performed, however, changed from 80 procedures of a low risk in June to 80 procedures being regarded as a high risk in July. The SIR adjusts the difference in risk between July and June accordingly and the SIR declines from 1.77 in June to 1.00 in July. This concludes that even though there were many more high risk operative procedures performed in July, the sum of the number of infections remained the same [14].

2.3 Logistic Regression

Logistic regression is a statistical method used when the dependent variable is a binary classifier, *i.e.* one wants to predict an outcome or response that is either a value of 1 ("yes") or a value of 0 ("no"). In its simplest form, the outcome is dependent on one independent variable only. Let the binary outcome variable be Y and the independent variable be X, then one wants to model the conditional probability $p(Y = 1 \mid X = x)$ as a function of x. In other words, what is the probability that the outcome is "yes", given that the independent variable has a value of x? This may be modelled by the so-called *log odds*.

Let p be the probability of an outcome of 1 ("yes"), then the probability of an outcome of 0 ("no") is $(1 - p)$. The ratio $p/(1 - p)$ is called the *odds* and the *log odds*, or in short the *logit*, is the logarithm of the odds. Thus, the logit function is

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right).$$

The simple logistic regression model is then

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X,$$

where the parameters β_0 and β_1 are called the *regression coefficients*.

After the regression coefficients are estimated using an existing data set, solving the regression model, one obtains the probability p of a positive occurence of the outcome, given that the independent variable X has the value x.

Logistic regression may be expanded such that the binary classifier are dependent on multiple independent variables. Again the conditional probability $p(Y = 1 \mid \mathbf{X} = \mathbf{x})$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is the values of the n independent variables, are modelled through the logistic regression model

$$\ln\left(\frac{p}{1-p}\right) = \text{logit}(p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n. \tag{2.1}$$

Again, p is the probability of a positive occurence of the outcome, given that the independent variables have the values $X_1 = x_1$, $X_2 = x_2$ until $X_n = x_n$.

CHAPTER 3

Data Collection

For this study a collective group of data was obtained from the healthcare facilities Head Office in Stellenbosch, South Africa. After a discussion with various computer and data scientists, clinical professionals and data warehousing specialists who are employed at the healthcare facilities Head Office, a specific group of variables were selected and data were obtained through and from the data warehouse department and team.

Hospitals collect a wide variety of categorical and numeric data on a daily basis and make use of information systems and tools for data collection and reporting purposes. Financial data (eg. billing and cost), human resource data (eg. hours worked, leave, vacancies and absenteeism) and clinical data (eg. demographics, the number of infections, surgical procedures, pharmaceutical prescriptions and patient experience surveys) as well as much more detailed data on specific diagnoses and procedures per patient within healthcare units are recorded. Every event that occurs on or for a patient within a hospital from admission and from there onwards, needs to be captured and recorded, so that these data may be analysed for reporting and improvement purposes.

Section 3.1 explains the admission process, hospital stay and discharge process of a patient and all the necessary data that is recorded and captured on a patient level. This is followed in section 3.2 by the process of clinical coding classifications and an explanation of each operative procedure based on the clinical coding classifications. The identification of the predictive risk variables to be used in the logistic regression model are discussed in section 3.3. Finally, in section 3.4 some data exploration are explained.

3.1 Patient data

Once a patient is admitted into hospital for a routine procedure and the patient has been shown to his/her room, by the porter, all the vital signs are taken and recorded. At this stage of the patient's visit to the hospital, all the patient's demographic and personal information (age, gender, name, street address, marriage status) doctors' notes, previous and current medical conditions as well as vital signs (blood pressure, weight, height, blood glucose, cholesterol) have been recorded. These are all recorded on the patient's hardcopy file as well as captured on the electronic Patient Admin System (PAS) to be used at the healthcare facilities Head Office.

Before the patient goes to theatre, the patient is briefed on the procedure as well as the process going forward. The anesthetist visits the patient before surgery and makes his or her own personal recordings or notes, including the patient's ASA score. Once the patient is wheeled to theatre a lengthy recording and data collection process takes place. This includes all the clinical information regarding the patient's diagnosis and the procedure the patient will encounter. The diagnosis of the patient as well as the procedure performed on the patient is recorded through a clinical coding classification system, to be explained in section 3.2. The data that is collected throughout the process of the patient being admitted to theatre includes the time the patient enters and leaves the theatre, all the doctor and nurses involved in the procedure, the slate in which the doctor performs the surgery as well as everything billed to the patient during surgery.

The drugs, bandages and ethicals used on the patient are recorded for billing, legal and stock purposes. The billing data on a patient's account include the length of stay (general ward, high care or the intensive care unit), plasters, gauze, drugs and any equipment and pharmaceutical fees used on the patient relevant to his/her specific procedure.

3.2 Clinical coding classifications

Clinical coding of a patient's medical state may be divided into two sections namely, the International Statistical Classification of Diseases and related health problems (ICD) and the Current Procedural Terminology (CPT). ICD and CPT codes are used within healthcare facilities to code and record every procedure, disease and medical condition of a patient during their stay in hospital.

The ICD-10 is the diagnostic coding scheme that was accepted by the National Health Information System of South Africa (NHISSA) in 1996, where the 10 indicates version 10 of the ICD classification system being used. In January 2005, the ICD-10 diagnostic coding scheme was implemented for all healthcare facilities in South Africa. The National Department of Health and the Council for Medical Schemes supported the implementation of ICD-10 in both the public and private health sector.

The ICD-10 coding structure was introduced by the World Health Organisation (WHO) in 1948 and is made up of a series of international classifications of the diagnoses or diseases. Comparison of ICD-10 coding is permitted as the ICD-10 structure is regarded as the international classification method and should be the coding system being used worldwide.

The second set of medical clinical codes, known as Current Procedural Terminology (CPT) codes, are used to capture and report clinical information regarding a patient and all surgical procedures a patient may undergo. These codes are numerical [7]. For example, all CPT codes for anesthesia range between 00100 and 01999 and 99100 and 99150. Another example, emphasizes that within each range of CPT codes are codes for various body parts such as head (00300 and 00352), neck and thorax (00400 and 00474). The CPT codes are used in line with the ICD codes.

For example, suppose a patient is admitted into hospital for a routine tonsillectomy procedure. This patient is first diagnosed with tonsillitis and will be recorded with an ICD-10 classification. Suppose the ICD-10 code is equal to J03.90 which may be described as *acute tonsillitis, unspecified*. When the patient is admitted for surgery, the CPT code recorded against the patient's account will equal 42820 (*tonsillectomy and adenoidectomy; under age 12*) or 42821 (*age 12 or over*).

The operative procedures in the CPT coding classification system may be grouped into operative procedure groups, where an operative procedure group consists of a group of surgeries or procedures which are grouped together based on CPT coding and the type of surgeries. For example, the operative procedure group HYST, is made up of twenty one CPT codes mainly consisting of hysterectomies and surgeries which share similar characteristics to a hysterectomy. Other procedures in the operative procedure group HYST are laparoscopic hysterectomies, and/or removal of ovaries or tubes, with or without lymphadenectomies and exenteration. An extract of the CPT codes that fall under the operative procedure HYST can be found in Table 3.1 below, with the complete list in Table A.1 and Table A.2 in Appendix A.

Operative Procedure Group HYST	
CPT Codes	**Code Description**
58150	Total abdominal hysterectomy (corpus and cervix), with or without removal of tube(s), with or without removal of ovary(s)
58152	Total abdominal hysterectomy (corpus and cervix), with or without removal of tube(s), with or without removal of ovary(s); with colpo-urethrocystopexy (eg, Marshall-Marchetti-Krantz, Burch)
58180	Supracervical abdominal hysterectomy (subtotal hysterectomy), with or without removal of tube(s), with or without removal of ovary(s)
58200	Total abdominal hysterectomy, including partial vaginectomy, with para-aortic and pelvic lymph node sampling, with or without removal of tube(s), with or without removal of ovary(s)
58210	Radical abdominal hysterectomy, with bilateral total pelvic lymphadenectomy and para-aortic lymph node sampling (biopsy), with or without removal of tube(s), with or without removal of ovary(s)
58240	Pelvic exenteration for gynecologic malignancy, with total abdominal hysterectomy or cervicectomy, with or without removal of tube(s), with or without removal of ovary(s), with removal of bladder and ureteral transplantations, and/or abdominoperineal resection of rectum and colon and colostomy, or any combination thereof

Table 3.1: Some CPT codes from the operative procedure group HYST

Within this study five operative procedure groups will be analysed and a SSI SIR model will be built for each operative procedure group. These groups are HYST as described above as well as BILI, SB, KPRO and CARD and the CPT codes for each operative procedure group can be found in Appendix A. The operative procedure group BILI consists of any open and laparoscopic liver, bile duct and pancreatic surgery, including resections, excisions, ablations, biopsies, repairs and ostomies, while the operative procedure group CARD consists of operative procedures on the heart valves or septum and does not include coronary artery bypass graft, surgery on vessels, heart transplantation, or pacemaker implantation. The operative procedure group SB includes all CPT codes describing any open or laparoscopic small bowel surgeries and resections including revisions and repairs, ileostomy and jejunostomy. Finally, the operative procedure group KPRO includes arthroplasty of the knee.

3.3 Predictive Risk Variables

A general list of *predictive risk factors* or *variables* was obtained through literature (see for example [22]) and guidelines provided by the NHSN, where predictive risk factors are variables that will aid in predicting the SSIs per operative procedure group. There are approximately twenty five unique predictive risk variables across all operative procedure groups, however, different combinations of the risk variables are used together per operative procedure group. Specific variables have a stronger association to a specific operative procedure group than other variables. Specific combinations of the predictive risk factors are grouped under each operative procedure group.

A discussion was held at the healthcare facilities Head Office between the researcher and medical doctors. The medical doctors provided valuable clinical insight on additional predictive risk variables that should be included when predicting SSIs. For example, the medical doctors stated that certain hospitals treat similar patients based on the patients diagnosis or procedure. The variable, **location code band** groups hospitals with similar patient profiles together. This variable was not suggested by the NHSN, but is a healthcare facilities specific variable.

The predictive risk factors or variables associated with the specific operative procedures studies in this research are diabetes, medical school affiliation, location code band, hospital bed size, scope, procedure duration, Body Mass Index (BMI), oncology, gender, age, anesthesia, emergency, trauma, wound closure technique and ASA. Most of these information from every patient that is admitted into a hospital are recorded on the Patient Admin System (PAS).

Type 2 **Diabetes** is a chronic condition that affects the way the body processes blood sugar (glucose). Type 1 Diabetes is a chronic condition in which the pancreas produces little or no insulin. Each patient admitted into a healthcare facility completes a clinical examination form which asks whether the patient is a diabetic or not. Therefore, a diabetic patient is flagged as 1 and a patient who does not have diabetes is flagged as 0.

The predictive risk factor **medical school affiliation** refers to whether or not a hospital is regarded as a teaching or academic hospital. Within the healthcare facility, Wits Donald Gordan Medical Centre provides academic support to the University of Witwaterstrand and is therefore the only hospital within the healthcare facilities group that is affiliated with a medical school.

The **location code band** variable refers to the grouping of certain hospitals with similar characteristics into a band. These characteristics include the type of patients and cases the hospital admits as well as the level of care the patients receive at the hospital (for example, intensive care, trauma and general).

Each hospital consists of a certain amount of beds, referred to as **hospital bed size**. The variable **scope** refers to whether a patient had laproscopic surgery or not. If a patient is billed for laproscopic surgery via chargemaster codes, they are regarded as having had a scope done, while **procedure duration** indicates the amount of minutes the patient spent in theatre during their surgery.

Body Mass Index (BMI) is the quotient of a patient's weight in kilograms (kg) and a patient's height to the power of two in metres (m^2). The patient's BMI is calculated during the admission process. **Oncology** indicates whether a hospital's sole focus is on oncology patients. Within the healthcare group, there are no solely focused oncology hospitals.

During the admission process the patients must indicate whether they consider themselves either as male or female and provide their date of birth, which is then captured in the **gender** and **age** variable, respectively. The binary variable **anesthesia** indicates whether the patient had received any form of drug or anesthetic, where all patients who went to theatre have a binary variable equal to 1 (Yes).

If a patient arrived as an emergency patient through the emergency centre and was then admitted to a hospital, it is indicated in the **emergency** variable. There are a list of clinical CPT codes, which are specific procedural codes linked to trauma cases. Therefore, if a patient has one of the trauma CPT codes on their patient file they are flagged as a **trauma** patient.

The **wound closure technique** is the technique how a laceration was repaired. These techniques include sutures, staples, adhesive tapes, or tissue adhesives. The wound closure technique is a poorly captured variable at healthcare facilities.

Finally, the American Society of Anesthesiologists (ASA) classification of Physical Health is a widely used grading system for preoperative health of the surgical patients. In a hospital setting, an anesthesiologist will visit the patient before they are admitted to theatre and grade the patient on an ASA scoring level of 1 to 6, 1 being scored to a healthy patient with no complications and 6 being scored to a patient with a very high mortality rate. Patients who are graded with an **ASA score** equal to 6 are typically organ donors and deceased. There are multiple variations between ASA scoring and an anesthetists assessment of a patient pre-surgery.

An ASA score per patient is regarded as a significant variable by the NHSN, to accurately calculate the predicted HAIs and more specifically SSIs for a patient and per hospital [9]. Since there is currently no data on ASA scores being captured in the electronic PAS for record keeping at healthcare facilities, a method to obtain ASA scores was developed for the purpose of this study.

3.3.1 ASA Logic

The logic behind the development of an ASA scoring system is available on online articles published on the ASA website [2]. These guidelines for an ASA scoring grade a patient from 1 to 6 based solely on the ICD-10 coding, thus the diagnosis that may lead to an operative procedure. Therefore, depending on the ICD-10 code or codes on a patient's account or file, an ASA score may be determined.

The more at risk of death a patients is when being admitted for a procedure in theatre, the higher their ASA score. The biggest portion of patients, being admitted for an operation, should be scored with an ASA score equal to 1, as they are normal healthy patients. The second biggest portion of patients, with mild diseases, will score an ASA score of 2. Patients with a severe disease will have an ASA score equal to 3, while patients with an ASA score of 4 or 5 are higher risk and more sick patients. The smallest portion of patients should score an ASA score of 6 as these patients are extremely sick and high risk patients (e.g. brain dead patients who are admitted for organ donor purposes) [2]. At the heath facilities Head Office, *clinicians* (i.e. doctors, nurses, health practitioners, etc.) and clinical coders provided insight and information on how to develop an internal ASA score for the healthcare group.

The first iteration of developing an ASA score entailed obtaining all the ICD-10 codes on a patient's account and grading the patient from 1 to 6 based on the specific ICD-10 coding of the patient. Based on the literature provided by the NHSN and the American Society of Anesthesiologists, certain patient profiles resulted in fixed ASA scoring without having to take into account additional factors of the patient [9]. For example, all maternity patients are to be scored with an ASA score equal to 2 based on their ICD-10 coding and are therefore classified as healthy low-risk surgical patients. The logic developed, stated that patients who did not have any clinical ICD-10 code linked to an ASA score of 2 to 6 were scored with an ASA score equal to 1.

The second iteration of developing the ASA score entailed determining whether a patient who scored an ASA equal to 1, also had a complication or *comorbidity* recorded on their account, where a comorbidity is the presence of a medical disorder in addition to their medical diagnosis. These patients would be graded an ASA score equal to 2 as a comorbidity or complication may result in the patient being a diabetic or having high blood pressure which increases the operative risk when the patient is going to theatre.

3.3.2 ASA Audit

Due to the importance of the ASA score in terms of the treatment of a patient, a coding audit process is conducted every four months on a sample of patient files at the healthcare facilities Head Office in Stellenbosch. The coding audit process entails auditing patient's files to ensure that the correct ICD-10 and CPT codes were recorded on a patient's file in relation to the procedures, medical conditions and or diseases the patient experienced during their hospital stay. The main aim of the audit process is to ensure and encourage correct coding practices at hospitals.

Within each patient's file is a hard copy page which is called "Anesthetist's Notes". This is where anesthetists write down any details or information about their patient before or whilst in surgery, as well as a specific section for an anesthetist to give their patient an ASA score. A total of 152 patient files that were readily available at the healthcare facilities Head Office were selected for an ASA auditing process by the researcher for this project. The patient files may range from 5 A4 pages up to 100 or more pages, depending on the complexity of the patient and their medical prognosis. The Anesthetist Notes which is usually 1 or 2 pages, was removed from the patient's files since this contains all the information to determine a patient's ASA score as indicated by the Anesthetist. This information is used by the researcher in connection with clinicians to determine what the ASA score should be, based on the ICD codes, etc. These ASA scores are then compared to the ASA score recorded on the "Anesthetist's Notes". In some case the anesthetist did score and record their patient's ASA score, while in other cases two clinical nurses scored the patient based on their medical condition and prognosis in their file. From the total of 152 patient files, there were 51 files for which neither the anesthetist nor the nurses recorded the patient's ASA score in the formal notes.

Of the 101 files for which the ASA scoring could be audited, 83% of the patient files had a perfect match between the ASA score recorded on the patient's file and the ASA score assigned during the auditing process. The 17% of patient files audited that did not have a perfect match were analysed on a per patient level in order to determine the reason for the mismatch. It was found that in the mismatched cases, the ASA score for the patient were higher than the ASA score recorded in the anesthetists notes in the patient file. For example, many maternity patients were scored an ASA score equal to one by their anesthetist, however, literature states that all healthy maternity patients should receive an ASA score equal to two.

During the once-off auditing process by the researcher, the ASA score were also recorded electronically, which had not been done previously. The ASA score per patient provided a new clinical indicator to be reported on and enrich the quality of reporting and improvements throughout the hospitals. It has now also become compulsory for anesthetists to record an patient's ASA score before an operative procedure. In conclusion, the final list of predictive risk variables that will be used in the logistic regression model in the research are: ASA, diabetes, medical school affiliation, location code band, hospital bed size, procedure duration, body mass index, oncology, gender, age, anesthesia, emergency and trauma.

3.4 Data Exploration

Once the data collection has been completed, the exploration of the data began. Data Exploration may be described as the process of understanding, exploring and analysing a certain dataset consisting of specific variables over a period of time. The analyses may include interactive graphs or visual representations of the data. Graphs and visual representations allow the researcher to better understand the dataset and better understand specific variables and the relationships between certain variables. Whilst exploring the data, the researcher made use of analytic tools and software consisting of a graphical user interfaces [29]. Throughout this research project the statistical software SAS is used.

Initially, the data was divided into a training and validation dataset. A *training data set* may be described as a portion or sample of the main dataset, where the data is used to fit the model. The *validation dataset* is the sample or portion of data that is used to provide an unbiased evaluation of a model fit to the training dataset. The training dataset consists of 70% and the validation dataset consists of 30% of the total dataset. The initial dataset was divided randomly in SAS, and the process is given in pseudo code in Algorithm 3.1. The binary variable called `Build` in Algorithm 3.1 is used to assign each row (patient) to either the training dataset or the validation dataset. A random seed number equal to 2019. The `ranuni` function in Step 1 returns a number that is generated from the uniform distribution on the interval (0,1). All patients for which a random number less than or equal to 0.7 are generated, are given a value of 1 for `Build` and these are the patient data that forms the training data set.

Algorithm 3.1 Training and Validation Data Split

Input: Patient Data with all data related to one patient in one row.

Output: The data set with a value for the flag `Build` for each patient used to indicate whether this patient's data will be used in the training set or the validation set.

1: For each patient a random number by using the `ranuni` function is obtained..
2: `If` the random number is less than or equal to 0.7 then `Build` = 1.
3: `If` the random number is greater than 0.7 then `Build` = 0.
4: All patients that are flagged as 1 within the `Build` column of the dataset are part of the training dataset.
5: The training and validation datasets are used to calculate the predicted SSIs per operative procedure, hospital and year. Therefore, where `Build` = 1 and 0.

It is also necessary to determine which variables within the dataset were either categorical or continuous. The parameter estimate from the logistic regression will reflect the nature of the relationship between the variable and the risk of an SSI. In the case of a categorical variable, the risk of the SSI in an individual category is compared to the risk of an SSI in the "referent" category. A positive parameter estimates indicates that the risk of an SSI in that category is higher than the risk of SSIs in the referent category. Whereas, a negative parameter estimate indicates that the SSI risk in that category is lower compared to the SSI in the referent category [10]. By default, SAS Viya predefines and distinguishes each variable between either a categorical or continuous variable. The continuous variables consisted of hospital bed size, procedure duration, Body Mass Index (BMI) and age. Whereas, diabetes, medical school affiliation, location code band, scope, oncology, gender, anesthesia, emergency, trauma, wound closure technique and ASA score are categorical variables.

In order to fully understand the variables within the dataset, *before* modeling the prediction of HAIs are done, different graphic techniques were used to understand the trends as well as frequencies across certain risk variables, periods, hospitals and regions within Southern Africa. Although these results are not incorporated into the model, it allows the researcher to make sense of and validate the final results from the HAI predictive model.

For this purpose two types of data exploration analyses will be done. Firstly, various relationships and characteristics of specific variables were visually explored through an analytics software program called Model Studio within SAS Visual Analytics (Viya) as discussed in §3.4.1. Secondly, the researcher analysed the frequency of some values of certain variables in SAS Enterprise Guide as illustrated in §3.4.2.

3.4.1 Exploring combinations of risk variables

Several combinations of risk variables such as ASA and Hospital Bed Size were manually selected and their interactions studied. In Figure 3.1 and Figure 3.2 are examples of two of the results obtained from Model Studio within SAS Visual Analytics within the operative procedure group HYST.

In the first case, in Figure 3.1, a numeric figure of 0.8 indicates that the 265 patients who had an ASA score greater than or equal to 3, but less than 6, who attended a hospital which was also a Medical School, while their minutes in theatre was greater or equal to 1200 minutes (20 hours) and who are less than 46 years old, would result in a probability of obtaining a SSI of 0.8. The second case depicted in Figure 3.1 indicates that there is a 0.41 probability that 286 patients would contract a SSI when these patients had an ASA score greater than 3, but less than 6, went to a hospital which also served as a Medical School, were in theatre for more than 20 hours and who were older than 45 years of age.

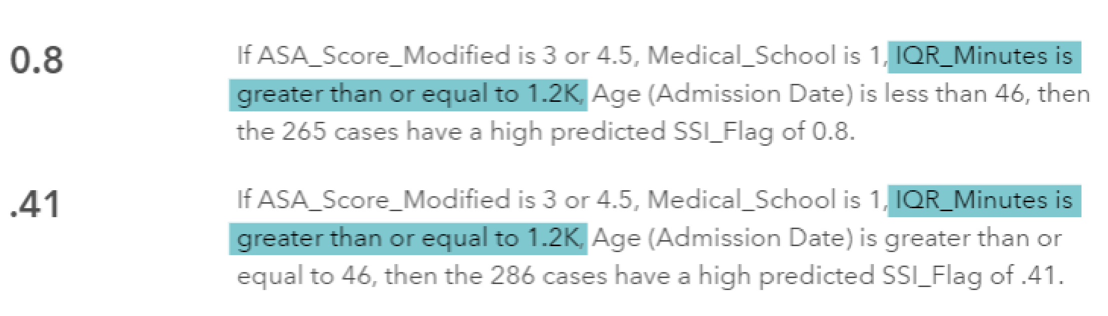
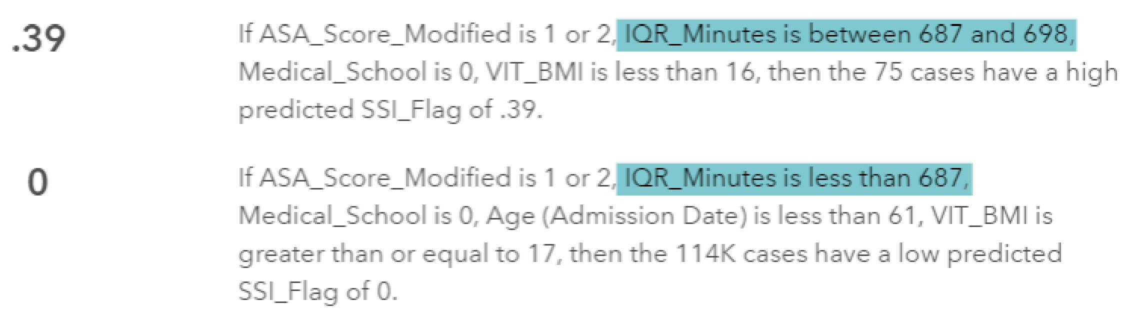


Figure 3.1: Results of exploring risk variables in the HYST operative procedure group

In Figure 3.2, the first case equal to 0.39 indicates that 75 patients who scored an ASA score equal to either 1 or 2, who were in theatre for between 687 and 698 minutes, who were not admitted to a hospital of medical school affiliation and all the patients had a BMI of less that 16, would result in a high probability of contracting an SSI of 0.39. The second case, in Figure 3.2, equal to 0 states that there were 114 000 patients that has an ASA score of 1 or 2, were in theatre for less than 687 minutes, were not admitted to a hospital with medical school affiliation, all the patients were less than 61 years of age and had a BMI of greater or equal to 17 would result in the very low predicted probability 0 of obtaining an SSI.



.39 — If ASA_Score_Modified is 1 or 2, IQR_Minutes is between 687 and 698, Medical_School is 0, VIT_BMI is less than 16, then the 75 cases have a high predicted SSI_Flag of .39.

0 — If ASA_Score_Modified is 1 or 2, IQR_Minutes is less than 687, Medical_School is 0, Age (Admission Date) is less than 61, VIT_BMI is greater than or equal to 17, then the 114K cases have a low predicted SSI_Flag of 0.

Figure 3.2: Results of exploring risk variables in the HYST operative procedure group

3.4.2 Study frequencies in risk variables

Figure 3.3 is an example of the analyses of the variable Age across all patients within the operative procedure group HYST. Within the operative procedure grouping HYST, the highest frequency of the variable age ranged between 45 and 48 years old.

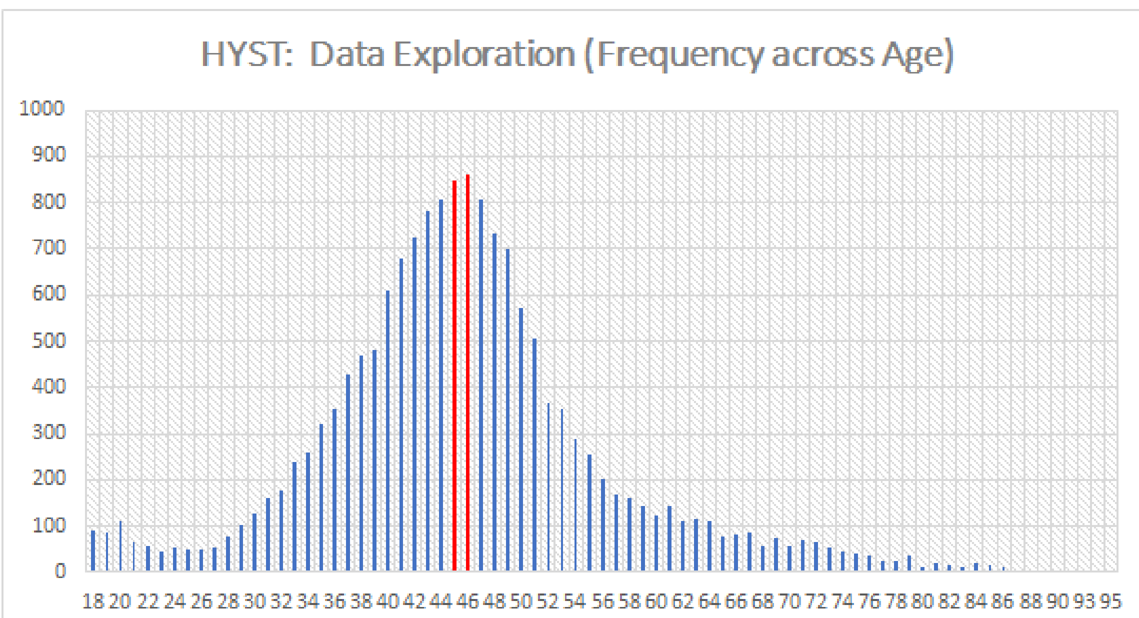


Figure 3.3: Histogram illustrating the frequency of the age of the patients in the operative procedure group HYST

Figure 3.4 is an example, within the operative procedure group HYST, of the variable ASA score across all patients. The highest frequency of patients of the variable ASA score is where the ASA score is equal to 2, followed by the second highest frequency of ASA score equal to 3 and the third highest frequency of ASA score equal to 1. The ASA score of 4 and 5 is grouped together, due to the individual low frequencies, and an ASA score equal to 4 and 5 has the lowest frequency across the operative procedure group HYST.

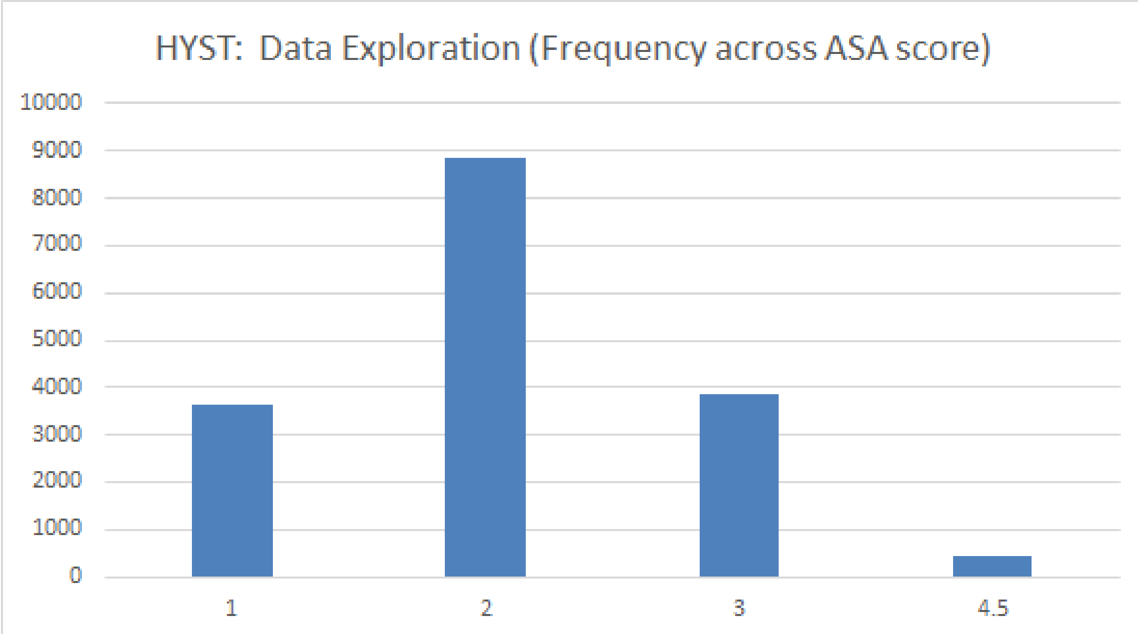


Figure 3.4: Frequency across ASA score of the patients in the operative procedure group HYST

3.4.3 Final data file

The final data file contain all the raw data collected, consisted of all theatre or surgical cases throughout healthcare facilities in South Africa for a period of 3 years as well as the values of predictive risk variables. Table 3.2 illustrates the total number of patients per operative procedure group, which make up the final data file. This file will be the input file for the logistic regression model and an extract of the data are distributed over Table 3.3 and Table 3.4.

	Number of patients:
HYST	13691
BILI	1687
SB	2163
CARD	3141
KPRO	12199

Table 3.2: Total number of patients per operative procedure group

Patient Key	Hospital	Year	CPT Code	ICD Code	Age	BMI	Diabetic	Emergency	Days in Hospital	Gender
171144797	Hospital A	2016	58150	D25.9	63	29	0	0	3	F
143606836	Hospital A	2016		K01.1	37	37	1	0	1	M
532419454	Hospital A	2016		K01.1	32	24	0	0	1	M
96077572	Hospital A	2016	58150	T81.0	33	33	0	1	7.5	F
422892482	Hospital A	2016	58150	N93.9	39	29	0	1	9	F
910162516	Hospital A	2016	58150	D25.1	46	33	0	0	4	F
692111099	Hospital A	2016	58150	D25.0	67	25	0	0	3	F
534130404	Hospital A	2016		I62.0	64	26	0	0	2.5	M
511091966	Hospital A	2016		K01.1	30	31	0	0	1	M
104198500	Hospital A	2016	58150	N92.1	49	29	0	0	4.5	F
870668083	Hospital A	2016	58150	D25.1	68	38	0	0	3.5	F
35858951	Hospital A	2016	58180	N81.3	67	35	0	0	3.5	F
761045954	Hospital A	2016	58150	D25.9	51	37	0	0	3	F
178936165	Hospital A	2016	58570	D26.1	50	38	0	0	2.5	F
126115289	Hospital A	2016	58150	N85.1	46	25	0	0	3.5	F
307107307	Hospital A	2016		I25.9	88	25	0	0	3.5	F
278474224	Hospital A	2016	58150	D25.1	46	33	0	0	3.5	F
96617359	Hospital A	2016	58150	D27	55	25	0	0	3	F
231105983	Hospital A	2016	58150	D26.1	53	26	0	0	3	F
647369918	Hospital A	2016	58150	D06.9	66	37	0	0	3.5	F
886335695	Hospital A	2016	58150	D25.1	54	21	0	0	3.5	F
621472533	Hospital A	2016		I48.9	43	42	0	0	1	M
992945167	Hospital A	2016		K01.1	29	42	0	0	1	F
289565429	Hospital A	2016		J96.99	60	33	0	1	25	F
832376689	Hospital A	2016	58150	D25.0	47	27	0	0	3.5	F
325005653	Hospital A	2016	58150	N92.0	43	42	1	0	3.5	F
495784372	Hospital A	2016	58150	N92.1	39	44	0	0	2.5	F
535705562	Hospital A	2016	58150	D25.1	44	25	0	0	3.5	F
785570870	Hospital A	2016	58150	D25.1	44	22	0	0	3.5	F

Table 3.3: An example of the data collected per patient and predictive risk variables for the logistic regression model

Patient Key	Procedure Duration	Operative Procedure Group	Hospital Bed Size	Oncology Hospital	Anesthesia	Scope	ASA Score	Trauma	Medical School	SSI
171144797	71	HYST	120	0	1	0	2	0	0	0
143606836	52	HYST	120	0	1	0	2	0	0	0
532419454	55	HYST	120	0	1	0	2	0	0	0
96077572	191	HYST	120	0	1	0	4.5	1	0	0
422892482	95	HYST	120	0	1	0	3	0	0	0
910162516	74	HYST	120	0	1	0	2	0	0	0
692111099	86	HYST	120	0	1	0	2	0	0	0
534130404	103	HYST	120	0	1	0	4.5	0	0	0
511091966	69	HYST	120	0	1	0	2	0	0	0
104198500	103	HYST	120	0	1	0	2	0	0	0
870668083	83	HYST	120	0	1	0	2	0	0	0
35858951	91	HYST	120	0	1	0	2	0	0	0
761045954	92	HYST	120	0	1	0	2	0	0	0
178936165	119	HYST	120	0	1	0	2	0	0	0
126115289	93	HYST	120	0	1	0	1	0	0	0
307107307	22	HYST	120	0	1	0	4.5	0	0	0
278474224	92	HYST	120	0	1	0	2	0	0	0
96617359	88	HYST	120	0	1	0	1	0	0	0
231105983	70	HYST	120	0	1	0	2	0	0	0
647369918	78	HYST	120	0	1	0	2	0	0	0
886335695	94	HYST	120	0	1	0	3	0	0	0
621472533	33	HYST	120	0	1	0	4.5	0	0	0
992945167	59	HYST	120	0	1	0	3	0	0	0
289565429	30	HYST	120	0	1	0	4.5	1	0	0
832376689	111	HYST	120	0	1	0	1	0	0	0
325005653	133	HYST	120	0	1	0	3	0	0	0
495784372	108	HYST	120	0	1	0	3	0	0	0
535705562	68	HYST	120	0	1	0	1	0	0	0
785570870	87	HYST	120	0	1	0	3	0	0	0

Table 3.4: An example of the data collected per patient and predictive risk variables for the logistic regression model

CHAPTER 4

Methodology

The statistical method, logistic regression, was selected to predict SSIs within the hospitals. The specific logistic regression model used in this study are described in §4.1. The predicted SSIs calculated by the logistic regression model and the actual SSIs that are reported on per hospital were used to calculate the SIR per operative procedure group per hospital as discussed in §4.2.

4.1 Logistic Regression Model

For the purpose of this study, the binary output variable in a logistic regression model as described in §2.3, is to state whether patient i contracted an SSI or not. Let Y be the response (dependent) variable and let $Y = 1$ if the particular patient obtain or contract an SSI and $Y = 0$ otherwise. A seperate logistic regression model was developed for each operative procedure group. The researcher therefore aims to determine the conditional probability $p_i(Y = 1|\mathbf{X}_i = \mathbf{x}_i)$, where the independent variable $\mathbf{X}_i$ is specific to the operative procedure group and have n significant variables. The value x_{i1} of the variable

$$\mathbf{X}_i = (x_{i1}, x_{i2}, ..., x_{in}) \tag{4.1}$$

is the specific value of patient i for the first significant risk variable in the model.

For example, say the only two significant risk variables in the operative procedure group HYST are BMI and ASA score, then the researcher is interested in the probability $p_i(Y = 1|\mathbf{X}_i = (29, 2))$ that patient i with a BMI of 29 and an ASA score of 2, will contract an SSI.

The logistic regression model per operative procedure group is thus

$$\text{logit}(p_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_n X_{in}, \tag{4.2}$$

where β_0, β_1,...,β_n are the regression coefficients of the significant independent risk variables for the particular operative procedure group and p_i is the probability that patient i will contracted an SSI. Each risk factor are associated with a parameter estimate which describes the relationship between the significant variable and the risk of a patient contracting a SSI. A positive parameter estimate in §4.2 indicates that the risk of a SSI occurring increases with increasing the value of the variable [9].

The process of determining the significant variables $\mathbf{X}_i$ per operative procedure group to be used in the model to predict SSIs, is illustrated in §4.1.1. The Goodness of Fit test between the actual and predicted SSI values per operative procedure group follows in §4.1.2.

4.1.1 Variable Selection

If several independent variables are identified as having a possible effect on the dependent variable, first, these variables need to be investigated to determine whether they do have a significant effect on the dependent variable. If an independent variable has a significant effect on the outcome of the dependent variable, this significant variable should be included in the logistic regression model. On the other hand, if an independent variable does not have a significant effect on the outcome of the dependent variable, this independent variable may be excluded from the logistic regression model.

In this study, possible independent risk factors per operative procedure were first selected from the literature since these factors were found to be significant variables based on the NHSN data of the population of the United States used for the NHSN's logistic regression model. However, the data used for the logistic regression model in this research project consisted of patients and their profiles within the South African private healthcare facilities. Certain variables were excluded and certain variables were included based on the significance of the variables and/or clinical insight from doctors and clinicians employed at the hospitals as described in §3.3. The final list of risk variables considered in this research project was also given in §3.3.

To investigate which risk variables per operative procedure group must be selected to be included in the logistic regression model to predict SSIs, a series of logistic regression models with one independent variable may be fit. Therefore, if the variable X_j in

$$\text{logit}(p) = \beta_0 + \beta_j X_j, \qquad j \in [1, 2, ..., n], \tag{4.3}$$

has an associated p value of less than 0.05, then X_j is significant and will be included in the logistic regression model to predict SSIs for the specific operative procedure.

The above process of selecting significant variables per operative procedure group can be done simultaneously in SAS. Figure 4.1 indicates the variables selected to be analysed in the operative procedure group HYST. Since a risk variable with an associated p value ≤ 0.05 is regarded as significant, the final logistic regression model for the operative procedure HYST will include four significant variables, which are Location Code Band, BMI, Diabetes and ASA score. Note that the p value, given in the last column, of medical school affiliation is much greater than 0.05 and will thus not be included in the final logistic regression model.

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Loc_Code_Band	3	30.7655	<.0001
BMI	1	13.3450	0.0003
Diabetic	1	5.5159	0.0188
ASA_Score_Modified	3	22.2399	<.0001
Medical_School	1	0.0095	0.9223

Figure 4.1: Determination of significant risk variables for the logistic regression model for operative procedure group HYST

A comparison of the variables that have been found significant by the NHSN versus the significant variables obtained from SAS based on the healthcare facilities data per operative procedure group can be found in Table 5.1 in Chapter 5.

4.1.2 Hosmer-Lemeshow Goodness-Of-Fit Test

The main purpose of any goodness-of-fit test is to determine whether the fitted model adequately describes the observed outcome experience in the data or follows a similar count to the actual SSIs being reported on. Thus, after a logistic regression model has been fitted, a global test of the goodness-of-fit of the resulting model should be performed [1]. The Hosmer-Lemeshow test calculates if the observed event rates match the expected event rates across the population subgroups or *ranks* [1]. The data, on a per patient level, is *ranked* based on their predicted *probability* of contracting an SSI. A pre-defined function in SAS, **Proc Rank**, assigned each row of data to one of the 10 ranks. The 10 ranks do not have an equal amount of patients, because the function **Proc Rank** assigned patients (rows) with equal probabilities into the same rank. In other words, patients who had a tied probability of contracting an infection will be grouped into the same rank and therefore an unequal amount of patients are grouped per rank. The Hosmer-Lemeshow (HL) test is essentially a chi-squared goodness-of-fit test for data that are grouped, usually into 10 equal groups. The HL test is also done automatically in SAS.

To investigate whether a good fit between the model and the desired outcomes is obtained, the average predicted SSIs, *i.e.* the average probability of a patient contracting a SSI is compared against the average actual SSIs or the average number of SSIs contracted by a subgroup of patients across 10 ranks. The average predicted SSIs for each rank i, α_i, is calculated as

$$\alpha_i = \frac{\sum_j A_{ij}}{T_i} \qquad i \in [0, ..., 9], \tag{4.4}$$

where A_{ij} is the predicted SSI for patient j in rank i and T_i is the total number of patients in rank i. Similarly, the average actual SSI for each rank i, γ_i, is calculated as

$$\gamma_i = \frac{\sum_j B_{ij}}{T_i} \qquad i \in [0, ..., 9], \tag{4.5}$$

where B_{ij} is the actual SSI for patient j in rank i. The sum of the average predicted SSIs across all 10 ranks as well as the sum of the average actual SSIs across all 10 ranks will also be reported per operative procedure group in Chapter 5.

The computer software SAS Enterprise Guide Version 8.1 and the computer programming language SQL were used to develop a logistic regression model per operative procedure grouping. The results will be given in Chapter 5.

4.2 Calculation of a SIR

The same method as done by the NHSN will be followed to determine a SIR per operative procedure group per hospital. The NHSN first determines the probability p_i of an infection for each patient i from the logit function for each operative procedure group. Summing over the probabilities of an infection being contracted by all patients, across all procedures in a given time frame, will result in the total number of predicted infections.

The sum of all the infections contracted by all the patients as recorded in the data, over the same given time frame, is the total number of actual infections contracted. The SIR for the operative procedure group is the ratio

$$\frac{\text{total number of actual SSIs}}{\text{total number of predicted SSIs}}. \tag{4.6}$$

Hospitals which achieve a low or good SIR of below 1 will have less actual infections than predicted infections, while hospitals achieving a high SIR of above 1 will have a higher number of actual infections than predicted infections within the hospital. Therefore, if the model predicted less infections to occur than the actual amount of infections that occurred, then these results indicate that infection prevention and intervention techniques need to be focused on at that specific hospital and operative procedure grouping.

CHAPTER 5

Results

This chapter discusses the results of the logistic regression model per operative procedure group across three hospitals over a period of three years. For each individual operative procedure group the significant variables are given in §5.1, while the predicted versus actual SSIs are listed in §5.2 to §5.6. Finally, in §5.7 the SIR per hospital per year per operative procedure group are given.

5.1 Variable Selection results

In Table 5.1 a comparison between the predictive risk variables suggested by the NHSN, the significant variables based on a non-specified selection method, the significant variables based on a forward selection method and the variables found to be significant in a backward selection method in the logistic regression model per operative procedure group, are listed. A forward selection method may be described as the process of selecting significant variables one by one from an empty model. Whereas, a backward selection method is the process of beginning the selection process with all variables selected and removing the least significant variables one by one.

As indicated in Table 5.1 within the operative procedure groups HYST, BILI, SB and CARD there is a difference in the variables recommended by the NHSN and the variables found to be significant via the logistic regression model. When investigating the data, it was found that by grouping certain hospitals (location code band) and hospital sizes (bed groups) the p value decreased which resulted in significant variables as well as an improvement in the fit of the model. A possible reason for the variable, Bed Size, being a significant variable is because within the United States of America the different facilities, patient profiles and location settings across the country in hospitals is not notable and therefore not an important variable to take into account when developing a logistic regression model. In South Africa, however, the patient profiles differ across regions of the country as well as different infections are more susceptible in certain areas of the country than others. The size of the hospital (bed size) ranges from small, medium and large sized hospitals across Southern Africa. The larger the hospital the more the hospital is at risk of any infection. The larger the hospital will also result in a wider variety of patient types.

Operative Procedure:	NHSN:	Non-specified:	Forward:	Backward:
HYST	Diabetes, ASA, Medical School Affiliation, Hospital Bed Size, Scope, Age, Procedure Duration, BMI, Oncology	Diabetes, ASA, BMI, Medical School Affiliation, Location Code Band	ASA, BMI, Location Code Band	BMI, ASA
BILI	Gender, Emergency, Trauma, Wound Class, Hospital Bed Size, Scope, Age, Procedure Duration	ASA, Diabetes, Medical School Affiliation, Hospital Bed Size, Age, Procedure Duration	Diabetic, Age, Procedure Duration	Diabetic, Age, Procedure Duration
SB	ASA, Emergency, Trauma, Medical School, Oncology	ASA, Emergency, Trauma, Medical School, Diabetes, Location Code Band	ASA, Diabetic, Medical School, Emergency	ASA, Diabetic, Medical School, Emergency
CARD	Emergency, Medical School, Age, Procedure Duration, BMI	Emergency, Medical School, Age, Procedure Duration, BMI, Location Code Band, Diabetes	Diabetes, Procedure Duration, BMI	Diabetes, Procedure Duration, BMI
KPRO	ASA, Medical School, Bed Size, Scope, Age, Procedure Duration, BMI, Oncology	ASA, Medical School, Bed Size, Scope, Age, Procedure Duration, BMI, Oncology	Emergency, Procedure Duration, Bed Size	Emergency, Procedure Duration, Bed Size

Table 5.1: Variable Selection per operative procedure group

In this study, a non-specified variable selection method was used and the significant variables were based off the model specified. A final set of variables used in the model per operative procedure was used based off the non-specified variable selection method and clinical knowledge from healthcare practitioners.

5.2 Results of the operative procedure group HYST

A total of 4108 patients had an ICD code that, grouped them within the HYST operative procedure group. The results of the HL goodness of fit test as obtained from SAS are given in Figure 5.1. Since $p = 0.4796$ there is no evidence of a lack of fit in the model. The number of patients, the average probability of predicted SSIs and average number of actual SSIs per rank over the 3 years (2016, 2017 and 2018) are listed in Table 5.2 and graphically displayed in Figure 5.2. As seen in Table 5.2, the average probability of predicted SSIs and the actual SSIs follow the same trend across the 10 ranks, therefore confirming that the model and the data are a good fit. The highest count of patients fell within rank 3 and 5. The **Total** row in Table 5.2, indicates the sum of the number of patients and the average predicted SSIs and average actual SSIs across the total sum of all the patients.

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
5.5155	6	0.4796

Figure 5.1: HL goodness of fit test for operative procedure group HYST

HYST			
Rank	**Count of patients**	**Average of Predicted SSIs**	**Average of Actual SSIs**
0	608	0.00000	0.00000
1	252	0.00000	0.00000
2	71	0.00001	0.00000
3	692	0.00003	0.00145
4	335	0.00017	0.00000
5	858	0.00031	0.00000
7	278	0.00080	0.00000
8	516	0.00189	0.00194
9	498	0.01991	0.02209
Total	**4108**	**0.00279**	**0.00316**

Table 5.2: HYST: Predicted VS Actual SSIs

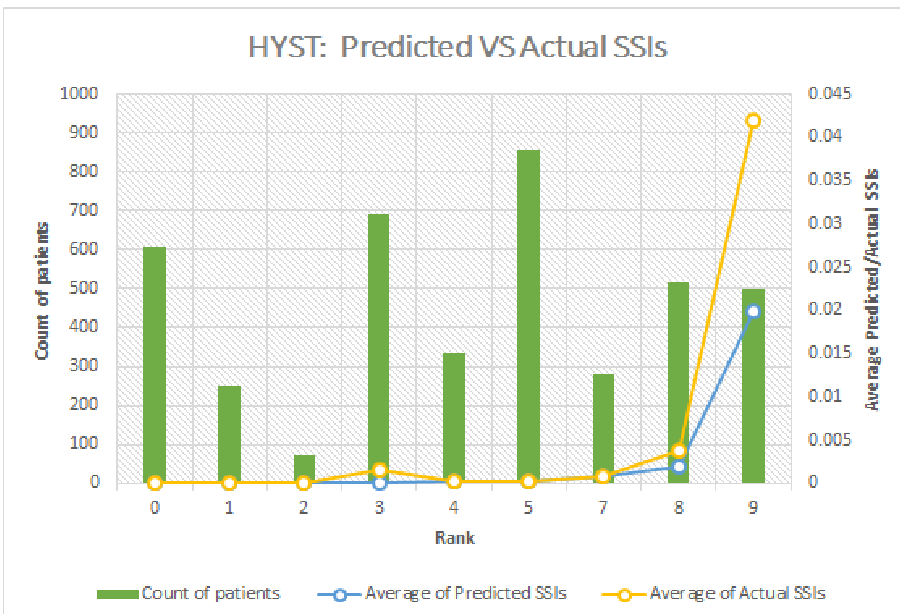


Figure 5.2: Predicted VS Actual SSIs for operative procedure group HYST

5.3 Results of the operative procedure group BILI

The operative procedure group BILI consists of 510 cases and an average predicted probability of contracting a SSI, across all ten ranks, equal to 0.10981. The average actual SSI equals 0.14706, seen in Table 5.3. In Figure 5.3, the HL goodness of fit test where, $p = 0.6464$, is displayed. Therefore this model also provides a good fit. The results of the average predicted probability of SSIs vs average actual number of SSIs may be viewed graphically in Figure 5.4. The **Total** row in Table 5.3, indicates the sum of the number of patients and the average predicted SSIs and average actual SSIs across the total sum of all the patients.

BILI			
Rank	**Count of patients**	**Average of Predicted SSIs**	**Average of Actual SSIs**
0	50	0.00821	0.00000
1	57	0.01789	0.05263
2	51	0.02837	0.07843
3	52	0.04139	0.13462
4	57	0.05838	0.08772
5	52	0.07761	0.07692
6	44	0.10423	0.13636
7	55	0.14093	0.14545
8	47	0.22102	0.27660
9	45	0.46413	0.55556
Total	**510**	**0.10981**	**0.14706**

Table 5.3: BILI: Predicted VS Actual SSIs

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
14.2917	17	0.6464

Figure 5.3: HL goodness of fit test for operative procedure group BILI

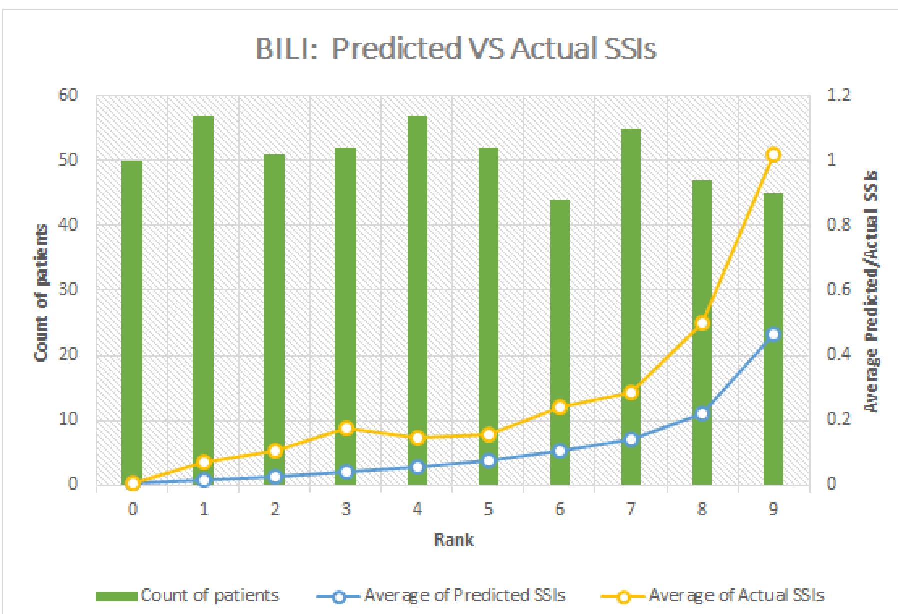


Figure 5.4: Predicted VS Actual SSIs for operative procedure group BILI

5.4 Results of the operative procedure group SB

The operative procedure group SB consisted of 662 patients. The results of the HL goodness of fit test is illustrated in Figure 5.5. The $p = 0.3240$ and therefore there is no evidence of a lack of fit of the model.

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
19.0788	17	0.3240

Figure 5.5: HL goodness of fit test for operative procedure group SB

The average probability of predicted SSIs and average number of actual SSIs followed a similar trend across rank 0 to rank 9, as seen in Table 5.4 and Figure 5.6. An outlier occurs at rank 4 where only 1 patient was randomly assigned and the model predicts a very low probability of SSI, however, the patient did actually contract an SSI. The **Total** row in Table 5.4, indicates the sum of the number of patients and the average predicted SSIs and average actual SSIs across the total sum of all the patients.

SB			
Rank	**Count of patients**	**Average of Predicted SSIs**	**Average of Actual SSIs**
0	72	0.00989	0.00000
1	39	0.01334	0.00000
2	61	0.01857	0.04918
3	148	0.02222	0.01351
4	1	0.02812	1.00000
5	43	0.03041	0.00000
6	107	0.03705	0.04673
7	65	0.04049	0.01538
8	37	0.06532	0.00000
9	89	0.10162	0.05618
Total	**662**	**0.03784**	**0.02568**

Table 5.4: SB: Predicted VS Actual SSIs

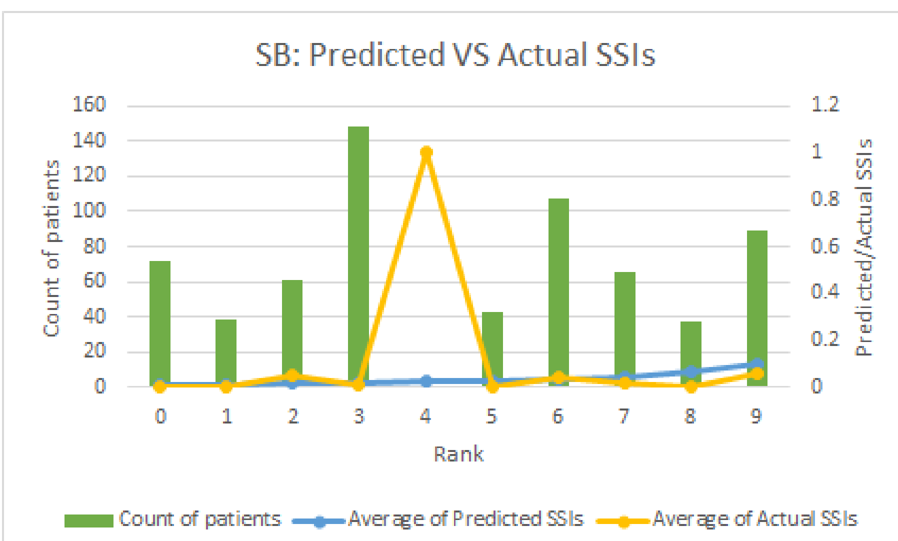


Figure 5.6: Predicted VS Actual SSIs for operative procedure group SB

5.5 Results of the operative procedure group CARD

The HL goodness of fit results, where $p = 0.2699$ are displayed in Figure 5.7 and thus there is no evidence of a lack of fit of the model. Table 5.5 and Figure 5.8 illustrate that across all 10 ranks a total of 961 patients were included in the CARD operative procedure group, and that a similar average probability of predicted SSIs than the average number of actual SSIs were obtained. The **Total** row in Table 5.5, indicates the sum of the number of patients and the average predicted SSIs and average actual SSIs across the total sum of all the patients.

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
20.0842	17	0.2699

Figure 5.7: HL goodness of fit test for operative procedure group CARD

CARD			
Rank	**Count of patients**	**Average of Predicted SSIs**	**Average of Actual SSIs**
0	94	0.00006	0.01064
1	98	0.00090	0.00000
2	90	0.00134	0.00000
3	92	0.00230	0.01087
4	92	0.00374	0.00000
5	104	0.00537	0.00000
6	104	0.00819	0.00000
7	89	0.01532	0.01124
8	102	0.03326	0.00000
9	96	0.12877	0.12500
Total	**961**	**0.02008**	**0.01561**

Table 5.5: CARD: Predicted VS Actual SSIs

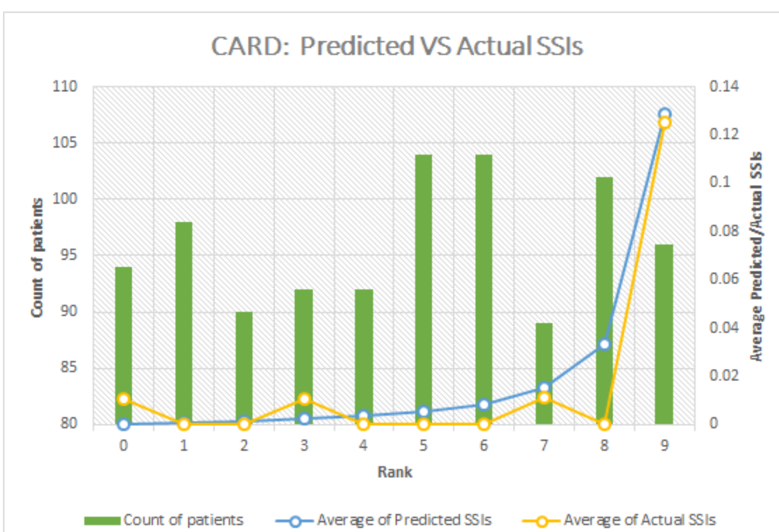


Figure 5.8: Predicted VS Actual SSIs for operative procedure group CARD

5.6 Results of the operative procedure group KPRO

The operative procedure group KPRO totaled to 3648 patients. Figure 5.9 illustrates the HL goodness of fit test, where $p = 0.3461$ and there is no evidence of a lack of fit in the model. As seen in Table 5.6, the average probability of predicted SSIs and the average number of actual SSIs had a relatively equal match and this trend can also be seen in Figure 5.10. The **Total** row in Table 5.6, indicates the sum of the number of patients and the average predicted SSIs and average actual SSIs across the total sum of all the patients.

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
8.9548	8	0.3461

Figure 5.9: HL goodness of fit test for operative procedure group KPRO

KPRO			
Rank	**Count of patients**	**Average of Predicted SSIs**	**Average of Actual SSIs**
0	393	0.00013	0.00254
1	348	0.00035	0.00287
2	358	0.00059	0.00000
3	374	0.00084	0.00000
4	374	0.00112	0.00000
5	387	0.00151	0.00258
6	364	0.00205	0.00000
7	384	0.00289	0.00000
8	321	0.00432	0.00000
9	345	0.01306	0.00870
Total	**3648**	**0.00259**	**0.00164**

Table 5.6: KPRO: Predicted VS Actual SSIs

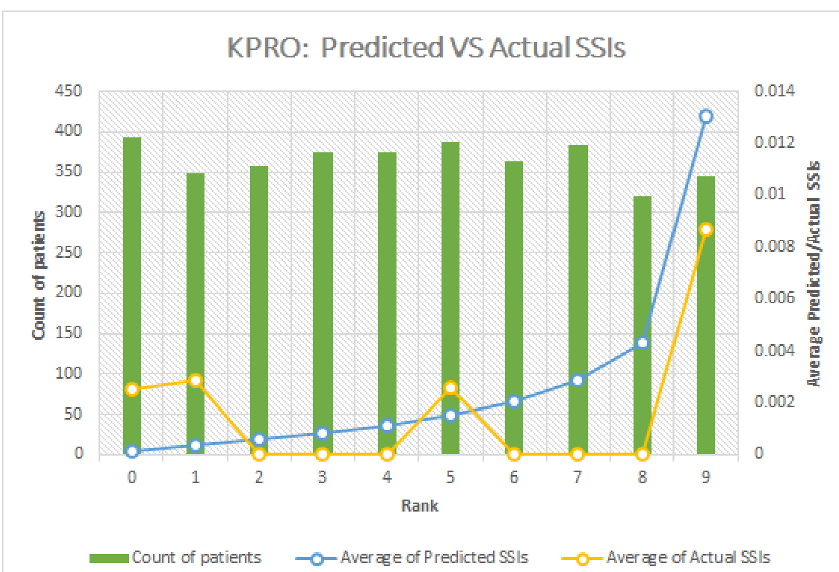


Figure 5.10: Predicted VS Actual SSIs for operative procedure group KPRO

5.7 Standard Infection Ratio (SIR)

Table 5.7 to Table 5.11 illustrates the predicted SSI and actual SSI per operative procedure group. The predicted SSIs are calculated by summing the predicted probabilities of each patient and the actual SSIs are summing the actual SSIs that were determined from the data. The SIR per operative procedure group is also determined per hospital. For this project, the researcher selected three hospitals to compare. For discussion purposes, these hospitals are referred to as Hospital A, Hospital B and Hospital C. The three hospitals are all regarded as large hospitals as well as hospitals that have a *high case mix index*, i.e. hospitals that are performing complex theatre cases and that are handling more intricate medical and surgical cases.

In the case of the operative procedure group HYST, Hospital A and Hospital C achieved a very good SIR across all three years whereas, Hospital B achieved an unacceptable SIR across all three years as indicated in Table 5.7. A SIR of 1 and above is regarded as an unacceptable result and stakeholders are expected to address the hospitals where the SIRs were above 1. Hospital A achieved a SIR of 0.61 in 2016, which is regarded as an acceptable SIR. In other words, the model predicted Hospital A to achieve more SSIs than Hospital A actually achieved. Hospital A's performance decreased slightly from 2016 to 2017, however, in 2018 Hospital A scored a SIR equal to 0 which is the best possible score a hospital could achieve.

Hospital B, improved slightly from 2016 to 2017 with the SIR dropping by 0.640, however, from 2017 to 2018 the SIR increased by more than double. Thus, special attention should be given to Hospital B with the aim to decrease their number of actual SSIs.

Hospital C achieved SIRs of 0 throughout the three years.

SIR Results: HYST				
Year	**Hospital**	**Hospital A**	**Hospital B**	**Hospital C**
	Predicted	3.28	1.12	0.74
2016	**Actual**	2.00	2.00	0.00
	SIR	**0.61**	**1.79**	**0.00**
	Predicted	3.04	0.87	0.88
2017	**Actual**	3.00	1.00	0.00
	SIR	**0.99**	**1.15**	**0.00**
	Predicted	3.08	1.14	0.4
2018	**Actual**	0.00	3.00	0.00
	SIR	**0.00**	**2.64**	**0.00**

Table 5.7: HYST: Predicted SSI, Actual SSI and SIR

In Table 5.8 the SIR results for the operative procedure group SB are given. Hospital A achieved an SIR of 0.99 in 2016, therefore the predicted SSI and actual SSI occurring at hospital A were very similar. In 2016, Hospital B did better than the predicted, SSI achieving a SIR of 0, whereas Hospital C obtained a negative SIR of 1.74. Both Hospital B and C achieved a SIR of 0 in the following year, while Hospital A scored a very high SIR of 3.86. During 2018 two high SIR results at Hospital A and B were obtained and Hospital C scored a positive SIR of 0. The actual count of SSIs for Hospital A for 2018 was lower than in 2017, suggesting a positive reduction in their SIR. As for Hospital B, their SIR increased from 0 actual SSIs in 2017 to 2 actual SSIs and a higher SIR for 2018. It is advised to report on the SIR results as well as the actual SSI results, Hospital B and Hospital C achieved a similar SIR, however, different actual SSI counts.

SIR Results: SB				
Year	**Hospital**	**Hospital A**	**Hospital B**	**Hospital C**
	Predicted	1.01	1.20	0.57
2016	**Actual**	1.00	0.00	1.00
	SIR	**0.99**	**0.00**	**1.74**
	Predicted	1.29	1.07	0.70
2017	**Actual**	5.00	0.00	0.00
	SIR	**3.86**	**0.00**	**0.00**
	Predicted	1.49	1.16	0.90
2018	**Actual**	4.00	2.00	0.00
	SIR	**2.69**	**1.73**	**0.00**

Table 5.8: SB: Predicted SSI, Actual SSI and SIR

The SIR results for the operative procedure BILI are illustrated in Table 5.9. Hospital A achieved a high SIR for the first year. The actual SSI that occurred at the hospital remained constant (actual SSIs = 1). The SIR, however, decreased dramatically from 2016 to 2017 but from 2017 to 2018 there was only a slight decrease in SIR for Hospital A. In 2016, Hospital A achieved a predicted SSI of 0.8 and 1 actual SSI which resulted in a high SIR of 1.25. In 2017 and 2018, the predicted SSIs were very similar to the actual SSIs that occured at Hospital A resulting in an adequate SIR of below 1.

Hospital B achieved the low SIR results of 0 for all three years. In 2016 the model's predicted SSI equalled 0.74, however, no actual SSI was reported by Hospital B and therefore their SIR resulted in 0.

In Table 5.9 Hospital C had a higher number of actual SSI cases for 2016 and 2017, the model predicted SSIs lower than the actual SSIs resulting in high SIRs for both 2016 and 2017. Hospital C scored a SIR of 0 in 2018, therefore showing an improvement from 2016 and 2017.

SIR Results: BILI				
Year	**Hospital**	**Hospital A**	**Hospital B**	**Hospital C**
	Predicted	0.8	0.74	0.4
2016	**Actual**	1.00	0.00	2.00
	SIR	**1.25**	**0.00**	**5.04**
	Predicted	1.23	0.57	1.21
2017	**Actual**	1.00	0.00	2.00
	SIR	**0.81**	**0.00**	**1.65**
	Predicted	1.27	0.43	0.83
2018	**Actual**	1.00	0.00	0.00
	SIR	**0.78**	**0.00**	**0.00**

Table 5.9: BILI: Predicted SSI, Actual SSI and SIR

Table 5.10 indicates the SIR results for the operative procedure group CARD. Across all three years, a high SIR was calculated for Hospital A and Hospital B. The actual SSIs remained constant across three years for Hospital A, however, the model always predicted the SSIs to be a lot lower than the number of actual SSIs at Hospital A and this resulted in a SIR greater than 1 for 2016, 2017 and 2018. Hospital B showed a very similar trend to Hospital A, however, Hospital B had a higher actual SSI count of 8 for 2017 in comparison to Hospital A. Hospital B is a hospital that admits sicker patients than Hospital A and C. Therefore, Hospital B's predicted values are higher than Hospital A. Hospital C achieved the best SIR result overall, by obtaining for all three years a SIR below 1.

SIR Results: CARD				
Year	**Hospital**	**Hospital A**	**Hospital B**	**Hospital C**
	Predicted	0.81	2.93	0.97
2016	**Actual**	2.00	5.00	0.00
	SIR	**2.47**	**1.7**	**0.00**
	Predicted	0.65	3.56	1.22
2017	**Actual**	2.00	8.00	1.00
	SIR	**3.09**	**2.25**	**0.82**
	Predicted	1.52	1.80	1.05
2018	**Actual**	2.00	2.00	0.00
	SIR	**1.32**	**1.11**	**0.00**

Table 5.10: CARD: Predicted SSI, Actual SSI and SIR

The SIR results for the operative procedure group KPRO are illustrated in Table 5.11. Hospital A obtained a SIR result below 1 for 2016, 2017 and 2018. There was a slight increase in Hospital A's SIR from 2016 and 2017 to 2018, however, the SIR remained below 1. In 2018, Hospital A had their highest number of actual SSIs (equal to 2) ,however, the model predicted them to contract 2.35 infections and therefore the SIR equals 0.85. Hospital B showed an increase in their actual SSIs and SIR from 2016 to 2018. Hospital B requires an intervention and needs to strive to decrease their actual SSIs and therefore SIR in the future. Hospital C has achieved no actual SSIs and subsequently a SIR equal to 0 for all three years being analysed.

SIR Results: KPRO				
Year	**Hospital**	**Hospital A**	**Hospital B**	**Hospital C**
	Predicted	2.42	2.87	0.49
2016	**Actual**	1.00	3.00	0.00
	SIR	**0.41**	**1.04**	**0.00**
	Predicted	2.41	2.88	0.42
2017	**Actual**	1.00	6.00	0.00
	SIR	**0.41**	**2.07**	**0.00**
	Predicted	2.35	3	0.54
2018	**Actual**	2.00	10.00	0.00
	SIR	**0.85**	**3.34**	**0.00**

Table 5.11: KPRO: Predicted SSI, Actual SSI and SIR

CHAPTER 6

Conclusion

In this chapter a summary of the research project is given in §6.1 with reference to achievement of Objectives I to III as set in §1.4. Recommendations made to the healthcare facilities, which is Objective IV of the research project as listed in §1.4, is the topic of §6.2. This is followed in §6.3 with some remarks on possible future work.

6.1 Summary of thesis

This research project aimed to first predict SSIs at three hospitals over three years across five different operative procedure groups. This was done via a logistic regression model to achieve Objective II listed in §1.4. The data required for the logistic regression model was obtained from the healthcare facilities Head Office. Futhermore, a thorough search in the literature as well as through conversations with clinicians at hospitals was done to obtain possible risk factors for the logistic regression model. All these steps addressed Objective I of this study as stated in §1.4.

The predicted SSIs from the logistic regression model and the actual count of SSIs were used to calculate the SIR as stipulated in Objective III of this study. The SIRs of three hospitals which were analsyed in the research project referred to as Hospital A, Hospital B and Hospital C and showed a unique trend across the three years. There were different trends and results showcased over the three year period and across the five operative procedure groups which were analysed in the research project namely: HYST, BILI, SB, CARD and KPRO. Finally, the hospitals were provided with SIR results to prevent SSIs at hospitals and aid in a safe environment for all patients to receive healthcare.

6.2 Recommendations

Overall, Hospital A achieved a SIR of below 1 and therefore good results across all three years and all operative procedure groups, except for the operative procedure group CARD where Hospital A scored a SIR of above 1 for all three years. The SIR for the operative procedure group BILI showed an improvement from 2016 to 2018, with the SIR decreasing from above 1 in 2016 to a SIR of below 1 for 2017 and 2018. It is recommended that Hospital A must investigate post-operative care after a CARD operative procedure was done.

Hospital B achieved SIRs of greater than 1 for KPRO, CARD and HYST for all three years being analysed. The SIR for BILI showed an improvement or a decrease in the SIR from 2016 to 2018 and Hospital B did achieve a SIR of below 1 for all three years for the operative procedure group SB.

Hospital C scored a SIR of below 1 for all three years across all five operative procedures and therefore achieved the best results compared to Hospital A and Hospital B.

In conclusion, based on the SIR results for Hospital B across certain operative procedures, Hospital B needs to improve their infection prevention control and surveillance. It is evident, that Hospital B is experiencing more SSIs than Hospital A and Hospital C. Hospital B's patients are also contracting more SSIs than those that are being predicted. Therefore Hospital B requires an intervention to improve the current number of SSIs being contracted by patients in the healthcare facility per operative procedure.

6.3 Future work

As previously mentioned, HAIs are made up of four groups of infections, namely: SSIs, CLABSI, CAUTI and VAP. A SIR may be developed for all three additional infection types (CLABSI, CAUTI and VAP) based on significant variables per infection and operative procedure groups. By calculating the predicted values and SIR results of the three additional HAI groups, the hospitals will benefit of an overall understanding of all the infection groups within their hospital.

List of References

[1] Archer KJ & Lemeshow S, 2006, *Goodness-of-fit test for a logistic regression model fitted using survey sample data*, The Stata Journal, **6(1)**, pp. 97–105.

[2] Aronson W, McAuliffe MS, Miller K *et al.*, 2003, *Variability in the american society of anesthesiologists physical status classification scale*, AANA journal, **71(4)**, pp. 265–276.

[3] Arredondo-García J, Diaz-Ramos R, Solórzano-Santos F, Sosa-González I & Beltrán-Zúñiga M, 1992, *Neonatal septicaemia due to k. pneumoniae. septicaemia due to klebsiella pneumoniae in newborn infants. nosocomial outbreak in an intensive care unit.*, Revista Latinoamericana de Microbiologia, **34(1)**, pp. 11–16.

[4] Bagnall NM, Vig S & Trivedi P, 2009, *Surgical-site infection*, Surgery-Oxford International Edition, **27(10)**, pp. 426–430.

[5] Bakar AA, Kefli Z, Abdullah S & Sahani M, 2011, *Predictive models for dengue outbreak using multiple rulebase classifiers*, Proceedings of the Electrical Engineering and Informatics (ICEEI), 2011 International Conference, pp. 1–6.

[6] Bonorchis R & Kew J, 2017, *Bloomberg daily newsletter; how South Africa stumbles on health care for all*, Available from `https://www.bloomberg.com/news/articles/2017-07-27/how-south-africa-stumbles`.

[7] Cummings Jr DD, *All care health management system, 1994*, uS Patent 5,301,105.

[8] Diekema DJ, BootsMiller BJ, Vaughn TE, Woolson RF, Yankey JW, Ernst EJ, Flach SD, Ward MM, Franciscus CL, Pfaller MA *et al.*, 2004, *Antimicrobial resistance trends and outbreak frequency in united states hospitals*, Clinical Infectious Diseases, **38(1)**, pp. 78–85.

[9] Dudeck MA, Horan TC, Peterson KD, Allen-Bridson K, Morrell G, Pollock DA & Edwards JR, 2011, *National healthcare safety network (nhsn) report, data summary for 2010, device-associated module*, American Journal of Infection Control, **39(10)**, pp. 798–816.

[10] Dudeck MA, Weiner LM, Allen-Bridson K, Malpiedi PJ, Peterson KD, Pollock DA, Sievert DM & Edwards JR, 2013, *National healthcare safety network (nhsn) report, data summary for 2012, device-associated module*, American Journal of Infection Control, **41(12)**, pp. 1148–1166.

[11] Elvy J & Colville A, 2009, *Catheter associated urinary tract infection: what is it, what causes it and how can we prevent it?*, Journal of Infection Prevention, **10(2)**, pp. 36–41.

[12] Ferris VM, McMullen KM, Russo AJ, Wood HA & Lane MA, 2011, *Critical care leads the way in the prevention of central line associated bloodstream infections*, American Journal of Infection Control, **39(5)**, p. E196.

[13] Fukuda H & Kuroki M, 2016, *The development of statistical models for predicting surgical site infections in japan: Toward a statistical model–based standardized infection ratio*, Infection Control & Hospital Epidemiology, **37(3)**, pp. 260–271.

[14] Gustafson TL, 2006, *Three uses of the standardized infection ratio (SIR) in infection control*, Infection Control Hospital Epidemiology, **27(4)**, p. 427430.

[15] Hlophe ST & McKerrow N, 2014, *Hospital-acquired klebsiella pneumoniae infections in a paediatric intensive care unit*, South African Journal of Child Health, **8(4)**, pp. 125–128.

[16] Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, Jha P, Mills A & Musgrove P, 2006, *Disease control priorities in developing countries*, World Bank Publications.

[17] Jun J, Jacobson SH & Swisher JR, 1999, *Application of discrete-event simulation in health care clinics: A survey*, Journal of the Operational Research Society, **50(2)**, pp. 109–123.

[18] Klevens RM, Edwards JR, Richards Jr CL, Horan TC, Gaynes RP, Pollock DA & Cardo DM, 2007, *Estimating health care-associated infections and deaths in US hospitals, 2002*, Public Health Reports, **122(2)**, pp. 160–166.

[19] Leanne B Gasink ELMSNOF Paul H Edelstein, 2009, *Risk factors and clinical impact of klebsiella pneumoniae carbapenemaseproducing k. pneumoniaea*, Hospital Infection Control, **12(1)**, pp. 1–3.

[20] Lobdell KW, Stamou S & Sanchez JA, 2012, *Hospital-acquired infections*, Surgical Clinics, **92(1)**, pp. 65–77.

[21] Mahieu L, Buitenweg N, Beutels P & De Dooy J, 2001, *Additional hospital stay and charges due to hospital-acquired infections in a neonatal intensive care unit*, Journal of hospital Infection, **47(3)**, pp. 223–229.

[22] McKibben L, Horan TC, Tokars JI, Fowler G, Cardo DM, Pearson ML & Brennan PJ, 2005, *Guidance on public reporting of healthcare-associated infections: recommendations of the healthcare infection control practices advisory committee*, Infection Control & Hospital Epidemiology, **26(6)**, pp. 580–587.

[23] Miholic J, Hudec M, Domanig E, Hiertz H, Klepetko W, Lackner F & Wolner E, 1985, *Risk factors for severe bacterial infections after valve replacement and aortocoronary bypass operations: analysis of 246 cases by logistic regression*, The Annals of Thoracic Surgery, **40(3)**, pp. 224–228.

[24] Narong MN, Thongpiyapoom S, Thaikul N, Jamulitrat S & Kasatpibal N, 2003, *Surgical site infections in patients undergoing major operations in a university hospital: using standardized infection ratio as a benchmarking tool*, American Journal of Infection Control, **31(5)**, pp. 274–279.

[25] Pillay R, 2009, *Work satisfaction of professional nurses in south africa: a comparative analysis of the public and private sectors*, Human Resources for Health, **7(1)**, p. 15.

[26] Plowman R, Graves N, Roberts J *et al.*, 1997, *Hospital acquired infection*, Series on Health, National Center for Biotechnology Information.

[27] Saman DM, Kavanagh KT & Abusalem SK, 2013, *Redefining the standardized infection ratio to aid in consumer value purchasing*, Journal of Patient Safety, **9(2)**, pp. 55–58.

[28] Scott RD, 2009, *Centers for Disease Control and prevention: The direct medical costs of healthcare-associated infections in US hospitals and the benefits of prevention.*

[29] Selfridge PG & Srivastava D, *Interactive data exploration apparatus and methods, 1996*, US Patent 5,999,192.

[30] Thom DH, Hall MA & Pawlson LG, 2004, *Measuring patients trust in physicians when assessing quality of care*, Health Affairs, **23(4)**, pp. 124–132.

[31] Umscheid CA, Mitchell MD, Doshi JA, Agarwal R, Williams K & Brennan PJ, 2011, *Estimating the proportion of healthcare-associated infections that are reasonably preventable and the related mortality and costs*, Infection Control & Hospital Epidemiology, **32(2)**, pp. 101–114.

[32] Van Jaarsveld A, 2012, *Mediclinic scenario award: Clinical pharmacy and antimicrobial stewardship: a symbiotic relationship*, SA Pharmaceutical Journal, **79(10)**, pp. 50–52.

APPENDIX A

Clinical coding classifications

In this Appendix all the CPT codes for the relevant operative procedures used in this study are given. In Table A.1 and Table A.2 all the CPT codes for the operative procedure group HYST are listed.

CPT Codes	**CPT Code Description**
58150	Total abdominal hysterectomy (corpus and cervix), with or without removal of tube(s), with or without removal of ovary(s)
58152	Total abdominal hysterectomy (corpus and cervix), with or without removal of tube(s), with or without removal of ovary(s); with colpo-urethrocystopexy (eg, Marshall-Marchetti-Krantz, Burch)
58180	Supracervical abdominal hysterectomy (subtotal hysterectomy), with or without removal of tube(s), with or without removal of ovary(s)
58200	Total abdominal hysterectomy, including partial vaginectomy, with para-aortic and pelvic lymph node sampling, with or without removal of tube(s), with or without removal of ovary(s)
58210	Radical abdominal hysterectomy, with bilateral total pelvic lymphadenectomy and para-aortic lymph node sampling (biopsy), with or without removal of tube(s), with or without removal of ovary(s)
58240	Pelvic exenteration for gynecologic malignancy, with total abdominal hysterectomy or cervicectomy, with or without removal of tube(s), with or without removal of ovary(s), with removal of bladder and ureteral transplantations, and/or abdominoperineal resection of rectum and colon and colostomy, or any combination thereof
58541	Laparoscopy, surgical, supracervical hysterectomy, for uterus 250 g or less
58542	Laparoscopy, surgical, supracervical hysterectomy, for uterus 250 g or less; with removal of tube(s) and/or ovary(s)
58543	Laparoscopy, surgical, supracervical hysterectomy, for uterus greater than 250 g
58544	Laparoscopy, surgical, supracervical hysterectomy, for uterus greater than 250 g; with removal of tube(s) and/or ovary(s)
58548	Laparoscopy, surgical, with radical hysterectomy, with bilateral total pelvic lymphadenectomy and para-aortic lymph node sampling (biopsy), with removal of tube(s) and ovary(s), if performed

Table A.1: CPT codes in the operative procedure group HYST

CPT Codes	CPT Code Description
58570	Laparoscopy, surgical, with total hysterectomy, for uterus 250 g or less
58571	Laparoscopy, surgical, with total hysterectomy, for uterus 250 g or less; with removal of tube(s) and/or ovary(s)
58572	Laparoscopy, surgical, with total hysterectomy, for uterus greater than 250 g
58573	Laparoscopy, surgical, with total hysterectomy, for uterus greater than 250 g; with removal of tube(s) and/or ovary(s)
58951	Resection (initial) of ovarian, tubal or primary peritoneal malignancy with bilateral salpingo-oophorectomy and omentectomy; with total abdominal hysterectomy, pelvic and limited para-aortic lymphadenectomy
58953	Bilateral salpingo-oophorectomy with omentectomy, total abdominal hysterectomy and radical dissection for debulking
58954	Bilateral salpingo-oophorectomy with omentectomy, total abdominal hysterectomy and radical dissection for debulking; with pelvic lymphadenectomy and limited para-aortic lymphadenectomy
58956	Bilateral salpingo-oophorectomy with total omentectomy, total abdominal hysterectomy for malignancy
59525	Subtotal or total hysterectomy after cesarean delivery
58575	Laparoscopy, surgical, total hysterectomy for resection of malignancy (tumor debulking), with omentectomy including salpingo-oophorectomy, unilateral or bilateral, when performed

Table A.2: CPT codes in the operative procedure group HYST

The CPT codes associated with the operative procedure group SB are illustrated in Table A.3 and Table A.4.

CPT Codes	CPT Code Description
43496	Free jejunum transfer with microvascular anastomosis
44010	Duodenotomy, for exploration, biopsy(s), or foreign body removal
44015	Tube or needle catheter jejunostomy for enteral alimentation, intraoperative, any method (List separately in addition to primary procedure)
44020	Enterotomy, small intestine, other than duodenum; for exploration, biopsy(s), or foreign body removal
44021	Enterotomy, small intestine, other than duodenum; for decompression (eg, Baker tube)
44120	Enterectomy, resection of small intestine; single resection and anastomosis
44121	Enterectomy, resection of small intestine; each additional resection and anastomosis (List separately in addition to code for primary procedure)
44125	Enterectomy, resection of small intestine; with enterostomy
44126	Enterectomy, resection of small intestine for congenital atresia, single resection and anastomosis of proximal segment of intestine; without tapering
44127	Enterectomy, resection of small intestine for congenital atresia, single resection and anastomosis of proximal segment of intestine; with tapering

Table A.3: CPT codes in the operative procedure group SB

Table A.5 includes all the CPT codes within the operative procedure group KPRO.

CPT Codes	CPT Code Description
44128	Enterectomy, resection of small intestine for congenital atresia, single resection and anastomosis of proximal segment of intestine; each additional resection and anastomosis (List separately in addition to code for primary procedure)
44186	Laparoscopy, surgical; jejunostomy (eg, for decompression or feeding)
44187	Laparoscopy, surgical; ileostomy or jejunostomy, non-tube
44202	Laparoscopy, surgical; enterectomy, resection of small intestine, single resection and anastomosis
44203	Laparoscopy, surgical; each additional small intestine resection and anastomosis (List separately in addition to code for primary procedure)
44300	Enterostomy-External Fistulization of Intestines Procedures
44310	Ileostomy or jejunostomy, non-tube
44312	Revision of ileostomy; simple (release of superficial scar) (separate procedure)
44314 44316	Revision of ileostomy; complicated (reconstruction in-depth) (separate procedure) Continent ileostomy (Kock procedure) (separate procedure)
44602	Suture of small intestine (enterorrhaphy) for perforated ulcer, diverticulum, wound, injury or rupture; single perforation
44603	Suture of small intestine (enterorrhaphy) for perforated ulcer, diverticulum, wound, injury or rupture; multiple perforations
44615	Intestinal stricturoplasty (enterotomy and enterorrhaphy) with or without dilation, for intestinal obstruction
44640	Closure of intestinal cutaneous fistula
44650	Closure of enteroenteric or enterocolic fistula
44800	Excision of Meckel's diverticulum (diverticulectomy) or omphalomesenteric duct
45136	Excision of ileoanal reservoir with ileostomy

Table A.4: CPT codes in the operative procedure group SB

CPT Codes	CPT Code Description
27438	Arthroplasty, patella; with prosthesis
27440	Arthroplasty, knee, tibial plateau
27441	Arthroplasty, knee, tibial plateau; with debridement and partial synovectomy
27442	Arthroplasty, femoral condyles or tibial plateau(s), knee
27443	Arthroplasty, femoral condyles or tibial plateau(s), knee; with debridement and partial synovectomy
27445	Arthroplasty, knee, hinge prosthesis (eg, Walldius type)
27446	Arthroplasty, knee, condyle and plateau; medial OR lateral compartment
27447	Arthroplasty, knee, condyle and plateau; medial AND lateral compartments with or without patella resurfacing (total knee arthroplasty)

Table A.5: CPT codes in the operative procedure group KPRO

The CPT codes associated with the operative procedure group BILI are illustrated in Table A.6, Table A.7 and Table A.8.

CPT Codes	CPT Code Description
47010	Hepatotomy; for open drainage of abscess or cyst, 1 or 2 stages
47015	Laparotomy, with aspiration and/or injection of hepatic parasitic (eg, amoebic or echinococcal) cyst(s) or abscess(es
47100	Biopsy of liver, wedge
47120	Hepatectomy, resection of liver; partial lobectomy
47122	Hepatectomy, resection of liver; trisegmentectomy
47125	Hepatectomy, resection of liver; total left lobectomy
47130	Hepatectomy, resection of liver; total right lobectomy
47140	Donor hepatectomy (including cold preservation), from living donor; left lateral segment only (segments II and III)
47141	Donor hepatectomy (including cold preservation), from living donor; total left lobectomy (segments II, III and IV)
47142	Donor hepatectomy (including cold preservation), from living donor; total right lobectomy (segments V, VI, VII and VIII)
47300	Marsupialization of cyst or abscess of liver
47350	Management of liver hemorrhage; simple suture of liver wound or injury
47360	Management of liver hemorrhage; complex suture of liver wound or injury, with or without hepatic artery ligatio
47361	Management of liver hemorrhage; exploration of hepatic wound, extensive debridement, coagulation and/or suture, with or without packing of liver
47362	Management of liver hemorrhage; re-exploration of hepatic wound for removal of packing
47370	Laparoscopy, surgical, ablation of 1 or more liver tumor(s); radiofrequency
47371	Laparoscopy, surgical, ablation of 1 or more liver tumor(s); cryosurgical
47379	Unlisted laparoscopic procedure, liver
47380	Ablation, open, of 1 or more liver tumor(s); radiofrequency
47381	Ablation, open, of 1 or more liver tumor(s); cryosurgical
47400	Hepaticotomy or hepaticostomy with exploration, drainage, or removal of calculus
47420	Choledochotomy or choledochostomy with exploration, drainage, or removal of calculus, with or without cholecystotomy; without transduodenal sphincterotomy or sphincteroplasty
47425	Choledochotomy or choledochostomy with exploration, drainage, or removal of calculus, with or without cholecystotomy; with transduodenal sphincterotomy or sphincteroplasty
47460	Transduodenal sphincterotomy or sphincteroplasty, with or without transduodenal extraction of calculus (separate procedure)
47560	Laparoscopy, surgical; with guided transhepatic cholangiography, without biopsy
47561	Laparoscopy, surgical; with guided transhepatic cholangiography with biopsy
47700	Exploration for congenital atresia of bile ducts, without repair, with or without liver biopsy, with or without cholangiography
47701	Portoenterostomy (eg, Kasai procedure)

Table A.6: CPT codes in the operative procedure group BILI

CPT Codes	CPT Code Description
47711	Excision of bile duct tumor, with or without primary repair of bile duct; extrahepatic
47712	Excision of bile duct tumor, with or without primary repair of bile duct; intrahepatic
47715	Excision of choledochal cyst
47760	Anastomosis, of extrahepatic biliary ducts and gastrointestinal tract
47765	Anastomosis, of intrahepatic ducts and gastrointestinal tract
47780	Anastomosis, Roux-en-Y, of extrahepatic biliary ducts and gastrointestinal tract
47785	Anastomosis, Roux-en-Y, of intrahepatic biliary ducts and gastrointestinal tract
47800	Reconstruction, plastic, of extrahepatic biliary ducts with end-to-end anastomosis
47802	U-tube hepaticoenterostomy
47900	Suture of extrahepatic biliary duct for pre-existing injury (separate procedure)
48000	Placement of drains, peripancreatic, for acute pancreatitis;
48001	Placement of drains, peripancreatic, for acute pancreatitis; with cholecystostomy, gastrostomy, and jejunostomy
48020	Removal of pancreatic calculus
48100	Biopsy of pancreas, open (eg, fine needle aspiration, needle core biopsy, wedge biopsy)
48105	Resection or debridement of pancreas and peripancreatic tissue for acute necrotizing pancreatitis
48120	Excision of lesion of pancreas (eg, cyst, adenoma)
48140	Pancreatectomy, distal subtotal, with or without splenectomy; without pancreaticojejunostomy
48145	Pancreatectomy, distal subtotal, with or without splenectomy; with pancreaticojejunostomy
48146	Pancreatectomy, distal, near-total with preservation of duodenum (Child-type procedure)
48148	Excision of ampulla of Vater
48150	Pancreatectomy, proximal subtotal with total duodenectomy, partial gastrectomy, choledochoenterostomy and gastrojejunostomy (Whipple-type procedure); with pancreatojejunostomy
48152	Pancreatectomy, proximal subtotal with total duodenectomy, partial gastrectomy, choledochoenterostomy and gastrojejunostomy (Whipple-type procedure); without pancreatojejunostomy
48153	Pancreatectomy, proximal subtotal with near-total duodenectomy, choledochoenterostomy and duodenojejunostomy (pylorus-sparing, Whipple-type procedure); with pancreatojejunostomy
48154	Pancreatectomy, total
48155	Pancreatectomy, proximal subtotal with near-total duodenectomy, choledochoenterostomy and duodenojejunostomy (pylorus-sparing, Whipple-type procedure); without pancreatojejunostomy

Table A.7: CPT codes in the operative procedure group BILI

CPT Codes	CPT Code Description
48160	Pancreatectomy, total or subtotal, with autologous transplantation of pancreas or pancreatic islet cells
48500	Marsupialization of pancreatic cyst
48510	External drainage, pseudocyst of pancreas; open
48520	Internal anastomosis of pancreatic cyst to gastrointestinal tract; direct
48540	Internal anastomosis of pancreatic cyst to gastrointestinal tract; Roux-en-Y
48545	Pancreatorrhaphy for injury
48548	Pancreaticojejunostomy, side-to-side anastomosis (Puestow-type operation)

Table A.8: CPT codes in the operative procedure group BILI

Table A.9, A.10, A.11 and A.12 are the CPT codes for the operative procedure group CARD.

CPT Codes	CPT Code Description
32658	Thoracoscopy, surgical; with removal of clot or foreign body from pericardial sac
32658	Thoracoscopy, surgical; with removal of clot or foreign body from pericardial sac
32659	Thoracoscopy, surgical; with creation of pericardial window or partial resection of pericardial sac for drainage
32661	Thoracoscopy, surgical; with excision of pericardial cyst, tumor, or mass
33020	Pericardiotomy for removal of clot or foreign body (primary procedure)
33025	Creation of pericardial window or partial resection for drainage
33030	Pericardiectomy, subtotal or complete; without cardiopulmonary bypass
33031	Pericardiectomy, subtotal or complete; with cardiopulmonary bypass
33050	Resection of pericardial cyst or tumor
33257	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), performed at the time of other cardiac
33258	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), extensive (eg, maze procedure), without cardiopulmonary bypass (List separately in addition to code for primary procedure)
33310	Cardiotomy, exploratory (includes removal of foreign body, atrial or ventricular thrombus); without bypass
33315	Cardiotomy, exploratory (includes removal of foreign body, atrial or ventricular thrombus); with cardiopulmonary bypass
33365	Transcatheter aortic valve replacement (TAVR/TAVI) with prosthetic valve; transaortic approach (eg, median sternotomy, mediastinotomy)
33366	Transcatheter aortic valve replacement (TAVR/TAVI) with prosthetic valve; transapical exposure (eg, left thoracotomy)
33390	Valvuloplasty, aortic valve, open, with cardiopulmonary bypass; simple (ie, valvotomy, debridement, debulking, and/or simple commissural resuspension)

Table A.9: CPT codes in the operative procedure group CARD

CPT Codes	CPT Code Description
32659	Thoracoscopy, surgical; with creation of pericardial window or partial resection of pericardial sac for drainage
32661	Thoracoscopy, surgical; with excision of pericardial cyst, tumor, or mass
33020	Pericardiotomy for removal of clot or foreign body (primary procedure)
33025	Creation of pericardial window or partial resection for drainage
33030	Pericardiectomy, subtotal or complete; without cardiopulmonary bypass
33031	Pericardiectomy, subtotal or complete; with cardiopulmonary bypass
33050	Resection of pericardial cyst or tumor
33130	Resection of external cardiac tumor
33130	Resection of external cardiac tumor; simple (ie, valvotomy, debridement, debulking, and/or simple commissural resuspension)
33692	Complete repair tetralogy of Fallot without pulmonary atresia;
33250	Operative ablation of supraventricular arrhythmogenic focus or pathway (eg, Wolff-Parkinson-White, atrioventricular node re-entry), tract(s) and/or focus (foci); without cardiopulmonary bypass
33251	Operative ablation of supraventricular arrhythmogenic focus or pathway (eg, Wolff-Parkinson-White, atrioventricular node re-entry), tract(s) and/or focus (foci); with cardiopulmonary bypass
33254	Operative tissue ablation and reconstruction of atria, limited (eg, modified maze procedure)
33255	Operative tissue ablation and reconstruction of atria, extensive (eg, maze procedure); without cardiopulmonary bypass
33256	Operative tissue ablation and reconstruction of atria, extensive (eg, maze procedure); with cardiopulmonary bypass
33782	Aortic root translocation with ventricular septal defect and pulmonary stenosis repair (ie, Nikaidoh procedure); without coronary ostium reimplantation
33783	Aortic root translocation with ventricular septal defect and pulmonary stenosis repair (ie, Nikaidoh procedure); with reimplantation of 1 or both coronary ostia
33786	Total repair, truncus arteriosus (Rastelli type operation)
33813	Obliteration of aortopulmonary septal defect; without cardiopulmonary bypass
33814	Obliteration of aortopulmonary septal defect; with cardiopulmonary bypass
33920	Repair of pulmonary atresia with ventricular septal defect, by construction or replacement of conduit from right or left ventricle to pulmonary artery
0051T	Implantation of a total replacement heart system (artificial heart) with recipient cardiectomy
0052T	Replacement or repair of implantable component or components of total replacement heart system (artificial heart), thoracic unit
0053T	Replacement or repair of implantable component or components of total replacement heart system (artificial heart), excluding thoracic unit

Table A.10: CPT codes in the operative procedure group CARD

CPT Codes	CPT Code Description
33390	Valvuloplasty, aortic valve, open, with cardiopulmonary bypass
33391	Valvuloplasty, aortic valve, open, with cardiopulmonary bypass; complex (eg, leaflet extension, leaflet resection, leaflet reconstruction, or annuloplasty)
33404	Construction of apical-aortic conduit
33405	Replacement, aortic valve, open, with cardiopulmonary bypass; with prosthetic valve other than homograft or stentless valve
33406	Replacement, aortic valve, with cardiopulmonary bypass; with allograft valve (freehand)
33410	Replacement, aortic valve, open, with cardiopulmonary bypass; with stentless tissue valve
33411	Replacement, aortic valve; with aortic annulus enlargement, noncoronary sinus
33412	Replacement, aortic valve; with transventricular aortic annulus enlargement (Konno procedure)
33413	Replacement, aortic valve; by translocation of autologous pulmonary valve with allograft replacement of pulmonary valve (Ross procedure)
33414	Repair of left ventricular outflow tract obstruction by patch enlargement of the outflow tract
33415	Resection or incision of subvalvular tissue for discrete subvalvular aortic stenosis
33416	Ventriculomyotomy (-myectomy) for idiopathic hypertrophic subaortic stenosis (eg, asymmetric septal hypertrophy)
33417	Aortoplasty (gusset) for supravalvular stenosis
33420	Valvotomy, mitral valve; closed heart
33425	Valvuloplasty, mitral valve, with cardiopulmonary bypass;
33426	Valvuloplasty, mitral valve, with cardiopulmonary bypass; with prosthetic ring
33427	Valvuloplasty, mitral valve, with cardiopulmonary bypass; radical reconstruction, with or without ring
33430	Replacement, mitral valve, with cardiopulmonary bypass
33697	Complete repair tetralogy of Fallot with pulmonary atresia including construction of conduit from right ventricle to pulmonary artery and closure of ventricular septal defect
33702	Repair sinus of Valsalva fistula, with cardiopulmonary bypass;
33710	Repair sinus of Valsalva fistula, with cardiopulmonary bypass; with repair of ventricular septal defect
33720	Repair sinus of Valsalva aneurysm, with cardiopulmonary bypass
33722	Closure of aortico-left ventricular tunnel
33732	Repair of cor triatriatum or supravalvular mitral ring by resection of left atrial membrane
33735	Atrial septectomy or septostomy; closed heart (Blalock-Hanlon type operation)
33736	Atrial septectomy or septostomy; open heart with cardiopulmonary bypass
33737	Atrial septectomy or septostomy; open heart, with inflow occlusion
33770	Repair of transposition of the great arteries with ventricular septal defect and subpulmonary stenosis; without surgical enlargement of ventricular septal defect

Table A.11: CPT codes in the operative procedure group CARD

CPT Codes	**CPT Code Description**
33774	Repair of transposition of the great arteries, atrial baffle procedure (eg, Mustard or Senning type) with cardiopulmonary bypass;
33776	Repair of transposition of the great arteries, atrial baffle procedure (eg, Mustard or Senning type) with cardiopulmonary bypass; with closure of ventricular septal defect
33780	Repair of transposition of the great arteries, aortic pulmonary artery reconstruction (eg, Jatene type); with closure of ventricular septal defect
33612	Repair of double outlet right ventricle with intraventricular tunnel repair; with repair of right ventricular outflow tract obstruction
33615	Repair of complex cardiac anomalies (eg, tricuspid atresia) by closure of atrial septal defect and anastomosis of atria or vena cava to pulmonary artery (simple Fontan procedure)
33617	Repair of complex cardiac anomalies (e.g., single ventricle by modified Fontan)
33619	Repair of single ventricle with aortic outflow obstruction and aortic arch hypoplasia (hypoplastic left heart syndrome) (eg, Norwood procedure)
33641	Repair atrial septal defect, secundum, with cardiopulmonary bypass, with or without patch
33645	Direct or patch closure, sinus venosus, with or without anomalous pulmonary venous drainage
33647	Repair of atrial septal defect and ventricular septal defect, with direct or patch closure
33660	Repair of incomplete or partial atrioventricular canal (ostium primum atrial septal defect), with or without atrioventricular valve repair
33665	Repair of intermediate or transitional atrioventricular canal, with or without atrioventricular valve repair
33670	Repair of complete atrioventricular canal, with or without prosthetic valve
33675	Closure of multiple ventricular septal defects;
33676	Closure of multiple ventricular septal defects; with pulmonary valvotomy or infundibular resection (acyanotic)
33677	Closure of multiple ventricular septal defects; with removal of pulmonary artery band, with or without gusset
33681	Closure of single ventricular septal defect, with or without patch;
33684	Closure of single ventricular septal defect, with or without patch; with pulmonary valvotomy or infundibular resection (acyanotic)
33688	Closure of single ventricular septal defect, with or without patch; with removal of pulmonary artery band, with or without gusset

Table A.12: CPT codes in the operative procedure group CARD