Patient level predictions in bowel surgery

comparing variable selections for rare outcome modelling on real surgery data

Gidius van de Kamp

Delft University of Technology

TU Delft

Patient level predictions in bowel surgery

comparing variable selections for rare outcome modelling on real surgery data

by

Gidius van de Kamp

Student Name	Student Number
van de Kamp	4593014

Chair:	T. Nane
Daily supervisor:	Ö. Şahin
Project Duration:	2, 23 - 10, 24
Faculty:	Faculty of Electrical Engineering, Mathematics and Computer Science , Delft

Cover:	Blurred zoom in of Table 6.6
Style:	TU Delft Report Style, with modifications by Daan Zwaneveld

TU Delft

Preface

Predicting surgery outcomes is important because it can show us which factors can contribute to undesired surgery outcomes. At the same time, it is no surprise that this is difficult. Doctors keep up with newly published studies on how to improve surgery outcomes. This self-aware system makes it difficult to see any strong relationships. In the literature, many studies show specific relations. In this thesis, we will see that we cannot always see these empirical results back in a strong way in the data from Medical Spectrum Twente (MST). We have data where many variables relate to the outcome to predict yet the models employed exhibit modest performance. We will take a statistical approach to predict surgery outcomes for bowel surgery. We show what challenges in accurately predicting surgery have to be countered.

Gidius van de Kamp
Delft, October 2024

Summary

We have studied the predictability of the complications after bowel surgery with the dataset provided by the MST. We see that the correlated data and rare event outcomes lead to the modest performance of the models employed. We have done an extensive preprocessing where we went from a total of 375 variables to 51 explanatory variables. We applied 3 variable selection methods to predict severe complications and any complications in 3 different scenarios (before surgery, right after surgery, and after primary stay). We will see that interoperative variables add valuable information (comparing Scenario 1 with Scenario 2). However, the interoperative variables become less important for predicting the complications post-primary stay (in Scenario 3). Generally, the most important variables are general health scores (WHO score and ASA class) except when predicting complications right after surgery then interoperative variables, for example blood loss, seem to contribute more to complications than variables indicating preoperative health. We will see that post-operative variables are selected for modelling complications after the primary hospital stay, but that these models do not perform adequately since not many complications are present after primary hospital stay. We compared a literature based selection, one resulting from performing marginal testing and an AIC forward variable selection. We see that in the literature many variables are considered important. Sadly we did not extract information about the timing or severity of the complications from the literature. The AIC forward seems prone to select correlated variables which often led to the resulting model to overfit. The marginal importance was more restrictive than the literature model. We have seen that the (unweighted) logistic regression severely underestimates the outcomes and that the weighted logistic regression will lead to a more informative model but with more misclassifications. We included the Threat score (TS) and the Matthews correlation coefficient (MCC) in our validation and confirmed that these are valid measures for scoring the performace of models with a rare outcome.

Contents

Preface i

Summary ii

Nomenclature v

1 Introduction 1
- 1.1 Overview of bowel surgeries . . . 2
- 1.2 Enhanced Recovery After Surgery (ERAS) protocols . . . 2
- 1.3 Data collection methods . . . 2
- 1.4 State of the field . . . 3
- 1.5 Research objectives . . . 4
  - 1.5.1 Overview of models to be developed . . . 4

2 Mathematical background and methodology 6
- 2.1 Descriptive statistics . . . 6
  - 2.1.1 Interpreting odds ratio . . . 6
  - 2.1.2 Kendall's tau and its applications . . . 7
  - 2.1.3 Key statistical tests . . . 7
  - 2.1.4 Examples . . . 8
- 2.2 Variable selections methods . . . 9
- 2.3 Model development techniques . . . 10
- 2.4 Model validation approaches . . . 11
  - 2.4.1 In-sample validation . . . 11
  - 2.4.2 Cross-validation strategies . . . 11
  - 2.4.3 Model validation measures . . . 12
  - 2.4.4 Assessing model performance for rare outcome . . . 14

3 Data description 15
- 3.1 Overview of patient journeys . . . 16
- 3.2 Patient characteristics . . . 17
- 3.3 Preoperative variables . . . 17
- 3.4 Intraoperative variables . . . 18
- 3.5 Postoperative variables . . . 18
- 3.6 Data from postoperative follow up . . . 19
- 3.7 Data collected for complications . . . 19

4 Feature engineering 21
- 4.1 Outcome choice . . . 22
- 4.2 Identification of data not suitable for statistical analyisis . . . 23
- 4.3 Ensuring similarity of the observations . . . 24
- 4.4 Exclusion of overparticular variables . . . 25
- 4.5 Classification of perioperative variables . . . 26
- 4.6 Removal of low-information variables . . . 27
  - 4.6.1 Daily measurements . . . 27
  - 4.6.2 Near-constant variables . . . 28
  - 4.6.3 Handeling variables with missing observations . . . 30
- 4.7 Preprocessing according to patient journey stages . . . 30
  - 4.7.1 Preprocessing patient characteristics . . . 30
  - 4.7.2 Preprocessing preoperative variables . . . 32
  - 4.7.3 Intraoperative variables . . . 32

4.7.4 Preprocessing postoperative variables . . . 34
4.8 Handeling Missing values . . . 36
4.9 Summary of feature engineering . . . 37

5 Explanatory data analysis (EDA) 38
5.1 Table general characteristics . . . 39
5.2 Analysis of preoperative variables . . . 41
5.3 Analysis of intraoperative variables . . . 43
5.4 Analysis of intraoperative anaesthesia and drugs variables . . . 44
5.5 Analysis of intraoperative surgery type . . . 45
5.6 Analysis of postoperative variables . . . 46
5.7 Analysis of postoperative drugs usage . . . 47
5.8 Analysis of Continous variables . . . 48
5.9 Comparision with previous data study . . . 49

6 Variable selections 50
6.1 Literature-based variable selection . . . 51
6.1.1 Summary of literature review findings . . . 53
6.2 Marginal importance assessment . . . 57
6.3 AIC forward selection method . . . 60
6.4 Comparing different variable selection techniques . . . 62

7 Model validations 63
7.1 Estimated coefficients . . . 63
7.1.1 Scenario 1 with any complications . . . 63
7.2 In-sample validation . . . 72
7.3 Cross-validation . . . 75

8 Exploring weighted logistic regression 80
8.1 Estimated coefficients in weighted logistic regresion . . . 80
8.2 In-sample validation . . . 84
8.3 Cross-validation . . . 86
8.4 Conclusion . . . 88

9 Discussion 89
9.1 Interpertation of key findings . . . 89
9.2 Limitations and future directions . . . 89
9.3 Other . . . 90

10 Conclusion 91
10.1 Summary of research contributions . . . 91
10.2 Final thoughts . . . 92

References 95

A Nested data 99

B Variable names of unpreprocessed data 101

C Summary of relations in bewteen variables from literature study by P. Kirchhoff et al. 109

D Variable meanings 111

E Overview of outcome related variables 113

F CD of complications complete information 114

G Modifications text to numbers 116

H Additional information of studies included in literature based variable selection 117

I Estimated coefficients of logistic regression models predicting any complications 118

J Estimated coefficients of weighted logistic regression models using literature based variables 124

Nomenclature

Abbreviations

Medical abbreviations	Definition
ERAS	Enhanced Recovery After Surgery
LOS	Length Of (hospital) Stay
CID	Colon surgery In Daycare setting
MST	Medisch Spectrum Twente
CCI	Charlson Comorbidity Index
VAS	Visual Aided Scale
POD	Post Operative Day (often followed by a specific number)
IV	Intra Venous (means in the veins directly, mostly referring to method of administering fluids)
CD	Clavien Dindo (scale for severity of complication based on method of treatment)
ADL	Activities of Daily Living (activities that one does to take care of oneself)
WHO	World Health Organisation
ASA	American Society of Anesthesiologists
BMI	Body Mass Index (kg/m^2)

Medical terms

Most medical terms are also explained in Chapter 3. This overview contains the ones that are occurring more often throughout the thesis.

Term	Meaning
Bowel	The small and large intestine and the rectum
Stoma	An artificial connection made by operation, connecting part of the bowel to the outside directly through a hole in the belly.
Primary stay	The hospital stay of the surgery included in the data (no rehospitalisation)
Anastomosis	The surgically made connection between the bowel, bowels are reconnected to themselfs after removing part of the bowel.
Abdomen	Part of the body, medical term for the belly
Abdominal	Concerning the abdomen
Comorbidity	The presence of one or more other conditions influencing health beside the disease of interest
Anaesthesia	Medication administered to patient to be unaware of the surgery
Anesthesiologist	Medical doctor in charge of anaesthesia
Perioperative	Throughout the whole process that is surgery (before during and after surgery)
Pre- inter- and post-operative	Before, during or after the surgery.
Abdominal cavity	Inside of the abdomen
Prophylaxis	A protective treatment, in this context, to avoid a (specific) complication
Thrombosis	The complication due to the solidification of blood (blood clots)
Thromboprophylaxis	The treatment(s) to avoid blood clots.
Anaemia	Condition of not having enough functioning red blood cells
Epidural	An injection in the spine often an anaesthesia
Metastasis	(Cancer) cells that have moved from their origin
Colorectal	Concerning the colon and rectum
Gastrointestinal	Concerning the stomach and/or the intestines
Ileus	Paralysis of bowels
Mortality	Refers to the number of deaths
Morbidity	Condition of being sick
Anastomotic leak	Leakage of bowel contents through the anastomosis to the abdominal cavity
Vasoactive drugs	Drugs influencing the health function and or veins and arteries
Colloids	Type of fluid that can be administered through Intra Venous (IV)
Conversion of surgery	Changing the surgery from laparoscopic or robotic to open surgery during the operation.

1 Introduction

All surgeries come with a risk. In medicine, the occurrence of an injury or disorder besides the treated disease is called a complication. A complication can differ in many aspects. Some complications are deadly, while others express discomfort. This thesis aims to mathematically model the risk of complications that comes with elective intestine surgery in the Medical Spectrum Twente (MST). Studies in medicine are often done in small randomised trials or with large observational data sets. We consider the data set that will be studied not to be large or small. However, we think the data set's strength lies in the many different types of information contained. While modelling, we find which variables indicate a patient's risk of complications. We should keep in mind the limits of observational studies [51].

In the data available to us, 562 surgeries on 530 patients are recorded. The data set used requires in-depth feature engineering since many explanatory variables can be considered as rare events (the variable differs with a low frequency). Surgery and the preparation for surgery are highly adjusted to the differences in the patients and illnesses. The average length of stay in the hospital for the operations is 5.83 nights, and the average length of primary stay is 4.80 nights. 31% of the patients get some sort of complication after surgery. These complications range from nausea, or vomiting to surgery-related death. Of all the complications, 31% is classified as a severe complication. This means that in the entire data set after 9.61% of the surgeries, severe complications are present. Severe complications can also be considered a rare outcome, leading to difficulty in modelling. A complication is referred to as severe in this thesis if a surgical intervention under general anaesthesia is necessary to treat the complication, if the patient suffers from organ failure or if the patient dies from a cause related to the surgery.

This thesis identifies variables useful for predicting complications in bowel surgery performed at the Medical Spectrum Twente under their ERAS protocol for bowel surgery. ERAS is a society publishing guidelines (for more information concerning ERAS, see Section 1.2). Most surgeries undergoing this protocol are bowel resections. The reason for performing a bowel resection can differ, some patients have a tumour that needs to be removed surgically, and some patients suffer from an inflammation that can be treated surgically. This thesis studies how to make a patient-level prediction model for complications and its predictive power. Since the available data set contains many variables compared to observations and the surgery is a process containing many steps, it is important to study the available literature on this topic. Much literature is available about bowel surgery and its complications, but few studies make prediction models [49][34][29][18].

1.1. Overview of bowel surgeries

The most common surgery in this thesis is called resection. In this surgery, part of the intestines or rectum is removed. Sometimes, the remaining parts of the bowels are reconnected during surgery, in other cases a stoma is placed. A stoma is a connection directly to the outside of the body through the belly. Some bowel resections are performed with open surgery and others are performed laparoscopically. The reversal, placement and moving of a stoma are surgeries also studied in this thesis.

A good recovery is not easily defined. Postoperative variables are not commonly used in prediction models. One study emphasises that postoperative recovery is an energy-requiring process that has four dimensions – physiological, psychological, social and habitual recovery [1]. We will look at the presence of complications post-surgery. Note that this does not mean a quick recovery.

1.2. Enhanced Recovery After Surgery (ERAS) protocols

The ERAS Society is a non-profit academic society based in Sweden [15]. It aims to improve healthcare practised around the globe. To achieve this goal, it publishes guidelines for different kinds of surgeries. It argues that recovery is a multifaceted process and therewith takes a holistic approach. The published guidelines promote a quick healthy recovery and include stress-avoiding measures, which is known to relate to bad recovery from surgery. In the MST, an ERAS protocol is implemented to improve recovery and reduce the length of hospital stay. In our data, all the patients underwent the ERAS protocol. A reduced length of hospital stay can be beneficial for both the hospital and for the patient's recovery. Patients tend to perform less beneficial physical movement while staying in a hospital compared to recovering at home. Since movement stimulates the recovery of bowel functions it is endorsed in the hospital [15]. Also by aiming for a quick release from the hospital, the patients' mobility is increased. Recovery at home is also perceived as more comfortable for the patient.

1.3. Data collection methods

There are different types of doctors and nurses, all working to improve surgery outcomes. The data we have access to, is collected from different sources in the hospital. In Table 1.1 we see which people perform tasks to improve surgery outcomes. Some data come from the anaesthesiologist and other data is provided by the surgeon. The anaesthesiologist records data with a computer application specifically designed for anesthesiologists. This leads to consistent data quality. The data coming from the surgeon is extracted from a text written postsurgerically. This extra step of extraction can lead to rare cases of errors or misinterpretations in the data.

Person	**Task**
Patient	undergoes surgery and recovery
Surgeon	decides on plan of action and performs the surgery
Assistant physician or specialised nurse	assists the surgeon and decides on a plan of action
Anaesthesiologist	coordinates perioperative (before, during and after the operation) pain treatment
Case manager oncology	coordinates the pathway of the patient in case of a cancer diagnosis.
Stoma nurse, oncology nurse	informs patients and other nurses
Nurse	informs other staff about the recovery of the patient
Holding / Recovery nurse	takes care of patients when going in or out of surgery

Table 1.1: Specialist working on improving surgery outcomes.

Some variables in the unpreprocessed data contain values further specified in another variable. For example, the variable named *Stomal Procedure* specifies which stoma procedure was performed. In some instances, there are errors in the data. Sometimes a stoma procedure is the main surgery, which is then stated in the variable named *Main surgery*. However, for one surgery it was specified that the main surgery was a stoma procedure and the variable stating what kind of stoma procedure was performed stated no stoma procedure was performed. These values contradict each other and we expect this to be some error in the data collecting.

Nurses are expected to follow protocols, whereas doctors also rely on decision-making they learned during their medical study. In the protocol, we find some decision rules. For instance when a patient can be released from the hospital or for which surgeries laxatives should be given preoperatively.

The patients provide some information about how they experience their health. A VAS is a Visual Aided Scale, which means that besides numbers there are some pictures of a face experiencing comfort or discomfort present. On the first 3 postoperative days, the patients state their nausea and pain experienced on a VAS. The patients also keep track of how much they ate and how many hours they were out of bed.

The original data set used in the thesis contains 375 variables for 562 observations. Many variables contain information about rare events, which complicates modelling. The frequency of outcome differs depending on which outcome is considered. We distinguish between severe or any complications and we distinguish between post-operational or after primary stay complications. After preprocessing, 51 explanatory variables are constructed, of which 19 preprocessed variables are preoperative, 18 preprocessed variables are intraoperative, and 14 preprocessed variables are postoperative.

1.4. State of the field

A lot of medical research has been conducted on bowel surgery. Sometimes prediction models are made for use in medical practice, mostly using a regression model or a Cox proportional hazards model [43] [42] [24][37]. A few studies were found predicting surgery outcomes using statistical models besides regression or Cox proportional hazards models.

Medical research can be done with observational studies, as in this thesis, or more conventionally with clinical trials. It is not uncommon for prediction models to be made with more observations than we have available [47]. The metastudy by Souwer et al. [47] included 25 studies that are used to develop a prediction model. The number of observations in the included studies ranged from 119 to 235407. Clinical trials often include fewer patients than observational studies [49]. The metastudy by Varadhan et al. [49] included 6 studies on colorectal surgery, the number of patients was between 25 to 103. The Meta study concluded that for open surgery, the length of stay could be shortened by 2.5 days and the number of postoperative complications can be reduced significantly with an ERAS program. Many medical studies use standard statistical methods, where one compares two groups of patients with statistical testing. However, p-values can be manipulated by trying different tests and significance levels until the result is as desired. Misinterpretation of p-values can lead to wrong conclusions. A p-value is the probability of seeing a value as observed or a value more extreme under the null hypothesis. A small p-value is sometimes considered as proof for dismissing the null hypothesis. However, it is no proof of the alternative hypotheses, and multiple hypotheses should be considered before drawing any conclusions [26].

Scholars [18] relate many risk factors, for example: 'Male gender is associated with increased anastomotic leakage rates after low rectal anastomosis.'. In this study specific variables are linked to specific complications. It implies that a complex dependency structure is present within different explanatory variables and averse surgery outcomes. Further, it states that the surgeon's intuition and gut feeling are good predictors of postoperative outcomes. The literature study refers to the study by Markus et al. [28] which compares the prediction of the surgeon right after surgery with the POSSUM score. Here 1077 planned and emergency surgeries were taken into account. The study concludes that the surgeon's gut feeling is more accurate than the POSSUM score. The POSSUM score (Physiologic and Operative Severity Score for the Study of Mortality and Morbidity) uses 12 Physiological Parameters and 6 operative parameters. It was shown that this score mostly struggles when assigning very high or low risks. Where the surgeon's highest predicted chance of morbidity was 60 for the possum score it was 90. It further demonstrated that the surgeon's predictions are better for elective surgery than emergency surgery. A literature study from 2010 by Kirchhoff, Clavien, and Hahnloser [18] identifies research questions that are difficult to answer with the available literature. For example, the lack of a uniform definition of nutritional status led to difficulty in comparing studies assessing nutritional status in patients. These discrepancies lead to difficulty in relating findings from other studies to the variables in our data. Other open questions from literature study by Kirchhoff, Clavien, and Hahnloser [18] are about risks in conversion from laparoscopic surgery to open surgery and preference for hand sewing

or stapling for anastomoses. To highlight the complexity of the relationships between the different variables, a summary of the relations as described by the study is made (see Appendix C).

Treatments before surgery are specific to both the patient and the disease, for example when treating anaemia preoperatively one has to take into account the reason for operating to ensure that the anaemia treatment is effective. It is known that oral iron medication works badly for patients with inflammatory bowel disease [15]. In the unprocessed data set, we have 1 patient with inflammatory bowel disease and anaemia (from the 12 patients with inflammatory bowel disease), this patient indeed received an IV with supplementary iron instead of oral iron treatment. This patient with anaemia and inflammatory bowel disease did not develop any complications after surgery. We further see that the patients with *severe heart disease* and anaemia get either no iron supplements or IV iron administered (12 observations have anaemia and *severe heart disease* both present, of these 12 observations 8 treatments are known, these treatments are 6 times no treatments, and 2 times an IV iron treatment).

Since comorbidities (the presence of one or more other conditions influencing health beside the disease of interest) are increasingly present in old age, studying the effect of old age and comorbidities separately is difficult. Expert elicitation by Arts [5] states that: 'They (referring to surgeons) decided that the age was not an important criterion, since the vitality of the patient is not based on the age, but is defined in the ASA class and WHO performance score.' A study by Kirchhoff, Clavien, and Hahnloser [18] states that laparoscopic surgery can be considered for anyone despite age, implying this is not the case for open surgery.

1.5. Research objectives

We aim to answer the questions:

1. What variables lead to a high risk of complications after bowel surgery?
2. Can complications be predicted before bowel surgery, after surgery or after hospital stay?

We aim to compare methods to predict the risk in surgery, within the available data set. This thesis aims to identify where patients' risk of complications lay under the current protocol. To do this we will first give an overview of the statistical methods that will be applied. After that, we study the available data and we process the data for a more efficient risk modelling. At last, we will fit the chosen models and we will evaluate and compare the results.

1.5.1. Overview of models to be developed

To answer the question in which case the complications can be predicted, we will model complications in 3 different scenarios. In Scenario 1, a prediction of complications after surgery is made before the surgery takes place. Here we see whether the patient's characteristics and previous treatment can explain the occurrence of complications. In the second scenario called Scenario 2, we make a prediction right after surgery of any postoperative complications. In this scenario, the information collected intraoperatively is included. When comparing Scenarios 1 and 2 we see the effect of intraoperative data on the prediction. In Scenario 3 a prediction is made of complications after the the primary hospital stay. Here we see which variables indicate a safe release from the hospital. In this scenario, we can include information about recovery during a hospital stay. We make a difference in outcomes after surgery, used in Scenarios 1 and 2, and outcomes after the primary stay, used in Scenario 3. In this way, we can study the relation of the post-operational variables during primary stay on complications without the explanatory variables and outcome arising simultaneously.

We will apply 3 different variable selection methods and 2 different prediction methods. We make different selections of variables by applying the AIC forward algorithm, by looking at marginal importance and by studying importance in other studies. We will use weighted and non-weighted logistic regression. This leads to many models that will be studied, and compared. We listed the different scenarios, variable selections and prediction models in Figure 3.2.

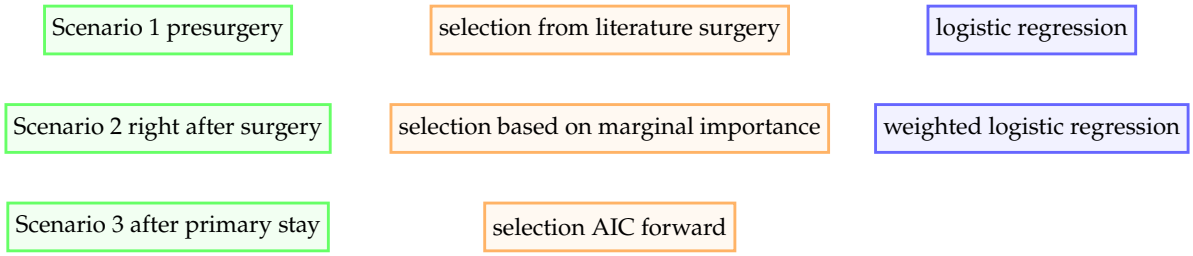

Figure 1.1: Combination of scenarios, variable selections and prediction models to be studied

We first define our mathematical framework (Chapter 2) and then preprocess our dataset so that we have no missing values and our variables are limited to useful information (Chapter 4). We make tables to study the preprocessed data (Chapter 5). In Chapter 6 we will apply different variable selection methods to the preprocessed data. Then, we make our models and study their predictive power in Chapter 7. We first study the in-sample validation. We do not keep one separate test data since there are many rare events and the influence of separating a single test set is therefore expected to be large. For this reason, we decided to do repeated cross-validation to study out-of-sample results and see the models' accuracy. With the cross-validation, we discover the robustness of the model. At the end of this chapter, we compare our findings for the different models made. We take a look to see whether the model performance increases by switching to the weighted logistic regression (Chapter 8). Afterwards, we limit the outline of our study in the discussion in Chapter 9 and finally, summarise our findings in the conclusion in Chapter 10.

2

Mathematical background and methodology

In this chapter, we will explain the mathematical background of the thesis. The following chapters focus on the results and their interpretation. We explain the variable selection methods, the data descriptives, the statistical models to be applied and the measures to validate the modelling. We start by explaining our notation.

We model the uncertainty in the outcome of interest by the variable Y, a binary value indicating the presence of complications (see more in Chapter 3). We aim to study Y with respect to other explanatory variables $X_1, \ldots, X_m$. We have n observations and m explanatory variables.

We will use a lowercase letter for the assigned values of the random variables. So y is a specific value. We aim to find and validate a model that estimates the probability of our outcome. We will denote $P(Y = y|X = x)$ for the estimated probability of Y being equal to y when given X equals x. When we want to emphasise the parameters used in the model we denote $P(y|x, \beta)$, where β specifies the used parameters.

We use $|\ |$ for the size of the set. For example, $|Y = 1|$ equals the number of observations for which the outcome is present.

2.1. Descriptive statistics

In our data, different variables are included, discrete and continuous variables. To quantify relationships between variables we will use odds ratios, Kendall's tau and statistical testing.

2.1.1. Interpreting odds ratio

The odds ratio is a measure of the association between two binary variables. We will always use it with an outcome of interest. So we use it for a pair (Y, X_i). The contingency table for a binary X_i is given in Table 2.3.

	$Y = 1$	$Y = 0$
$X_i = 1$	$P_{11} := \|Y = 1 \ \& \ X_i = 1\|$	$P_{01} := \|Y = 0 \ \& \ X_i = 1\|$
$X_i = 0$	$P_{10} := \|Y = 1 \ \& \ X_i = 0\|$	$P_{00} := \|Y = 0 \ \& \ X_i = 0\|$

Table 2.1: Definitions used for defining of the odds ratio (formula 2.1). $|Y = 1 \ \& \ X_i = 1|$ is the number of observations for which both $Y = 1$ and $X_i = 1$.

Now the definition of the odds ratio is given by:

$$\text{odds ratio} := \frac{\frac{P_{11}}{P_{01}}}{\frac{P_{10}}{P_{00}}}. \tag{2.1}$$

For non-binary categorical variables, we can use multiple odds ratios by considering a binary variable stating whether the categorical variable takes on a specific value. We will use the odds ratio mostly with complications and an explanatory variable.

2.1.2. Kendall's tau and its applications

Kendall's tau is a correlation based on the number of concordant pairs of observations. Say we have the pair of explanatory variables (X_1, X_2). We then call the observations from the variable $\mathbf{X}_1 := (x_1^1, x_2^1, \cdots x_n^1)^T$, and use $x_1^2, x_2^2, \cdots x_n^2$ for the observations of X_2.

The pair of observations (x_i^1, x_i^2) and (x_j^1, x_j^2) are called concordant when $x_i^1 < x_j^1$ and $x_i^2 < x_j^2$ or when $x_i^1 > x_j^1$ and $x_i^2 > x_j^2$. So a pair of two observations is called concordant when for one pair of observations both values are bigger than the values of the other pair. When the number of concordant pairs for the variables X_i and X_j is large, there is a positive relationship between the variables.

Kendall's tau is calculated by:

$$\tau = \frac{\text{number of concordant pairs} - \text{number of not concordant pairs}}{\text{number of pairs}}. \tag{2.2}$$

When there are ties in the data, the following is used (τ_b) :

$$\tau_b = \frac{\text{number of concordant pairs} - \text{number of not concordant pairs}}{\sqrt{(n_p - n_{t1}) \cdot (n_p - n_{t2})}}, \tag{2.3}$$

where we have

$$n_p := \text{number of pairs}, \tag{2.4}$$
$$n_{t1} := \text{number of pairs with ties in first quantity}, \tag{2.5}$$
$$n_{t2} := \text{number of pairs with ties in second quantity}. \tag{2.6}$$
$$\tag{2.7}$$

Kendall's tau does not assume the shape of the relationship between the variables, unlike the Pearson correlation coefficient, which measures whether the points follow a linear relationship. Since we want to compare numbers from different units in this thesis, we use Kendall's tau. For example, we use Kendall's tau to relate intraoperative blood loss and the number of nights spent in the hospital.

2.1.3. Key statistical tests

P-values will be used for testing an association between variables in our data set. We mostly use p-values to see up to which degree a relationship proven in other studies is also visible in our data set. A p-value represents the chance of seeing a difference at least as large as the one observed under the null hypothesis. Often the null hypothesis states that two variables are independent.

In the table below we see which testing method we decide to use to test for different combinations of variables:

Variable 1	**Variable 2**	**Test**
Binary	Binary	Fisher exact test
	Continuous	Two-sided student T-test
	Discrete	Chi-squared test
Discrete	Discrete	Chi-squared test
Continous	Continous	Kendall's tau test.

Table 2.2: Testing methods used for different combinations of variables, to test for independence.

Further, a few binomial tests are performed to determine whether the patient population could be considered a random sample of the Dutch population. This test is not used to study any dependency between variables.

The null hypothesis of the Fisher exact test states that the observed values in the contingency table are independent. We used the Fisher exact test for two binary variables. It assumes that the variables were independently sampled from the observed marginal distribution. The null hypothesis of the Fisher exact test is that two variables are independent, this leads to a known distribution for two binary variables. For two binary variables with fixed margins, one number in the contingency table defines all other numbers in the contingency table. Therefore the chance that one number in the table is a certain value also sets the other numbers in a contingency table. The test calculates the chance of seeing the observed number in one of the cells in the contingency table assuming independence. The used distribution is called hypergeometric distribution.

The null hypothesis of the chi-squared test is that two discrete variables are independent. The null hypothesis is similar to the Fisher exact test. However, the method differs. We used the chi-squared test for two discrete variables, that are not both binary variables. The chi-squared test assumes independence between two discrete variables. It uses the observed marginal distributions to find the expected value of each cell in the contingency table of the two discrete variables. For each cell in the contingency table, the observed number of times is normalised, then the normalised numbers are squared and summed to obtain the test statistic. Squared standard normal variables that are added follow a chi-squared random variable with degrees of freedom equal to the number of standard normal variables. Afterwards, the p-value is found by comparing the test statistic with the chi-squared distribution with equal degrees of freedom as cells in the contingency table.

The Welch's t-test null hypothesis states that the true means between two groups is the same. We use Welch's t-test to see if a continuous variable is related to a binary variable, by seeing if the mean for the continuous variable differs when the binary variable is present and when the binary variable is not present. It assumes that the variables are normal. The Welch's t-test uses the t-distribution (with n-1 degrees of freedom) to see if the population means are similar. This test does not study a difference in sample variances.

The null hypothesis of Kendall's tau is that the variables are independent. We use the Kendall's tau test for two continuous variables. For two independent variables, one expects the number of pairs to be concordant and the number of pairs that are discordant to be similar, and thus the true Kendall's tau to equal zero. For two independent variables, the expected value of the Kendall's tau is zero. The variance of Kendall's tau for independent variables can be calculated. We then compare the Kendall's tau with the normal distribution to obtain the p-value.

2.1.4. Examples

Say we want to see whether gender and smoking are related in our data. We thus make the contingency table (see Table 2.3)

	Female	**Male**	
Non-smokers	220	212	432 non-smoker
Smoker	61	66	127 smoker
	281 woman	278 man	559 observations

Table 2.3: Contigency table of gender and smoker status. limited to the pair-wise complete observations. Here we included the patients who stopped smoking because of surgery in the smoker group.

If we fix the number of observations (559), the number of smokers (432) and the number of female patients (281), we could still fill in the contingency table in different ways. The Fisher test counts all the possible ways in which we can fill in the table and it then calculates the probability of seeing a table more unlikely than observed.

For the female non-smokers the expected value under the null hypothesis is (432/559)(281/559)559 = 217.2. So we want to know the probability of the female non-smokers to be more or equal to 220. The p-value from the Fisher test is twice this probability. The distribution used to find this probability is called hypergeometric distribution.

If we apply the chi-squared test on Table 2.3, we use that assuming independence we have the expected values as in Table 2.4.

	female	male	total
no smokers	0.503* 0.773 * 559= 217.3	0.497*0.773* 559 = 214.8	432 no smoker
smoker	0.503 *0.277*559= 63.8	0.497* 0.227*559 = 63.1	127 smoker
	281 man	278 woman	559 observations

Table 2.4: Expected values under the observed marginal distribution with independence.

The chi-squared test now does the following; we look at the differences between the observed and the expected values. These are

$$217.3 - 220 = -2.7, \tag{2.8}$$
$$214.8 - 212 = 2.8, \tag{2.9}$$
$$63.8 - 61 = 2.8, \tag{2.10}$$
$$63.1 - 66 = -2.9. \tag{2.11}$$

The test statistic is the sum of these differences squared over their expected value, so

$$\frac{2.7^2}{217.3} + \frac{2.8^2}{214.8} + \frac{2.8^2}{77.9} + \frac{3^2}{63.1} = \frac{7.3}{217.3} + \frac{7.8}{217.8} + \frac{7.8}{77.9} + \frac{8.4}{63.1} = 0.3. \tag{2.12}$$

the test statistic should follow a chi-squared distribution with 1 degree of freedom. The degrees of freedom equal (the number of columns in the contingency table -1) ·(the number of columns in the contingency table -1). In this case, this leads to a p-value of 0.57. We conclude that smoking and gender are not related in our data.

The t-test looks at whether population means differ. On average the patients that develops a complication lose 98 ml more blood than the patients who do not suffer from complications. Say we want to see if this difference in blood loss (X_1) tests statistically relevant with any complications (X_2). We use the test statistic:

$$T = \frac{\text{mean}(X_1; X_2 = 1) - \text{mean}(X_1; X_2 = 0)}{\sqrt{(\frac{\text{var}(X_1; X_2=1)}{|X_2=1|} + \frac{\text{var}(X_1; X_2=0)}{|X_2=0|})}} = \frac{167.9 - 69.5}{\sqrt{\frac{186447}{174} + \frac{31542}{388}}} = 2.9. \tag{2.13}$$

We now compare the test statistic with the T distribution with the corresponding degrees of freedom. This leads to the p-value of 0.0043. We conclude that blood loss is significantly higher for patients who will develop a complication, and the difference in mean blood loss is not coincidental.

The Kendall's tau test shows us whether we can consider the calculated τ a coincidence of independent variables. Say we want to apply the Kendalls tau test to the age and length of stay. The Kendall's tau equals 0.050 for this example. We have 556 complete observations for length of stay and patient's age. To compare Kendall's tau with the standard normal variable we first have to divide by its standard deviation as it would be under the null hypothesis. This leads to the test statistic equalling $Z = 1.61$. Since $P(Z \geq 1.61) = 0.054$, we have a p-value of twice this amount; 0.108. This since the event that the Kendall's tau is more extreme than 0.050 is when $\tau > 0.050$ or $\tau < -0.050$. We conclude that age and length of stay are not related in our data.

2.2. Variable selections methods

We will use 3 different variable selection methods, namely literature study, AIC forward and marginal importance.

Literature study

We will make a list of 21 studies. We tabulate accordingly whether a study sees a variable as an important risk factor. Then we will select all variables that are concluded as an important risk factor in at least one of the studies.

AIC forward
This algorithm adds repeatedly one variable to the variable selection if it leads to a model with a lower AIC (the definition of AIC is given in Formula 2.31). In each iteration, it fits as many models as variables not included in the selection at the current step. In the first step, the number of models fit is equal to the number of possible variables. Each model fitted included the variables already selected and one of the variables not included at the step. Then it compares the in-sample AIC of the fitted models. If this finds a model resulting in a lower AIC it adds the variable of the model with the lowest AIC to the variable selection and repeats the process. So at each step, it looks at which variable to include that leads to the best model with respect to the AIC, if including this variable leads to improvement considering the previous variable selection. If it does not find a better AIC the variable selection stops and keeps the selection as it is.

Marginal Importance
This variable selection tests whether the variables and the outcomes are related. The statistical test as explained in Section 2.1.3 will be used. If the p-value is lower than 0.05, we add the variable to the variable selection. The tests used are the first three stated in Table 2.2.

2.3. Model development techniques
We will use a logistic regression and a weighted logistic regression model.

The models we will use return a probability of the outcome (i.e., complication). In other words, we will use probability models as a classification model. Models that return a class without assigning a probability of the outcome are not desirable in this setting. When making a classification based on the observed probability, one could take into consideration which risk is acceptable. The probability of the outcome is considered relevant. After the model returns a probability of the outcome for a certain observation, we classify the outcomes by comparing the probability with 0.5. If the probability under a model is larger than 0.5, we decide the model predicts the outcome to be present.

Logistic regression
Logistic regression models the chance of an outcome, by assuming the probability applied to the inverse of the logistic function is equal to a linear combination of explanatory variables. The probabilities under the model are given by:

$$P(y|x) = \frac{1}{1 + \exp(\beta_0 + \sum_{i=1}^{m} \beta_i \cdot x_i)}, \tag{2.14}$$

where β_i is a parameter for the variable i and β_0 is a parameter not associated with an explanatory variable and x_i is the value the variable X_i. Both the weighted and the non-weighted logistic regression are of this form, only the way the β's are found differs for the methods. Fitting a logistic regression with maximum likelihood estimates the parameters β, with

$$\arg\max_{\beta} \sum_{i=1}^{n} \log P(Y = y_i | X = x^i, \beta). \tag{2.15}$$

Weighted logistic regression
Weighted logistic regression is a way to counter the underestimation due to the outcome being less present in the data [36].

When one uses a weighted logistic regression one estimates β differently from the (unweighted) logistic regression. This method adds a different amount of importance to different classes. In this thesis, the class is equal to the outcome. We denote c for the values the outcome can take, so $c \in \{0, 1\}$. The class weights we denote by w_c. We estimate β by solving:

$$\arg\max_{\beta} \sum_{c \in \{0,1\}} \sum_{i|Y=c}^{n} w_c \log P(Y = c | X = x^i, \beta). \tag{2.16}$$

The balanced class weights are defined by 1 over the size of the class. In our case:

$$w_0 := 1/|Y = 0|, \tag{2.17}$$
$$w_1 := 1/|Y = 1|. \tag{2.18}$$

Since one can multiply the loss function (the function maximised in 2.16) with a constant, we see that equivalent with the balanced class weights are the weights:

$$w_0 := 1, \tag{2.19}$$
$$w_1 := |Y = 0|/|Y = 1|. \tag{2.20}$$

These are the weights we will use. We refer to $|Y = 0|/|Y = 1|$ as the balanced weight. In this thesis, we will also use a fixed class weight. The fixed weights will be of the shape:

$$w_0 := 1, \tag{2.21}$$
$$w_1 := 2. \tag{2.22}$$

This is equivalent to including every observation with a complication present twice. We suspect that weights too large will lead to a too big influence of the outcome group and include a bias on the outcome group.

2.4. Model validation approaches

In this section, we explain the different method used to validate the models that will be made in this study.

2.4.1. In-sample validation

With in-sample validation, we see how much the data used to estimate the model's coefficients, supports the resulting model. In-sample validation does not indicate how well the model will perform when applying the model to new observations. When we have a model that performs well in-sample, but not in the out-of-sample we say this model overfits.

2.4.2. Cross-validation strategies

In cross-validation one fits multiple models to evaluate the effect of keeping some data separate, has on the resulting model. One can obtain multiple out-of-sample validations that are calculated on multiple test sets. We will use a cross-validation that ensures that each observation has been in a test set the same times as the other observations. We consider this to be a good choice due to the presence of rare events in the outcome and to a lesser extent also in the explanatory variable.

In our cross-validation, we create 20 times a 10-fold on the dataset. For each 10-fold we fit 10 models, all these models exclude one different fold for external validation. This way 200 models were evaluated in the cross-validation. For each 10-fold the entire data set has been in the test set one time. Since we repeat this 20 times, each data point is used in an external validation exactly 20 times. We calculate the average and standard deviation of our validation measures out-of-sample.

2.4.3. Model validation measures

In this thesis, the following scores are calculated to evaluate any models:

- Misclassification rate
- False positives (FP)
- False negatives (FN)
- Brier score (BS)
- Log-likelihood (LL)
- Matthews correlation coefficient (MCC)
- Threat score (TS)
- False omission rate (FOR)
- Balanced accuracy (BA)
- Area under the Receiver Operating Curve (auc ROC)
- Bayesian information criterion (BIC)
- Akaike information criterion (AIC)

In Table 2.5 we see the definition of true positive (TP), false positive (FP), true negative (TN) and false negative (FN). This table is a contingency table with the true outcome and the outcome predicted under a model.

	Predicted to have the outcome	Predicted to have no outcome
Actually have the outcome	TP := Number of predictions that are predicted to have the outcome present and have the outcome	FN := number of predictions that are predicted to have no outcome present but do have the outcome
Do not actually have the outcome	FP := number of predictions that are predicted to have the outcome present but not have the outcome	TN := number of predictions that are predicted to not have the outcome present and do not have the outcome

Table 2.5: Definitions of true positive (TP), false positive (FP), true negative (TN) and false negative (FN).

Misclassification rate

The misclassification rate is the percentage of mistakes in all the predicted outcomes:

$$\frac{FN + FP}{TP + FN + FP + TN}. \tag{2.23}$$

Brier Score (BS)

The Brier score is the mean of the squared distance between modelled probability and outcome:

$$BS := \frac{1}{n}\sum_{i=0}^{n}(P(Y = y_i|X = x^i) - y_i)^2. \tag{2.24}$$

(2.25)

Log-likelihood

Expresses the probability of seeing the data in the model. The higher the log-likelihood the better the data fits the model.

$$\hat{ll} := \sum_{i}^{n} \log(P(Y = y_i|X = x^i)). \tag{2.26}$$

Matthews Correlation Coefficient (MCC)

$$MCC := \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{2.27}$$

The Mattews correlation coefficient takes values in [-1,1]. MCC equals 1 in case only correct predictions are made. It equals -1 in case there are only wrong predictions. It performs well only if the predictions are well for the group with the outcome and the group without the outcome. A study from 2021 concludes that the MMC should be considered in many different fields [8].

Threat Score (TS)

$$TS := \frac{TP}{TP + FN + FP}. \tag{2.28}$$

The threat score takes values in [0,1]. It is the ratio of true positive predictions overall true positive and any false predictions. So when there are no wrong predictions (FP + FN = 0) then the TS equals 1. If there are no true positive predictions the TS = 0. In other cases, the results are somewhere in between.

False Omission Rate (FOR)

The false omission rate is the false negatives over the number of negative predictions.

$$FOR := \frac{FN}{TN + FN}. \tag{2.29}$$

Balanced Accuracy (BA)

$$BA := \frac{TPR + TNR}{2} = \frac{1}{2}(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}). \tag{2.30}$$

The balanced accuracy is formed by the average of the true positive rate and the true negative rate. It takes an unbalanced outcome into account in a straightforward way.

Area under the Receiver Operating Curve (auc ROC)

The c statistic, or area under the ROC (auc ROC), is a measure of goodness of fit. It can be considered as the chance of a random pair having concordant predictions under the model. Say we have the pair X_i, X_j where $Y_i > Y_j$, then we wish to see in the model $P(Y = 1|X_i) > P(Y = 1|X_j)$. The AUC is the number of pairs where this indeed happens, over all the pairs X_i, X_j where $Y_i > Y_j$.

This measure does not consider any classification made, but the order of the probabilities. This measure does not take into account the rarity of the outcome.

Akaike Information Criterion (AIC)

The AIC is given by;

$$AIC := 2k - 2\hat{l}l, \tag{2.31}$$

where k is the number of parameters. A small AIC is preferred. The AIC creates a trade-off between the number of parameters and the likelihood. Often a better likelihood is obtained when using more parameters. However, having more parameters is not always desired. A small AIC means that the likelihood is considered sufficiently large to compensate for the increase in the number of parameters.

Bayesian Information Criterion (BIC)

The BIC is given by;

$$BIC := k\ln(n) - 2\hat{l}l, \tag{2.32}$$

where k is the number of parameters. A small BIC means a likelihood big enough for the number of parameters. The BIC is similar to the AIC they both use the $\hat{l}l$ but, penalise for the number of parameters differently. The BIC penalises more for an increase in the number of parameters than the AIC does.

2.4.4. Assessing model performance for rare outcome

Since some of the outcomes studied can be considered rare events, a model that predicts everyone without any complications could perform relatively well. A model like this is of course not informative and not desired. For this reason, we include measures that take the rarity of the outcomes into account. Matthews correlation coefficient (MCC) and Threat score (TS) work well for imbalanced outcomes [16]. The balanced accuracy accounts for imbalanced outcomes too. We will also report on the number of false negatives and false positives, these numbers should be compared to the number of observations and the number of outcomes. The number of outcomes can differ over the models made in this study.

3

Data description

The goal of this chapter is to sketch the character of the available data. We start this chapter by giving the context of the data-generating process. Then we go over the data collected chronologically. We start with patient characteristics which consists of information about the patient, separated from the disease or treatment, it indicates the status of their general health. Afterwards, we see preoperative variables, we learn that preoperative measures exist to improve recovery. The data collected during surgery is listed. These include many actions from the surgeons. Postoperative variables mostly include an indication of the rate of recovery. At last, we investigate the data collected about the complications.

The data set we will study contains information about patients who underwent surgery performed under the ERAS protocol in the hospital: Medisch Spectrum Twente (MST). MST is a large hospital in Enschede, The observations used are collected in the years 2020-2022 and contain information about several types of planned bowel surgeries. As the implementation of the ERAS protocol in the MST reduces the length of hospital stay, the MST wishes to explore whether surgeries can be performed safely as ambulatory surgery. This means the patient is sent home on the day of surgery. This can lead to a sped-up recovery for a patient. However, it also brings different risks since patients will be monitored in a way different from usual.

The data set provided by the MST consists of information about 562 different surgeries on 530 different patients. For these surgeries, 375 different variables were collected. The list of all variables in the unpreprocessed data set is given in Appendix B. The data includes general information about the patient such as: gender, age, whether they smoke, whether they have diabetes, and information about the surgery, for example, the type of surgery (laparoscopic or open) or the blood loss in ml during the operation. Furthermore, the data contains information about recovery after surgery, for example how much liquid a patient drank or the usage of opioids. The latest collected information for each surgery should be from a follow-up 30 days after surgery. The patient's physical mobility is recorded with the WHO score before and after surgery. The WHO score measures mobility, it increases when a patient is required to spend more time in bed resting while awake.

Not all variables contain information concerning all patients. For example, in the case of a cancer diagnosis, scores about the stage of the cancer are included in the data set. Since not all patients are suffering from cancer, these scores are not present for all patients. More variables are only relevant in specific cases. For example is a patient can only have a screening instrument assigned in case the nutritional status is assessed. A list of these variables that further specify another variable is given in Appendix A.

During the majority of the surgeries, a part of the bowel is removed and the remaining parts are reconnected. The artificially created connection is referred to as anastomosis. In 79% (446 obs.) of the surgeries an anastomosis is created. Some other surgeries remove part of the bowel but do not place an anastomosis but instead create a stoma. A stoma is a surgically made opening in the abdominal wall to the outside of the belly through which the end of the intestine is connected. Bowel content is collected in a bag connected to the stoma. A stoma can be a temporary measure in case the bowels recover without the passage of food through the affected part of the bowel. A stoma can be relocated when more bowel needs to be removed or in case of a stoma complication. For each main procedure, different additional procedures can be performed during the surgery. For example, 88 patients (17% from 562 observations) underwent a surgery called anterior resection of the rectum. During this surgery, the rectum is removed completely or partially. For 60 of the anterior resection surgeries (68% of 88 anterior resections), no stoma was placed or closed (the bowels were reconnected, an anastomosis was placed). For 16 anterior resection surgeries (18% of 88 anterior rectum resections), a colostomy was placed (a stoma from the colon), for 11 anterior resection surgeries (13% of 88 anterior rectum resections) an ileostomy was placed (a stoma from the ileum, the small intestine), and for 1 anterior resection surgery (1.1% of 88 anterior rectum resections), a colostomy was closed during the procedure. We see that during one type of surgery, different stomal procedures can be performed. Further different non-stomal related additional procedures can be performed during the surgery, for example, a hernia repair.

3.1. Overview of patient journeys

We will make predictions with the variables limited to the information gathered up to 3 certain time points. We gain insights about the time a prediction can be made and we make different models for the 3 different scenarios. In this thesis, 2 outcomes of interest are considered; severe and any complications. A prediction of complications after surgery is made before the surgery takes place. In the second scenario called Scenario 2, we make a prediction right after surgery of any postoperative complications. In Scenario 3, a prediction is made of complications after the primary hospital stay. We make a difference in outcomes after surgery, used in Scenarios 1 and 2, and outcomes after the primary stay, used in Scenario 3. In this way, we can study the relation of the post-operational variables during primary stay on complications without these arising simultaneously. The 3 scenarios we study are summarised in Figure: 3.1 and 3.2. The variables in the unpreprocessed variables are not ordered chronologically. We have to classify them as preoperative, intraoperative or postoperative. Some variables will need preprocessing for the variable to fit one of the sets made in Figure: 3.1 and 3.2.

Figure 3.1 shows the patient's path through the hospital, stating the different occasions information is collected from the patient. In red, there are 4 sets of explanatory variables. These are patient characteristics, preoperative variables, intraoperative variables, and post-operative variables. Left of these is denoted in which scenario the information is used. On the right hand side the information about complications is stated. Complications can arise during or after the primary stay. On their right are denoted in which scenarios the information about the complications is used. In Figure 3.2 the 3 scenarios are also summarised. Here we see the explanatory data on top and the outcomes below. In the middle we see the scenarios, the arrows connect scenarios showing the data used within this scenario.

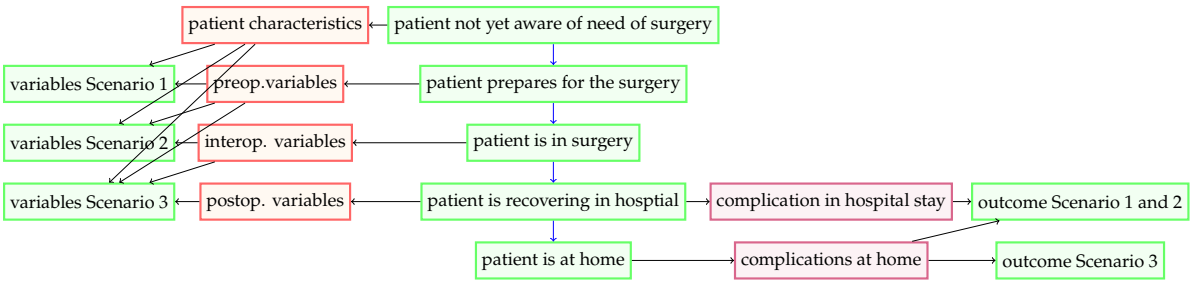

Figure 3.1: Subsets of variables made. Explanatory data is on the left, outcomes on the right and in the middle the pathway of the patient.

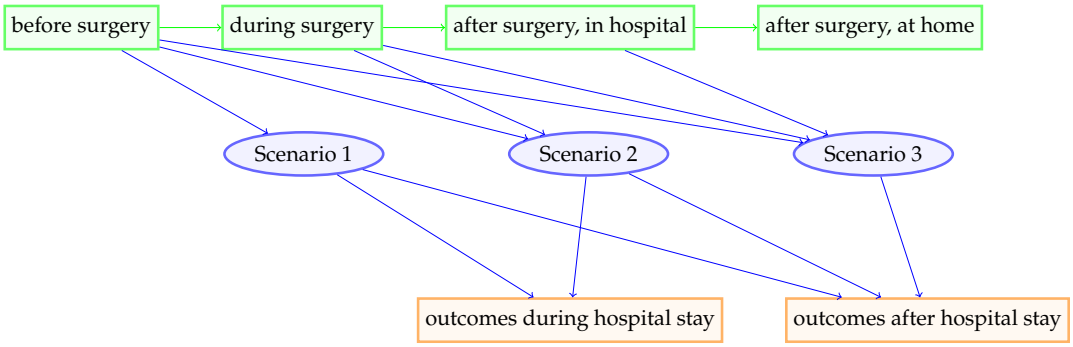

Figure 3.2: Overview of the 3 models that will be made and studied. We have above the explanatory data and below the timing of the outcome.

3.2. Patient characteristics

Patient characteristics are used to indicate the health situation of the patient separated from the illness or cause for operation. Patient characteristics include non-modifiable factors such as age or comorbidities, as well as lifestyle factors like smoking or alcohol consumption. In our data, some general characteristics are included about the patients such as gender and age. Also, some general health factors such as BMI, or whether the patient smokes or drinks alcohol. For some patients, the quantity of alcoholic drinks taken per week is specified. Comorbidities such as heart disease, lung disease or diabetes, are recorded. The BMI correlates to the amount of fat stored in the body. However, the BMI also fluctuates with the patient's fluid balance, which in turn can be distorted due to the administration of medication.

3.3. Preoperative variables

Information collected before the surgery can concern recent patient changes in their well-being, how the surgery is planned or any preoperative treatment. The preoperative treatment information can concern treatment for the disease or treatment to improve surgery outcome. Since stress is a general health risk, the hospital staff is asked to give social support to the patient when possible and is instructed to inform the patient through conversation. Being ill and in need of surgery is stressful in itself. Information about the surgery is presented on paper for the patient to read when desired. Hospital staff is asked to show human interest in the patients' experience [10].

Preoperative body weight change, nutritional status and whether the patient stopped smoking or drinking before the surgery is recorded. Whether the operation is planned laparoscopic, open, through stoma or robotic is known. Data about previous treatments is available. The data set contains a variable indicating whether the patient had another surgery also recorded in this data set. There is information about whether this surgery is the first for the patient in the specific area of the abdomen. A variable indicates whether the patient received immune system suppressive treatment and a variable indicates whether the patient had any chemotherapy and how long ago is present. It is known whether the patient had radiotherapy in the area of operation. If the patient had any long-working sedative medication is known. In the data set not much information about the preoperative fluids intake is available. We know the patient is invited to urinate before the surgery [10].

Before surgery, the doctor performs screenings and takes measures to improve the surgery outcome. If the patients are set on a special diet, whether the patient is tested for anaemia and whether and how they received iron supplements is known. Some medications can be administered to decrease risks. For example, blood thinners can be given to avoid blood clots. We have information concerning whether laxatives or carbohydrates are given before surgery. Laxatives can be given in order to empty bowels during surgery, to avoid bowel spillage. In the nurses' protocol, a list of operations that require laxatives is present. Carbohydrates can be given to improve recovery by supplying the body with nutrients. A set of measurements to avoid a complication is called prophylaxis. In the data set is information about the prophylaxis to avoid infections, thrombosis and postoperative nausea and vomiting (PONV).

3.4. Intraoperative variables

During the surgical procedure, different data is collected. For example: the location of the surgery within the bowels and the surgical name of the surgery. Also, the time passed during the surgery is recorded. A main procedure and sometimes an additional major procedure are specified. Additional procedures include partial removal of the liver (liver resection), removal of cancerous growth (debulking) or hernia repair. We have information about whether a stoma was placed and in which part of the patient's intestine. If an anastomosis was done and if it was made with sewing or stapling the tissue. Information about the type of anaesthesia and narcotics is collected. It is recorded whether general anaesthesia was administered through the air the patient inhales or administered via the veins (IV). Noted is whether a narcotic block was successful in the procedure and whether the recovery of muscle function was monitored. Sometimes an injection is administered in the back, two possible injecting locations are recorded in the data set. Most surgeries were performed with the use of a drug type called nerve blocks. Nerve blocks can be administered locally or by IV. Anaesthesia affects postoperative nausea and vomiting and post-operative pain control [15].

During the surgery, 3 types of fluids can be administered by IV: crystalloids, colloids and blood products. Crystalloids are saline solutions. Colloids refer to a solution with bigger molecules. The costs of colloids are crystalloids are generally lower than colloids [25]. Blood loss is recorded in ml. It is known whether the skin was cleaned before the operation or not. The method of oxygen supply during the surgery is known. Further, whether a drain is placed at the site of surgery or whether a urinary drain was placed during surgery are both recorded.

3.5. Postoperative variables

After operating the patient continues to be monitored. In this section, we explain which variables concerning recovery are collected in the data set. The number of nights for which the patient was connected to a drip is recorded. The number of nights spent with a nasogastric tube if placed during surgery is known. If a nasogastric tube or IV was reconnected is recorded. Furthermore, the day the urinary drain is removed is noted. The weight change of the patient for the first 3 postoperative days is recorded. The Kcal intake of the patient for the first 3 postoperative days is recorded by the patients themselves. The patient also reports pain and nausea experience on the first 3 postoperative days. Nausea or retching and pain scores are awarded by the patient on a VAS with values from 1 to 10. For the first 3 days, it is recorded whether vomiting or nausea was observed by hospital staff. It is known whether laxatives or painkillers were administered, which type of painkillers and sometimes up to which day these are given. For some events the first night they occured is recorded. For example, the first night the patient could eat independently or the number of nights until the patient passed gas or faeces for the first time postoperatively (two different variables). The total amount of IV fluids on the day of surgery is known. For the first 3 postoperative days, we have limited information. Weight change on the first 3 postoperative days is available, but fluid intake is not. Furthermore, the number of nights the patient spent connected to an IV is known as well. A minimum amount of fluids the patient should drink each day during the hospital stay is set. In case the patient fails to take in the prescribed minimum of fluids, the doctor will be alarmed by a nurse and necessary steps to avoid dehydration are taken [10].

Whether the patient was released from the hospital within 30 days and if the patient was sent to a private home or nursing home is also denoted in the data set. Under normal circumstances, the patient is released when they meet the following conditions:

1. Flatus or peristaltic movement is present.
2. Patient takes in a sufficient amount of food and fluids.
3. Patient is mobile.
4. Patient shows no signs of infection or fever.

3.6. Data from postoperative follow up

Some data is present about the patient after release from the hospital. We have information at our disposal about complications after surgery. A phone call somewhere around 30 days after surgery is made with the patient. In case, this phone call is made it usually takes place 28.6 days after surgery (standard deviation 4.9). The WHO score 30 days after surgery is collected. Even though information after the primary stay can provide an important indication of the patient's health, we do not focus on these variables, since we cannot use these for any prediction modelling, due to the simultaneous timing of the outcome.

3.7. Data collected for complications

The outcome of the prediction models represents the presence of complications. For the complications, no exact recording of the time the complications occurred is available. Whether a complication occurred during the hospital stay or after the patient was released from the hospital is present in the available data. The complications are scored for severity with the Clavien Dindo score (CD score). The CD scale takes the values in {I, II, IIIa, IIIb, IVa, IVb, V}. The meaning of each value is listed in Table 3.1. The CD measures severity objectively by measuring the intensity of the necessary treatment for the complication [9]. In this thesis, we call a complication severe when the CD score is equal to or higher than IIIb. We study both severe and any complications as an outcome. Damaskos et al. [9] states that using the Clavien Dindo score (CD) for measuring surgery outcomes is an over-simplification. According to this study, the main flaw according is that it does not assess the status of the patient itself. A critical patient and a relatively healthy patient with a similar complication have the same CD score but have different prospects of recovery and should be treated differently. There is no risk assessed in the CD score, for it is solely based on treatment actions.

Number	Meaning	Example from dataset
I	No treatment	Urinary retention without medication
II	Medication	Urinary tract infection treated with medication
IIIa	Complications treated with surgery without general anaesthesia	Abscess cut open and drained surgically without general anaesthesia
IIIb	Complications treated with surgery with general anaesthesia	Anastomotic leak that is reoperated (without radiotherapy)
IVa	Single organ failure	Single organ failure due to severe pneumonia
IVb	Multiple organ failure	Severe bleeding and reoperation for anastomotic leak
V	Death	Death due to sepsis

Table 3.1: Meaning and example of different values of the Clavien Dindo scale measuring severity of complications.

Some complications were monitored but not present in the data set. These complications include pancreatitis (inflammation of the pancreas) and portal vein thrombosis (a blood clot in the vein between the intestine and liver). In Table 3.2, we see a list of complications that actually were present. We have separate variables for the complications during and after the primary stay (hospital stay of surgery of interest.)

Most complications have self-explanatory variable names in Table 3.2. However, some need further explanation. We briefly explain the following complications; paralytic ileus, anastomotic leak, abscess, pneumonia, haematoma, dehiscence, mechanical bowel obstruction, arrhythmia and sepsis. An ileus is the interruption of normal bowel movement. There are two types of ileus called paralytic ileus, and mechanical ileus (here called mechanical bowel obstruction). An anastomotic leak is a leakage of bowel contents from the anastomosis into the abdominal cavity. The anastomosis is the place where the bowels are reconnected to each other surgically. A drain could be placed to remove the bowel content from the cavity. An abscess is a collection of build-up pus. Pneumonia is an infection of the lungs. Hematoma is an internal bleeding. This is not necessarily the result of trauma. The complication named dehiscence, is a wound that opens up after it has been closed, finalising the surgery. Cardiac arrhythmia is a complication where the patient's heart gains an irregular heartbeat. Sepsis is a result of an infection. It refers to the body reacting extremely to the infection. Sepsis can lead to a septic shock, where the body reacts extremely severely, this can be fatal.

variable name
Surgical complications
Infectious complications
Renal hepatic pancreatic and gastrointestinal complications
Postoperative paralytic ileus
Anastomotic leak
Wound infection
Other complication
Intraperitoneal or retroperitoneal abscess
Pain
Respiratory complications
Urinary retention
Other surgical technical complication or injury
Postoperative excessive haemorrhage
Obstipation or diarrhoea
Nausea or vomiting
Pneumonia
Hematoma
Primary cause of death
Deep wound dehiscence
Other respiratory complication
Renal dysfunction
Cardiovascular complications
Gastrointestinal haemorrhage
Heart failure
Intraoperative excessive haemorrhage
Mechanical bowel obstruction
Other infectious complication
Other organ dysfunction
Pleural fluid
Psychiatric complications
Sepsis
Urinary tract infection
Cardiac arrhythmia
Septic shock
Urinary tract injury

Table 3.2: List of variables that indicate a type of complication, ordered by prevalence in unpreprocessed data.

Some variables directly relate to the complication like in case a reoperation took place. Reoperation is surgery after a surgery that has not had the desired outcome, for instance, if a complication develops that needs surgical treatment. In this dataset, there 28 resurgeries (5.0%) occurred during the primary stay. Other variables that directly relate to the complication are readmissions to the hospital, as well as variables about receiving intensive care and a variable about the cause of death. Some patients diseased as a result of a complication from the surgery. Information about if the patient survived up to 30 days after surgery is available. Also, the cause of death is listed in the data set, this is important since some mortalities are not related to the surgery.

4

Feature engineering

In this chapter, various pre-processing steps will be taken in order to prepare our data for the modelling in Chapter 6. We refer to these steps as feature engineering. When studying these variables we encounter different ways to include these in the feature-engineered data. This chapter results in a complete dataset suitable for modelling. Different variable selection methods will be applied in Chapter 6 on the feature-engineered data set. We aim to include many variables in the feature-engineered dataset to better study the difference in variable selection methods and make sure we do not miss important information. After this chapter, no changes will occur within the variables However, different subsets of the variables will be created using different variable selection methods. The feature-engineering requires many steps. We start the feature engineering process by arguing our outcome choice (Section 4.1). Then we remove the absolute uninformative data (Section 4.2). Afterwards, we make sure our observations are similar enough to be used in one model (Section 4.3). Then we remove some non-sensical variables (Section 4.4). We modify some variables for the information to fit into the corresponding scenario (presurgery, right after surgery and at the time of hospital release) (Section 4.5). We remove variables with little information (Section 4.6). Subsequently, we limit the values the variables can take (Section 4.7). Finally, we solve our problem of the missing values (Section 4.8).

We model the complications for three different scenarios. These scenarios are the presurgery prediction of complications, the prediction of complications right after surgery and the prediction of complications after a hospital stay. We should be careful not to fit outcomes on variables collected after the presence of the outcome. We decided to use multiple outcomes. We differentiate between any and severe complications. We call complications severe in case they need to be treated with a surgical procedure under general anaesthesia or in case they lead to organ failure or death. The severe complications are more clinically relevant. However, since they are quite rare we also include the outcome of any type of complication and see if this leads to predictions of higher quality. We further need to make modifications to some variables for them to fit a classification based on time. For example, the IV fluid intake on the day of the operation will be separated into fluids administered during the operation and fluids administered after the operation. In this way, we do not miss available information due to the timing of the variable.

All variables should have a clear meaning and so repetitions will be avoided. Variables will be separated into multiple features in case the information stored concerns information not mutually exclusive. Similarly, values are merged when having similar meanings. For example the values 'No, contraindicated' and 'No, other reason' can both be treated as 'No'. Another reason for preprocessing is to avoid overfitting from occurring. This can happen when variables are present in a small number of observations. To avoid this, we remove redundant information while aiming to keep the relevant parts.

In Appendix D, E and F information is listed about the data. We made an overview of the meaning of the variables in the feature-engineered dataset (Appendix D). Variables about adverse surgery outcomes are summed up in appendix E. In Appendix F the full prevalence and severity (CD score) for different complications is given.

4.1. Outcome choice

We have selected any and severe complications as the model outcome. In this section, we argue the outcome choice. We supply some more information and context about some specific complications. We study infectious complications and urinary complications since these are influenced by many variables. We briefly study ileus and anastomotic leak, these complications are typical for colon surgery. Ileus and anastomotic are relatively common compared to other complications (see Table 3.2). In the patient selection occur 15 anastomotic leaks, and 29 ileus (mechanical or paralytic ileus).

Two patients developed a urinary tract infection, one of these patients had a urinary drain for 1 night and the other patient for 6 nights. Both patients developing urinary tract infections were female. Urinary retention was present 15 times during the hospital stay. For one of these 15 patients (7.1%), urinary retention was also present after the hospital stay. Male gender is known to be a risk factor for urinary retention [15]. Of the patients with urinary retention 9 patients (60% of 15) were male and 6 patients (40% of 15) were female (binomial test p-value: 0.61, 9 from 15, with chance = 0.5). We do not see a clear connection between *More than one night with urinary drain* with *Urinary retention* (fisher test odds ratio 1.5, p-value 0.93). We see that the number of patients with urinary complications is too small to predict.

The ERAS protocol says that prescribing antibiotics and oral bowel preparation prevent the risk of wound infections [15]. However, we do not see this back in our data. After 22 surgeries the patient developed a wound infection. We see a p-value of 0.79 for *oral bowel preparation* with wound infections. Laparoscopic patients are less at risk for contracting wound infections. We see that the variable *Surgical approach group* equalling to open surgery (144 of 562 observations) tests significant with wound infections using the Fisher test (p-value 0.00070). In the data, for 43 occasions is an infectious complication present after surgery. The t-test on *BMI* and infectious complications leads to a p-value of 0.0038. The *BMI* of those with infectious complications is 28.7 versus 26.3 for those without.

Diet is said to be impacting the presence of infectious. However, in the data set *Preoperative nutritional treatment* equalling any type of diet prescribed (561 observations with 103 diets present (18%)) does not relate to infectious complications (fisher p-value: 0.21) [18]. On average patients with infectious complications are 3.1 years older (p-value: 0.084 t-test). The ASA class relates to the prevalence of infections. However, in our data an *ASA physical status class* being equal to 3 or 4 (224 times in 562 observations) does not relate to infectious complications (Fisher test results in a p-value of 0.62). Seemingly infections are easier to predict than urinary problems.

Anastomotic leak is the leakage of bowel contents into the patient's abdominal cavity after an anastomosis is created. This complication is present in 15 patients in the data. Obviously, only patients with an anastomosis placed can get an anastomotic leak. These are 446 people, so in 3.4 % of the surgeries where a leak could be present a leak was actually present. The average time of surgery for patients obtaining an anatomic leak is 36 minutes longer than for those who do not develop this complication. This was tested with the student t-test, the result was a p-value equalling 0.13. Previous surgery is said to be important for this complication, but in our data, it did not test significant (Fisher test, p-value = 0.44, with an odds ratio of 1.6). Of the anastomotic leak patients 60% are female (p-value Fisher: 0.60). In the literature, it is said that especially men who undergo surgery close to the rectum are in danger of anastomotic leak [18].

The interruption of normal bowel movement is called ileus. An important difference is the cause of ileus, this can be a mechanical obstruction or a paralysis. The bowel movement is stopped in both cases so food does not pass through the intestines. Postoperative mechanical bowel obstruction happened for 2 observations in the data. The paralytic ileus occurred 27 times. The operation time for patients who later on develop a paralytic ileus is 20 minutes longer on average (student t p-value of 0.18). Blood loss is on average 82 ml more for patients that get this complication (p-value 0.22 t.test). However, This is not surprising when realising that it occurs less with the laparoscopic approach, the laparoscopic approach tests significant with the occurrence of ileus (p-value 0.00013). Further, we notice from the data a relation with opioid usage (p-value 0.038). The usage of a nasogastric tube is related to the occurrence of ileus (p-value $2.5 \cdot 10^{-8}$).

Some variables do not relate to a specific complication. For example: different types of fluids relate to complications in general in different degrees. In Table 4.1 we see the p-values of the t-test for the ml of different fluids for the patients with and without severe complications. A recommended intraoperative

	Mean ml for surgeries with severe complication	Mean ml for surgeries without severe complication	p-value of t-test
Crystalloids	1760	1152	0.0041
Colloids	194	84	0.015
Bloodproducts	116	9	0.047

Table 4.1: Mean of different fluids administered intraoperatively, also given in the last row is the t-test for the variable separated by severe complications. Table made using the observations which are not missing these are 547 observations for crystalloids, 550 observations for colloids and 550 observations for blood products.

fluid intake is 1-4 ml/kg/h [15]. Assuming all patients are 1.7 meters tall, this recommendation would be 152-611 ml on average (based on 512 observations). In 7.5% of the surgeries (42 times), the patient received more than 2500 ml of liquids through IV during surgery. The Fisher test for fluid intake of more than 2500 ml intraoperatively and severe complications results in a p-value of $1.1 \cdot 10^{-6}$, (odds ratio 6.8). Also a fluid intake of more than 1500 ml tests significant with the Fisher test. This leads to a p-value of 0.0028 (odds ratio 2.4). In the data, most values are close to a multiple of 500 ml. This relates to the fact that 500 ml is a standard size for a bag of IV fluids.

In the literature, blood loss is related to complications [18]. In our data, patients in the observation selection who have no missing values assigned for blood loss and get a severe complication lose on average 345 ml of blood. For the patients who do develop any complications whilst not having a missing value assigned, the average blood loss is 74 ml. This means a difference of 270 ml (t-test p-value: 0.010). The average blood loss for all surgeries in the data is 101 ml (2 missing observations removed). We see whether blood products are given (Table 4.1) or whether blood is being lost, both contain relevant information.

We have seen in this section that some variables relate to specific complications and some variables relate to complications in general. We stress that not all relations that are shown in the available literature can be seen in our data. We will predict severe (Clavien Dindo more than IIIa) as well as any complications, since the number of the more specific complications is too small to be considered an outcome to model. The severe complications are relevant since these contain information about the treatment needed. However, these are not commonly present. Therefore we will also model any complications and see if this leads to a model that also finds the more interesting subset, namely severe complications.

4.2. Identification of data not suitable for statistical analyisis

The first change to the data set is removing the variables not suitable for statistical analysis. We remove empty columns. We further remove columns that contain only one value or a missing value. The original data set has 374 columns, after deleting the unusable columns, 248 columns remain. The data has been altered in the following steps:

1. Deleting the unusable data.
   (a) Deleting the empty columns. (18 columns)
   (b) Deleting columns that are constant variables in the data. (76 columns)
   (c) Deleting the columns indicating missing values. (32 columns)
2. Change the data to numbers.
   (a) Changing the columns given in hours and minutes to only to minutes. (1:30→ 90)
   (b) Delete text following the numbers. (1- Male → 1)
   (c) Change text to numbers. (Yes, renal failure → 1, see Appendix G)

The 126 features deleted are not useful for any type of statistical analysis. Most data removed in this section did not contain any information differing between individual patients. The effect of these variables can therefore not be studied with this data. Many columns contain information about complications after surgery. It is important to note that these complications are possible and monitored, but not present in any of the actual patients. Therefore we cannot study these complications within this data set. In Appendix D there is an overview of the complications not present.

We briefly state some informative variables that are constants in the data set. These variables could be important for the frequency of complications in this study, however we cannot study the effect of these variables with the data available to us. All observations are from an ERAS protocol therefore we cannot study the relevance of this protocol, but should look for varying factors that are of importance within this protocol, for example, whether the promoted mobilisation of the patient succeeded. An upper body forced air heating cover and a heated IV was used during all surgeries. Further, all patients received a peripheral opioid receptor antagonist (a type of narcotic). Since these variables are constants in the data we do not study these.

The change from *Length of operation* to *Length of operation Minutes* is made in Excel. The changes from text to numbers are shown in Appendix G. The cells that contained an "-" were split. The string after the "-" was then deleted. This way only a number remained.

We have removed the columns that are not useful for any statistical analysis. Then 248 variables remain, and from these 114 contain information about the complications or about the health check-up performed after approximately 30 days post-surgery. These variables are not considered explanatory variables since they contain information about the outcome itself or are collected after the chosen outcome being present. This leaves 134 variables that we should study to see if these could be used as explanatory variables.

Table 4.2 shows the number of variables before preprocessing and after preprocessing for different subsets of variables. The table further states the number of perioperative variables. These variables contain information that can not be classified as pre-, inter- or postoperative. Therefore we preprocess them in order to separate the information so that it can be classified in our subsets of features. For example, a variable stating at which moment a type of medication is given is preprocessed into two variables. One containing information about if this medication was given preoperationally and one containing information about whether the medication was given interoperationally. Table 4.3 shows how many patients contracted a complication in the not feature-engineered data set. This is the outcome we aim to predict.

set of variables	**no. of var. before feature engineering**	**no. of var. after feature engineering**
patient characteristics	13	11
preoperative variables	23	8
intraoperative variables	34	18
postoperative variables	59	14
perioperative variables/ can not be assigned subset without preprocessing	5	0

Table 4.2: Number of variables in the different subsets of variables based on time of occurrence for before feature engineering and after feature engineering.

complications occuring	**number of severe comp.**	**number of any comp.**
during primary stay	51	165
after primary stay	21	82

Table 4.3: Number of complications before preprocessing data. The number of observations with the outcome for the different models before preprocessing. In Table 4.12 we see the number of observations with the different outcomes after the preprocessing.

4.3. Ensuring similarity of the observations

We select patients to include in the preprocessed data to ensure the variable of interest is recorded sufficiently, and the observations are comparable enough to be used in one model. Later on, we select observations based on the number of missing values per observation.

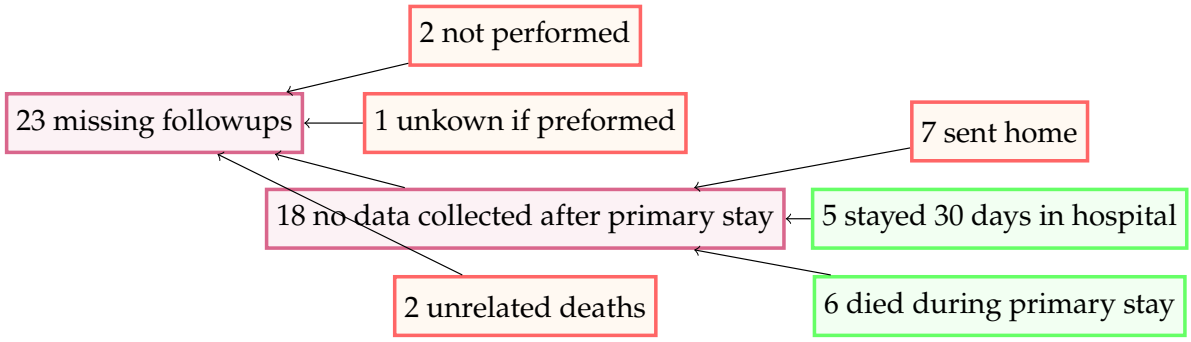

Figure 4.1: Information about the 23 values for which no 30-day follow-up information was available. In green are included observations and in red are excluded observations.

To ensure the quality of the outcome only patients on whom have a 30-day follow-up is performed are included, unless they died before or never left the hospital within 30 days after surgery. This is performed in order to filter the cases, where contact is lost with the patient or the patient died after primary stay due to causes unrelated to the surgery. This results in the exclusion of 12 surgeries from this research (see Figure 4.1).

We have to be alert with surgeries if the patient underwent surgery multiple times, since this results in repeated measurements. To be more specific; due to the repeated values, the correlation between the patients' characteristics is distorted. In the entire data set, 26 observations are recorded where the surgery is not the first operation for the patient in the data. For patients that underwent multiple surgeries in this data set only the earliest surgery is included in the preprocessed data. As a result the variable, *Number of operation*, is a constant in our selection and therefore removed. Now a surgery and a patient both refer to a unique observation.

Further, we exclude 1 observation where the patient underwent a surgery called exploratory laparoscopic. This surgery took 192 minutes. Exploratory laparoscopic surgeries are be performed in order to diagnose a patient. For our models, we exclude this type of surgery and add the variable called 'final diagnosis' as an intraoperative variable. Since the diagnosis is not always known or completely established presurgically.

Further, we removed 4 surgeries executed through the patients' stoma. This does not mean all stoma surgeries are excluded. However, the surgeries for which no cut was made in the belly to access the intestines (directly or laparoscopic) are excluded. These surgeries are less invasive and therefore have a quicker recovery. We do not have sufficient surgeries, performed via the stoma to include this information in a feature. The 4 patients who underwent surgery through the stoma experienced an average operation time of 32 minutes. This is almost a quarter of the mean operation time in the entire data set (118 minutes).

After preprocessing the observations in this section, 515 observations from the initial 562 remain. After the feature selection, we will study the observations for which many values are missing and subsequently we will consider which observations are to be further removed before imputting the missing values.

4.4. Exclusion of overparticular variables

In this section, we remove some variables that intrinsically do not relate to the surgery outcome as well as some information that further specifies other variables. First, we have to study if the more general variables could be used for modelling.

We remove the variable *Year ok* for being unrelated to surgery outcome. This variable states the year the surgery took place. We further remove the binary variable *Preoperative nutritional status assessment*, and only keep the treatment for malnutrition. This is stored in the variable called *Preoperative nutritional treatment*. Whether a screening took place should not be a risk factor, we rather include the outcome of a screening. The variable *Was the patient screened for anaemia preoperatively* is not removed, since this variable contains the result of the screening.

Some variables are a further specification of another variable. These specifications often contain information about one certain value in the other variable. This leads to very specific information, in many cases too specific to include. For example, the variable *Last HBA1c value mmolmol* contains blood values for diabetic patients. Even though this information is useful and the predictive power may increase when we change the variable about diabetes to include a restriction on the last HBA1c value, we do so not since we have only few diabetes patients (13.2% of the entire data set). First, we investigate whether the more general information needs preprocessing. In Section 4.6.2, some of the information removed here is included again. In Table 4.4 are listed the variables (left) that further specify values from another variable (right). There are 24 variables specifying other variables.

Variable	**Specified variable**
- Screenings intrument	*- Preoperative nutritional status assessment*
- Termination of smoking no weeks before surgery	*- Smoker*
- Standard units *- Termination of alcohol no weeks before surgery*	*- Alcohol usage*
- Last HBA1c value mmolmol	*- Diabetes smellitus*
- Days between admission an the last chemo therapy	*- Preoperative chemotherapy*
- Was Iron replacement treatment given	*- Was the patient screened for anaemia preoperatively*
- When was the first Anticoagulant prophylaxis done *- What was the duration of anticoagulant prophylaxis*	*- Thrombosis prophylaxis*
- Type of bowel anastomosis *- Anastomotic technique*	*- Bowel anastomosis*
- Ensure full reversal of Neuromuscular block	*- Deep neuromuscular blockade*
- Other main postoperative analgesia *- Successful block* *- Time to termination of epidural analgesia nights* *- Strong opioids given within 48 hrs postoperatively*	*- Postoperative epidural analgesia*
- T Primary Tumour *- N Regional Lymph Nodes* *- M Distant Metastasis*	*- Final diagnosis*
- Other free notes	*- Stomal Procedure*
- Level of insertion	*- Epidural or spinal aneasthesia*
- Lumbar supplementary analgesia	*- Main procedure name*
- Time to termination of urinary drainage nights	*- Urinary drainage post op*
- Nasogastric tube nights	*- Nasograstric tube inserted*

Table 4.4: Variables removed (left) for specifying specific values in other variables (right).

4.5. Classification of perioperative variables

Here we explain the preprocessing of the perioperative variables. This is performed in order that all variables can be divided into the groups namely; pre-, inter- or postoperational. The variables that require modification for this reason are: *Thrombosis prophylaxis, Total length of stay nights* and *Total IV volume of fluids day zero*. We will explain the preprocessing steps made on these variables one by one.

Prevention of thrombosis can be started at different instances. In the data, no patient had any thrombosis complications after surgery. Therefore we do not expect this variable to be of importance. However, there is a possibility that the time when the medication is given is due to a risk assessment of the doctor that translates to other complications as well. For almost all patients anticoagulants, a type of medication avoiding thrombosis is given. For 4 operations (0.78%) it is not given. Sometimes anticoagulant medication is given together with compression. Since compression usage is not common (29 times 5.6%), we limit our feature to only anticoagulant medication. The time the anticoagulants are given is stated in the variable *When was the first anticoagulant prophylaxis done*. We will add the variables: *Preoperative thrombosis prophylaxis* and *Intraoperative thrombosis prophylaxis* to the dataset. We decide not to include post-operative thrombosis prophylaxis. If we would, the variable selection could

select all these variables, which can lead to overfitting on the 4 observations that did not receive any anticoagulant medication. In the variable *What was the duration of anticoagulant prophylaxis* is stated whether the medication was given until the end of the hospital stay or also after the hospital stay. We do not include any variables after the primary stay and do not include this information.

The variable *Total length of stay nights* stores the total number of nights the patient spent in the hospital. This includes any readmissions. The variable *Length of stay nights in hospital after primary operation* stores the length of the primary stay. In the post-operational variable, *Length of stay nights in hospital after primary operation*, we impute the missing observations in with the values from *Total length of stay nights*. *Total length of stay nights* can be seen as an upper bound for *Length of stay nights in hospital after primary operation*. Afterwards, we remove the variable *Total length of stay nights*, since we do not keep variables concerning information from after the primary stay.

There are two variables about IV fluids on the day of operation. One states the fluids taken during the operation (*Total IV volume of fluids intraoperatively*) and the one other states the amount of IV fluids taken during the day of operation (*Total IV volume of fluids day zero*.). This results in the fluids administered inter-operational being counted twice. Separating this information will avoid unnecessary correlations within these variables. So for the number of IV fluids on the day of surgery, we keep an inter-operational and a post-operational variable. These are called *IV volume post operational on day of Surgery* and *Total IV volume intraoperatively*. We add the feature *IV volume post operational on day of Surgery* by subtracting *Total IV volume of fluids intraoperatively* from *Total IV volume of fluids day zero*.

4.6. Removal of low-information variables

In this section, we will remove variables that do not contain much information. This can occur because they are taking the same value many times or when many values are missing. We first look at the variables with similar information measured on different postoperative days. Since these variables have many missing data, we see if we can form one variable about the information contained in these variables. Then we take a look at which variables are close to a constant and make modifications if possible. Afterwards, we look at variables with many missing observations.

4.6.1. Daily measurements

Here we analyse the variables that measure the same information on different postoperative days, such as *Weight change on postoperative day 1, Weight change on postoperative day 2* and *Weight change on postoperative day 3*. We have 30 variables of this kind, measuring 8 different aspects. Many observations are missing for these variables. First, we denote how many observations are missing on all the days it should have been recorded. Then we can consider ignoring the missing values by combining similar variables from different days into one feature. If more than 52 observations (10% of the current selection) are missing on all days we do not include a feature based on those variables. In Table 4.5, we see the variables that are measured on multiple days and how many observations have not even one record on any day.

variable	**how many days it should have been recorded**	**# patients for which no observation is present on all days**
Weight change	3	124
Observed nausea retching and vomiting	4	4
On day of surgery	3	139
Opioid use	4	2
Oral fluids total volume taken	4	153
Oral nutritional supplements energy intake	4	87
Patient reported maximum nausea VAS	4	88
Patient reported maximum pain VAS	4	48

Table 4.5: The different daily measurements (left), the number of days it was recorded (middle), and the number of observations with variables missing on all days (right).

Besides the variables named *Observed nausea retching and vomiting*, *Opioid use* and *Patient reported maximum pain VAS*, no variables are present sufficiently to include them in a feature (less than 52 observations (5%) missing).

The variable nausea *Observed nausea retching and vomiting On day of surgery* is replaced with the variable, *PONV observed In 3 days*. This variable takes the value 'yes' when any nausea, retching or vomiting was observed, after surgery up to the third postoperative day.

Except for 19 observations (3.7%), all other patients use opioids on the day of surgery. This means if we make the binary variable containing any opioid used during the first 3 postoperative days, we get an almost constant variable. The number of missing observations about opioid use increases over time. On the day of surgery, 3 observations (0.58%) are missing. On postoperative day 1, 12 observations (2.3%) are missing. On postoperative day 2, 85 observations (17%) are missing. And on postoperative day 3, 226 observations (44%) are missing. Most commonly the patient starts taking opioids on the day of surgery, but not necessarily. In a few cases, the patients start taking opioids some days after. This event, the event where a patient used opioids and did not use opioids the previous day is seen in the data only a few times (26, 5.0%) during the first 3 postoperative days. Whereas for 190 patients (37%) it is recorded that they used no opioids but did the previous day, during the first 3 postoperative days. We decide to only include *Opioid use On postoperative day 1* since it is relatively complete and not as much as a constant as *Opioid use On day of surgery*. If the length of stay after surgery was 0 nights (for 3 surgeries (0.58%)) we impute *Opioid use On postoperative day 1* with 'no'.

For *Patient reported maximum pain VAS on day of surgery* a lot of data is missing. On the day of surgery 144 observations (28%) are missing. On the postoperative day 1, 149 observations (29%) are missing. On postoperative day 2, 273 observations (53%) are missing. And on postoperative day 3, 386 observations (75%) are missing. The means on the different days do not differ much (1.8, 2.4, 2.3, 2.0, on a VAS scale of 1 to 10). For simplicity, we assume that for which day an observation is missing does not matter. We decide to take the maximum of the known values. We call this variable *Max Pain VAS in 3 days*.

From the 30 variables recorded on multiple days, we keep 3 features. In this section, we removed 27 variables.

4.6.2. Near-constant variables

We remove a variable if it is almost a constant. We consider a variable almost a constant if it takes the same values at least 90% (464 times). We picked 90% over the more conventional 95%. Since there are many variables with rare values we fear that using these in one model leads to overfitting. In the case of missing observations, we decide to subtract the number of missing observations from this 464. If the variable is the same for at least 464 times minus the number of missing observations. We consider it close to a constant. In other words, we assume that missing values are actually the most common value, we delete variables with many similar values and missing values. Here 27 variables are close to a constant, these are listed in Table 4.6.

In Table 4.6 we see the similar variables, *Preoperative chemotherapy, Recent immunosuppressive treatment* and the *Any radiotherapy to operating field*, these variables state if any nonsurgical preoperative treatment is given. The variable about previous treatments called *Previous surgery to same abdominal region* is not considered to be almost constant. We make a binary variable called *Any nonsurgical Preoperative Treatment* taking the value 1 if any of the variables *Preoperative chemotherapy, Recent immunosuppressive treatment* or *Any radiotherapy to operating field* equals 'yes'. This variable is not almost constant with 58 times (11%) any nonsurgical preoperative treatment being recorded.

Operation converted and blood products given are rare. A total of 34 operations (6.6%) were converted and 16 patients (3.1%) received blood products during operation. We know that the variable stating the amount of any blood products given during operation, *IV volume of blood products intraoperatively*, can be considered a risk factor [4]. We delete the variable *IV volume of blood products intraoperatively* but still include some information from it, via the variable about blood loss. This variable is called *If bloodloss* and will be explained in Section 4.7.3.

Name of variable
Antibiotic prophylaxis before incision
Any radiotherapy to operating field
Preadmission patient education given
Preoperative chemotherapy
Preoperative long acting sedative medication
Recent immunosuppressive treatment
Time between admission and primary operation
Airway control
Fluid administration guidance
IV volume of blood products intraoperatively
Nasogastric tube used postoperatively
Nitrousoxide used
Operation converted
Skin preparation used
Urinary drainage postop
Recreational druguse
At all on day of surgery
Discharged to
Discharged with in 30 post op days
Nasogastric tube reinserted
PONV prophylaxis administered
Postoperative epidural analgesia
Stimulation of gut motility
Systemic opioids given
Use of 09 Na Cl
Deep neuromuscular blockade
Artificial nutrition

Table 4.6: List of almost constant variables, two variables are altered and the other are deleted.

Whether the conversion of surgery from closed to open forms a risk factor, is an interesting topic. A study by Masoomi et al. [29] states that converted surgeries have an increased risk of complications compared to closed surgery, but a risk as high as open surgery. A study by Gorgun et al. [14] states that a slight increase of surgical site infections for converted surgeries compared to planned open surgery occurs, but leads to a shorter length of stay. In a new feature called *If open or converted* the information about the converted surgeries is included. This variable is a binary that equals 1 in the case that the surgery ended as an open surgery. This feature is made with *Surgical approach group* and *Operation converted*. These variables are then excluded from the feature-engineered data set. The variable *Surgical approach group* shows the type of surgery: open, laparoscopic or robotic. So we do not keep the difference between the two types of closed surgery, robotic and laparoscopic surgeries because only few surgeries are robotic surgeries.

We have some variables that contain more information specified in another variable in the list of almost constant variables (Table 4.6). The information from *Days between admission and the last chemo therapy* cannot be used to make the variable *Preoperative chemo therapy* less close to a constant. The variable is close to constant since not many people received any chemotherapy. Similarly, for *Nasograstric tube reinserted* we have a variable specifying the rare values further. Luckily for *Urinary drainage post op* we can use the variable *Time to termination of urinary drainage nights*, to make this variable more informative. We know from the protocol that with the use of a urinary drain the risk of urinary tract infections increases over time [15]. It also hinders the mobilisation of the patient [15]. We decided not to remove *Urinary drainage nights* but transform it into *More than one night with urinary drain*.

We delete 27 almost constant variables. These variables are shown in Table 4.6. We constructed the variables *1 or more night with urinary drain* and *Any Nonsurigical preoperative Treatment*.

4.6.3. Handeling variables with missing observations

We removed 6 variables because they miss at least 52 observations (10%). These are: *Preoperative weight change*, *Previous PONV or motion sickness*, *Minimum core body temperature during operation*, *On day of surgery postoperatively* (kcal intake), *Time to passage of flatus nights* and *Time to passage of stool nights*.

An observational study on 464 patients [21], shows that having the first bowel movement on or after postoperative day 3 is related to complications. We tried to use *Time to passage of flatus nights* and *Time to passage of stool nights* together to form a feature about whether any bowel movement on or before the second night after surgery is present. If one variable is recorded and takes a value present before the third night, then it doesn't matter whether the other variable is missing. This way we can reduce the missing values in this new feature. However, 124 (24%) and 146 (30%) observations are missing for *Time to passage of stool nights* and *Time to passage of flatus nights* respectively. In 38 observations (7.4%) both the variables *Time to passage of flatus nights* and *Time to passage of stool nights* are missing, these would need imputation. For 24 observations (4.6%) one variable takes the a value of 3 or more nights, and the other variable is missing. These observations would also need imputing. So in total 62 observations (12%) would require imputing. Therefore, we decide not to include a feature about any stool or flatus before the second night after surgery.

4.7. Preprocessing according to patient journey stages

In this section, we study whether the variables currently selected need any preprocessing. We order this section by subsets of the variables, based on the moment the variable is collected. We start each section with a table of alterations executed (Tables 4.7 - 4.10) and, we argue the reason for the modification in the text.

4.7.1. Preprocessing patient characteristics

Variable name	Value	Meaning
ASA class	2	ASA class <=2
	3	ASA class >=3
alcohol usage	1	alcohol drinker or drinker that stopped less than 5 weeks before surgery
	0	not an alcohol drinker or a drinker that stopped for 5 or more weeks before surgery
smoker	1	smoker or smoker that stopped less than 9 weeks before surgery
	0	non smoker or a smoker that stopped for 9 weeks or more before surgery
BMI	0	smaller than 21.5
	1	in the range (21.5-24.9)
	2	preobese (25.0–29.9)
	3	obese (30 >)
Diabetes mellitus	1	yes
	0	no
if predisease	0	no comorbidity
	1	comorbidity

Table 4.7: List of altered values in the patient characteristics.

ASA class (American Society of Anesthesiologists class) is a scale used by anaesthesiologists to score the fitness of a patient. The number increases as the health of the patient decreases. In the data set it attains the values from 1 up to 4. The values 1 and 4 are not present often; each of these values we see 18 times. We change the variable ASA class to a binary variable, we take together the ASA classes 1 and 2 as one value, and the ASA classes 3 and 4 as the other value.

Variables *Smoker* and *Alcohol usage* have a value indicating if the patient stopped using these harmful substances, to improve the surgery outcome. We do not want to include these values separately since these values are assigned to a few observations (14 (2.7%) for smoking and 6 (1.2%) for alcohol usage).

We will assign a patient being a smoker when the patient stopped for less than 9 weeks before surgery. Similarly, we assign a patient as an alcohol consumer in case the patient stopped drinking alcohol less than 5 weeks before the surgery took place. This division is as the ERAS protocol recommends [15] even though ERAS admits that information on this topic is limited.

A low BMI or a high BMI are both risk factors. A study by Amri et al. [3] shows that wound healing complications differ over the different BMI classes defined by the WHO (World Health Organization). These WHO BMI classes are Underweight (<=18.5), Healthy (18.5–24.9), Preobese (25.0–29.9), Class I obesity (30–34.9), Class II obesity (35–39.9), Class III obesity (40>=). If we apply this division to our data set we see 9 patients (1.7%) classified as underweight, 192 (37%) are considered healthy, 182 (36%) is classified as preobese, 61 (12%) is classified as class I obese, 23 (4.5%) is classified as class II obese, and 6 (1.2%) is class III obese. The groups underweight and obese class II and III, are very small. We apply an alternative classification since we do not want any group to contain less than 52 observations (10%). We keep the class preobesity. We make one class for obesity (class I-III) and change the underweight and healthy classes to: '21.4 or below' and 'between 21.5 and 24.9'. In this way we can include a low BMI class with sufficient observations. So we have 59 patients (11%) with a BMI smaller than 21.5 and 142 patients (28%) have a BMI in the range (21.5-24.9) whereas 182 (35%) are classified as Preobese (25.0–29.9) another 90 (17%) can be classified as obese (30 >=). We deviate from conventional WHO BMI classifications since we have seen the relevance of lower BMI in different studies. We argue that the conventional BMI categories are frequently studied however the choice of classification is not well argued. Frankenfield et al. [13] shows that obesity defined by body fat percentage differs from obesity defined by BMI. It found that 30% of men and 46% of women with a BMI below 30 have body fat levels qualifying for obesity. If we separate the patients by gender and take the mean of the BMI, we see a mean of 27.1 for male patients and 25.8 for female patients. These numbers both classify as preobese. For the missing values for BMI, we will use the variable *Preoperative nutritional status assessment* to make imputations. The variable *Preoperative nutritional status assessment* states for 3 patients with missing BMI they are malnourished, and for an other 3 patients with missing BMI they are at risk of malnutrition. The average BMI of the patients assigned 'malnourished' is 24.9. We assign these patients in the range (21.5-24.9). The average BMI of the patients assigned 'risk of malnourished' is 25.0, and we impute the missing values for these patients with preobese (25-29.9). The 30 patients with missing BMI and that were assessed as 'no risk of malnutrition' we impute with preobese, since the average of the patient assest with 'no risk of malnutrition' is 26.9. There remain 3 patients with a missing BMI and a missing *Preoperative nutritional status assessment*.

Diabetes can be treated with medication or by diet. For 10 observations (1.9%), the patients control their diabetes with a diet. We do not distinguish how diabetes is treated in the variable *Diabetes mellitus*.

A variable is made by taking the comorbidities we have in the data set together as one variable. These comorbidities are from the variables: *Diabetes mellitus*, *Severe heart disease* and *Severe pulmonary disease*. We name this variable *If predisease*.

4.7.2. Preprocessing preoperative variables

Variable name	Value	Meaning
Preoperative nutritional treatment	0 1	no normal food yes, for example: immunonutrition, supplements, parenteral or tube feeding.
Preoperative oral carbohydrate treatment	1 0	yes no
Was the anaemia found	2 1	yes screened, and iron or non iron anaemia found not screened or not found

Table 4.8: List of altered values in the preoperative variables.

The variable *Preoperative nutritional treatment* shows the type of special nutrition the patient was taking (if any). The values the variable can take are, 'supplements' (58 times, 11%), 'immunonutrition' (25 times, 4.9%), 'parenteral nutrition' (2 times, 0.4%) or 'tube feeding' (2 times). Since most values are rare we change this variable to a binary variable, indicating whether any special nutritional treatment was prescribed.

For *Preoperative oral carbohydrate treatment*, we do not distinguish between 'No, any other reason' and 'No, contraindicated'. Contraindicated is the medical term for advising against a treatment due to some other reason than the treatment itself (for example, medications that interfere with each other, or allergies to substances in a treatment). The result of both values is the same. There are two missing values in *Preoperative oral carbohydrate treatment*. These observations have assigned 'Anterior resection of rectum' and 'Other' as the main operation. In the protocol of the MST ([10]), we see that conventionally during an anterior resection, oral bowel preparation is used only in case a stoma is placed. For this patient, we assume that the protocol was followed and impute with 'no' since no stoma was placed. For the other surgery, the main operation assigned is 'other', and we do not impute a value now.

The variable *Was the patient screened for anaemia preoperatively* distinguishes the type of anaemia. We do not differentiate for anaemia whether it refers to an iron or a non-iron blood deficiency. So we merge the values: 'non-iron def. anaemia found' and 'iron deficiency anaemia found'. For 9 observations (1.7%) we do not know if the patient was screened. For 29 patients (5.6%) we know they were not screened. Adding a value unknown would create a new small group. We rather assume the doctor had a reason not to screen these patients for anaemia and impute them with 'no anaemia'. The 9 missing observations we retain as missing for now. It is unclear if the screening took place and the value was not properly recorded or if these patients were not screened. We rename the variable *Was the patient screened for anaemia preoperatively* to *Was aneamia Found*.

4.7.3. Intraoperative variables

We change additional major procedures to a binary value. There are many different additional major medical procedures possible. Here 11 different additional major procedures were denoted. The most common additional procedure is likely to be a liver resection. This is recorded 19 times (3.6%). Since an 'other' additional procedure is recorded 23 times (4.5%) in the variable, we decide not to differentiate between any additional major procedures in the variable.

In the data different diagnoses occur. Some rare diagnoses are; 'other primary malignancies', 'surgery specifically for metastasis or recurrence of any malignant disease', 'Other benign disease or disorder', 'unknown', 'benign tumours including polyps', 'inflammatory bowel disease', 'complicated diverticular disease' and 'uncomplicated diverticular disease'. These diagnoses are present in only 7 (1.4% of 515 obs.), 3 (0.58%), 36 (7.0%), 3 (0.58%), 12 (2.3%), 12(2.3%), 7 (1.4%) and 25 (4.9%) patients respectively. We decide to make the categories: 'any cancer diagnosis', 'functional disorder' and 'other' for the diagnoses. Making a separative value for diverticular disease would lead to only 32 observations (6.2%). We include diverticular disease in the category: other. Since the group cancer is large (345 diagnoses (67%)) and we do have more information concerning these patients (variables: *T primary Tumour, N regional lymph Nodes* and *M distant metastasis*), we add a postoperative variable stating more information about the cancer in Section: 4.7.4.

Variable name	Value	Meaning
Additional major procedures	0 1	None yes
Final diagnosis	0 1 2	functional disorder cancer diagnosis: primary adenocarcinoma, other primary malignancy, surgery specifically for metastasis or recurrence of any malignant disease other: Other benign disease or disorder, unknown, benign tumour including polyps, inflammatory bowel disease or any type of diverticular disease, complicated or Uncomplicated.
If spinal anaesthesia	5 0	Spinal (intrathecal) no spinal anaesthesia
Intraoperaive blood loss	1 0	yes more than 100ml of blood loss. no, less or equal than 100ml of blood loss.
If colloids	1 0	yes no
Main procdeure name	7 3 8 6 9	stoma procedures ileocaecal/ right hemicolectomy, anterior resection of the rectum, sigmoid resection uncommon procedure
Stomal procedure	0 1	none any stomal procedure
If lidocaine	1 0	yes Lidocaine administered no lidocaine administered

Table 4.9: List of altered values in the intraoperative variables.

For *Epidural or spinal anaesthesia* we do not distinguish between 'No, any other reason' and 'No, contraindicated'. This variable shows that 40 patients (7.8%) did not receive spinal anaesthesia or an epidural. Of the patients 38 (7.4%) received an epidural. We limit this variable further so that it only keeps the information about spinal anaesthesia and renames this variable to *If spinal anaethesia*. *IV volume of blood products intraoperatively* and *IV volume of colloids intraoperatively* equal zero often (488 times (95%) for blood products and 416 times (81%) for colloids). We keep the total amount of IV fluids as a continuous variable and specify what kind of fluids this could have been with binary variables. We make the variable *If colloids are given* stating if *IV volume of colloids intraoperatively* equals more than zero. We delete the variable *IV volume of crystalloids intraoperatively* since this variable is very close to *Total IV volume of fluids intraoperatively*.

For blood products given during operation and intraoperative blood loss, we create a separate feature. Since 16 patients (3.1%) received blood products we do not include this information in one variable. Of those 16 patients (3.1%) only 1 did not lose any blood during surgery. The other 15 patients (2.9%) all lost at least 100 ml of blood. We make a binary value indication if more than 100 ml of blood was lost or if any blood products are given during the operation. This variable is called *If blood*. In the study by Rasilainen et al. [43] blood loss over 100 ml is linked to Ileus and anastomotic dehiscence.

The main surgical procedure can be one of the 11 different types. We take some procedures together in the following way. We make 5 groups, These are 'stoma procedure', 'Ileocaecal/ right hemicolectomy', 'anterior resection of the rectum', 'sigmoid resection' and 'uncommon procedure'. The group 'uncommon procedure' contains the values: 'left hemicolectomy', 'total/subtotal colectomy', 'small bowel resection', 'proctocolectomy with the anus', 'abdominoperineal resection' and 'other large small bowel surgery'. For all the different surgeries included in the variable *Main procedure* a stoma procedure could be performed.

The variable called *Stoma procedure* indicates which kind of stoma procedure is performed. The variable can be equal to: 'placement of colostomy', 'removal of colostomy', 'placement of ileostomy', 'removal of ileostomy', 'moving of stoma' or 'other'. Two logical ways to simplify this variable are one to differentiate between colostomy or ileostomy procedure, or two to differentiate between placement and removal of a stoma. However, in both cases, we have 9 observations equaling 'moving a stoma' or 'other'. The variable is preprocessed into a binary variable indicating whether any stoma procedure was performed.

The variable *Nerve blocks or local anaesthesia* takes the value 'no' 8 times (1.6%), the value 'Local infiltration' 430 times (83%), and 'IV lidocaine' 74 times (14%). The group that takes the value 'no' is not desirable to include in the variable. The variable seems to have little direct effect on the complications. (p-value equals 0.75 from a chi-squared test with any complication, and p-value = 0.25 with complications after the primary stay, limiting to where the variable is not missing). We decided not to remove the 8 observations (1.6%) with the value 'no' for a variable that does not seem to be of high influence on the outcome. In the protocol is stated that it is proven that lidocaine reduces pain after surgery better than in comparison to a placebo [15]. However limited studies explore possible adverse effects or compare lidocaine with other analgesia (pain relief). We will include a variable called *If lidocaine*. Testing directly on the outcome we see no direct relation with the outcome. We do not expect this variable to be important directly to the outcome. However protocol states it influences the use of other painkillers.

4.7.4. Preprocessing postoperative variables

Variable name	Value	Meaning
Duration of IV fluid infusion nights	0	on day of surgery
	1	on first postoperative day
	2	2 or more nights
Time to tolerating solid food nights	0	on day of surgery
	1	on first post operative day or after
Time to pain control with oral analgesics nights	0	on day of surgery or first day after surgery
	2	on second post-operative day or after
CD at least 2	1	yes complications present before patient was released from the hospital
	0	no complications during hospital stay.
Mobilised on POD 1 or 2	1	Recovery of ADL ability on day of surgery or on POD1.
	2	Recovery of ADL ability on POD2
	3	Recovery of ADL ability on or after POD 3
High cancer Stage	1	yes distant metastasis is present
	0	no known distant metastasis or no cancer

Table 4.10: List of altered values in the postoperative variables.

Duration of IV fluid infusion nights is a skewed variable. This positive variable has a mean of 1.51 and an empirical variance of 20.5. For 265 patients (51%) the IV was removed on the day of operation. For 126 (24%) it was the first operational day, and for 117 (23%) it was removed later than this day. A total of 18 patients (3.5%) needed IV fluids for more than one week. We decide to let the variable take the values '0', '1' and '2 or more'.

Furthermore 379 (74%) patients could tolerate solid food on the day of surgery, for 103 patients (20%) this was not the case. A number of 16 patients (3.1%) needed 2 or more nights inorder to tolerate solid food. We change the variable *Time to tolerating solid food nights* into a binary variable stating if the patient could tolerate solid food on the day of surgery.

The variable *Time to pain control with oral analgesics nights* has some high values with a maximum of 48 nights. Yet 53 patients (10%) took only oral pain medication on the of surgery, 209 (41%) on the first postoperative day, 125 (24%) on the second postoperative day and 59 patients (11%) sometime after the second postoperative day. A total of 8 patients (1.6%) needed at least a week before the experienced pain was controlled with the use of oral analgesics only. We decide to take all values over 2 as one value. So now this variable is transformed into a binary variable. The variable *Time to pain control with*

oral analgesics nights is bounded by *Length of stay nights in hospital after primary operation*. This is useful information when imputing the missing values. On average the patient takes only oral analgesics for pain control until 3 nights before being released from the hospital. However, patients who are released from the hospital before POD 4, control their pain with oral analgesics on average 1 day before being released. We impute with 'on day of surgery or POD1' if the length of stay was 0, 1, or 2 nights. If the length of stay after surgery was more than 3 nights we impute with 'on or after POD 2' for time to pain control with oral analgesics nights.

We also add a feature about any complications arisen during the primary stay as a postoperative variable. A study by Miyamoto et al. [34] shows that complications with a Clavin Dindo (CD) scale equaling 3 or more lead to other adverse patient outcomes and increase in mortality. However since 23 patients developed a postoperative complication (4.5%) with a CD of 3 or higher during their primary stay, we added a variable stating if any complications during stay with Clavin Dindo scale 2 or more are present. There are 64 observations (12%) with postoperative complications during primary stay with a CD of 2 or higher. This variable is called *CD at least 2*. It takes the values: 'Yes complication' or 'No complication'. We emphasise that the patients who died during their the hospital are in this group (5 observations (0.97%)). The patients who died during their hospital stay do not have any possibility of developing a complication after their primary stay. However, this can not be determent from the variable.

A paper by Rasilainen et al. [43] stresses the importance of early mobilisation for its effect on decreasing the prevelance of complications. We change the variable *Time to recovery of ADL ability nights*, to a feature called *Mobilised on POD 1 or 2* (Post Operative Day). This positive variable has a mean of 2.9 and an empirical variance of 28.9. The variable *mobilised on POD 1 or 2* has three possible values one for on the day of surgery or POD 1 (201, (39%)), one for on POD 2 (137 (27%)) and one for any day after POD 2 (149 (29%)), while 28 observations (5.4%) are missing. For 1 patient the ADL abilities were returned after being released from the hospital, therefore we decided not to use the length of stay to impute the missing values for this variable. Possibily the missing values in *Time to recovery of ADL ability nights* indicate that the patient was sent home before mobilisation. For the sake of completeness, we decide to impute them. *Mobilised on POD 1 or 2* differs greatly within the ASA classes. Luckily no ASA classes are missing, with 43% of patients with ASA class 1 or 2 mobilised on POD 1, 30% on POD 2, and 26% after POD 3. For the patients with ASA class 3 or 4, 36% mobilised on the first POD, 27% on the second POD and 36% on or after the third POD. We impute the missing data with 'mobilised on POD 2' in case the ASA class equals 3 or 4. This is assigning them between the larger groups. If the ASA class is 1 or 2 we impute with 'mobilised on POD 1', since this is the most common value for these patients.

We want to further differentiate within the 345 (67%) cancer patients. In the literature often a classification of cancer that is used is named "cancer stage". There are 5 stages determined from the T, N and M score (in our data set called: *T primary tumour, N regional lymph nodes* and *M distant metastasis*). The first 3 stages; stage 0, stage 1 and stage 2, are based on the size of the tumour (T score). For stage 3 nearby lymph nodes contain cancerous growth (N score > 0). During the last stage, stage 4, distant metastasis is present (M score > 0). Different studies include this score in different ways. A study by Bakker et al. [6] includes all 5 stages. A study by Sluis et al. [46] includes it as a variable, by differentiting in: no cancer or stage 1-2 cancer and cancer stage 3-4. A study by Ortiz-López et al. [37] and Zhang et al. [52] includes the last 4 stages separately, probably because stage 0 is rarely diagnosed. In our data 2 stage 0 diagnoses are present, as well as one possible stage 0 diagnosis, which has an incomplete M-score. A study by Warps et al. [50] does not differentiate between stages 2 and 3. This study does not include any information based on the N-score. A study by Sparreboom et al. [48] does not include the cancer stage but uses the variables T- and M-score as two separate variables directly. The relevance of the different cancer scores is not similar in all studies. It is considered important, in the studies by Sluis et al. [46], Ortiz-López et al. [37] and Warps et al. [50]. In a study by Sparreboom et al. [48] the T and M-score are not considered useful for predicting the complication ileus. We include one variable for indicating a high-risk group for the cancer patients. This variable includes distant metastases, However since distance metastasis is often missing and rare in the data, we also include the patients with two close by lymph nodes containing cancerous growth in this variable. This leads to a group of 56 patients (11%). In the data, 63 cancer patients (12%) have missing metastases and 23 cancer patients (4.5%) have distant metastases diagnosis. We see 39 patients (7.6%) have two closeby lymph nodes containing cancerous growth and 6 patients (1.2%) have an N score of 2 as well as distant metastases present.

4.8. Handeling Missing values

In the previous section, we ensured that the variables can be used for modelling. However, we still have to deal with missing values in the data set. Therefore we will remove some observations with many missing variables, and impute some missing values.

The applied observation selection is limited further. Right now it contains 515 observations. The observations containing many missing values are removed. Table 4.11 shows the number of observations containing multiple missing values. We decide to remove observations when 3 or more variables are missing, we see in Table 4.11 that this means we remove 19 observations. We keep 496 observations (88% of the original dataset).

number of observations	having at least this many values missing
105	1
35	2
19	3
13	4
10	5
8	6
5	7

Table 4.11: Number of observations with multiple values missing. We see that we have 105 observations with at least 1 data point missing, of these 105 observations 35 have at least 2 data points missing.

The number of complications in the new patient selection is shown in Table 4.12.

complications occuring	number of severe comp.	number of any comp.
any time	50	159
after primary stay	21	77

Table 4.12: Number of complications after preprocessing data. The number of observations with the outcome for the different models after preprocessing. In Table 4.3 we see the number of the observations with the different outcomes before the preprocessing.

In Table 4.13, we see the variables with missing values. We see that often 1 to 6 values are missing. For *Time to tolerating solid food nights* and *Max pain VAS in 3 days* we have to deal with many missing values. The missing values we impute with a random value from the marginal distribution (without missing values). For example, for the 3 patients for which we do not know whether they smoke, we assign them 'smoker' at random. This is with the probability equal to the ratio of smokers in the data ignoring the missing values.

Note that in previous sections possibly, more missing values were stated. Some are removed in the previous section when removing the observations with more than 3 values missing.

Variable	Missing
Smoker	3
Alcohol usage	5
BMI	3
Preoperative WHO performance score	1
Previous surgery to same abdominal region	2
Was the patient screened for anaemia preoperatively	6
Preoperative nutritional treatment	2
If Blood	2
Depth of anaesthesia monitored	2
Epidural or spinal anaesthesia	1
Opioid use On postoperative day 1	5
Time to tolerating solid food nights	21
more than one night with urinary drain	6
Time to pain control with oral analgesics nights	2
Length of stay nights in hospital after primary operation	5
Duration of IV fluid infusion nights	1
Core body temperature at end of operation	3
Max Pain VAS in 3 days	39

Table 4.13: Number of random imputations made for different variables.

4.9. Summary of feature engineering

In Table 4.14, we see how the number of variables changes in this chapter. Many changes were made during the preprocessing. Most changes were to limit the values one variable can take. Much information was deemed being not useful for modelling. We have 51 preprocessed variables, this should be enough to study the differences between variable selection methods.

For the data preprocessing a deep understanding of the variables is needed. This is not always feasible and we expect some lack of medical interpretations. We stress the need for an interdisciplinary approach. Sadly, the approach of imputation is not consistent over the variables. Some imputations were based on a relation to another variable which was sometimes found from a statistical/ personal approach and are not medically argued. Some variables that could attain many values are preprocessed and there with lose information. The reason for missing data is not clearly stated, which can lead to wrong interpretations, which in turn can lead to wrong imputations.

Preprocessing step	# of vars.	# vars. eliminated	# vars. added
before feature engineering	134		
4.3: Ensuring similarity of the observations	133	-1	
4.4: Exclusion of overparticular variables	107	-26	
4.5: Classification of perioperative variables	107	-3	+3
4.6.1: Daily measurements	80	-29	+2
4.6.2: Near-constant variables	55	-27	+2
4.6.3: Handeling variables with missing observations	49	-6	
4.7.1: Preprocessing patient characteristics	49		+1
4.7.2: Preprocessing preoperative variables	50		
4.7.3: Intraoperative variables	50	-1	
4.7.4: Preprocessing postoperative variables	49		+ 2
Final # features	51		

Table 4.14: Change in the number of variables over the sections in this chapter (Chapter 4).

5

Explanatory data analysis (EDA)

In this Chapter we will conduct an exploratory data analysis, by providing several descriptive statistics regarding the variables selected for analysis. Most information about the variables will be tabulated. Some information from literature will be compared in the text underneath these tables.

We explain what the Tables 5.1 - 5.10 tell us; these tables follow a classification of variables for the tables to fit one page of this thesis. In Figure 5.1 we see the subsets of our variables. Of the ten columns in the tables, the first column shows the variable in italics and a short description of the variable thereafter, below the variable's name are the values listed this variable can attain. Subsequently the second column shows the numbers of the patients with the value assigned. The third column contains the number of patients with this value as well as any complications. The fourth column shows the number of patients with this value as wel as a severe complication. The fifth column shows the odds ratio with the complications in this group, the sixth column contains the odds ratio for the severe complication for patients with this value. If the odds ratio equals a number lower than 1 for a certain value of a variable, it means that one is less likely to have the outcome when having that value. Similarly, if the odds ratio is larger than 1, one is more likely to develop a complication if the corresponding value is present. The seventh column contains two numbers, above the p-value when testing for this variable with any complication (in blue), and below testing with severe complications (in red). The ninth column shows the mean length of stay in this group, and the last column contains the p-value of the t-test for the length of hostpital stay.

Note that the tables for the postoperative variables are performed with the outcomes that occur after the primary hospital stay, and the tables for the preoperative variables and intraoperative variables show only occurance of complications after surgery. We will also use the preoperative variables and intraoperative variables to model the complications that occur after the primary stay. However, we do not study this by means of creating a new table.

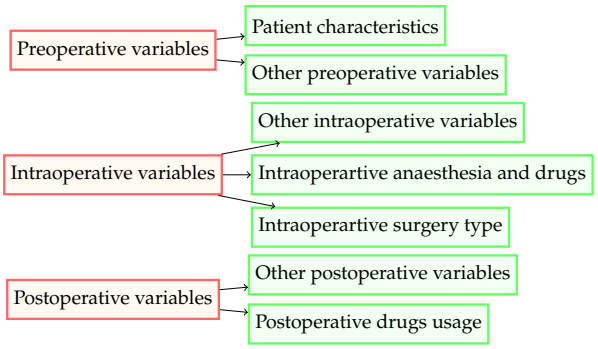

Figure 5.1: Subsets of variables used for making the Tables 5.1 -5.10.

5.1. Table general characteristics

***variable*, explanation of variable** **Value**	**#**	**#**	**#**	**odds ratio**	**odds ratio**	**p-value**	**mean LOS**	**p-value LOS**
Gender								
male	246	82	25	1.12	1.02	0.56	4.74	0.28
female	250	77	25	0.89	0.98	1	5.44	0.28
Smoker								
no, smoker (or stopped for at least 9 weeks)	388	119	38	0.75	0.87	0.24	4.72	0.13
yes, smoke (or stopped less than 9 weeks)	108	40	12	1.33	1.15	0.72	6.44	0.13
Alcohol usage								
yes, alcohol drinker	265	77	27	0.74	1.03	0.15	5.18	0.77
no, no alcohol drinker	231	82	23	1.34	0.97	1	4.99	0.77
BMI, kg/m^2								
BMI in (25-29.9) preobese	214	71	19	1.09	0.79	0.79	4.48	0.09
BMI in (21.5-24.9)	137	40	10	0.83	0.63	0.09	5.07	0.95
BMI >= 30 obese	86	30	10	1.17	1.22		4.72	0.5
BMI $\leq$ 21.4	59	18	11	0.92	2.34		7.92	0.11
If predisease, any comorbidity recorded								
no	332	99	30	0.74	0.72	0.15	5	0.7
yes	164	60	20	1.36	1.4	0.27	5.27	0.7
Diabetes mellitus disease affecting blood sugar								
no, diabetes	429	140	45	1.22	1.45	0.57	5.32	0
yes, with or with out medication	67	19	5	0.82	0.69	0.66	3.64	0
Severe heart disease								
no	440	139	43	0.83	0.76	0.55	4.97	0.46
yes	56	20	7	1.2	1.32	0.48	6.07	0.46
Severe pulmonary disease, disease affecting lung function								
no	428	128	39	0.51	0.52	0.01	4.8	0.12
yes	68	31	11	1.96	1.92	0.08	6.94	0.12
ASA physical status class, health score made by anaesthesiologist								
1 or 2	292	77	27	0.53	0.8	0	4.79	0.25
3 or 4	204	82	23	1.88	1.25	0.54	5.52	0.25
Preoperative WHO performance score, scores mobility								
0	434	120	38	0.23	0.4	0	4.51	0.02
1	62	39	12	4.44	2.5	0.02	9.19	0.02

Table 5.1: Information of the general characteristics. Blue cells contain information about patients with any complications and the red cells are for patients with severe complications. Columns named # indicate the number of surgeries with the variable p-values in the first column named p-value is from the Fisher test or the chi-squared test depending on the data type. The second column named p-value contains Kendall's tau test for the length of stay.

In Table 5.1, we see information from the patient characteristics. 33% of the surgeries (164 times) is performed on patients with a recorded comorbidity. As stated in the chapter 1.4, old age is not always considered a risk factor. However, it increases the chance of having comorbidities and bad mobility which are known risk factors. Older patients are more likely to have comorbidities and health-related issues. In our data, we see this back if we use the Fisher test on the variable *Preoperative WHO performance score* and severe complications, we see a p-value of 0.021 (Table 5.1). If we test *Age* with *Preoperative WHO performance score* with the student t-test we see a p-value of $4.6 \cdot 10^{-6}$. So our data confirms that these variables are related. However, if we use the t-test for *Age* and severe complications we see a p-value of 1. This is in line with the belief that age is less important for predicting complications than mobility or comorbidities.

In Table 5.2, We see the relation of *Age* with different comorbidities. We see that *Severe pulmonary disease* is not related to age in our data.

	Data separated by Severe complications	**Predisease**	**Lung disease**	**Hearth disease**	**Diabetes**
Mean age of those with the variable present	65.7	68.7	64.8	72.0	71.3
Mean age of those without the variable present	65.7	64.2	65.9	64.9	64.9
p-value (Student t-test)	0.99	$3.3 \cdot 10^{-4}$	0.56	$2.6 \cdot 10^{-5}$	$6.5 \cdot 10^{-6}$

Table 5.2: Mean age separated by different variables (above) and the t.test for age with this variable (below).

We see in Table 5.1, the test results of *Severe pulmonary disease* and *Severe heart disease* with complications. It seems that *Severe heart disease* does not relate to complications. A possible explanation can be that cardiovascular complications are not common in the data. In the data set, 3 surgeries with cardiovascular complications are present.

We test to see if *Severe pulmonary disease* is related to pneumonia, pleural fluid and other respiratory complications. We see the p-values for the Fisher test to be respectively: 0.00032, 0.14 and 0.092. We further test for respiratory complications with *severe pulmonary disease* and see a p-value of 0.000062. These complications are not frequent with 9, 1, 4 and 13 times in the processed observations respectively.

We see in Table 5.1, that *Alcohol usage* does not seem to be a reliable predictor for complications. Alcohol consumption of more than 2 units a day increases the risk of infections. The consequences for patients drinking less than 2 units a day are unclear [15]. In the data set, 35 patients of the 201 patients for which we know how high their alcoholic consumption is per week, stated they drank 14 units or more per week (17.4%). Of these drinkers, 8 developed complications of which 1 was a severe complication.

We will compare some variables of the preprocessed data with values of the entire Dutch population in Table 5.3. We tabulate the percentage for some patient characteristics for our patients and underneath this percentage for the entire population of the Netherlands. We see that except for *Gender*, we cannot think of our patient population as a random sample of the Dutch population (in Table 5.3). Interestingly we see that alcohol usage is lower for the patients than for the entire Dutch population, maybe this is related to the age of the patients or the patients feel ashamed to admit they drink alcohol in a hospital setting.

	Male Gender	**Alcohol drinker or stopped because of surgery**	**Smoker or stopped because of surgery**	**Diabetic**
In data set	49.4%	53.6%	22.0%	13.5%
In Netherland	49.7% [31]	80% [20]	19%[20]	6.6% [12]
p-value of binomial test	0.96	$< 2.2 \cdot 10^{-16}$	0.097	$3.4 \cdot 10^{-8}$

Table 5.3: Table containing percentages for the preprocessed patients and the entire Dutch population and p-values of binomial tests. These binomial tests are performed to see if the percentages in our data can be considered a random result of sampling from the Dutch population.

5.2. Analysis of preoperative variables

***variable*, explanation of variable Value**	**#**	**#**	**#**	**odds ratio**	**odds ratio**	**p-value**	**mean LOS**	**p-value LOS**
Previous surgery to same abdominal region								
yes	246	86	29	1.3	1.46	0.18	6.19	0
no	250	73	21	0.77	0.69	0.23	4.02	0
Any nonsurgical Preoperative Treatment								
no	448	148	45	1.66	0.96	0.19	5.13	0.6
yes	48	11	5	0.6	1.04	1	4.73	0.6
Preadmission stoma counseling								
yes	172	58	24	1.12	1.86	0.61	5.88	0.09
no	324	101	26	0.89	0.54	0.04	4.67	0.09
Was aneamia Found								
no anaemia found	427	130	42	0.6	0.83	0.07	4.88	0.1
anaemia found	69	29	8	1.66	1.2	0.67	6.43	0.1
Preoperative nutritional treatment								
no	409	128	40	0.82	0.83	0.45	4.93	0.2
yes	87	31	10	1.22	1.2	0.69	5.87	0.2
Preoperative oral carbohydrate treatment, nutritional substance given before surgery								
yes	430	140	46	1.19	1.86	0.57	5.32	0
no	66	19	4	0.84	0.54	0.38	3.64	0
Oral bowel preparation laxatives								
no	381	117	36	0.77	0.75	0.26	5.01	0.58
yes	115	42	14	1.3	1.33	0.38	5.37	0.58
Preoperative thrombosis prophylaxis								
no	382	115	36	0.69	0.74	0.11	4.99	0.53
yes	114	44	14	1.46	1.35	0.38	5.44	0.53

Table 5.4: Information of the preoperative variables. Blue cells contain information about patients with any complications and the red cells are for patients with severe complications. Columns named # indicate the number of surgeries with the variable p-values in the first column named p-value is from the Fisher test or the chi-squared test depending on the data type. The second column named p-value contains Kendall's tau test for the length of stay.

In Table 5.4, we see that *Preadmission stoma counseling* relates to severe complications. This is surprising, as we have not found any studies stressing the importance of counselling. We see that there are more patients with *Preadmission stoma counseling* present than those with a *Stoma procedure* performed (See Table 5.8). Perhaps counselling is only offered for patients with risky surgery. Or maybe the risk relates to the surgeries for which the procedure to be done is uncertain beforehand. A goal of counselling is also to avoid stress, a general health risk factor. Sadly we do not have any indication of stress experienced in the data.

In Table 5.4, *Preoperative oral carbohydrate treatment* does not seem directly related to complications. Carbohydrate loading (*Preoperative oral carbohydrate treatment*) is a nutritional strategy to avoid the patient entering the surgery in a fasted state. Entering surgery in a fasted state has adverse metabolic consequences. However, from a surgical approach bowel contents should not be present, for this could increase the risk of infection. Carbohydrate loading is used to improve the outcome of surgery, by

supplying the body with nutrients [15]. However, this strategy is not necessary for every type of surgery. Carbohydrate prevents PONV, in our data we test *Preoperative oral carbohydrate treatment* and *PONV observed in 3 days* with the Fisher test. This results in the p-value equaling 0.15.

Nutritional status is important for a good recovery, especially for patients with cancer [15]. We do not see a strong relation between a *Prescribed nutritional treatment* and the *Final diagonsis* equalling cancer (Fisher test p-value 0.13). Neither a strong relation between *High cancer stage* and *Prescribed nutritional treatment* was found, we see a p-value of 0.34. We do see a relation between *BMI* and the *Prescribed nutritional treatment* with the chi-squared test on Table 5.5, this resulted in a p-value of 0.0039.

	***BMI* '<= 21.4'**	***BMI* '[21.5-24.9]'**	***BMI* in '[25-29.9]'**	***BMI* '>= 30'**
No *Preoperative nutritional treatment* present	40	110	182	77
Preoperative nutritional treatment present	19	28	30	10

Table 5.5: contigency table of *BMI* and *prescribed nutritional treatment*. The chi-squared test on this contingency table returns a p-value of 0.0039.

Anaemia is the lack of functioning red blood products or haemoglobin in the body. This is commonly treated with iron supplements. Oral iron is less expensive than IV-administered iron. However, IV iron is more effective [15]. Anaemia and complications are known to be related [18][33].

According to the website of the WHO health organisation, 12.8% of the women in the Netherlands aged 15-49 have anaemia [41]. In our data set, we have 32 women aged between 15-49. Of these patients, 1 tested positive for anaemia. So 1 of the 32 tested patients were tested positive for anaemia (3.1%). Of the 27 men aged between 15-49, 14.8% tested positive for anaemia (4 positive). We expected fewer male patients to have anaemia than female patients. However, we do not conclude anything due to the small number of observations. We see no link between *Was aneamia Found* and *Severe heart disease* (p-value 0.16).

5.3. Analysis of intraoperative variables

***variable*, explanation of variable Value**	**#**	**#**	**#**	**odds ratio**	**odds ratio**	**p-value**	**mean LOS**	**p-value LOS**
Final diagnosis								
functional disorder	70	21	6	0.89	0.81	0.49	5.21	0.84
other	94	35	13	1.33	1.58	0.4	5.67	0.37
cancer	332	103	31	0.87	0.79		4.9	0.37
If colloids any colloids given								
none	409	126	34	0.73	0.4	0.21	4.44	0
yes	87	33	16	1.37	2.49	0.01	8.16	0
If blood								
no or less than 100ml	348	93	22	0.45	0.29	0	3.93	0
yes more than 100ml	148	66	28	2.21	3.46	0	7.82	0
Resection site drainage If a drain was placed								
no	440	132	38	0.46	0.35	0.01	4.84	0.02
yes	56	27	12	2.17	2.89	0.01	7.09	0.02

Table 5.6: Information on the intraoperative variables. Blue cells contain information about patients with any complications and the red cells are for patients with severe complications. Columns named # indicate the number of surgeries with the variable p-values in the first column named p-value is from the Fisher test or the chi-squared test depending on the data type. The second column named p-value contains Kendall's tau test for the length of stay.

In our data no significant relation between the variables *if blood* and *Preoperative thrombosis prophylaxis* is present, when doing the t-test we see a p-value of 0.64.

5.4. Analysis of intraoperative anaesthesia and drugs variables

***variable*, explanation of variable Value**	**#**	**#**	**#**	**odds ratio**	**odds ratio**	**p-value**	**mean LOS**	**p-value LOS**
General anesthesia								
inhalation (volatiles)	111	34	13	0.92	1.25	0.82	5.71	0.32
total intravenous (TIVA)	385	125	37	1.09	0.8	0.59	4.91	0.32
Depth of anaesthesia monitored								
no	100	32	11	1	1.13	1	4.84	0.61
yes	396	127	39	1	0.88	0.71	5.16	0.61
If spinal anaesthesia								
any type non spinal	76	27	9	1.2	1.24	0.51	6.25	0.1
spinal	420	132	41	0.83	0.81	0.54	4.88	0.1
Infusion of vasoactive drugs medicine influencing hearth function								
no	125	42	16	1.1	1.45	0.66	4.54	0.2
yes	371	117	34	0.91	0.69	0.23	5.28	0.2
Intraoperative thrombosis prophylaxis								
no	357	121	41	1.36	1.87	0.17	5.18	0.67
yes	139	38	9	0.73	0.53	0.13	4.86	0.67
If lidocaine type of anastesia								
no	422	138	44	1.23	1.32	0.5	4.92	0.35
yes	74	21	6	0.82	0.76	0.68	6.09	0.35

Table 5.7: Information on the intraoperative anaesthesia and drugs. Blue cells contain information about patients with any complications and the red cells are for patients with severe complications. Columns named # indicate the number of surgeries with the variable p-values in the first column named p-value is from the Fisher test or the chi-squared test depending on the data type. The second column named p-value contains Kendall's tau test for the length of stay.

Anticoagulants, also called blood thinners, are drugs that avoid blood clots. It can be given before, during or after surgery. No thrombosis complications, the complication formed by a blood clot, are present in the data set. We see that *Preoperative thrombosis prophylaxis* means a small increase in the frequency of complications in Table 5.7. However, this did not test significant. In the ERAS guidelines risk factors for thrombosis are mentioned. These include high cancer stage, advanced age, obesity and steroid usage [15]. We see no clear relation between *Intraoperative thrombosis prophylaxis* and *Age* (t-test p-value 0.55). For the *BMI* we do see a relation between *BMI* and the *Intraoperative thrombosis prophylaxis* with the (chi-squared test p-value 0.0077).

5.5. Analysis of intraoperative surgery type

***variable*, explanation of variable Value**	**#**	**#**	**#**	**odds ratio**	**odds ratio**	**p-value**	**mean LOS**	**p-value LOS**
Main procedurename								
stoma procedure	60	19	6	0.98	0.99	0.66	5.42	0.61
ileocaecal/ right hemicolectomy	168	53	13	0.97	0.66	0.09	5.55	0.41
uncommon procedure	80	31	11	1.42	1.54		5.35	0.7
anterior resection of the rectum	84	27	14	1	2.09		5.45	0.58
sigmoid resection	104	29	6	0.78	0.48		3.67	0
Additional major procedures								
no	395	106	31	0.33	0.37	0	4.1	0
yes	101	53	19	3.01	2.72	0	8.99	0
Stomal procedure								
yes	124	41	14	1.06	1.19	0.82	5.95	0.07
no	372	118	36	0.94	0.84	0.61	4.81	0.07
Bowel anastomosis								
yes	406	133	40	1.2	0.87	0.53	4.99	0.4
no	90	26	10	0.83	1.14	0.7	5.58	0.4
If open surgery Or Converted								
yes	116	51	19	1.98	2.21	0	7.97	0
no	380	108	31	0.51	0.45	0.01	4.21	0

Table 5.8: Information on the intraoperative variables about surgery type. Blue cells contain information about patients with any complications and the red cells are for patients with severe complications. Columns named # indicate the number of surgeries with the variable p-values in the first column named p-value is from the Fisher test or the chi-squared test depending on the data type. The second column named p-value contains Kendall's tau test for the length of stay.

We see a clear relationship in the data between blood loss and the surgery ending as open surgery with an odds ratio of 2.51 (*If open surgery or converted*, *If blood*, p-value $4.2 \cdot 10^{-5}$). In our data, we see an even clearer relation between open surgery and previous surgery to the same region in the body (*If open surgery or converted* and *Previous surgery to same abdominal region* odds ratio 3.93, p-value $1.573 \cdot 10^{-9}$).

5.6. Analysis of postoperative variables

***variable*, explanation of variable** **Value**	**#**	**#**	**#**	**odds ratio**	**odds ratio**	**p-value**	**mean LOS**	**p-value LOS**
PONV observed in 3 days postop. nausia or vomiting								
yes	239	47	16	1.85	3.62	0.02	6.52	0
no	257	30	5	0.54	0.28	0.01	3.77	0
Mobilised on POD 1 or 2 return of daily living activities								
on or before POD 1	204	25	5	0.64	0.43	0.24	3.54	0
on POD 2	146	27	5	1.36	0.74	0.06	3.84	0
on or after POD3	146	25	11	1.18	2.77		8.51	0
Intravenous fluid infusion restarted reconnected to drip								
no	429	64	19	0.73	1.51	0.36	4.35	0
yes	67	13	2	1.37	0.66	0.75	9.85	0
CD at least 2 complications during primary stay								
no comp with CD2 or more	433	62	18	0.53	0.87	0.06	3.5	0
yes comp with CD2 or more	63	15	3	1.87	1.15	0.74	16.02	0
Time to tolerating solid food nights								
tolerated solid food on day of surgery	392	56	12	0.66	0.33	0.17	4.61	0.03
did not tolerate solid foods on day of surgery	104	21	9	1.52	3	0.02	6.89	0.03
More than one night with urinary drain								
no	330	45	11	0.66	0.54	0.12	4.14	0
yes	166	32	10	1.51	1.86	0.16	6.98	0
High cancer stage high risk cancer patients								
no	442	67	19	0.79	1.17	0.55	4.98	0.37
yes	54	10	2	1.27	0.86	1	6.04	0.37
Duration of IV fluid infusion nights nights with drip								
removed on day of surgery	261	29	7	0.49	0.44	0.02	3.48	0
removed on POD 1	121	26	5	1.74	0.97	0.06	5.08	0.98
removed on or after POD 2	114	22	9	1.42	2.64		8.79	0

Table 5.9: Information on the postoperative variables. Blue cells contain information about patients with any complications and the red cells are for patients with severe complications. Columns named # indicate the number of surgeries with the variable p-values in the first column named p-value is from the Fisher test or the chi-squared test depending on the data type. The second column named p-value contains Kendall's tau test for the length of stay.

The low p-values for the test with the length of stay in Table 5.9 do not come as a surprise, since some variables contain implicit information about whether a patient was present in hospital on a certain day. For example, if the patient mobilised on POD 3, the patient had to be still in hospital on that day. If one wants to see the effect of postoperative data on the length of stay more research should be performed.

We see that *More than one night with urinary drain* and *Mobilised On POD 1 or 2* are related we see a p-value of $4.8 \cdot 10^{-13}$ with the chi-squared test. We also see a relation between *Mobilised On POD 1 or 2* and respiratory complications with a p-value of 0.0041 from the chi-squared test. One should keep in mind that only 13 respiratory complications were present, of these complications 9 mobilised on POD 3.

In the literature, diet post-surgery is related to the rate of recovery [18]. We indeed see, that *Time to tolerating solid food nights* relates to complications and LOS in Table 5.9.

5.7. Analysis of postoperative drugs usage

***variable*, explanation of variable** **Value**	**#**	**#**	**#**	**odds ratio**	**odds ratio**	**p-value**	**mean LOS**	**p-value LOS**
Opioid use on postoperative day 1								
yes	379	67	20	2.3	6.46	0.02	5.73	0
no	117	10	1	0.44	0.15	0.04	3.04	0
Postoperative use of NSAIDS type of drug								
yes	129	25	10	1.46	2.72	0.16	4.97	0.78
no	367	52	11	0.69	0.37	0.04	5.14	0.78
Time to pain control with oral analgesics nights								
on POD 1 or day of surgery	255	35	12	0.75	1.27	0.27	3.17	0
on POD 2 or after	241	42	9	1.33	0.79	0.66	7.12	0

Table 5.10: Information on postoperative drugs usage. Blue cells contain information about patients with any complications and the red cells are for patients with severe complications. Columns named # indicate the number of surgeries with the variable p-values in the first column named p-value is from the Fisher test or the chi-squared test depending on the data type. The second column named p-value contains Kendall's tau test for the length of stay.

We see a general trend in Table 5.10 indicating that the fewer painkillers needed, the better a patient recovers. However multiple reasons can exist for an increase in painkiller usage for the group developing severe complications. It could be that the complications themself hurt or that the medication eliminates pain but increases the risk of developing a complication. Another reason could be that they both are the result of the same cause, a more severe surgery leading to an increased risk of complications and an increased need for painkillers.

5.8. Analysis of Continous variables

Variable	Minimum value	Mean value	Maximum value
Total IV volume of fluids intraoperatively	0	1344.44	9030
IV volume Postoperational	0	237.23	3499
Core body temperature at end of operation	34.5	36.39	37.8
Length of operation minutes	5	121.96	475
Age	20	65.73	92
Max pain VAS in 3 days	0	2.84	8
Length of stay nights in hospital after primary operation	1	4.87	83

Table 5.11: The minimum, mean and max value for continuous variables in the feature-engineered data.

In Table 5.11 we see that the length of operation time can be short, with the minimum equaling 5 minutes. We test if these values are similar to other studies. In the studies by Rencuzogullari et al. [44] and by Jurt et al. [17] the surgery time and standard deviation for colectomy surgery and colorectal surgery are denoted. These studies find a mean surgery time of 177.48 and 180 minutes with standard deviations of 96.98 and 90 respectively. For our data, we see a mean surgery time of 121.96 minutes with a standard deviation of 58.77. This seems to differ from what we have in our study. We use the chi-squared test to see whether our observations can be from a normal variable with a mean of 180 with a standard deviation of 90. We obtain our test statistic by first normalising the surgery times with the values from the literature. We subtract 180 minutes and divide by 90, then square and sum all values. The obtained test statistic equals 416.8, $P(\chi^2(495) \leq 416.8) = 0.0045$. We consider this an unlikely tail event and dispose of the null hypothesis. If we repeat this for the values from the second study specifying surgery times we see a smaller p-value (test statistic 343.8, p-value $3.6 \cdot 10^{-8}$).

However, we are hesitant to alter the surgery times. Possible reasons for the shorter surgery times can due to a different patient selection. Not all surgeries in our data are classified as colorectal or colectomy surgeries but include stoma procedures and ileo procedures as well. Another reason for short surgery times could be the hospital itself. The MST is considered being the largest non-academic hospital in the Netherlands. With the size of the hospital, we can assume its surgeons perform many surgeries. Patients needing experimental or very specialised surgery would likely not be operated at the MST but in an academic hospital instead. Furthermore, both studies denote a bigger intraoperative blood loss (30% more than 100 ml in our data set, in study [17] 30% more than 200 ml. Study [44] denotes a bloodloss in need of transfusions of 8.8%, in our data this is 3.1%).

In Table 5.12 different values about the complications and continuous variables are shown. In the first column, we see Kendall's tau with severe complication, the second column shows Kendall's tau with any complication. Subsequently, the third and fourth columns show the p-value t-test with the severe complication and with any complication. Then in the last two columns, we have the Kendall's tau with the length of stay and the p-value for the Kendall's tau test with the length of stay. We see that *Age* and *Core body temperature at end of operation* don't test significantly with the two different outcomes. However, *Age* does seem to be related to the length of stay.

When comparing pre and intraoperative variables in Table 5.12 and Table 5.13, we see that these variables relate less to complications post-primary stay. This is not surprising since there are fewer post-primary stay complications than post-operational complications. An exception is the variable *Age* which interestingly seems to relate more to post-primary stay complications than to any complications.

Hypothermia
When the core body temperature is less or equal to 35 degrees, it is a serious medical condition called hypothermia (Greek for below heath). The ERAS protocol states that a median intraoperative core body temperature of 35.6 degrees or below can lead to adverse health-related effects. In the ERAS guidelines the importance of accurately measuring the core body temperature is stressed [15]. Hypothermia can lead to complications, mostly cardiovascular complications, cardiac arrhythmia, blood loss, infections and increased length of hospital stay [15]. In the data only one patient developed cardiac arrhythmia, this person had a core body temperature during the end of the operation equal to 36.1.

variable	k. tau	k. tau	p-value	p-value	k. tau	p-value
Total IV volume of fluids intraoperatively	0.12	0.13	0	0	0.19	0
Core body temperature at end of operation	-0.02	0.03	0.77	0.22	0.05	0.14
Length of operation minutes	0.08	0.1	0.03	0	0.08	0.01
Age	0.01	0.04	1	0.43	0.08	0.01

Table 5.12: Kendall's tau and p-value from tau test (blue for any complication and red for a severe complication), with post-operational complications.

variable	k. tau	k. tau	p-value	p-value	k. tau	p-value
IV volume post operational	-0.03	0.06	0.11	0.71	0.1	0
Max pain VAS in 3 days	0.01	0.19	0.54	0.09	0.2	0
Length of stay nights in hospital after primary operation	0.03	0.42	0.61	0.19	1	0
Pre- and intraoperative variables						
Total IV volume of fluids intraoperatively	0.05	0.13	0.19	0.05	0.19	0
Core body temperature at end of operation	-0.01	0.03	0.9	0.22	0.05	0.14
Length of operation minutes	0.01	0.1	0.73	0.13	0.08	0.01
Age	-0.08	0.04	0.09	0.38	0.08	0.01

Table 5.13: Kendall's tau and p-value from tau test (blue for any complication and red for a severe complication), with postprimary stay complications.

When using the t-test on the variable *Core body temperature at the end of operation* separated by *If blood*, we see a p-value of 0.038. We do not see any clear relation between the temperature at the end of surgery with the length of stay (Kendall's tau test $\tau = 0.049$, p-value 0.14) or infectious complications during primary stay (p-value t.test 0.17) in our data.

5.9. Comparision with previous data study

In the study by F. Raijmakers the same data set as in this thesis is used [42]. Since our study is performed after the study of F. Raijmakers, some additional observations are included. We do not see many differences appear in the shape of the data. The variables seem to be distributed similarly and the severe complications seem to relate similarly to the variables.

A difference between the studies is the significance of the BMI-related variables. In Raijmakers study, a binary variable is made for obesity. This variable had a p-value of 0.178. We suspect that the increase in the number of observations made the BMI test significant in this study. Another explanation could be that we include the BMI-related variable as a discretised variable including a lower body weight class. We also see some improvement in the p-value for the length of operation in our study.

We see more laxatives being administered in the study of F. Raijmakers. This may be due to the different observation selections, the study by F. Raijmakers limits to laparoscopic or robotic-assisted surgeries. In the study by F. Raijmakers 32% (61 of 118) received some oral bowel preparation, and in our study, we see 23% (115 of 496) with oral bowel preparation.

Blood loss seems to be twice as much in this study compared to the study by F. Raijmakers. This is likely also to be explained by the study of F. Raijmakers limiting to laparoscopic and robotic-assisted surgeries.

6

Variable selections

In this chapter, the three different variable selection methods, as explained in Chapter 2, will be applied to the feature-engineered data. We briefly describe the 3 variable selection methods again at the beginning of this chapter. We will apply these methods and list the selected variables in this chapter. We compare the selected variables and determine which variables are considered for modeling in the 3 different variable selections.

Literature-based variable selection
The first variable selection method applied is based on related studies. A variable selection is made from the variables used in the literature to predict adverse surgery outcomes. Much information is available in different studies. We want to study whether this information is useful for predicting in our setting. We include a variable if seen as useful in at least one study in our selection of studies. The studies we included are listed in Table 6.1. We make a table stating whether the variables we have in our feature-engineered data are also studied in a different study. We denote whether the variables are concluded to be risk factors or not (Tables 6.2 - 6.5). The included studies differ in many aspects. Statistical models, initial variables, thresholds, selected patient population and outcomes can all differ within these studies. We aim to include many different studies in order to obtain an overview of all that could be of importance. We briefly state variables studied in our literature selection, but happen not to be available to us.

Marginal importance assessment
In the second variable selection method, we test each variable with the outcome directly. We test for a significant difference within the explanatory variable for the patients who did and did not get a complication. The test calculates the chance of seeing the observed difference in the explanatory variable over the outcome groups (complication group) assuming the variables are independent of the complication. The Fisher test, the chi-squared test and the t-test depending on the type of variable are used. If the test returns a p-value of less than 0.05 we decided to include the explanatory variable. Since the outcome for Scenario 3 differs from Scenarios 1 and 2, the selections are expected to differ here as well. The variables selected in Scenario 1 are also selected in Scenario 2 since these are based on the same tests. However, Scenario 2 will also include intraoperative variables. Furthermore, the variable selections made with marginal importance differ for the severe and the non-severe complications outcomes.

AIC forward selection method
AIC forward refers to a greedy algorithm selecting variables based on the best AIC. During each step multiple models are fit, and one variable is added that leads to the best better AIC, in case one model with a better AIC exists. The variable selection with BIC forward was also tried but seen as too restrictive on this data, leading to models with fewer variables.

6.1. Literature-based variable selection

In this section, we select variables based on their perceived importance in different studies. We make a table of which variables are perceived by the study as important (see Table 6.2 - 6.5). We limit this table of studied variables to the variables for which we have similar information in the feature-engineered data set. Our data only contains planned ERAS protocol surgeries. Variables about emergency or planned surgery, ERAS protocol adherence conversion to open surgery, and risk assessment from an exception. These variables are included in the table because even though they are not included through a variable in our data set, we are still interested in this information. In feature engineering, we merged the information about whether surgery was started as open or closed surgery and the information about whether the surgery was converted from closed surgery to open together into one variable. Since most studies include these variables separately, we decided to list them separately in the table. Table 6.1 summarises the studies we included. More information about the type of studies included in the literature selection is in Appendix H. The results of the literature are summarised in Tables 6.6 and 6.7. Here we see the frequency a variable was included in a study, the percentage of times it was considered to be important (percentage of times important overall variable selections) and the percentage of times it was considered important when it was studied (percentage important overall variable selections where this variable was explicitly studied).

no. of study	**ref. no.**	**type of pop. studied**	**outcome studied**	**# of obs.**	**% with outcome**	**# risk factors**
1	[15]	colorectal	LOS and recovery	*	*	24
2	[18]	colorectal	post- and intraoperative complications	*	*	13
3.1	[43]	colon	Failure of ERAS, anastomic dehiscence (3.1)	908	4.2	4
3.2		i.d.	and ileus (3.2)	908	11.8	5
4	[42]	colorectal	severe complications	188	18	15
5	[11]	general	postoperative complications	400	31.5	11
6	[45]	abdominal cancer	complications within 30 days	308	34	6
7.1	[6]	colon cancer	leakage needing treatment (7.1)	15667	7.5	6
7.2		i.d.	leakage leading to death (7.2)	15667	1.2	5
8.1	[38]	different types 22% abdominal	health-related quality of life: pain 8.1,	336	**	5
8.2		i.d.	physical functioning 8.2	356	**	4
8.3		i.d.	perceived recovery 8.3	359	**	5
8.4		i.d.	mental health 8.4	373	**	4
8.5		i.d.	vitality 8.5	378	**	6
9	[24]	colorectal cancer	mortality	39000	43.3	10
10.1	[48]	colorectal cancer	anastomotic Leakage, early (10.1)	36929	2.3	7
10.2		i.d.	and late (10.2) leakage	36929	1.8	7
11	[37]	colorectal cancer	affect of complications on mortality	604	4.1	4
12	[52]	colorectal cancer	cost, LOS and mortality during primary stay (12)	10271	0.89	1
13	[27]	colorectal	readmission, reoperation or different type of complications	3552	30***	3
14	[44]	colon	postoperative ileus	29201	13	10
15	[19]	colorectal	intra- and postoperative complications (15)	1316	22	5
16	[46]	colorectal	in-hospital mortality	185000 (±)	9.1	5
17	[47]	colorectal cancer in older patients	postoperative morbidity or postoperative mortality	25	*	*
18	[23]	colorectal	anastomotic leakage	739	8.7	2
19	[17]	colorectal	postoperative respiratory complications	1298	9.2	5
20	[30]	colorectal	early versus late readmission	69222	10.8	9
21	[2]	different types 45% gastro intestinal	quality of recovery	182	13.2	1

Table 6.1: Information of studies included in the variable selection based on the literature. * is not present since it is a literature review. ** none binary outcome, but the outcome is the percentage of recovery, *** upper bound, exact number was not stated.

We see in Table 6.1 that the number of observations, percentage of the studied outcome and the number of risk factors differ greatly over the diverse studies. We aimed to include many studies to minimise the effect of the discrepancies within the selected studies. If a study (from in Table 6.1) includes multiple outcomes separately we indicate which outcome is included by denoting it with a number after the outcome. For example, study 3 is included twice (3.1 and 3.2), where 3.1 is for the outcome of ileus, and 3.2 is for the outcome of an anastomotic leak from study 3. This study also looked at risk factors for failure to follow the ERAS protocol. However, we did not include the variable selection for the failure of ERAS since this is not a complication.

Studies with more observations included can be considered more reliable. Strangely, there does not seem to be a relation between the number of observations and the number of variables concluded to be important in a study. Postoperative variables were often considered a result of surgery and not a risk factor for further complications and accordingly were not commonly studied as possible explanatory variables.

Most studies are executed with more restricted inclusion criteria for their observations than ours. In other words, they limit to a more specific patient population based on diagnoses or surgery type. In Table 6.1, shows that we included studies on different parts of the intestine and on different diseases that are also included in the data we will study. Sadly, information about surgery on the small intestine is lacking.

Often a logistic regression was fitted. Typically the model was fitted on a selection of variables that were considered important by looking at a univariate analysis (testing if 1 explanatory variable is differently present separated by the outcome). This selection of variables was then limited to avoid the inclusion of highly correlated variables. Then from a multivariate model (mostly logistic regression see Appendix H) is concluded which variables form risk factors. Which variables were omitted in the model to avoid confounding is often not argued. Why a model that assumes independence between explanatory variables is chosen is unclear.

In medical studies, mathematics is not the main focus. We observed some conventions for handling certain variables in the literature. We are not certain that these are medically well-founded or merely a product of mathematical convenience. For example, BMI and age are commonly studied in a discretised way, using similar cut-offs throughout studies. The reason for discretion is often not mathematically or explicitly motivated. An age above 65 is related to increased risk of complication, and increased time of recovery [7] (this particular study is not selected for variable selection from literature). However, no studies are found that explicitly consider where to place this threshold for the categorisation of age, merely its significance. A study by Pirrera et al. [39] claims that there is no direct risk obtained from old age, but rather that the suggested link between surgery outcome and age that is found in different studies is obtained by the prevalence of comorbidities in the higher age group (this study is not selected for variable selection from literature).

BMI is commonly discretised in the way suggested by the World Health Organisation (WHO). A commonly used indication for a healthy weight is by categorizing BMI based on population means and the association of mortality in low and high BMIs. Similar tables were used in insurance policies at the beginning of the 20th century [22]. During this time period, tuberculosis and pneumonia affected greatly the survival of persons with a lower weight. With the common availability of antibiotics, the recovery rates of the different weight classes have changed greatly for these diseases.

One study concluded from the univariate analysis which variables were the risk factors, and a model was fitted in the study for extra proof of the variables' importance [11]. Another 2 studies ([52] and [2]) state only 1 variable to be a risk factor. These studies only studied the importance of this one variable. Study 9 researches mortality in colorectal cancer and has a mortality rate of 43 % which is very high [52]. The time the patients were followed in this study is between 2 and 8 years. One study compared the quality of postoperative recovery and assessed health status before and after surgery [2]. The study concluded that a bad recovery was more noticeable at 2 weeks after surgery than at 3 months after surgery and concluded that a bad recovery is predictable from the patient's characteristics. However, the study did report that not many variables were studied and stated that there could be a significant difference after 3 months when using a larger sample size (182 observations were used).

6.1.1. Summary of literature review findings

In the ERAS guidelines, some measures are recommended even though the evidence of these measures is weak, mostly this is for measures that are at worst harmless when applied. We assume their predictive power not to be great. We denote these kinds of findings, those that are deemed as important while stating to have insufficient proof, with (+) in the Tables 6.2 - 6.5.

The different studies include variables differently. For example for a variable about comorbidity, often the Charlson Comorbidity Index (CCI) was used to give a quick indication about what is considered a comorbidity. One study included high BMI and weight loss as a comorbidity [35]. In Table 6.1 the number of variables, as they appear in the study, is denoted. For this reason, the numbers of variables do not appear equal in Table 6.1 as in Tables 6.2 - 6.5.

The variables diagnosis and procedures differ greatly over the studies. They can vary from only differentiating between a tumour in the colon or a tumour in the rectum ([48]), to differentiating between 10 seperate procedures ([27]). If a study included a specific heart disease as a variable it is denoted as severe heart disease. Similarly, any lung disease was considered as a severe pulmonary disease. For the studies that studied a variable about preoperative functional status or dependency, we include this as the WHO score since the WHO score measures the ability to take care of oneself.

On average 11 variables that are present in our data are studied. From these variables, the studies classified 53% as a risk factor.

no. study	1	2	3.1	3.2	4	5	6	7.1	7.2	8.1	8.2	8.3	8.4	8.5	9	10.1	10.2	11	12	13	14	15	16	17	18	19	20	21
patient characteristics																												
Gender		+	-	-	+	-	-	+	-	-	-	-	-	-	+	+	+	-		-	+	+	-		-	+	+	-
Smoker	+				+		-														+				-	-		
Alcohol usage	+				-																					-		
BMI (obesity or malnutrition)		+			+	+	-	-	-							+	-			+	-	-			+	-		-
If predisease (comorbidities)			-	-		+	-	-	+						+	-	+	+	+	-	-	-						
Diabetes mellitus					+		-												+		-	-			-	+		
Severe hearth disease (any or specific hearth disease inc. hypertension)		+			+		-												+		-	-	+		-			
Severe pulmonary disease (any or specific pulmonary disease)					+		-												+		-	-	+		-			
ASA physical status class		+	-	-	+	+	+	+	+	+	+		+	+		-	+			-	+	+		+	-	+		-
Preoperative WHO performance score (or functional status/dependency)					+		-														-					+		
Age		+	-	-	-	-	+	-	+	-	-	-	-	-	+	+	-	+		+	+	+	+	+	-	+		-

Table 6.2: The importance of patient characteristics for the studies listed in Table 6.1.

no. study	1	2	3.1	3.2	4	5	6	7.1	7.2	8.1	8.2	8.3	8.4	8.5	9	10.1	10.2	11	12	13	14	15	16	17	18	19	20	21
preoperative variables																												
Previous surgery to same abdominal region					+		-	-	-							-	-									-		
Any nonsurgical preoperative treatment							-									-	+	-							-	-		
Preoperative stoma counseling	(+)																											
Was anaemia found (preoperative anaemia or haemoglobin level)	+	+	+	-			+															-						
Preoperative nutritional treatment (or nutritional status)	+	+																										
Preoperative oral carbohydrate treatment	+																											
Oral bowel preparation	+				+																+				-			
Preoperative thrombosis prophylaxis																												

Table 6.3: The importance of preoperative variables for the studies listed in Table 6.1.

no. study	1	2	3.1	3.2	4	5	6	7.1	7.2	8.1	8.2	8.3	8.4	8.5	9	10.1	10.2	11	12	13	14	15	16	17	18	19	20	21
interoperative variables																												
Final diagnosis		+	+	+																-	-	+	+		-	-	+	
If colloids							+																					
If blood (blood loss)			-	-	-	+	+													-*			-			+		
Surgical/resection site drainage (surgical site drain)	+																								+			
General anaesthesia					-		-																					
Depth of anaesthesia monitored	+																											
If spinal anaesthesia																												
Infusion of vasoactive drugs							+																					
Interoperative thrombosis prophylaxis																												
If lidocaine	+																											
Main procedure (surgical site/ cancer site/ procedure)			-	+	+	+	-	+	+	-	-	+	-	-		+	+	-		+	-	-	-		-	+		
Additional major procedures								+	-							-	+											
Stomal procedure (could be a specific one)					+			+	-							+	-									-	+	
Bowel anastomosis																												
If open surgery of converted (open surgery)	+		-	+		+										+	-	-		-	+				-	+		
Total IV volume of fluids intraoperatively (fluid balance/ IV fluids)	+				-		-																					
Core body temperature at end of operations (or during)	+		-	-	-																							
Length of operation minutes		+	-	-	+	+	-			+	-	+								+	+	-			-	+		-

Table 6.4: The importance of interoperative variables for the studies listed in Table 6.1, * bloodloss was estimated.

no. study	1	2	3.1	3.2	4	5	6	7.1	7.2	8.1	8.2	8.3	8.4	8.5	9	10.1	10.2	11	12	13	14	15	16	17	18	19	20	21
postoperative variables																												
PONV observed in 3 days (or any day)																											-	
(*Mobilised on POD 1 or 2* (or any day)	+		+	+																								
Intravenous fluid infusion restarted																												
CD at least 2 (postoperative complications)			(+)*			+												+		-	+						+	
(*time to tolerating solid food nights* (postoperative oral intake)	+																											
More than one night with urinary drain (or any usage)	+																											
High canser stage (metastasis or cancer stage)							-	-	-						+	-	-	+	+			-	+	+				
Time to pain control with oral analgesics nights	+																											
IV volume postoperational	+																											
Opioid use on postoperative day 1 (or any day)	+																											
Postoperative use of NSAIDS																												
Max pain VAS in 3 days (any postoperative discomfort)										+	+	+															-	
length of stay nights in hospital after primary operation						+															-	-						+
no variable directly in the data set																												
Thrombosis prophylaxis	+				+																							
Intraoperative blood given		+					-															-						
Risk assessment	(+)		(+)*																									
Emergency surgery						+		+	+							+	-			-			+	+	-	+		
ERAS adherence	(+)		(+)	(+)																								
Conversion of closed surgery to open			+	-																	-					+		

Table 6.5: The importance of postoperative variables for the studies listed in Table 6.1, $*$ study uses the risk of other complications as a risk factor.

We select variables when seen as important in at least one study. This leads to 41 variables. The variables that are selected are in blue in the Tables 6.6 - 6.7.

name variable	% important when studied	# studied	% found to be of importance
patient characteristics			
ASA physical status class	0,71	21	0,54
Smoker	0,5	6	0,11
Severe heart disease (any or specific hearth disease inc. hypertension)	0,5	8	0,14
Preoperative WHO performance score (or functional status/dependency)	0,5	4	0,071
Age	0,48	25	0,43
BMI (obesity or malnutrition)	0,43	14	0,21
If predisease (comorbidities)	0,43	14	0,21
Diabetes mellitus	0,43	7	0,12
Severe pulmonary disease (any or specific pulmonary disease)	0,43	7	0,12
Gender	0,4	25	0,36
Alcohol usage	0,33	3	0,036
preoperative variables			
Peopertive nutritional treatment (or nutritional status)	1	2	0,071
Preoperative oral carbohydrate treatment	1	1	0,036
Oral bowel preparation	0,75	4	0,11
Was anaemia found (Preoperative anaemia, or haemoglobin level)	0,67	6	0,14
Any nonsurgical preoperative treatment	0,17	6	0,036
Previous surgery to same abdominal region	0,14	7	0,036
Preoperative stoma counseling	0	1	0
Preoperative thrombosis prophylaxis		0	0

Table 6.6: Summary of variable selection from importance in literature selection. In blue are the selected variables.

Unavailable variables from literature

In the literature, many variables are studied which are not present in our data. The following variables are studied at least twice: the experience of the surgeon (patients treated or times specific surgery performed) ([18],[27],[19]), data about the hospital ([45],[6],[24]), psychological factors (self-perceived health and long-time fear) ([11],[38]), albumin level (class of protein found in blood) or other blood values([43],[11],[44],[47]), bowel spillage (peritoneal soiling) ([46],[17]), hearth rate ([11],[46]), preoperative tumour complications ([45],[48]), steroid usage (hormone) ([23],[44]), other specific medications (for example oxycodone) ([43],[23]) and more information about the resection (extensive resection or if multiple resections were made) ([6],[48]).

Some variables are in the ERAS guidelines but not in our data. These include; prehabilitation (this includes prehabilitation exercises, protein supplements and relaxation strategies.), preoperative blood transfusion, routine sedative medication and preoperative fluids [15]. Variables mentioned in the ERAS guidelines and present in the data set but not included in the feature-engineered data are: Antibiotic prophylaxis and postoperative laxatives. The last two mentioned variables are almost constant and were therefore not included in our data. The recommendations are almost always followed and therefore not studied here [15].

Interpretation of variables selected based on literature

We see that this method leads to many variables being selected. We see that nonmodifiable patient characteristics and intraoperative variables are studied most often. We see a few variables that are studied but not considered important in any of the studies. These variables are *Preoperative stoma counselling*, *General anaesthesia* and *PONV observed in 3 days*. These variables were only studied once except for *General anaesthesia* which was studied twice. Furthermore, a lot of the variables not selected in the literature-based variable selection concern medication.

We have seen that in the literature-based selection method both the variable *If predisease* and the variables used to create this variable are selected. Similarly, we have a causality with *Stomal procedure* and *Main procedure*. A stoma procedure can be the main procedure, however a stoma procedure can also be performed during other types of surgeries. One has to keep this in mind when interperting any estimated coefficients for these variables.

name variable	% important when studied	# studied	% found to be of importance
Interoperative variables			
If colloids usage	1	1	0,036
Resection site drainage (surgical site drain)	1	2	0,071
Depth of anaesthesia monitored	1	1	0,036
Infusion of vasoactive drugs	1	1	0,036
If lidocaine	1	1	0,036
Final diagnosis	0,6	10	0,21
Stomal procedure (could be a specific one)	0,57	7	0,14
If open or surgery or converted (open surgery)	0,55	11	0,21
Length of operation minutes	0,53	15	0,29
Additional major procedures	0,5	4	0,071
Main procedure (surgical site/ cancer site/ procedure)	0,48	21	0,36
If blood (blood loss)	0,38	8	0,11
Total IV volume of fluids intraoperatively (fluid balance/ IV fluids)	0,33	3	0,036
Core body temperature at end of operations (or during)	0,25	4	0,036
General anaesthesia	0	2	0
If spinal anaesthesia		0	0
Interoperative thrombosis prophylaxis		0	0
Bowel anastomosis		0	0
postoperative variables			
Mobilised on POD 1 or 2 (or at all)	1	3	0,11
Time to tolerating solid food nights (postoperative oral intake)	1	1	0,036
More than one night with urinary drain (or any usage)	1	1	0,036
Time to pain control with oral analgesics nights	1	1	0,036
IV volume Postoperational	1	1	0,036
Opioid use on postoperative day 1	1	1	0,036
Max pain VAS in 3 days (any post-op pain)	0,75	4	0,12
CD at least 2 (any postoperative discomfort)	0,67	6	0,14
Length of stay nights in hospital after primary operation	0,5	4	0,071
High cancer stage (metastasis or cancer stage)	0,45	11	0,18
PONV observed in 3 days (or at all)	0	1	0
Intravenous fluid infusion restarted		0	0
Postoperative use of NSAIDS		0	0
Duration of IV fluid infusion nights		0	0
No variable directly in the data set			
Thrombosis prophylaxis	1	2	0,071
Emergency surgery	0,7	10	0,25
Conversion of closed surgery to open	0,5	4	0,071
Intraoperative blood given	0,33	3	0,036
Risk assessment	0	2	0
ERAS adherence	0	3	0

Table 6.7: Summary of variable selection from importance in literature selection. In blue are the selected variables.

6.2. Marginal importance assessment

The marginal variable selection we will apply is based on statistical testing. The Fisher test, the chi-squared test and the t-test are used depending on the variable type. We use the Fisher test with binary variables, the chi-squared test with other discrete data and the t-test with continuous variables. When a p-value less or equal to 0.05 is returned, we include the variable in this selection. We report on the picked variables for each scenario and outcome in Tables: 6.8 - 6.13. The variables are ordered by the lowest p-value first.

	Scenario 1, any complicaitons	
	Variable	**p-value**
1	*Preoperative WHO performance score*	1.3 e-07
2	*ASA physical status class*	0.0013
3	*Severe pulmonary disease*	0.012

Table 6.8: Variables selected by marginal importance for Scenario 1 (preoperative) with the outcome; any complications, ordered by p-values.

	Scenario 2, any complicaitons	
	Variable	**p-value**
1	*Preoperative WHO performance score*	1.7e-07
2	*Additional major procedures*	2.3 e-06
3	*Total IV volume of fluids intraoperatively*	0.00013
4	*If blood*	0.00014
5	*ASA physical status class*	0.0013
6	*If open surgery or converted*	0.0021
7	*Length of operation minutes*	0.0036
8	*Resection site drainage*	0.0092
9	*Severe pulmonary disease*	0.012

Table 6.9: Variables selected by marginal importance for Scenario 2 (directly after surgery) with the outcome; any complications, ordered by p-values.

	Scenario 3, any complications	
	Variable	**p-value**
1	*Preoperative WHO performance score*	0.0012
2	*Additional major procedures*	0.0031
3	*If open surgery or converted*	0.012
4	*Duration of IV fluid infusion nights*	0.015
5	*PONV observed in 3 Days*	0.018
6	*Opioid use on postoperative day 1*	0.019
7	*Total IV volume of fluids intraoperatively*	0.047
8	*Resectionsite drainage*	0.049

Table 6.10: Variables selected by marginal importance for Scenario 3 (after primary hospital stay), predicting any complications, ordered by p-values.

Interpretation of variables selected by marginal importance

With the variable selection method based on marginal importance, not many variables are selected.

We see that in Scenario 1 selection variables for the severe complications by marginal importance lead to only 2 variables being selected (Table 6.11). Surprisingly, *Preadmission stoma counselling* is selected for predicting severe complications.

We see that *Preoperative WHO performance score* is selected often with this method while age is not *Age*. This is in line with the findings from the literature (See Section 1.4). *ASA physical status class* and *Preoperative WHO performance score* are both selected. These variables both indicate preoperational health or mobility.

In Scenario 2 (Table 6.9 and 6.12) we see some variables that are not frequently present. More specifically *If open surgery or converted*, *Resectionsite drainage* and *Additional major procedures* are binary variables present 23%, 11.2% and 20% respectively in the data.

In Scenario 3 (Table 6.10 and 6.13) we see fewer variables. This is not surprising, considering the low number of postoperative complications present. We see that the preoperative and intraoperative selection in Scenario 3 forms a subset of the variables selected in Scenario 2. We would expect to see the intraoperative variables to be less frequently selected in Scenario 3. Since the intraoperative information is more distant time-wise from the complication postprimary stay. However, we would not expect zero intraoperative variables to be picked for severe complications. For Scenario 3 with any complications, 4 intraoperative variables are selected.

Interestingly the *Duration of IV fluid infusion nights* relates to any complications where whereas it does not relate to severe complications. *Postoperative use of NSAIDs* and *Time to tolerating solid food nights* are both selected for severe complications but not for any complications.

	Scenario 1, severe complicaitons	
	Variable	**p-value**
1	*Preoperative WHO performance score*	0.021
2	*Preadmission stoma counseling*	0.042

Table 6.11: Variables selected by marginal importance for Scenario 1 (preoperative), predicting severe complications, ordered by p-values.

	Scenario 2, severe complicaitons	
	Variable	**p-value**
1	*If blood*	5.9e-05
2	*Total IV volume of fluids intraoperatively*	0.0020
3	*Additional major procedures*	0.0025
4	*Resection site drainage*	0.0073
5	*If colloids*	0.0094
6	*If open surgery or converted*	0.014
7	*Preoperative WHO performance score*	0.021
8	*Length of operation minutes*	0.033
9	*Preadmission stoma counseling*	0.042

Table 6.12: Variables selected by marginal importance for Scenario 2 (directly after surgery), with the outcome; severe complications, ordered by p-values.

	Scenario 3, severe complicaitons	
	Variable	**p-value**
1	*PONV observed In 3 Days*	0.012
2	*Time to tolerating solid food nights*	0.024
3	*Preadmission stoma counseling*	0.035
4	*Opioid use on postoperative day 1*	0.036
5	*Postoperative use of NSAIDS*	0.038

Table 6.13: Variables selected by marginal importance for Scenario 3 (after primary hostpital stay) with the outcome; severe complications, ordered by p-values.

6.3. AIC forward selection method

The variables selected with the AIC forward selection method, are listed in Tables 6.14-6.19. The variables are listed according to the order in which they are selected by the algorithm.

	Variables selected
1	*Preoperative WHO performance score*
2	*Preadmission stoma counseling*
3	*Preoperative oral carbohydrate treatment*

Table 6.14: Variables selected with AIC forward for Scenario 1 (preoperation) with any complications as the outcome in the order as selected by the algorithm.

	Variables selected
1	*Total IV volume of fluids intraoperatively*
2	*If blood*
3	*Infusion of vasoactive drugs*
4	*Intraoperative Thrombosis prophylaxis*
5	*BMI*

Table 6.15: Variables selected with AIC forward for Scenario 2 (right after surgery) with any complications as the outcome in the order as selected by the algorithm.

	Variables selected
1	*PONV observed In 3 Days*
2	*Preadmission stoma counseling*
3	*Opioid use on postoperative day 1*
4	*Intraoperative thrombosis prophylaxis*
5	*Time to tolerating solid food nights*
6	*IV volume postoperational*
7	*Time to pain control with oral analgesics nights*
8	*Mobilised On POD 1 or 2*
9	*If colloids*
10	*Total IV volume of fluids intraoperatively*

Table 6.16: Variables selected with AIC forward for Scenario 3 (right after primary stay) with any complications as the outcome in the order as selected by the algorithm.

Interpretation of variables selected by AIC forward

In Scenario 1 the selection consists of information mostly about general health and preoperative treatments. In Scenario 1 we see more patients' characteristics for the severe complications than for any complications. We see only 1 variable that is selected with both the severe and any complications outcomes. We see that the *Preoperative WHO performance score* is an important predictor in Scenario 1.

With the outcome of any complication, we have no variables selected in Scenario 2 which were also selected in Scenario 1. However, when the outcome equals severe complication we see that Scenario 2 consists of all variables selected in Scenario 1 except *Diabetes mellitus*. The variable *Diabetes mellitus* was selected last in Scenario 1 with the severe complications. We see that for both severe and any complications in Scenario 2 the variables *Total IV volume of fluids intraoperatively* and *Infusion of vasoactive drugs* are selected. Surprisingly in Scenario 2, *Infusion of vasoactive drugs* is selected. This comes as a surprise since no thrombosis complications are present. It could be selected since this variable contains implicit information about other variables for example *Age* and *BMI*. However *BMI* is also selected in this variable selection.

In Scenario 3 we see half of the selected variables are postoperative variables. We see the variables *Intraoperative thrombosis prophylaxis* and *Total IV volume of fluids* are picked when the outcome equals any complication in Scenario 2 and Scenario 3. We see that in Scenario 3 with the severe complications, fewer variables and only 2 postoperational variables are selected. We see that *PONV observed In 3 Days* is an important postoperative variable, for both any and severe complications in Scenario 3.

We consider *Total IV volume of fluids intraoperatively* being more important for short-term effects, since it was selected in Scenario 2 for both outcomes whilst selected in Scenario 3 with any complications.

When the outcome is severe complication, we see *Additional major procedures* is selected in Scenarios 2 and 3, we do not see this variable being selected when the outcome is any complications.

	Variables selected
1	*Preoperative WHO performance score*
2	*ASA physical status class*
3	*Any nonsurgical preoperative treatment*
4	*Oral bowel preparation*
5	*Diabetes mellitus*

Table 6.17: Variables selected with AIC forward for Scenario 1 (preoperation) with severe complications as the outcome in the order as selected by the algorithm.

	Variables selected
1	*Preoperative WHO performance score*
2	*Additional major procedures*
3	*Total IV volume of fluids intraoperatively*
4	*ASA physical status class*
5	*Infusion of vasoactive drugs*
6	*Any nonsurgical preoperative treatment*
7	*Resection site drainage*
8	*Intraoperative thrombosis prophylaxis*
9	*Epidural or spinal anaesthesia*
10	*Oral bowel preparation*

Table 6.18: Variables selected with AIC forward for Scenario 2 (right after surgery) with severe complications as the outcome in the order as selected by the algorithm.

	Variables selected
1	*Preoperative WHO performance score*
2	*Additional major procedures*
3	*PONV observed In 3 days*
4	*Opioiduse on postoperative day 1*
5	*If colloids*
6	*Resection site drainage*
7	*Preadmission stoma counseling*
8	*Bowel anastomosis*

Table 6.19: Variables selected with AIC forward for Scenario 3 (right after primary stay) with severe complications as the outcome in the order as selected by the algorithm.

6.4. Comparing different variable selection techniques

Variable selection	Scenario 1		Scenario 2		Scenario 3	
	severe	not severe	severe	not severe	severe	not severe
From literature	17	17	31	31	41	41
By marginal importance	2	3	9	9	5	8
With AIC forward	5	3	10	5	8	10

Table 6.20: Number of variables selected per model and variable selection.

In Table 6.20 we compare the number of variables selected by the different variable selection methods for the different models. We see that the number of variables from the literature does not depend on whether the outcome is severe or not. In the literature, more variables are seen as important than the number of variables that have a marginal importance in the data.

We see that the number of variables and the selected variables differ more within the AIC forward selection method. We see that with method AIC forward more variables are picked in Scenario 3 than with the marginal importance.

We see that mostly WHO score or ASA class are important preoperative variables in Scenario 1. We see that in Scenario 2 more variables are selected in with all methods. Here we see that *Total IV volume of fluids intraoperatively* and *Additional major procedures* are important variables. We see that in Scenario 3 intraoperative variables are not so important as in Scenario 2. We see that *PONV observed In 3 Days* and *Opioid use On postoperative day 1* are important postoperative variables indicating recovery without complications after primary stay.

7

Model validations

In this chapter, the observations from the preprocessed data are used to fit models with the different variable selections made in Chapter 6. We take a look at the estimated coefficients for the different models. The models are ordered first by scenario, subsequently by complication and finally by the variable selection method. We afterwards study the in-sample validation and cross-validation.

We have 4 outcomes; severe for Scenarios 1 and 2, any complication for Scenarios 1 and 2, severe complications for Scenario 3 and any complications for Scenario 3. The scenario variable selections for Scenarios 1 and 2. In Scenario 3 (after the primary stay) we can not include the complications present during the primary hospital stay in the outcome. Scenario 2 is similar to Scenario 1, but Scenario 2 contains more variables (also intraoperative data).

7.1. Estimated coefficients

In this section we study the made models. In Tables I.1-I.9 in the appendix and in Tables 7.1-7.9, we see the coefficients of a logistic regression fitted on all the preprocessed observations. In the appendix are the tables with the outcome equalling any complication, and in this section are the tables with severe complications as the outcome. After the variable name, we see the value between quotation marks of the variable corresponding to the coefficient. The value for which no coefficient is estimated is considered as the baseline. The next column states the estimated coefficient corresponding to the value in quotation marks. We decided to put the tables concerning any complications in Appendix I.

We state the changes in signs of the estimated coefficients in Scenario 2 compared to the sign in Scenario 1. With the inclusion of more variables in Scenario 2, the sign of a coefficient can change due to introducing a correlated variable. A positive coefficient means that under the model, a patient with the value assigned to the corresponding coefficient has an increased risk of complications, than an identical patient without this value assigned. On the other hand, a negative coefficient means that the value of the coefficient decreases the risk of complications. A coefficient further from zero means a stronger effect on the risk under the model than a coefficient close to zero.

7.1.1. Scenario 1 with any complications

In Scenario 1, predictions of postoperative complications are made before the surgery takes place.

The largest influence under the model with literature variables in Scenario 1 predicting any complications (Table I.1) is the variable *Preoperative WHO performance score*. It was in the literature study considered 4 times and from those times it was 2 times concluded to form a risk factor. It is not surprising that it is important since it indicates preoperative health. We see that the estimated coefficient for the *ASA physical status class* is smaller than the one for the *WHO performance score* (0.427 versus 1.465), even though the *ASA physical status class* was more commonly studied and more frequently selected as a predictor in our selection of studies (21 times studied and of these 15 times used as a predictor).

We see some estimated coefficients that we would expect to have a different sign. These include: *Severe heart disease*, *Diabetes mellitus*, *Alcohol usage* and *Preoperative oral carbohydrate treatment*. These are not as one would expect following the statements of the ERAS protocol.

Preopretrative nutritional treatment and *Any nonsurgical preoperative treatment* both seem to decrease the risk of complications under this model. Indicating that getting treatment previously would benefit surgery outcomes and not be an indication of a high-risk patient. Deciding whether the patient should undergo surgery is different in situations where other treatments are available besides surgery.

In the literature-based model predicting any complications in Scenario 1, we see that a *BMI* below 21.5 leads to the fewest complications (Table I.1).

We see that the *Preoperative WHO performance score* has the largest influence in the model using marginal important variables in Scenario 1 predicting any complications (see I.2). The order of the distance from zero for the estimated coefficient is not the same as the reverse order of the p-values we found (Tables 6.8-6.13). We see that *Severe pulmonary disease* has a bigger estimate coefficient than *ASA physical status class* while the p-value for *Severe pulmonary disease* was larger. So a smaller p-value does not correspond to a more influential estimated coefficient in the model. This is not surprising since the p-value aims to represents how much the data supportes a difference, but not how large this difference is.

We see from Table I.3, with the selection from the AIC forward and any complications, that the biggest influence is given by *Preoperative WHO performance score* for Scenario 1.

The order of the estimated coefficients is not similar to the order in which the variables are selected within the AIC forward algorithm. *Preadmission stoma counselling* was selected before *Preoperative oral carbohydrate treatment* but the estimated coefficient for *Preadmission stoma counselling* is smaller than the one for *Preoperative oral carbohydrate treatment*.

We have found that *Preadmission stoma counselling* and *Preoperative oral carbohydrate treatment* are suggested to increase the risk of complications under the model from Table I.3. These measures are performed to improve surgery outcomes, so we suspect these variables not to increase risk but to be correlated with some information about the disease or treatment of the patient that leads to an increased risk of adverse surgery outcomes. Perhaps Stoma counselling is closely related to the presence of a stoma procedure. In the data, 14 patients got a stomal procedure and no stomal counselling, and 62 patients got stoma counselling without getting a stomal procedure.

Scenario 1 with severe complications

We see that the biggest influence under the model predicting severe complications in Scenario 1 with the variables from the literature (see in Table 7.1) is *Preoperative oral carbohydrate treatment*. We would expect this treatment to decrease the risk instead of increasing the risk of complications as it does in the model. We see that the *Preoperative WHO performance score* is further from zero than the *ASA physical status class* just like it was in Scenario 1 with the literature-based variable selection predicting any complication (Table I.1). However, this time we see that the estimated coefficient for *ASA physical status class* is negative, and this is not as one would expect. Furthermore, we expected the variables *Smoker*, *Severe heart disease*, *Age*, *Preoperative oral carbohydrate treatment* and *Was anaemia found* to have a different sign in their estimated coefficients.

The model of the marginal important variables with severe complications in Scenario 1 (Table 7.2) contains two binary variables. It seems that a general indication of preoperative health; *Preoperative WHO performance score* is the most important variable in this model. Again, strangely we see that the *Preadmission stoma counseling* increases the risk of complications.

In the model using the AIC forward variable selection predicting severe complications in Scenario 1 (Table 7.3), we see that *Preoperative WHO performance score* has the highest predicted coefficient. We see that in this model, *Diabetes melitus* and *Oral bowel preparation* have a sign being different from what we would expect.

Literature variables, Scenario 1, Severe complications	
Variable	**Estimated coeficient**
(Intercept)	-3.402
ASA physical status class '3 or 4'	-0.0103
Smoker 'yes'	-0.124
Severe heart disease 'yes'	-0.436
Preoperative WHO performance score '1'	1.08
Age	-0.00222
BMI '(21.5-24.9)'	-1.125
BMI '(25-29.9)'	-0.944
BMI '(30 or more)'	-0.618
If predisease 'yes'	0.581
Diabetes mellitus 'yes'	0.574
Severe pulmonary disease 'yes'	0.0118
Gender 'female'	-0.120
Alcohol usage 'yes'	0.273
Preoperative nutritional treatment 'yes'	-0.119
Preoperative oral carbohydrate treatment 'yes'	1.646
Oral bowel preparation 'yes'	0.335
Was aneamia found 'yes aneamia'	-0.125
Any nonsurgical preoperative treatment 'yes'	-0.114
Previous surgery to same abdominal region 'yes'	0.318

Table 7.1: Coeficients of logistic regression with literature-based variables predicting severe complications in Scenario 1.

Marginal variables, Scenario 1, Severe complications	
Variable	**Estimated coeficient**
(Intercept)	-2.625
Preoperative WHO performance score '1'	0.975
Preadmission stoma counseling 'yes'	0.670

Table 7.2: Coeficients of logistic regression with marginal important variables predicting severe complications in Scenario 1 (before surgery).

AIC forward variables, Scenario 1, Severe complications	
Variable	**Estimated coeficient**
(Intercept)	-2.386
Preoperative WHO performance score '1'	0.974
ASA physical status class '3 or 4'	0.0279
Any nonsurgical preoperative treatment 'yes'	-0.0880
Oral bowel preparation 'yes'	0.338
Diabetes mellitus 'yes'	-0.479

Table 7.3: Coeficients of logistic regression with variables selected with AIC forward predicting severe complications in Scenario 1 (before surgery).

Conclusion Scenario 1

An important influence for predicting in Scenario 1 is the variable *Preoperative WHO performance score*. The variable is the largest estimated coefficient in most models in Scenario 1 except when predicting severe complications with literature variables (in this case the largest influence was *Preoperative oral carbohydrate treatment*).

We take a quick look at comparing the models with severe outcomes and non-severe outcomes. Since the outcomes change, we can not conclude changes to be due to correlations within the data. However, an interesting observation is, that the lowest BMI group (the baseline) forms the highest risk group when predicting severe complications and when predicting any complications it is the lowest risk group in the models with the literature variables. Further, we see that *Diabetes mellitus, Alcohol usage* have a decreasing effect on the risk of any complication, but indicate an increased risk of severe complications in the literature models.

We do not find interesting differences in the severe and non-severe models in Scenario 1 with the marginal and AIC forward variable selections. However, not many variables were selected for the severe and the non-severe cases in these variable selections.

Scenario 2 with any complications
Under the model with the literature-based selection predicting any complications in Scenario 2 the variable *Preoperative WHO performance score* has the highest estimated coefficient.

The variables that we would have expected to have different signs are *Severe heart disease, Diabetes mellitus, Alcohol usage, Previous surgery to same abdominal region, If colloids, Depth of anaesthesia monitored, Infusion of vasoactive drugs, Stomal procedure, Main procedure name* equalling 'stoma procedure' and *Core body temperature at end of operation*. We think these estimated coefficients are counterintuitive and expect that these results are due to correlations among variables.

The estimated coefficient for *Stomal procedure* is -0.274 and the estimated coefficient for *Main procedure name* equalling 'stoma preocedure' is 0.778. All observations where *Main procedure name* takes the value 'stomal procedure', also have the variable *Stomal procedure* present, but also when other values of the variable *Main procedure name* are present the observation can have the variable *Stomal procedure* present. This suggests surgeries placing an anastomosis instead of a stoma lead to an increased risk and that the purely stoma procedures are the most risky.

The models use intraoperative fluids in ml so the coefficient of 0.000226 is not considered small. We have seen that the minimum, the mean and the maximum for *Total IV volume of fluids intraoperatively* equal, 0 ml, 1344 ml and 9030 ml respectively.

We see that the estimated coefficient for *Diabetes mellitus* equals -0.921. However, this variable is only present when the variable *If predisease* is present as well. Therefore we can consider the effect of diabetes under the model if we add the coefficients of *If predisease* and *Diabetes mellitus* to each other. So, -0.921+ 0.288 = -0.633, we see that under this model the variable *Diabetes mellitus* does lead to a lower predicted risk of any complications. The variable *If predisease* was constructed in Section 4.7.1 with the variables *Severe heart disease, Diabetes mellitus* and *Severe pulmonary disease*. One has to keep this in mind when studying the estimated coefficients of these variables.

For the model using the marginal important variables predicting any complications in Scenario 2 (see Table I.5) the largest influence is *Preoperative WHO performance score*. In this model, we see no signs of the estimated coefficients which we expected to be different.

We see that within the selection from the AIC forward and any complications for Scenario 2 (Table I.6), the biggest influence is given by the variable called *If blood*. We see under this model the variable *Intraoperative thrombosis prophylaxis* leads to an increased risk of complications. Since in Table I.6 we see one negative sign for the estimated coefficients for the variable BMI, we can conclude that under this model a *BMI* in the range '(21.5-24.9)', leads to the lowest risk of any complications.

Scenario 2 with severe complications

Literature variables, Scenario 2, Severe complications	
Variable	**Estimated coeficient**
(Intercept)	-0.144
ASA physical status class '3 or 4'	0.0746
Smoker 'yes'	0.000880
Severe heart disease 'yes'	-0.447
Preoperative WHO performance score '1'	1.027
Age	0.008
BMI '(21.5-24.9)'	-1.228
BMI '(25-29.9)'	-1.217
BMI '(30 or more)'	-0.645
If predisease 'yes'	0.680
Diabetes mellitus 'yes'	0.614
Severe pulmonary disease 'yes'	-0.327
Gender 'female'	-0.0793
Alcohol usage 'yes'	0.319
Preoperative nutritional treatment 'yes'	-0.478
Preoperative oral carbohydrate treatment 'yes'	1.74
Oral bowel preparation 'yes'	0.0690
Was aneamia found 'yes aneamia'	-0.237
Any nonsurgical preoperative treatment 'yes'	-0.280
Previous surgery to same abdominal region 'yes'	-0.136
If colloids 'yes'	0.291
Resectionsite drainage 'yes'	0.550
Depth of anaesthesia monitored 'yes'	0.086
Infusion of vasoactive drugs 'yes'	-1.083
If lidocaine 'yes'	-0.664
Final diagnosis 'cancer'	0.155
Final diagnosis 'no cancer or functional disorder'	0.396
Stomal procedure 'yes'	-0.944
Length of operation minutes	0.0000715
Additional major procedures 'yes'	0.394
If open surgery or converted 'no'	-0.0709
Main procedure name 'sigmoid resection'	-0.375
Main procedure name 'stoma procedure'	1.704
Main procedure name 'anterior resection of the rectum'	1.324
Main procedure name 'uncommon procedure, no ileocaecal'	0.553
If blood 'yes'	0.905
Total IV volume of fluids intraoperatively	0.000314
Core body temperature at end of operation	-0.118

Table 7.4: Coeficients of logistic regression with literature-based variables predicting severe complications in Scenario 2 (right after surgery).

The largest influence under the model predicting severe complications with the literature-based variables in Scenario 2 is given by the variable *Preoperative oral carbohydrate treatment*. This was also the case when predicting the severe complications with the literature-based selection in Scenario 1. We find it unexpected that a preoperative variable is the largest influence for a model in Scenario 2. We know from the protocol that whether this treatment can be given or not is decided by the type of procedure performed. This could form a possible explanation for being the largest influence under this model.

The variables that we would have expected to have different signs in the model from Table 7.4 are *Severe heart disease, Severe pulmonary disease, Was anaemia found, Previous surgery to same abdominal region* and *Depth of anaesthesia monitored*. *Stomal procedure* decreases the risk of complications in the model and we suspect this is due to the absence of risk of an anastomotic leak if no anastomosis was made but a stoma was placed.

We see that the variable *If blood* has the largest estimated coefficient for the model with variables with marginal importance in Scenario 2 predicting severe complications (Table 7.5). We see that *Preadmission stoma counselling* increases the risk of complications under this model (Table 7.5).

We see in Table 7.6 that the greatest influence for the model predicting severe complications with AIC forward variables in Scenario 2 is given by the variable *Infusion of vasoactive drugs*. Under this model, the presence of *If spinal anaesthesia* increases the risk of severe complications.

Marginal variables, Scenario 2, Severe complications	
Variable	**Estimated coeficient**
(Intercept)	-3.216
Preoperative WHO performance score 'yes'	0.523
Preadmission stoma counseling 'yes'	0.518
Additional major procedures 'yes'	0.303
If open surgery or converted 'no'	-0.152
If colloids 'yes'	0.0885
If blood 'yes'	0.632
Resection site drainage 'yes'	0.292
Total IV volume of fluids intraoperatively	0.000213
Length of operation minutes	0.001

Table 7.5: Coeficients of logistic regression with marginal important variables predicting severe complications in Scenario 2 (right after surgery).

AIC forward variables, Scenario 2, Severe complications	
Variable	**Estimated coeficient**
(Intercept)	-2.977
Preoperative WHO performance score 'yes'	0.751
Additional major procedures 'yes'	0.595
Total IV volume of fluids intraoperatively	0.000369
ASA physical status class '3 or 4'	0.0653
Infusion of vasoactive drugs 'yes'	-0.894
Any nonsurgical preoperative treatment 'yes'	-0.208
Resectionsite drainage 'yes'	0.803
Intraoperative thrombosis prophylaxis 'yes'	-0.769
If spinal anaesthesia 'yes'	0.616
Oral bowel preparation 'yes'	0.353

Table 7.6: Coeficients of logistic regression with variables selected with AIC forward predicting severe complications in Scenario 2 (right after surgery).

Conclusion Scenario 2

We see that in Scenario 2 both the intraoperative variables and variables indicating baseline health are important.

Furthermore, we see that 2 of the 3 models using *BMI*, assigned the interval (21.5-24.9) as the lowest risk group (models AIC forward with any complication and literature variables with severe complications). In the model with any complications using the literature variables the lower BMI, (21.4 or below) was assigned the lowest risk of complications.

Since the outcome is different when predicting any complication or severe complications it is no surprise to see changes in the estimated coefficients for the same variables. However, we consider it to be clinically relevant to state which estimated coefficients changed their signs, for this could be an indication of a variable decreasing the risk of any complications but increasing the risk of severe complications or vice versa. These variables are *Diabetes mellitus, Severe pulmonary disease, Severe heart disease, Alcohol usage, Preoperative oral carbohydrate treatment, Was anaemia found, If colloids* and *Core body temperature at end of operation*. Here, the variable *Preoperative oral carbohydrate treatment* makes the greatest change from -0.203 for any complications to 1.74 for severe complications in the literature based model.

It is interesting that the estimated coefficient for the variable *If blood* is larger when used for modelling severe complications than when modelling any complications in the marginal and literature-based variable selection 0.033 versus 0.632 in the models with marginal important variables, 0.203 versus 0.905 with the literature based variables. We also see that the variable *If blood* was not selected for the AIC forward selection with severe complications. According to the logisitic regression modelling it seems that blood loss leads to severe complications more than to non-severe complications.

Comparing Scenario 1 and 2
We saw that in both Scenarios 1 and 2, that the variable *Preoperative WHO performance score* was of great influence.

We do not see any changing signs of the estimated coefficients between Scenarios 1 and 2 when modelling any complication or when modelling a severe complication for both the variable selection from marginal importance and for AIC forward variable selection. However, in the AIC forward variable selection, no variables were selected in both Scenarios.

For the literature variables, we see some changing signs in the estimated coefficients between Scenarios 1 and 2 when using any complication as the outcome. These are the variables *Age, Preoperative carbohydrate treatment* and *Previous surgery to the same abdominal region*. We expect these changes to derive from the introduction of correlated variables. We know that *Age* and *Oral bowel preparation* are present differently over the main procedures. The negative sign for the variable *Age* in Scenario 1 may be due to the inclusion of preoperative WHO. We have discussed that it is not straightforward whether it is age or reduced mobility that increases the risk of complications. It is interesting to see how the different models use these variables.

For the literature based variables with severe outcomes, we also see some variables with changing signs from Scenario 1 to Scenario 2. These are *ASA physical status class, Smoker, Age, BMI, Severe pulmonary disease* and *Previous surgery to same abdominal area*. We consider this to be an unexpected result. We do not see why these estimated coefficients change when looking at the meaning of the variables. We expect it to be the result of the introduction of correlated variables.

Scenario 3 with any complications
We see the largest influence for predicting any complication in Scenario 3 with the literature-based variables is *Main procedure* equaling to 'stoma procedure' (see Table I.7). The estimated coefficients that we would have expected to have different signs when predicting any complications in this model are: *Smoker, Diabetes mellitus, Severe pulmonary disease, Alcohol usage, Preoperative oral carbohydrate treatment, Was anaemia found, Precious surgery to same abdominal region, If colloids, If blood, Mobilised on pod 1 or 2, IV volume postoperational* equalling 'after POD 2', *Time to tolerating solid food nights, Time to pain control with oral analgesics nights* and *High cancer stage*.

We see in Table I.8 that the variable *Preoperative WHO performance score* is the biggest influence in Scenario 3 with the marginal important variables for predicting any complications. We see that opioid usage on the first POD seems to increase the risk of any complication. Surprisingly, we see that the *Duration of IV fluid infusion nights* has a higher coefficient for 'removed on POD 1' than for 'removed on or after POD 2'. This implies it best would be to remove the IV on the day of operation and after that on or after the second day of operation. We consider this result to be unexpected. We see that under the model more *Total IV volume of fluids intraoperatively* would lead to fewer complications.

We see the largest influence for predicting any complications with the AIC forward selection in Scenario 3 is *If colloids* (see Table I.9). Surprisingly the presence of *If colloids* seems to decrease risk in this model. Other variables which we expected to have different signs in the model predicting any complications are: *Preadmission stoma counselling* and *Time to pain control with oral analgesics nights*. We would have expected that the estimated coefficient for 'after POD 2' would be higher than 'on POD 2' for the variable *Mobilised on POD 1 or 2*. However under the model mobilisation after POD 2 or before POD 2 is similar. In the model we see a negative sign for the variable *IV volume postoperational* and a positive sign for *Total IV volume of fluids intraoperatively*. This implies that fewer fluids should be given during surgery and more after surgery.

Scenario 3 severe complications

Literature variables, Scenario 3, Severe complications	
Variable	**Estimated coeficient**
(Intercept)	-12.487
ASA physical status class '3 or 4'	0.399
Smoker 'yes'	0.297
Severe heart disease 'yes'	-0.266
Preoperative WHO performance score 'yes'	-0.293
Age	-0.00749
BMI '(21.5-24.9)'	-1.852
BMI '(25-29.9)'	-1.799
BMI '(30 or more)'	-0.641
If predisease 'yes'	0.913
Diabetes mellitus 'yes'	1.389
Severe pulmonary disease 'yes'	-0.0866
Gender 'female'	0.291
Alcohol usage 'yes'	0.121
Preoperative nutritional treatment 'yes'	-0.0931
Preoperative oral carbohydrate treatment 'yes'	1.892
Oral bowel preparation 'yes'	-1.503
Was aneamia found 'yes aneamia'	0.395
Any nonsurgical preoperative treatment 'yes'	1.017
Previous surgery to same abdominal region 'yes'	-0.278
If colloids 'yes'	-1.542
Resection site drainage 'yes'	0.365
Depth of anaesthesia monitored 'yes'	0.108
Infusion of vasoactive drugs 'yes'	-0.581
If lidocaine 'yes'	-2.202
Final diagnosis 'cancer'	-0.590
Final diagnosis 'no cancer or functional disorder'	-0.449
Stomal procedure 'yes'	-2.035
Length of operation minutes	-0.00490
Additional major procedures 'yes'	0.295
If open surgery or converted 'no'	1.403
Main procedure name 'sigmoid resection'	1.319
Main procedure name 'stoma procedure	3.18
Main procedure name 'anterior resection of the rectum'	2.191
Main procedure name 'uncommon procedure, no ileocaecal'	1.094
If blood 'yes'	0.0869
Total IV volume of fluids intraoperatively	0.000703
Corebody temperature at end of operation	0.125
Mobilised on POD 1 or 2 'on POD 2'	0.597
Mobilised on POD 1 or 2 'after POD 2'	2.109
Time to tolerating solid food nights '1 or more nights'	1.24
More than one night with urinary drain 'yes'	0.985
Time to pain control with oral analgesics nights '2 or more nights'	-2.04
IV volume postoperational	-0.00149
Opioid use on postoperative day 1 'yes'	2.89
Max pain VAS in 3 days	-0.0750
CD at least 2 'yes'	-1.296
Length of stay nights in hospital after primary operation	0.0299
High cancer stage'yes'	-0.119

Table 7.7: Coeficients of logistic regression with literature-based variables predicting severe complications in Scenario 3 (right after primary stay).

We see the biggest influence for predicting severe complications in Scenario 3 with the literature-based variables is *Main procedure* equaling to 'stoma procedure' (see Table 7.7). The factors that we would have expected to have different signs when predicting severe complications are: *Severe heart disease, Preoperative WHO score, Age, Severe pulmonary disease, Gender, Alcohol usage, Preoperative oral carbohydrate treatment, Oral preparation, Previous surgery to abdominal region, If colloids, Depth of anaesthesia monitored, Length of operation minutes, Open or converted, Core body temperature at end of operation, Time to pain control with oral analgesics, IV volume postoperational, Max pain VAS in 3 days, CD at least 2* and *High cancer stage*. We see many variables where the signs are not as we expected in this model. Compared with other models, the model from Table 7.7 has many large estimated coefficients.

Marginal variables, Scenario 3, Severe complications	
Variable	**Estimated coeficient**
(Intercept)	-6.33
Preadmission stoma counseling 'yes'	0.895
Opioid use on postoperative day 1 'yes'	1.748
Postoperative use of NSAIDS 'yes'	0.975
PONV observed in 3 days 'yes'	1.014
Time to tolerating solid food nights '1 or more nights'	0.944

Table 7.8: Coeficients of logistic regression with marginal important variables predicting severe complications in Scenario 3 (right after primary stay).

For predicting severe complications with marginal important variables in Scenario 3, *Opiod use on postoperative day 1* seems to be the greatest influence under this model (see Table 7.8). In Table 7.8, we see that *Opioid use on postoperative day 1* seems to increase the risk of any complications. We found that *Preadmission stoma counselling* is a risk factor for severe complications under the model from Table 7.8.

AIC forward variables, Scenario 3, Severe complications	
Variable	**Estimated coeficient**
(Intercept)	-6.145
Preoperative WHO performance score '1'	-0.217
Additional major procedures 'yes'	0.622
PONV observed In 3 days 'yes'	1.279
Opioid use on postoperative day 1 'yes'	1.528
If colloids 'yes'	-1.293
Resection site drainage 'yes'	0.821
Preadmission stoma counseling 'yes'	1.051
Bowel anastomosis 'yes'	0.395

Table 7.9: Coeficients of logistic regression with variables selected with AIC forward predicting severe complications in Scenario 3 (after primary stay).

We see that in the model with variables selected with AIC forward predicting severe complications in Scenario 3, the biggest influence is *Opioid use on postoperative day 1* (see Table 7.9). We would have expected the estimated coefficient for the variable *Preoperative WHO performance score*, *If colloids* and *Preadmission stoma counselling* to have a different sign.

Conclusion Scenario 3

We see many variables where the signs are not as we expect. With 23 of the 49 estimated coefficients we found a different sign when predicting severe complications than when predicting non-severe complications for Scenario 3 with the literature-based selection. For the models in Scenario 3 using the AIC forward selection method or the marginal important variables, we see no signs of coefficients changing between predicting any complications and predicting severe complications. However, not many variables are selected with these methods for predicting both severe complications and any complications.

The greatest postoperative influences when predicting any or severe complications are given by the variables *PONV observed in 3 days* and *Opioid use on postoperative day 1*.

7.2. In-sample validation

In this section, we study the in-sample validation. In-sample validation shows up to what degree the data used to fit the model, supports the model. Definitions of the validation measures are given in Chapter 2. In this section, we explain their interpretation again briefly.

The measures misclassification rate, balanced accuracy, MCC and TS take into account the rarity of the outcome. False omission rate and auc ROC do not take the rarity of the outcome into account. Neither the number of false negatives nor the number of false positives take into account the rarity of the outcome. However, since these are whole numbers one can compare it with the number of observations and outcomes directly. The number of outcomes and the number of observations can be found in the captions of the tables using these measures. These measures explicitly depend on the number of observations used to calculate them. The log-likelihood and the Brier score are not based on a classification of the model but on the assigned probability under the model.

MCC is a correlation coefficient of the predicted outcome with the true outcome, therefore we want it close to 1. TS is the amount of correct positive predictions over the amount of correct positive predictions plus all wrong predictions, so a TS close to 1 is desired. The balanced accuracy is the average of the true positive rate and true negative rate. The false omission rate is the percentage of mistakes in the positive predictions, in other words, the false positives over the total of positive predictions. In an accurate prediction model the MCC, TS, auc ROC, log-likelihood, and balanced accuracy are large. In an accurate prediction model the Brier score, misclassification rate, BIC, false negatives, false positives and false omission rate are small.

Our in-sample results are shown in the Tables 7.10-7.13.

In-sample validation Scenario 1 and 2

Any complications	Scenario 1			Scenario 2		
	literature	marginal	AIC forward	literature	marginal	AIC forward
Brier score	0.196	0.202	0.203	0.182	0.191	0.205
log-likelihood	-608.286	-592.512	-589.451	-647.677	-618.846	-590.699
BIC	1340.704	1241.399	1241.399	1452.422	1278.638	1266.225
auc ROC	0.601	0.589	0.589	0.646	0.609	0.552
false negatives	118	120	120	98	111	136
false positives	19	23	23	31	28	14
misclassification rate	0.276	0.288	0.288	0.26	0.28	0.302
false omission rate	0.271	0.276	0.276	0.243	0.264	0.296
balanced accuracy	0.601	0.589	0.589	0.646	0.609	0.552
MCC	0.288	0.25	0.25	0.35	0.283	0.183
TS	0.23	0.214	0.214	0.321	0.257	0.133

Table 7.10: In-sample validation logistic regression on the whole dataset (496 observations with 159 complications) for Scenario 1 (presurgery prediction) and Scenario 2 (right after surgery prediction). The coloured numbers indicate the best score in the scenario, The red colour indicates the best over both scenarios.

In Table 7.10 we see that for Scenario 1 with any complications, the literature-based variables seem to perform better on most measures than the marginal selection, except for the log-likelihood and BIC. The variable selection method based on the literature, selected 20 variables, whereas the selection based on marginal importance, selected 4 variables. When taking this into account the better performance of the literature-based model over the marginal selection is not that impressive. The model with the marginal-based variable selection and the AIC forward selection perform very similarly for Scenario 1. Only a small difference is visible when looking at the Brier Score and log-likelihood in Table 7.10.

In Scenario 2 we again see that the literature selection led to a better in-sample performance for any complications. In this scenario, the marginal and AIC forward selections differ more than in Scenario 1. In Scenario 2 we see the least false positives. These are the patients which would not develop a complication but are predicted to get a complication by the model. These are clinically less relevant than false negatives since false negatives under the model would lead to the absence of the necessary extra care, whereas false positives would lead to unnecessary extra care. We see that the addition of the intraoperative variables did not lead to a more informative model for the AIC forward variable selection method. The addition of more variables did lead to an improvement for the literature-based variables and the marginal important variables.

Comparing the selection from the literature in Scenario 2 and Scenario 1, we see some improvement in Scenario 2. This is not surprising, since more variables are included. Except for the loglikelihood for the AIC forward selection, the fit is worse in Scenario 1 than in Scenario 2 for the different selection methods. Interestingly, the loglikelihood-based scores are better in Scenario 1. This is due to the rare outcome. A model underestimating a rare outcome can lead to a better likelihood but contains no relevant information.

Severe complications	**Scenario 1**			**Scenario 2**		
	literature	**marginal**	**AIC forward**	**literature**	**marginal**	**AIC forward**
Brier score	0.086	0.089	0.089	0.074	0.082	0.081
log-likelihood	-1219.217	-1179.849	-1168.131	-1414.87	-1250.48	-1270.93
BIC	2562.566	2457.054	2475.673	2674.284	2500.5	2506.706
auc ROC	0.5	0.5	0.5	0.567	0.528	0.548
false negatives	50	50	50	43	47	45
false positives	0	0	0	3	2	2
misclassification rate	0.101	0.101	0.101	0.093	0.099	0.095
false omission rate	0.101	0.101	0.101	0.088	0.096	0.092
balanced accuracy	0.5	0.5	0.5	0.567	0.528	0.548
MCC	X	X	X	0.285	0.167	0.244
TS	0	0	0	0.132	0.058	0.096

Table 7.11: In-sample validation logistic regression on the whole dataset (496 observations with 50 severe complications) for Scenario 1 (presurgery prediction) and Scenario 2 (right after surgery prediction). The X indicates that there were no correct positive predictions either. X denotes it is not defined (dividing by zero). The coloured numbers indicate the best score in the scenario, The red colour indicates the best over both scenarios.

In Table 7.11, we see the in-sample validation from the models predicting severe complications in Scenarios 1 and 2. We see a slight improvement in predicting in Scenario 2 compared to Scenario 1.

We see that for Scenario 1, predicting severe complications the best performing model in-sample is the literature based model. We conclude this from the Brier score (Table 7.11). We see all models in Scenario 1 with severe complications do not predict any observation to developing a severe complication. The model underestimates the outcome greatly.

Similarly, in Scenario 2 with the severe complications, the literature-based variables selection performs better in-sample. Here we can see this in most measures of fit. Note that this model only predicted 4 patients with complications correctly, whereas the other two models in Scenario 2 with severe complications predicted 2 or 3 complications occurring correctly.

In-sample validation Scenario 3

Any complications	Scenario 3		
	literature	marginal	AIC forward
Brier score	0.115	0.121	0.125
log-likelihood	-1077.419	-1006.496	-988.246
BIC	2458.961	2216.905	2229.318
auc ROC	0.53	0.517	0.503
false negatives	72	74	76
false positives	2	2	3
misclassification rate	0.149	0.153	0.159
false omission rate	0.147	0.151	0.154
balanced accuracy	0.53	0.517	0.503
MCC	0.185	0.124	0.024
TS	0.063	0.038	0.013

Table 7.12: In-sample validation logistic regression on the whole dataset. (496 observations with 77 complications) for Scenario 3 (post-primary stay prediction).

In Table 7.12, we see that in Scenario 3 when predicting any complications the literature performs the best in-sample. We are not satisfied with the in-sample fit in Scenario 3, since the number of false negatives is close to the number of severe complications. This means that only a few complications are predicted correctly under the model.

Severe complications	Scenario 3		
	literature	marginal	AIC forward
Brier score	0.031	0.037	0.038
log-likelihood	-2424.532	-1879.127	-1861.375
BIC	5153.187	4886.304	4904.924
auc ROC	0.57	0.5	0.5
false negatives	18	21	21
false positives	1	0	0
misclassification rate	0.038	0.042	0.042
false omission rate	0.037	0.042	0.042
balanced accuracy	0.57	0.5	0.5
MCC	0.317	X	X
TS	0.136	0	0

Table 7.13: In-sample validation logistic regression on the whole dataset (496 observations with 21 severe complications) for Scenario 3 (post-primary stay prediction). X denotes it is not defined (dividing by zero).

In Table 7.13, we see that in Scenario 3 the literature performs better for the severe complications in-sample. Nonetheless, we see a very bad fit in all models predicting severe complications in Scenario 3. The marginal and the AIC forward variable selection both lead to no severe complications being predicted. Whereas the literature-based model predicted correctly one complication and predicted one complication falsely.

Conclusion insample validation

We have seen that the literature variable selection leads to a better in-sample fit for Scenarios 1, 2 and 3. This is not surprising, with the number of variables in the literature-based selection. We saw some models not predicting any complications at all. We see that the variable selection methods struggle with rare outcomes. Any complications in Scenario 2 appear to be relatively the easiest to predict.

We see in Scenario 2 with any complications that the variable selection from the literature performs well in-sample, but we expect this will change in the cross-validation, since we saw many estimated coefficients with signs that are inconsistent with what we saw in the literature. This will be investigated next.

7.3. Cross-validation

In this section, we will study the results of the cross-validation. We keep the variable selections fixed in this section. The Tables 7.14-7.17 summarise the outcome of the cross-validation. As explained in Chapter 2, we ensure that every observation is used 20 times in a validation set. We make 20 times a 10-fold on the dataset. For each 10-fold we fit 10 models, all these models exclude one fold, the validation set, for external validation. A number of 200 models are fitted in the cross-validation. In this way each data point was used in the external validation exactly 20 times. We picked this method for cross-validation since we have rare outcomes and explanatory variables. We calculate the average and standard deviation of our out-of-sample validation measures. In the Tables 7.14 - 7.17, we see the standard deviation in brackets.

Cross-validation Scenario 1 and 2, fixed variable selection

Any complications	Scenario 1			Scenario 2		
	literature	marginal	AIC forward	literature	marginal	AIC forward
Brier score	0.215 (0.031)	0.206 (0.028)	0.207 (0.027)	0.218 (0.035)	0.2 (0.031)	0.212 (0.028)
log-likelihood	-61.176 (3.521)	-59.316 (2.58)	-58.991 (2.381)	-65.68 (5.351)	-62.006 (3.525)	-59.163 (2.768)
BIC	200.431 (7.099)	134.247 (5.174)	133.599 (4.773)	279.711 (10.786)	163.052 (7.075)	149.557 (5.557)
auc ROC	0.58 (0.054)	0.586 (0.052)	0.586 (0.053)	0.591 (0.066)	0.596 (0.057)	0.54 (0.045)
false negatives	12.15 (3.104)	12.145 (3.035)	12.15 (3.002)	11.005 (3.217)	11.37 (3.036)	13.9 (3.202)
false positives	2.67 (1.62)	2.28 (1.446)	2.29 (1.434)	4.47 (2.164)	3.26 (1.717)	1.635 (1.224)
misclassification rate	0.299 (0.065)	0.291 (0.064)	0.291 (0.063)	0.312 (0.07)	0.295 (0.063)	0.313 (0.064)
false omission rate	0.281 (0.069)	0.279 (0.068)	0.279 (0.067)	0.273 (0.075)	0.271 (0.07)	0.302 (0.067)
balanced accuracy	0.58 (0.054)	0.586 (0.052)	0.586 (0.053)	0.591 (0.066)	0.596 (0.057)	0.54 (0.045)
MCC	0.223 (0.139)	0.243 (0.137)	0.242 (0.137)	0.215 (0.153)	0.245 (0.14)	0.143 (0.149)
TS	0.204 (0.086)	0.208 (0.085)	0.208 (0.085)	0.244 (0.094)	0.238 (0.087)	0.116 (0.074)

Table 7.14: Out of sample performance from cross-validation of logistic regression models predicting any complications for Scenario 1 (prediction preoperational) and Scenario 2 (prediction right after surgery) (496 observations with 159 complications). The coloured numbers indicate the best score in the scenario, The red colour indicates the best over both scenarios.

As one expects we see a worse performance in the cross-validation than we have seen in-sample. The log-likelihood, the false negatives and the false positives are a little higher than a tenth of the in-sample results. This is what one would expect since the measures from Table 7.14 are calculated out-of-sample with a sample size of one-tenth of the observations. We see that the literature-based variable selection does not lead to the better performing model. We conclude that the model with the literature variable selection overfits the data. In Scenario 1, with any complications, we see that the marginal important variables and the variables selected with the AIC forward method perform similarly. We see that the AIC forward has better likelihood-based scores.

In Scenario 2 when predicting any complications we see that the marginal important variables lead to the fewest misclassifications. In Scenario 2 with any complications, the literature-based model seems to predict the most false positives, and not underestimate as much as the model with the AIC forward selection.

In Table 7.14, we see the fewest misclassifications in Scenario 1 with the variable selected with marginal importance or variables selected with AIC forward. We see that the addition of the intraoperative variables did not lead to a more predictive model for the AIC forward variable selection method. This was also the case in the in-sample validation. In Table 7.14 the best balanced accuracy for predicting any complications in Scenario 1 or 2 is for the model with marginal important variables in Scenario 2. We see that the performance of the AIC forward did not improve in Scenario 2 compared to Scenario 1, whereas the other two methods did improve.

In Table 7.15, we see the results of the cross-validation when predicting severe complications in Scenarios 1 and 2. We denote a missing validation measure with an X. This is for the MCC in Scenario 1 with marginal selection. This is a result of two values in the contingency table of the outcome and the prediction equalling zero (except if these two values are diagonal, so in case of perfect prediction or perfectly wrong predictions). We see that the model predicted zero severe complications and the MCC could not be calculated. We consider this model to be unsuitable for this analysis since it does not predict any outcomes at all. This model underestimates severely. The TS equal to zero happens when there are no correct true predictions. The model in Scenario 1 with the literature-based variables

Severe complications	Scenario 1			Scenario 2		
	literature	marginal	AIC forward	literature	marginal	AIC forward
Brier score	0.096 (0.034)	0.09 (0.033)	0.092 (0.034)	0.098 (0.036)	0.089 (0.033)	0.088 (0.034)
log-likelihood	-123.381 (6.063)	-118.159 (3.886)	-117.286 (3.892)	-144.982 (11.254)	-125.642 (5.755)	-127.874 (6.502)
BIC	324.84 (12.168)	248.03 (7.779)	257.995 (7.802)	438.313 (22.598)	290.323 (11.541)	298.691 (13.039)
auc ROC	0.499 (0.003)	0.5 (0)	0.5 (0)	0.514 (0.049)	0.513 (0.038)	0.535 (0.068)
false negatives	5 (2.074)	5 (2.074)	5 (2.074)	4.75 (2.049)	4.83 (2.04)	4.635 (2.033)
false positives	0.07 (0.275)	0 (0)	0 (0)	0.97 (0.992)	0.25 (0.519)	0.275 (0.567)
misclassification rate	0.102 (0.041)	0.101 (0.042)	0.101 (0.042)	0.115 (0.046)	0.102 (0.042)	0.099 (0.042)
false omission rate	0.101 (0.042)	0.101 (0.042)	0.101 (0.042)	0.098 (0.042)	0.098 (0.041)	0.095 (0.041)
balanced accuracy	0.499 (0.003)	0.5 (0)	0.5 (0)	0.514 (0.049)	0.513 (0.038)	0.535 (0.068)
MCC	-0.041 (0.015)	X	X	0.065 (0.184)	0.158 (0.218)	0.26 (0.245)
TS	0 (0)	0 (0)	0 (0)	0.041 (0.079)	0.03 (0.07)	0.073 (0.129)

Table 7.15: Out of sample performance from cross-validation of logistic regression models predicting severe complications for Scenario 1 (presurgery prediction) and Scenario 2 (right after surgery prediction). X denotes one time it was not defined (dividing by zero). (496 observations with 50 severe complications) The coloured numbers indicate the best score in the scenario, The red colour indicates the best over both scenarios.

sometimes predicted an outcome falsely. In Scenario 1 with severe complications, we see no model accurately predicting severe complications.

In Scenario 2, the AIC forward selection seems to perform best. Note that the marginal selection leads to a better BIC and log-likelihood and false positives than the AIC forward. However, these can be explained through the model's underestimation. We see improvement in Scenario 2 compared to Scenario 1 for all variable selection methods, so Scenario 2, is overall better for predicting complications.

Cross-validation Scenario 3, fixed variable selection

Any complications	Scenario 3		
	literature	marginal	AIC forward
Brier score	0.155 (0.038)	0.128 (0.031)	0.132 (0.032)
BIC	413.241 (15.092)	240.904 (9.443)	245.458 (8.985)
log-likelihood	-110.974 (7.495)	-100.932 (4.704)	-99.305 (4.478)
auc ROC	0.502 (0.042)	0.514 (0.031)	0.502 (0.021)
false negatives	7.31 (2.213)	7.445 (2.277)	7.6 (2.26)
false positives	1.905 (1.499)	0.255 (0.501)	0.33 (0.627)
misclassification rate	0.186 (0.049)	0.155 (0.047)	0.16 (0.046)
false omission rate	0.154 (0.046)	0.152 (0.046)	0.155 (0.046)
balanced accuracy	0.502 (0.042)	0.514 (0.031)	0.502 (0.021)
MCC	0.012 (0.144)	0.152 (0.193)	0.038 (0.168)
TS	0.04 (0.061)	0.032 (0.059)	0.012 (0.038)

Table 7.16: Out of sample performance from cross-validation of logistic regression models predicting any complications for Scenario 3 (post-primary stay prediction). The coloured number indicates the best score. (496 observations with 77 complications)

In Table 7.16 we see that the marginal selection leads to the best performance for predicting any complications in Scenario 3. We see that the literature-based selection more commonly predicts wrongly (misclassification rate) but is also less likely to miss a complication (false negatives). Where the AIC forward seems to score better on the log-likelihood, it scores worse on the Brier score than the marginal important variables. This indicates that overall the predictions under the model are more acurate but for a few observations way worse than, the marginal important variables.

We see that the literature based variable selection performs best in the cross-validation for the severe complications in Scenario 3 (Table 7.17). The marginal selection does not predict any outcomes and the AIC forward selection sometimes leads to a severe complication being predicted wrongly. We see that the literature variables are more informative than the marginally important variables or the AIC forward. However, there are fewer errors with the marginal selection.

Severe complications	Scenario 3		
	literature	marginal	AIC forward
Brier score	0.06 (0.036)	0.039 (0.025)	0.04 (0.026)
log-likelihood	-341.537 (187.245)	-206.781 (62.479)	-210.338 (67.5)
auc ROC	0.526 (0.099)	0.5 (0)	0.5 (0.001)
BIC	874.366 (374.542)	436.985 (124.964)	455.812 (135.007)
false negatives	1.915 (1.392)	2.1 (1.456)	2.1 (1.456)
false positives	1.43 (1.321)	0 (0)	0.015 (0.122)
misclassification rate	0.067 (0.042)	0.042 (0.029)	0.043 (0.029)
false omission rate	0.04 (0.029)	0.042 (0.029)	0.042 (0.029)
balanced accuracy	0.526 (0.099)	0.5 (0)	0.5 (0.001)
MCC	0.079 (0.235)	X	-0.025 (0.006)
TS	0.055 (0.139)	0 (0)	0 (0)

Table 7.17: Out of sample performance from cross-validation of logistic regression models predicting severe complications for Scenario 3 (post-primary stay prediction). X denotes one time it was not defined (dividing by zero). The coloured number indicates the best score. (496 observations with 21 severe complications)

Cross-validation with repeated variable selection process Scenario 1 and 2

Since the variable selection based on marginal importance and the AIC forward variable selection are tailored to the entire dataset, we repeat the cross-validation but now without the variable selections being fixed. Now the variable selection is tailored to the data the model is fitted with, and therefore the model's performance is expected to improve.

Any complication	Scenario 1		Scenario 2	
	Marginal	AIC forward	Marginal	AIC forward
Brier score	0.207 (0.027)	0.215 (0.027)	0.201 (0.03)	0.213 (0.032)
log-likelihood	-59.305 (2.629)	-59.911 (2.949)	-62.083 (3.446)	-63.502 (3.753)
BIC	134.87 (6.41)	140.494 (8.155)	163.597 (8.151)	163.683 (10.635)
auc ROC	0.584 (0.058)	0.581 (0.059)	0.594 (0.066)	0.577 (0.066)
false negatives	12.175 (2.932)	12.305 (2.927)	11.33 (2.906)	11.57 (2.872)
false positives	2.28 (1.375)	2.17 (1.404)	3.33 (1.862)	4.025 (2.16)
misclassification rate	0.291 (0.061)	0.292 (0.06)	0.296 (0.061)	0.314 (0.066)
false omission rate	0.279 (0.065)	0.28 (0.064)	0.271 (0.065)	0.28 (0.067)
balanced accuracy	0.584 (0.058)	0.581 (0.059)	0.594 (0.066)	0.577 (0.066)
MCC	0.235 (0.148)	0.23 (0.154)	0.238 (0.157)	0.193 (0.161)
TS	0.205 (0.095)	0.199 (0.1)	0.236 (0.105)	0.219 (0.095)

Table 7.18: Out of sample performance from cross-validation of the logistic regression models predicting any complication for Scenario 1 (presurgery prediction) and Scenario 2 (right after surgery prediction). Here the variable selection was done on the corresponding validation data set. The coloured numbers indicate the best score in the scenario, The red colour indicates the best over both scenarios. (496 observations with 159 complications)

We see that the performance of the marginal selected variables with tailored variable selection (Table 7.18) is similar in the cross-validation with fixed variables (Table 7.14). This implies that the variable selection based on marginal importance does not change much in the cross-validation. The performance of the model improved a little since the models are being fitted only on the relationships that are present in the data used to fit the model. In other words, there are fewer instances where the model fitted a variable that was only seen as important in the validation set. However the AIC forward selection with any complications performed worse in Scenario 1 but a bit better in Scenario 2 in the adjusted variable cross-validation compared to the fixed variable selection cross-validation.

In Table 7.19 we see that the results of the cross-validation with adjusted variables selection are similar when the variable selection was fixed for Scenario 1. In Scenario 2 however, we see a small loss in performance compared to the fixed variable selection. Sometimes no variables were selected at all. This resulted in the model returning the frequency of the outcome as the estimated probability for all observations.

Severe complication	Scenario 1		Scenario 2	
	Marginal	AIC forward	Marginal	AIC forward
Brier score	0.094 (0.034)	0.093 (0.033)	0.092 (0.033)	0.094 (0.034)
log-likelihood	-117.564 (4.584)	-120.018 (5.279)	-126.543 (6.78)	-131.877 (8.581)
BIC	246.213 (11.829)	257.113 (13.432)	292.301 (19.124)	297.288 (23.799)
auc ROC	0.5 (0)	0.5 (0)	0.515 (0.041)	0.524 (0.065)
false negatives	5 (2.069)	5 (2.069)	4.83 (2.067)	4.755 (2.094)
false positives	0 (0)	0 (0)	0.26 (0.533)	0.495 (0.68)
misclassification rate	0.101 (0.042)	0.101 (0.042)	0.103 (0.042)	0.106 (0.044)
false omission rate	0.101 (0.042)	0.101 (0.042)	0.098 (0.042)	0.097 (0.043)
balanced accuracy	0.5 (0)	0.5 (0)	0.515 (0.041)	0.524 (0.065)
MCC	X	X	0.164 (0.23)	0.145 (0.241)
TS	0 (0)	0 (0)	0.033 (0.079)	0.052 (0.119)

Table 7.19: Out of sample performance from cross-validation of logistic regression models predicting severe complications for Scenario 1 (presurgery prediction) and Scenario 2 (right after surgery prediction). X denotes one time it was not defined (dividing by zero). The coloured numbers indicate the best score in the scenario, The red colour indicates the best over both scenarios. (496 observations with 50 severe complications)

Cross-validation with repeated variable selection process Scenario 3

In Table 7.20 we see that the cross-validation with adjusted variables selection performs overall worse than the cross-validation with fixed variables in Scenario 3. We see in Table 7.20 that 7.5 patients who develop complications are not predicted to develop a complication in the cross-validation for the model with the marginal variable selection predicting any complication after primary stay. At first, this seems to be low but for an average validation set in the cross-validation 7.7 patients actually get a complication. We see that this model underestimates the severe complications.

any complication	Scenario 3	
	marginal	AIC forward
Brier score	0.132 (0.034)	0.139 (0.034)
log-likelihood	-100.684 (4.617)	-103.822 (6.655)
BIC	234.922 (12.485)	242.683 (17.565)
auc ROC	0.512 (0.034)	0.503 (0.032)
false negatives	7.505 (2.443)	7.53 (2.402)
false positives	0.26 (0.56)	0.725 (0.896)
misclassification rate	0.157 (0.05)	0.166 (0.052)
false omission rate	0.153 (0.049)	0.155 (0.049)
balanced accuracy	0.512 (0.034)	0.503 (0.032)
MCC	0.143 (0.213)	0.024 (0.168)
TS	0.028 (0.064)	0.022 (0.055)

Table 7.20: Out of sample performance from cross-validation of logistic regression models predicting any complication for Scenario 3 (post-primary stay prediction). The coloured number indicates the best score. (496 observations with 77 complications)

We see a slightly more informative model being fitted for the adjusted AIC forward model as compared to the fixed variable selection in Scenario 3 with the cross-validation with adjusted variable selection (Table 7.21). We mean that the misclassification is a little worse, but we see that the balanced accuracy improved compared to the marginal variable selection. For the marginal important method, we see very similar performance in Table 7.21 as in Table 7.17.

Severe complication	Scenario 3	
	marginal	**AIC forward**
Brier score	0.043 (0.027)	0.049 (0.027)
log-likelihood	-205.166 (67.63)	-277.646 (171.731)
BIC	433.347 (137.009)	601.302 (352.51)
auc ROC	0.5 (0.001)	0.504 (0.035)
false negatives	2.1 (1.456)	2.035 (1.358)
false positives	0.005 (0.071)	0.5 (0.88)
misclassification rate	0.042 (0.029)	0.051 (0.033)
false omission rate	0.042 (0.029)	0.042 (0.028)
balanced accuracy	0.5 (0.001)	0.504 (0.035)
MCC	X	0.046 (0.182)
TS	0 (0)	0.015 (0.061)

Table 7.21: Out of sample performance from cross-validation of the logistic regression models predicting severe complications for Scenario 3 (post-primary stay prediction). X denoted one time it was not defined (dividing by zero). The coloured number indicates the best score. (496 observations with 21 severe complications)

Conculsion validation

We have seen difficulties in predicting rare outcomes in all models. In Scenario 1 and 2, the models predicting severe complications are performing poorly. In Scenario 3, the models predicting any complications or severe complications are performing poorly.

We have seen that in-sample the variables from the literature study performed well, however in the cross-validation, this was not the case. We conclude that this variable selection contains too many variables and leads to overfitting in the model.

In the cross-validation, we have seen that when comparing Scenarios 1 and 2, Scenario 2 performs generally better. Most of the time the BIC and the log-likelihood did not improve in Scenario 2 compared to Scenario 1 but the other measures did improve. We have seen that the AIC model can perform worse with the extra variables when predicting any complications, we suspect this shows that the AIC forward overfits the data.

We have seen that the cross-validation with adjusted variable selection performs better for any complications. However with the rarer outcomes sometimes no variables were selected and the performance of the fitted model became worse.

8

Exploring weighted logistic regression

In Chapter 6, we have seen that models commonly underestimate when the outcome is rare. We have seen that the strategy to predict any complications in order to find severe complications is not a good strategy. Therefore we will limit this study now to only use the severe complications as the model's outcome. At the same time, we have seen that the models struggle with rare outcomes. We will use a weighted logistic regression to predict severe complications in Scenario 1, Scenario 2 and Scenario 3, with the 3 different variable selections made on the entire dataset in Sections 6.1 - 6.3. We did not rerun the AIC forward algorithm but used the same selection of variables as the models were fit with in Chapter 6.

We use a weighted logistic regression with balanced weights (proportional to the frequency of outcome) and a fixed weight equal to 2. The weights are added to the observations with severe complications, and these weights make these observations more influential when fitting the model. We first study the change in the estimated coefficients when including the weights (See Tables 8.1 - 8.6 and Tables J.1 - J.3 in Appendix J). Subsequently we study the in-sample performance (Tables 8.7 - 8.10) and, thereafter, cross-validation (Tables 8.11 - 8.14).

8.1. Estimated coefficients in weighted logistic regresion

In Tables J.1-J.3 in Appendix J are the estimated coefficients of the weighted logistic regressions for the models using the literature based variable selection.

We see in Table J.1, that the estimated coefficients for the fixed-weighted model are close to the non-weighted model in Scenario 1 with the variables selected from the literature. Except for the intercept, we see that most estimated coefficients are further from zero, as the weights increase. For the following variables, the influence decreased in the model with the literature-based variables in Scenario 1 as the weights increased: *Severe heart disease*, *Preoperative WHO performance score*, *Age*, *BMI*, *If predisease*. *Alcohol usage*, *Preoperative nutritional treatment* and *Previous surgery to same abdominal region*. We see that in the balanced weighted model the sign of the variables *ASA physical status class* and *Smoker*, is in line with the literature, where this is not the case for the unweighted and the fixed-weighted models.

Marginal important variables, Scenario 1, severe complications			
Variable	**Estimated coeficient**		
	equal (or no) weights	**fixed weights**	**balanced weights**
weight	1	2	8.92
(Intercept)	-2.63	-1.94	-0.459
Preoperative WHO performance score '1'	0.975	0.987	1.02
Preadmission stoma counseling 'yes'	0.67	0.679	0.705

Table 8.1: Estimated coefficients of logistic regression (right), weighted logistic regression with fixed weight (middle) and balanced weighted logistic regression (right) on Margianl important variables predicting severe complications in Scenario 1 (before surgery).

We see that the intercept is the largest for the balanced weights in Scenario 1 with the marginal important variables (Table 8.1). The non-weighted model leads to the smallest intercept. The estimated coefficients, besides the intercept, increase as the weights increase.

AIC forward variables, Scenario 1, severe complications			
Variable	**Estimated coeficient**		
	equal (or no) weights	**fixed weights**	**balanced weights**
weight	1	2	8.92
(Intercept)	-2.39	-1.71	-0.263
Preoperative WHO performance score '1'	0.974	0.967	0.963
ASA physical status class '3 or 4'	0.0279	0.0518	0.116
Any nonsurgical preoperative treatment 'yes'	-0.088	-0.0613	0.0222
Oral bowel preparation 'yes'	0.338	0.35	0.382
Diabetes mellitus 'yes'	-0.479	-0.461	-0.417

Table 8.2: Estimated coefficients of logistic regression (right), weighted logistic regression with fixed weight (middle) and balanced weighted logistic regression (right) on variables selected with AIC forward predicting severe complications in Scenario 1 (before surgery).

For the model with AIC variables in Scenario 1 (Table 8.2), we do not see when the weights increase the estimated coefficients increase as well. We see that the estimated coefficients for the variables *Preoperative WHO performance score*, and *Diabetes mellitus* decrease as the weights increase. The estimated coefficient for *Any nonsurgical Preoperative Treatment* becomes larger as the weights increase.

In the models with the literature based variables for Scenario 2 (Table J.2), we see that, generally, coefficients move further from zero as the weight increases while the intercept decreases. Exceptions are *Severe pulmonary disease*, *Alcohol usage*, *Preoperative nutritional treatment*, *Preoperative oral carbohydrate treatment*, *Was aneamia Found*, *Previous surgery to same abdominal region*, *If colloids*, *Resection site drainage*, *Depth of anaesthesia monitored*, *Infusion of vasoactive drugs*, *Stomal procedure*, *Additional major procedures*, *If open surgery or converted*, *Main procedure name*, *If blood* and *Core body temperature at end of operation*. The variable, *If open surgery or converted*, is no longer a risk factor in the balanced weighted model and a larger *Length of operation minutes* leads to an increased risk in the weighted models for the literature based models in Scenario 2.

Marginal important variables, Scenario 2, severe complications			
Variable	**Estimated coeficient**		
	equal (or no) weights	**fixed weights**	**balanced weights**
weight	1	2	8.92
(Intercept)	-3.22	-2.59	-1.29
Preoperative WHO performance score '1'	0.523	0.449	0.325
Preadmission stoma counseling 'yes'	0.518	0.531	0.556
Additional major procedures 'yes'	0.303	0.285	0.289
If open surgery or converted 'no'	-0.152	-0.128	-0.0207
If colloids 'yes'	0.0885	0.116	0.174
If blood 'yes'	0.632	0.597	0.465
Resectionsite drainage 'yes'	0.292	0.252	0.237
Total IV volume of fluids intraoperatively	0.000213	0.000221	0.000262
Length of operation minutes	0.000757	0.00125	0.00209

Table 8.3: Estimated coefficients of logistic regression (right), weighted logistic regression with fixed weight (middle) and balanced weighted logistic regression (right) on Marginal important variables predicting severe complications in Scenario 2 (right after surgery).

Under the weighted models using the marginal important variables in Scenario 2 (Table 8.3), we see that the estimated coefficient for the variables *Preadmission stoma counselling, If colloids, Total IV volume of fluids intraopretaively* and *Length of operation* become more influential in the models as the weights increase.

AIC forward variables, Scenario 2, severe complications			
Variable	**Estimated coeficient**		
	equal (or no) weights	**fixed weights**	**balanced weights**
weight	1	2	8.92
(Intercept)	-2.98	-2.36	-1.04
Preoperative WHO performance score '1'	0.751	0.599	0.364
Additional major procedures 'yes'	0.595	0.62	0.75
Total IV volume of fluids intraoperatively	0.000369	0.000395	0.000453
ASA physical status class '3 or 4'	0.0653	0.101	0.162
Infusion of vasoactive drugs 'yes'	-0.894	-0.854	-0.81
Any nonsurgical preoperative treatment 'yes'	-0.208	-0.167	-0.11
Resection site drainage 'yes'	0.803	0.723	0.589
Intraoperative thrombosis prophylaxis 'yes'	-0.769	-0.78	-0.782
If spinal anaesthesia 'yes'	0.616	0.624	0.653
Oral bowel preparation 'yes'	0.353	0.415	0.506

Table 8.4: Estimated coefficients of logistic regression (right), weighted logistic regression with fixed weight (middle) and balanced weighted logistic regression (right) on variables selected with AIC forward predicting severe complications in Scenario 2 (right after surgery).

We see that the influence of the variables, *Additional major procedures, ASA physical status class, Intraoperative thrombosis prophylaxis, If spinal anaesthesia* and *Oral bowel preparation* increases as the weights increases for the weighted models using the variables from the AIC forward in Scenario 2 (Table 8.4). Furthermore, we see that the estimated coefficients for *Intraoperative thrombosis prophylaxis, If spinal anaesthesia* and *Infusion of vasoactive drugs* do not change much in the different models in Table 8.4.

For 26 of the 49 fitted coefficients in the models with the literature based variables in Scenario 3 (Table J.3), we see that the influence does not increase in the model as the weights increase. A number of 9 fitted coefficients change their sign for one of the models with literature based variables in Scenario 3.

Marginal important variables, Scenario 3, severe complications			
Variable	**Estimated coeficient**		
	equal (or no) weights	**fixed weights**	**balanced weights**
weight	1	2	22.6
(Intercept)	-6.33	-5.61	-3.03
Preadmission stoma counseling 'yes'	0.895	0.866	0.63
Opioid use on postoperative day 1 'yes'	1.75	1.74	1.78
Postoperative use of NSAIDS 'yes'	0.975	0.982	1.08
PONV observed In 3 days 'yes'	1.01	1	0.93
Time to tolerating solid food nights '1 or more nights needed'	0.944	0.922	0.716

Table 8.5: Estimated coefficients of logistic regression (right), weighted logistic regression with fixed weight (middle) and balanced weighted logistic regression (right) on marginal important variables predicting severe complications in Scenario 3 (after primary stay).

We see that the influence of *Opioiduse on postoperative day* and *Post operative use of NSAIDS* increases in the models for Scenario 3 with the marginal important variables as the weights increase (Table 8.5). The other estimated coefficients seem to become less important as the weights increase in Scenario 3 with the marginal important variables.

AIC forward variables, Scenario 3, severe complications			
Variable	**Estimated coeficient**		
	equal (or no) weights	**fixed weights**	**balanced weights**
weight	1	2	22.6
(Intercept)	-6.14	-5.47	-3.22
Preoperative WHO performance score '1'	-0.217	-0.167	0.216
Additional major procedures 'yes'	0.622	0.614	0.592
PONV observed In 3 days 'yes'	1.28	1.26	1.33
Opioid use on postoperative day 1 'yes'	1.53	1.52	1.51
If colloids 'yes'	-1.29	-1.28	-1.44
Resectionsite drainage 'yes'	0.821	0.82	0.781
Preadmission stoma counseling 'yes'	1.05	1.04	1.01
Bowel anastomosis 'yes'	0.395	0.443	0.63

Table 8.6: Estimated coefficients of logistic regression (right), weighted logistic regression with fixed weight (middle) and balanced weighted logistic regression (right) on variables selected with AIC forward predicting severe complications in Scenario 3 (after primary stay).

We see that the estimated coefficients for *Preoperative WHO performance score* are no longer in contrast with the literature in the balanced weighted model in Scenario 3 with the AIC forward variables (Table 8.6).

Conclusion

We have seen that the intercept becomes lower as the weight increases, except in Scenario 3 with the literature-based variables. Generally the higher the weights, the further the estimated coefficients (except the intercept) are from zero. Some unexpected signs in the estimated coefficients switched to a sign more in line with the literature in the different scenarios. However, this also happened a few times the other way around.

8.2. In-sample validation

In this section, we will look at the in-sample performance of the weighted models. We will first study the in-sample validations for the fixed weights in Scenarios 1 and 2, and then study these for the balanced weights in Scenarios 1 and 2. Afterwards, we study the in-sample validation for fixed weights in Scenario 3 and then for balanced weights in Scenario 3.

Fixed weights, severe complications						
	Scenario 1			**Scenario 2**		
	Literature	Marginal	AIC forward	Literature	Marginal	AIC forward
weight of outcome observations	2	2	2	2	2	2
Brier score	0.093	0.096	0.096	0.081	0.089	0.088
log-likelihood	-920.121	-883.943	-871.745	-1106.155	-951.022	-971.449
BIC	1964.374	1858.862	1877.482	2076.092	1902.308	1908.514
auc ROC	0.536	0.5	0.5	0.62	0.552	0.59
false negatives	46	50	50	37	44	40
false positives	4	0	0	9	7	9
misclassification rate	0.101	0.101	0.101	0.093	0.103	0.099
false omission rate	0.094	0.101	0.101	0.078	0.091	0.084
balanced accuracy	0.536	0.5	0.5	0.62	0.552	0.59
MCC	0.17	X	X	0.351	0.197	0.282
TS	0.074	0	0	0.22	0.105	0.169

Table 8.7: In-sample validation fixed-weighted logistic regression for Scenario 1 (presurgery prediction) and Scenario 2 (right after surgery prediction). The X for the MCC indicates that this measure was not always well defined (dividing by zero). The coloured numbers indicate the best score in the scenario, The red colour indicates the best over both scenarios. (496 observations with 50 severe complications)

In Table 8.7 we see the results for the fixed-weighted logistic regression in Scenarios 1 and 2. We see the results improve in Scenario 2 when compared to Scenario 1 for the different variable selections. For the marginal and for the AIC forward variables selections, the results of the in-sample validation are quite similar for the fixed-weighted and the unweighted models. We see that the Brier score is worse for the fixed weighted model but the log-likelihood and the BIC improve. We see that the literature based variables improve on most measures for the fixed-weighted compared to the unweighted model, except on the Brier score and the false positives. The improvement for the fixed weighted models is better visible in Scenario 2. We see that the misclassification rate for the literature based variable selection is similar for the fixed-weighted and the unweighted models in Scenario 2. We see that the fixed weighted model with the literature variables in Scenario 2 performs better on all measures except the on Brier score and the false positives. We see that the fixed-weighted model leads to a more informative model than the unweighted model.

Balanced weights, severe complications						
	Scenario 1			**Scenario 2**		
	Literature	Marginal	AIC forward	Literature	Marginal	AIC forward
weight of outcome observations	8.92	8.92	8.92	8.92	8.92	8.92
Brier score	0.226	0.235	0.239	0.184	0.208	0.203
log-likelihood	-407.78	-378.982	-368.19	-579.586	-428.554	-452.155
BIC	939.692	834.18	852.8	1051.411	877.626	883.833
auc ROC	0.636	0.614	0.59	0.711	0.677	0.692
false negatives	19	18	26	15	19	16
false positives	155	184	134	124	119	132
misclassification rate	0.351	0.407	0.323	0.28	0.278	0.298
false omission rate	0.061	0.064	0.077	0.045	0.055	0.048
balanced accuracy	0.636	0.614	0.59	0.711	0.677	0.692
MCC	0.169	0.138	0.116	0.272	0.232	0.245
TS	0.151	0.137	0.13	0.201	0.183	0.187

Table 8.8: In-sample validation balanced-weighted logistic regression for Scenario 1 (presurgery prediction) and Scenario 2 (right after surgery prediction). The coloured numbers indicate it is the best score in the scenario, The red colour indicates the best over both scenarios. (496 observations with 50 severe complications)

Moreover, we see that the Brier score and the misclassification rate perform worse for the models with the balanced weights compared to the fixed weighted models for Scenario 1 and 2 on any variables set used (Table 8.8). On the other hand, we see the balanced accuracy improve in the balanced weights models. The lowest balanced accuracy for the models with the balanced weights is with the AIC forward variables in Scenario 1, which equals 0.59. The highest balanced accuracy for the models using the fixed weights in Scenario 1 and 2 also equals 0.59. In Scenario 2 with the balanced weights we see the highest balanced accuracy is for the literature based variables selection. We see the most difference from comparing the fixed weights with the unweighed model even more clearly when comparing the balanced weighted model with the unweighted model. However, for the fixed-weighted model, the MCC performs better for some models with the fixed weights than with the balanced weights.

Fixed weights, severe complications			
	Scenario 3		
	Literature	Marginal	AIC forward
weight of outcome observations	2	2	2
Brier score	0.034	0.04	0.04
log-likelihood	-2091.481	-1551.426	-1533.246
BIC	4487.084	4220.201	4238.821
auc ROC	0.638	0.5	0.499
false negatives	15	21	21
false positives	5	0	1
misclassification rate	0.04	0.042	0.044
false omission rate	0.031	0.042	0.042
balanced accuracy	0.638	0.5	0.499
MCC	0.376	X	-0.009
TS	0.231	0	0

Table 8.9: In-sample validation fixed-weighted logistic regression for Scenario 3 (after primary stay). The X for the MCC indicates that this measure was not always well defined (dividing by zero). The coloured numbers indicate the best score in the scenario. (496 observations with 21 severe complications)

When comparing the 3 sets of variables in Scenario 3 for the models using the fixed weights we see the model with the literature based variable selection performing the best in-sample (Figure 8.9). We see the models with the marginal selection and the AIC forward selection predict very few severe complications in Senario 3 with the fixed weights.

Balanced weights, severe complications			
	Scenario 3		
	Literature	Marginal	AIC forward
weight of outcome observations	22.619	22.619	22.619
Brier score	0.131	0.194	0.198
log-likelihood	-1768.454	-574.357	-579.978
BIC	3841.03	3574.148	3592.767
auc ROC	0.884	0.705	0.683
false negatives	1	6	6
false positives	88	145	165
misclassification rate	0.179	0.304	0.345
false omission rate	0.003	0.018	0.019
balanced accuracy	0.884	0.705	0.683
MCC	0.374	0.176	0.154
TS	0.183	0.09	0.081

Table 8.10: In-sample validation balanced-weighted logistic regression for Scenario 3 (after primary stay). The coloured numbers indicate the best score in the scenario. (496 observations with 21 severe complications)

We see for the models using the balanced weights in Scenario 3 that the literature variables perform better compared to the marginal and AIC forward selection (Table 8.10). When comparing the marginal selection and the AIC forward selection with the balanced weights to the models with the fixed weights we see a large increase in the number of false positive predictions. The misclassification rate increased from 0.042 to 0.304 for the marginal selection, and from 0.044 to 0.345 for the AIC forward selection. However, the balanced accuracy improved. We see that the balanced weighted models predicting severe complications have a large weight equalling 22.619 (Table 8.10), With the high number of false positives, we see that this model no longer underestimates but overestimates the severe complications.

8.3. Cross-validation

The Tables 8.11 - 8.14 show the results of the cross-validation for the weighted models. For the fixed weighted models we denoted (0) beside the weights, this to stress that these weights did not change during the cross-validation, whereas the balanced weights did change depending on the number of outcomes present in the training set.

Fixed weights, severe complications						
	Scenario 1			**Scenario 2**		
	Literature	Marginal	AIC forward	Literature	Marginal	AIC forward
weights	2 (0)	2 (0)	2 (0)	2 (0)	2 (0)	2 (0)
Brier score	0.108 (0.027)	0.098 (0.024)	0.1 (0.025)	0.112 (0.03)	0.097 (0.026)	0.097 (0.027)
log-likelihood	-93.637 (5.503)	-88.536 (3.367)	-87.698 (3.392)	-114.447 (10.039)	-95.593 (4.863)	-98.014 (5.695)
BIC	265.352 (11.039)	188.784 (6.742)	198.82 (6.8)	377.244 (20.125)	230.225 (9.749)	238.971 (11.416)
auc ROC	0.505 (0.037)	0.501 (0.008)	0.5 (0.002)	0.545 (0.085)	0.538 (0.076)	0.565 (0.092)
false negatives	4.87 (1.965)	4.995 (2.001)	5 (1.995)	4.3 (1.913)	4.525 (1.957)	4.275 (1.902)
false positives	0.615 (0.812)	0.025 (0.354)	0.01 (0.141)	2.285 (1.548)	1.1 (1.103)	0.975 (1.01)
misclassification rate	0.111 (0.041)	0.101 (0.04)	0.101 (0.04)	0.133 (0.047)	0.113 (0.043)	0.106 (0.043)
false omission rate	0.1 (0.04)	0.101 (0.04)	0.101 (0.04)	0.092 (0.041)	0.094 (0.041)	0.089 (0.04)
balanced accuracy	0.505 (0.037)	0.501 (0.008)	0.5 (0.002)	0.545 (0.085)	0.538 (0.076)	0.565 (0.092)
MCC	0.042 (0.182)	X	X	0.111 (0.194)	0.145 (0.22)	0.228 (0.241)
TS	0.022 (0.065)	0.001 (0.009)	0 (0)	0.095 (0.111)	0.08 (0.112)	0.125 (0.152)

Table 8.11: Cross-validation fixed-weighted logistic regression for Scenario 1 (presurgery prediction) and Scenario 2 (right after surgery prediction). The X for the MCC indicates that this measure was not always well defined (dividing by zero) . The coloured numbers indicate the best score in the scenario, The red colour indicates the best over both scenarios. (496 observations with 50 severe complications)

The results of the cross-validation for the models with the fixed weights are slightly better than the results of the cross-validation for the unweighted models for Scenario 1 and 2 (Table 8.11 and Table 7.15). However, the Brier score and the misclassification rate are worse for the model using the fixed weights. We see that the models with the fixed weights are still underestimating in Scenario 1 and 2 however in a lesser extent than the unweighted models.

Balanced weights, severe complications						
	Scenario 1			**Scenario 2**		
	Literature	Marginal	AIC forward	Literature	Marginal	AIC forward
weights	8.944 (0.498)	8.944 (0.498)	8.944 (0.498)	8.944 (0.498)	8.944 (0.498)	8.944 (0.498)
Brier score	0.247 (0.023)	0.236 (0.016)	0.244 (0.015)	0.226 (0.038)	0.217 (0.023)	0.213 (0.024)
log-likelihood	-42.239 (3.65)	-37.986 (1.888)	-37.155 (1.657)	-63.647 (10.43)	-43.308 (2.957)	-45.898 (3.838)
BIC	162.556 (7.315)	87.683 (3.784)	97.734 (3.329)	275.644 (20.9)	125.656 (5.927)	134.738 (7.685)
auc ROC	0.498 (0.124)	0.62 (0.11)	0.551 (0.115)	0.551 (0.127)	0.607 (0.124)	0.618 (0.121)
false negatives	3.325 (1.728)	1.8 (1.215)	3.11 (1.689)	2.985 (1.599)	2.605 (1.533)	2.38 (1.529)
false positives	16.15 (3.43)	18.4 (3.167)	13.04 (3.748)	13.9 (3.631)	12.205 (3.295)	13.085 (3.438)
misclassification rate	0.393 (0.07)	0.407 (0.065)	0.326 (0.075)	0.34 (0.072)	0.299 (0.067)	0.312 (0.066)
false omission rate	0.105 (0.055)	0.065 (0.044)	0.09 (0.049)	0.088 (0.048)	0.074 (0.044)	0.07 (0.044)
balanced accuracy	0.488 (0.123)	0.614 (0.116)	0.548 (0.116)	0.546 (0.128)	0.606 (0.125)	0.618 (0.121)
MCC	-0.017 (0.138)	0.135 (0.131)	0.058 (0.133)	0.057 (0.15)	0.135 (0.148)	0.15 (0.144)
TS	0.078 (0.059)	0.136 (0.071)	0.104 (0.066)	0.106 (0.075)	0.139 (0.081)	0.145 (0.081)

Table 8.12: Cross-validation balanced-weighted logistic regression for Scenario 1 (presurgery prediction) and Scenario 2 (right after surgery prediction). The coloured numbers indicate the best score in the scenario, The red colour indicates the best over both scenarios. (496 observations with 50 severe complications)

We see the BIC and the log-likelihood increase further in the balanced weights models compared to the fixed weighted models in Scenario 1 and 2 (Table 8.12). However, again we see that the misclassification rate is higher with the balanced weights than the fixed weights. We see that the balanced accuracy improved for most variable selections in Scenario 1 and 2 when comparing the balanced weighted model to the fixed weighted model, except for the balanced accuracy for the literature based variable selection in Scenario 1.

Fixed weights, severe complications			
	Scenario 3		
	Literature	Marginal	AIC forward
weights	2 (0)	2 (0)	2 (0)
Brier score	0.071 (0.03)	0.042 (0.019)	0.044 (0.02)
log-likelihood	-291.131 (118.082)	-175.131 (64.69)	-177.498 (68.109)
BIC	773.555 (236.187)	373.685 (129.383)	390.132 (136.223)
auc ROC	0.525 (0.105)	0.502 (0.028)	0.5 (0.009)
false negatives	1.89 (1.251)	2.085 (1.322)	2.095 (1.313)
false positives	2.23 (1.486)	0.185 (0.438)	0.095 (0.294)
misclassification rate	0.083 (0.039)	0.046 (0.026)	0.044 (0.027)
false omission rate	0.04 (0.027)	0.042 (0.027)	0.042 (0.026)
balanced accuracy	0.525 (0.105)	0.502 (0.028)	0.5 (0.009)
MCC	0.051 (0.185)	0.026 (0.159)	-0.01 (0.079)
TS	0.048 (0.102)	0.005 (0.039)	0.001 (0.015)

Table 8.13: Cross-validation fixed-weighted logistic regression for Scenario 3 (after primary stay). The coloured numbers indicate the best score in the scenario. (496 observations with 21 severe complications)

We consider in the cross-validation for the fixed-weighted logistic regression in Scenario 3 the literature variable selection, the better performing selection as it was in Scenario 3 for the unweighted model. In scenario 3 we see more misclassifications, more false positives and a higher balanced accuracy when comparing the fixed-weighted model to the unweighted model (Table 8.13 and Table 7.17). The MCC was higher in the cross-validation for the literature variables in the unweighted model in Scenario 3.

Balanced weights, severe complications			
	Scenario 3		
	Literature	Marginal	AIC forward
weights	22.759 (1.912)	22.759 (1.912)	22.759 (1.912)
Brier score	0.18 (0.041)	0.198 (0.026)	0.207 (0.028)
log-likelihood	-309.595 (186.778)	-77.924 (62.748)	-81.698 (65.302)
BIC	810.483 (373.56)	179.272 (125.499)	198.531 (130.603)
auc ROC	0.521 (0.158)	0.692 (0.169)	0.619 (0.189)
false negatives	1.64 (1.276)	0.68 (0.849)	0.9 (1.032)
false positives	9.375 (3.1)	14.55 (3.346)	17.095 (3.615)
misclassification rate	0.222 (0.053)	0.307 (0.067)	0.363 (0.068)
false omission rate	0.04 (0.03)	0.02 (0.025)	0.028 (0.031)
balanced accuracy	0.521 (0.158)	0.692 (0.169)	0.619 (0.189)
MCC	0.017 (0.143)	0.166 (0.142)	0.097 (0.151)
TS	0.039 (0.056)	0.084 (0.065)	0.062 (0.056)

Table 8.14: Cross validation balanced-weighted logistic regression for Scenario 3 (post-primary stay). The coloured numbers indicate the best score in the scenario. (496 observations with 21 severe complications)

We see more misclassifications and more false positives when comparing the balanced-weighted model to the unweighted model (Table 8.14 and Table 7.17). We further see a higher balanced accuracy for the balanced weights than for the fixed weights in Scenario 3, except for the literature based variable selection. When comparing the balanced weighted and the fixed weighted models in Scenario 3 (Table 8.14 and Table 8.13), we see that the misclassification rate and the Brier score are higher for the balanced weighted models.

We consider the misclassification rate very high in the balanced weighted models. However, there seems to be more information in this model. In-sample the literature variable selection performed well, in the cross-validation this is no longer true.

8.4. Conclusion

We have seen that generally, the estimated coefficients increase their influence in the model and the intercept moves to zero as the weights increase.

We have seen that a weighted logistic regression leads to a larger Brier score and an increase in misclassification. However, the weighted models also lead to a better balanced accuracy and more false positives. The balanced weighted models do not seem to underestimate however they make a lot of errors in the classifications. The fixed weighted model underestimates less than the unweighted model but makes more mistakes. The balanced weighted model does not underestimate at all but makes many mistakes.

9

Discussion

9.1. Interpertation of key findings

The results of the cross-validation are dependent on the observations selected in the random folds made. These folds were not fixed and used for all cross-validations performed. This makes the cross-validations harder to compare in between tables. The usage of 200 repetitions in the cross-validation led to a quick outcome but made the numbers fluctuate.

Some speculations are made about the variable-generating process, these thoughts are not from a medical perspective. This should be considered when interpreting the findings of this study.

We concluded that the AIC forward selection was most prone to overfit the data; however, no steps were taken to account for the dependence within the data. We expect the method of AIC forward would improve significantly when one makes a preselection or when one uses another method to reduce highly correlated variables. Be aware that if one applies methods like oversampling or a weighted model, like in this thesis, these methods will seemingly alter the correlations within the data.

Some variables are known to influence a certain type of complication. These variables were not always selected since the influence on all complications was not clear. This might be countered with a different choice of model. We have seen some promising results in a study by Merath et al. [32], where a decision tree was used to predict complications after liver pancreatic and colorectal surgery.

In order to obtain findings with clinical relevance, the different outcome groups should be better adjusted to the specific treatment in order to be studied properly for use in clinical practice. For instance, we have seen that vasoactive drugs reduce the risk of complications. However, if this is due to comparing different procedures that use different anaesthesia or if this is due to the treatment itself, this has not been studied. It is hard to interpret coefficients of variables like *Infusion of vasoactive drugs*, *Any nonsurgical preoperative treatment* and *Preoperative nutritional treatment* from a mathematical perspective, since which decisions lead to the need for treatments is unknown to us.

9.2. Limitations and future directions

The conclusions on which variables are important greatly depend on their preprocessing. So, with different decisions on variable engineering, different variables might become more useful for modeling.

The information about anaesthesia and analgesia is hard to interpret. Perhaps one post-operative painkiller usage variable can be constructed. This variable can be compared with the expected amount of painkillers for the surgery. However, a reasonable amount of opioids is hard to define, for a non-medical professional. A variable indicating whether more than appropriate painkillers were needed would be interesting.

Variables for which we suspect that different preprocess steps could lead to better performance of the variables are *Alcohol usage, BMI, Mobilised on POD 1 or 2, Duration of IV fluid infusion nights, length of stay night in hospital after primary operation*.

The variable *Alcohol usage* could be more significant if a selection was made on the number of consumed alcoholic units per week. We think *BMI* should be discretized into 3 groups based on the data and not the WHO classifications. The variable *Mobilised On POD 1 or 2* should not have been differentiating on 3 values. A binary variable would have been better. We suspect transforming the IV fluids into a distance from the recommended or healthy amount of fluids could relate to the improvement of the importance of this variable since the signs for these variables changed in different models. In the ERAS protocol, a recommendation of intraoperative fluids based on the length of surgery and weight of the patient is given. Length of stay after operation contained some outliers and discretising this variable would have been better.

Perhaps the performance of the models will improve by altering the outcome choice. One could aim to predict the exact Clavien Dindo score of a complication or the length of stay. When looking in Chapter 5 at Tables 5.1- 5.13 the length of stay seems to be a predictable outcome.

It would be interesting to study non-severe outcomes to study the differences in risk factors for severe and non-severe outcomes, more specifically.

When comparing our performance with the study by Raijmakers [42], in Section 5.9, we think that additional limitations based on medical grounds in the observation selections would improve the performance of the models. Most studies in the literature start with defining their cohort in a more restrictive way.

We saw the importance of both the WHO score and the ASA class. ASA class was more commonly studied in our selection of literature. These variables are often used as explanatory variables for the preoperative health of the patient. In our data set, the WHO score seemed to be more informative. This is perhaps due to this variable being more mobility-based which is of importance for bowel function. We hope future research will look into the differences between these scores and see if these can be combined into one heath score indicating baseline health, especially for preoperative baseline bowel functionality.

Unexplained findings from this thesis could be further studied; for example, the importance of preadmission stoma counselling. Possibly, some correlations with other variables explain the importance of this variable. Perhaps this counselling is done for surgeries with uncertainty about the location of the resection necessary to treat the patient.

9.3. Other

Possibly removing discrepancies within the studies included in our literature based variable selection will improve the performance of this variable selection. We have seen that there are differences in risk factors for the different scenarios and for the different severity of the outcome. The literature based selection did not take these differences into account, even if the included studies perhaps did.

All the results are performed using one data set, so the conclusions here might not translate to different hospital situations or different patient populations. We did not confirm whether the findings here would translate to other hospitals. A study by Sluis et al. [46] uses data from a Dutch hospital to make a model and validates these on data from a Spanish hospital. This is a smart way since no data had to be transferred, only a model.

In Section 6.1.1 some variables are listed that are commonly studied but not in this data set. Further, some additional information about the resections could be of use. Some studies in the literature differentiated between an extensive and a non-extensive resections. Perhaps information about the weight or length of tissue removed can be a variable of importance.

10

Conclusion

10.1. Summary of research contributions

We have seen that modelling rare outcomes is difficult. We conclude that predicting any complication and severe complications should be approached differently. It is not a reliable approach to predict any complication in order to also find severe complications. We have seen that some variables, for example, diabetes or a low BMI, indicate a higher risk of severe complications but a lower risk of any complications. We suspect some extra care for these patients is present which leads to a lower frequency of any complication, but because their health is more fragile they do lead to an increase in severe complications. We think doctors are very aware of the general health of their patients and know which patients are likely to suffer from a severe complication if a complication arises.

We have summarized the importance of the variables in Tables 10.3 and 10.4. Table 10.3 summarises the important variables for any complications and Table 10.4 summarises the variables important for severe complications. We removed *Age, If predisease, Time to tolerating solid food night, Depth of anaesthesia monitored* and *Length of stay nights in hospital after primary operation* from literature selection for predicting severe complications, since these had a low estimated coefficient in the model. We further removed *If predisease* since the more precise comorbidities are more informative from the literature-based selection. *ASA physical status class* is removed for having a small coefficient from the variables important in the AIC forward selection for severe complications.

Oral bowel preparation and *More than one night with urinary drain* were removed from the important variables in the literature-based selection for having low estimated coefficients when predicting any complications. Further, *If predisease* was removed from the important variables from the literature, we think it would be more informative to add the comorbidities separately. The variables; *Preoperative thrombosis prophylaxis, General anaesthesia* (inhalation or intravenous) and *Intravenous fluid infusion restarted* are the only variables of the 51 preprocessed variables that were not selected in any variable selection.

In Table 10.1 and 10.2, we summarised which models had the best MCC and the best balanced accuracy for the different scenarios and outcomes. We conclude that the weighted model adds more information to the models for severe complications. However, we think that the models with balanced weights lead to too many misclassifications compared to the other models.

We see in Scenario 1 that the most important variable was the *Preoperative WHO performance score*. In the literature *ASA physical status class* was commonly used. Perhaps this is related to the importance of mobility for gut function. We have seen preoperative health has the biggest influence on postoperative complications in Scenario 1. We have seen that when including preoperative information in Scenario 2, intraoperative variables add useful information. However, the AIC forward selection method was more prone to overfit with the addition of the new variables. In Scenario 3 we see the intraoperative variables becoming less important, due to the complications in Scenario 3 being further away from the moment in time the surgery took place. In the literature, almost all variables are shown to relate to complications in any way. This made modelling using the literature-based variable selection difficult. We have seen that commonly explanatory variables are excluded from studies when they correlate to other explanatory variables.

Some variables should be interpreted by a medical professional and if not explained, be studied further in order to explain their influence within the models. These variables are *Intraoperative thrombosis*, *Preadmission stoma counselling*, *Oral bowel preparation*, *Carbohydrate loading* and *Infusion of vasoactive drugs*. These variables were selected in different models. Perhaps these variables are in reality encoding something correlated, which might not be in our available data or overlooked in the feature engineering.

10.2. Final thoughts

We have studied the predictability of the complications in the dataset. It has been a difficult task. We have preprocessed 3 variables that were not picked in any of the variable selection methods. We have seen the difference in predicting severe complications and any complications, and we have seen the difference in the importance of variables in the 3 different scenarios studied. We have seen that the standard logistic regression severely underestimates the rare severe complications and that the weighted logistic regression will lead to a model with more errors, but also more information. We have applied the TS and the MCC to our problem and seen that these do indeed form valid scores for scoring rare event models. We have seen that the p-value does not translate to influence under a model, that the AIC forward method does not select variables in order of coefficient size in the final model and that importance in the literature selection does not lead to higher influence in a model. We have seen that the AIC forward selection algorithm is more prone to overfitting when predicting any complications as the outcome rather than when predicting severe complications. When comparing the results of the variable selection method AIC forward and the marginal important variable selection, we better understand why variables are commonly selected using an univariate analysis in the studies selected for the literature based variable selection.

Even though the predictive power of the models was not great for the literature-based variable selection the method gave us an understanding of the applied strategies and variables used to predict adverse surgery outcomes. Tables 6.6 - 6.7 showing importance as perceived in different studies contain useful information. Furthermore, the literature study has led to insights about the state of prediction models for surgery outcomes and showed which variables are commonly studied but not present in our data.

	Scenario 1	**Scenario 2**	**Scenario 3**
Severe complications	Balanced weights marginal (0.135)	Fixed weighted AIC forward (0.228)	Balanced weights marginal (0.166)
Any complications	Marginal (0.243) (and AIC forward (0.242))	Marginal (0.245)	Marginal (0.152)

Table 10.1: Method with the best MCC for different scenarios from the cross-validations. The weighted models were not applied to predict any complications. In between brackets is the MCC from the cross-validation. For some models the MCC was not well defined, we consider these models to perform badly.

	Scenario 1	**Scenario 2**	**Scenario 3**
Severe complications	balanced weights Marginal (0.614)	balanced weights AIC forward (0.618)	Balanced weights marginal (0.692)
Any complications	Marginal and AIC forward (0.586)	Marginal (0.596)	Marginal (0.514)

Table 10.2: Methods applied that led to the highest balanced accuracy for different scenarios in the cross-validation. The weighted models were not applied to predict any complications. In between brackets is the balanced accuracy from the cross-validation.

For severe complications			
Variable selection	**Literature -based**	**Marginal**	**AIC forward**
Patient characteristics			
ASA physical status class	yes		
Smoker	yes		
Severe heart disease	yes		
Preoperative WHO performance score	yes	yes	yes
BMI	yes		
Diabetes mellitus	yes		scen 1
Severe pulmonary disease	yes		
Gender	yes		
Alcohol usage	yes		
Preoperative variables			
Preoperative nutritional treatment	yes		
Preoperative oral carbohydrate treatment	yes		
Oral bowel preparation	yes		scen1 and 2
Was aneamia found	yes		
Any nonsurgical preoperative treatment	yes		scen 2
Previous surgery to same abdominal region	yes		
Preadmission stoma counseling		yes	scen 3
Intraoperative variable			
If colloids	yes	scen 2	scen 3
Resectionsite drainage	yes	scen 2	yes
Infusion of vasoactive drugs	yes		scen 2
If lidocaine	yes		
Final diagnosis	yes		
Stomal procedure	yes		
Length of operation minutes	yes	scen 2	
Additional major procedures	yes	scen 2	yes
If open surgery or converted	scen 3	scen 2	
Main procedure name	yes		
If blood	yes	scen 2	
Total IV volume of fluids intraoperatively	yes	scen 2	scen 2
Core body temperature at end of operation	yes		
Bowel anastomosis			yes
Intraoperative thrombosis prophylaxis			scen 2
If spinal anaesthesia			scen 2
Postoperative variables			
mobilised on POD 1 or 2	yes		
Time to tolerating solid food nights	yes	yes	
more than one night with urinary drain	yes		
Time to pain control with oral analgesics nights	yes		
IV volume postoperational	yes		
Opioid use on postoperative day 1	yes		
Max pain VAS in 3 days	yes		
CD at least 2	yes		
Length of stay nights in hospital after primary operation	yes		
High cancer stage	yes		
Opioid use on postoperative day 1		yes	yes
Postoperative use of NSAIDS		yes	
PONV observed in 3 days		yes	yes

Table 10.3: Importance of different variables in the different scenarios for severe complications.

For any complications			
Variable selection	**Literature -based**	**Marginal**	**AIC forward**
Patient characteristics			
ASA physical status class	yes	scen 1 and 2	
Smoker	yes		
Severe heart disease	yes		
Preoperative WHO performance score	yes	yes	scen 1
Age	yes		
BMI	yes		scen 2
Diabetes mellitus	yes		
Severe pulmonary disease	yes	yes	
Gender	yes		
Alcohol usage	yes		
Preoperative variables			
Preoperative nutritional treatment	yes		
Preoperative oral carbohydrate treatment	yes		scen 1
Was aneamia found	yes		
Any nonsurgical preoperative treatment	yes		
Previous surgery to same abdominal region	yes		
Preadmission stoma counseling			scen 1 and 3
Intraoperative variable			
If colloids	yes		scen 3
Resectionsite drainage	yes	yes	
Depth of anaesthesia monitored	yes		
Infusion of vasoactive drugs	yes		scen 2
If liocaine	yes		
Final diagnosis	yes		
Stomal procedure	yes		
Length of operation minutes	yes	scen 2	
Additional major procedures	yes	yes	
If open surgery or converted	yes	yes	
Main procedure name	yes		
If blood	yes	scen 2	scen 2
Total IV volume of fluids intraoperatively	yes	yes	yes
Core body temperature at end of operation	yes		
Intraoperative thrombosis prophylaxis			yes
Postoperative variables			
Mobilised On POD 1 or 2	yes		yes
Time to tolerating solid food nights	yes		yes
Time to pain control with oral analgesics nights	yes		yes
IV volume Postoperational	yes		yes
Opioid use on postoperative day 1	yes	yes	yes
Max pain VAS on 3 days	yes		
CD at least 2	yes		
High cancer stage	yes		
PONV observed In 3 days		yes	yes
Duration of IV fluid infusion nights		yes	
Length of stay nights in hospital after primary operation	yes		

Table 10.4: Importance of different variables in the different scenarios for any complications.

References

[1] Renée Allvin et al. "Postoperative recovery: a concept analysis". In: *Journal of Advanced Nursing* 57.5 (2007), pp. 552–558. ISSN: 0309-2402. DOI: https://doi.org/10.1111/j.1365-2648.2006.04156.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2648.2006.04156.x.

[2] Renée Allvin et al. "Postoperative recovery: a concept analysis". In: *Journal of Advanced Nursing* 57.5 (2007), pp. 552–558. ISSN: 0309-2402. DOI: https://doi.org/10.1111/j.1365-2648.2006.04156.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2648.2006.04156.x.

[3] Ramzi Amri et al. "Obesity, outcomes and quality of care: body mass index increases the risk of wound-related complications in colon cancer surgery". In: *The American Journal of Surgery* 207.1 (2014), pp. 17–23. ISSN: 0002-9610. DOI: https://doi.org/10.1016/j.amjsurg.2013.05.016. URL: https://www.sciencedirect.com/science/article/pii/S0002961013005114.

[4] Christopher T. Aquina et al. "Association Among Blood Transfusion, Sepsis, and Decreased Long-term Survival After Colon Cancer Resection". In: *Annals of Surgery* 266.2 (2017). ISSN: 0003-4932. URL: https://journals.lww.com/annalsofsurgery/fulltext/2017/08000/association_among_blood_transfusion,_sepsis,_and.18.aspx.

[5] L.A.P. Arts. *Towards Colon Surgery in Daycare in the Medisch Spectrum Twente*. Report. 2022.

[6] I. S. Bakker et al. "Risk factors for anastomotic leakage and leak-related mortality after colonic cancer surgery in a nationwide audit". In: *BJS (British Journal of Surgery)* 101.4 (2014), pp. 424–432. ISSN: 0007-1323. DOI: https://doi.org/10.1002/bjs.9395. URL: https://bjssjournals.onlinelibrary.wiley.com/doi/abs/10.1002/bjs.9395.

[7] Dedrick Kok Hong Chan et al. "Age is an independent risk factor for increased morbidity in elective colorectal cancer surgery despite an ERAS protocol". In: *Langenbeck's Archives of Surgery* 405.5 (2020), pp. 673–689. ISSN: 1435-2451. DOI: 10.1007/s00423-020-01930-y. URL: https://doi.org/10.1007/s00423-020-01930-y.

[8] D. Chicco, N. Tötsch, and G. Jurman. "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation". In: *BioData Min* 14.1 (2021). 1756-0381 Chicco, Davide Orcid: 0000-0001-9655-7142 Tötsch, Niklas Jurman, Giuseppe Journal Article England 2021/02/06 BioData Min. 2021 Feb 4;14(1):13. doi: 10.1186/s13040-021-00244-z., p. 13. ISSN: 1756-0381 (Print) 1756-0381. DOI: 10.1186/s13040-021-00244-z.

[9] C. Damaskos et al. "Natural Ending or Surgical Complication: Is It the Time to Reconsider the Clavien-Dindo Classification System?" In: *Maedica (Bucur)* 17.4 (2022). 2069-6116 Damaskos, Christos Garmpis, Nikolaos Psilopatis, Iason Dimitroulis, Dimitrios Editorial Romania 2023/02/24 Maedica (Bucur). 2022 Dec;17(4):939-947. doi: 10.26574/maedica.2022.17.4.939., pp. 939–947. ISSN: 1841-9038 (Print) 1841-9038. DOI: 10.26574/maedica.2022.17.4.939.

[10] *Darmoperatie (ERAS), verpleegkundige zorg bij*. Government Document. Jan. 2023.

[11] S. B. Dharap, P. Barbaniya, and S. Navgale. "Incidence and Risk Factors of Postoperative Complications in General Surgery Patients". In: *Cureus* 14.11 (2022). 2168-8184 Dharap, Satish B Barbaniya, Priya Navgale, Shantanu Journal Article United States 2022/12/06 Cureus. 2022 Nov 1;14(11):e30975. doi: 10.7759/cureus.30975. eCollection 2022 Nov., e30975. ISSN: 2168-8184 (Print) 2168-8184. DOI: 10.7759/cureus.30975. URL: https://www-ncbi-nlm-nih-gov.tudelft.idm.oclc.org/pmc/articles/PMC9714582/.

[12] *Diabeter mellitus | Leeftijd en geslacht*. Web Page. 2023. URL: https://www.vzinfo.nl/diabetes-mellitus/leeftijd-en-geslacht.

[13] David C. Frankenfield et al. “Limits of body mass index to detect obesity and predict body composition”. In: *Nutrition* 17.1 (2001), pp. 26–30. ISSN: 0899-9007. DOI: https://doi.org/10.1016/S0899-9007(00)00471-8. URL: https://www.sciencedirect.com/science/article/pii/S0899900700004718.

[14] E. Gorgun et al. “Conversion in laparoscopic colorectal surgery: Are short-term outcomes worse than with open surgery?” In: *Techniques in Coloproctology* 20.12 (2016), pp. 845–851. ISSN: 1128-045X. DOI: 10.1007/s10151-016-1554-z. URL: https://doi.org/10.1007/s10151-016-1554-z.

[15] U. O. Gustafsson et al. “Guidelines for Perioperative Care in Elective Colorectal Surgery: Enhanced Recovery After Surgery (ERAS®) Society Recommendations: 2018”. In: *World Journal of Surgery* 43.3 (2019), pp. 659–695. ISSN: 1432-2323. DOI: 10.1007/s00268-018-4844-y. URL: https://doi.org/10.1007/s00268-018-4844-y.

[16] Steven A. Hicks et al. “On evaluation metrics for medical applications of artificial intelligence”. In: *Scientific Reports* 12.1 (2022), p. 5979. ISSN: 2045-2322. DOI: 10.1038/s41598-022-09954-8. URL: https://doi.org/10.1038/s41598-022-09954-8.

[17] Jonas Jurt et al. “Respiratory Complications After Colorectal Surgery: Avoidable or Fate?” In: *World Journal of Surgery* 42.9 (2018), pp. 2708–2714. ISSN: 1432-2323. DOI: 10.1007/s00268-018-4699-2. URL: https://doi.org/10.1007/s00268-018-4699-2.

[18] Philipp Kirchhoff, Pierre-Alain Clavien, and Dieter Hahnloser. “Complications in colorectal surgery: risk factors and preventive strategies”. In: *Patient Safety in Surgery* 4.1 (2010), p. 5. ISSN: 1754-9493. DOI: 10.1186/1754-9493-4-5. URL: https://doi.org/10.1186/1754-9493-4-5.

[19] Philipp Kirchhoff, Selim Dincler, and Peter Buchmann. “A Multivariate Analysis of Potential Risk Factors for Intra- and Postoperative Complications in 1316 Elective Laparoscopic Colorectal Procedures”. In: *Annals of Surgery* 248.2 (2008), pp. 259–265. ISSN: 0003-4932. DOI: 10.1097/SLA.0b013e31817bbe3a. URL: https://journals.lww.com/annalsofsurgery/fulltext/2008/08000/a_multivariate_analysis_of_potential_risk_factors.16.aspx.

[20] Rianne Kloosterman et al. *(On)gezonde leefstijl 2022: opvattingen, motieven en gedragingen*. Web Page. 2023. URL: https://www.cbs.nl/nl-nl/longread/rapportages/2023/on--gezonde-leefstijl-2022-opvattingen-motieven-en-gedragingen/7-alcoholgebruik.

[21] Kelsey E. Koch et al. “Male sex, ostomy, infection, and intravenous fluids are associated with increased risk of postoperative ileus in elective colorectal surgery”. In: *Surgery* 170.5 (2021), pp. 1325–1330. ISSN: 0039-6060. DOI: https://doi.org/10.1016/j.surg.2021.05.035. URL: https://www.sciencedirect.com/science/article/pii/S0039606021004955.

[22] M. Komaroff. “For Researchers on Obesity: Historical Review of Extra Body Weight Definitions”. In: *J Obes* 2016 (2016). 2090-0716 Komaroff, Marina Historical Article Journal Article Review United States 2016/06/18 J Obes. 2016;2016:2460285. doi: 10.1155/2016/2460285. Epub 2016 May 30., p. 2460285. ISSN: 2090-0708 (Print) 2090-0708. DOI: 10.1155/2016/2460285.

[23] Niels Komen et al. “After-hours colorectal surgery: a risk factor for anastomotic leakage”. In: *International Journal of Colorectal Disease* 24.7 (2009), pp. 789–795. ISSN: 1432-1262. DOI: 10.1007/s00384-009-0692-4. URL: https://doi.org/10.1007/s00384-009-0692-4.

[24] Yung-Heng Lee et al. “Effect of length of time from diagnosis to treatment on colorectal cancer survival: A population-based study”. In: *PLOS ONE* 14.1 (2019), e0210465. DOI: 10.1371/journal.pone.0210465. URL: https://doi.org/10.1371/journal.pone.0210465.

[25] S. R. Lewis et al. “Colloids versus crystalloids for fluid resuscitation in critically ill people”. In: *Cochrane Database Syst Rev* 8.8 (2018). 1469-493x Lewis, Sharon R Pritchard, Michael W Evans, David Jw Butler, Andrew R Alderson, Phil Smith, Andrew F Roberts, Ian 13/89/16/DH/*DepartmentofHealth/UnitedKi**AnalysisResearchSupport*, *Non–U.S.Gov'tSystematicReviewEngland*2018/08/04*CochraneDatabaseSystRe**CD*000567.*doi* : 10.1002/14651858.*CD*000567.*pub*7., p. Cd000567. ISSN: 1361-6137. DOI: 10.1002/14651858.CD000567.pub7.

[26] D. A. Ludwig. “Use and misuse of p-values in designed and observational studies: guide for researchers and reviewers”. In: *Aviat Space Environ Med* 76.7 (2005). Ludwig, David A Journal Article Review United States 2005/07/16 Aviat Space Environ Med. 2005 Jul;76(7):675-80., pp. 675–80. ISSN: 0095-6562 (Print) 0095-6562.

[27] E. Manilich et al. "Key Factors Associated With Postoperative Complications in Patients Undergoing Colorectal Surgery". In: *Diseases of the Colon and Rectum* 56.1 (2013), pp. 64–71. ISSN: 0012-3706. DOI: 10.1097/DCR.0b013e31827175f6. URL: https://journals.lww.com/dcrjournal/fulltext/2013/01000/key_factors_associated_with_postoperative.11.aspx.

[28] P. M. Markus et al. "Predicting postoperative morbidity by clinical assessment". In: *BJS (British Journal of Surgery)* 92.1 (2005), pp. 101–106. ISSN: 0007-1323. DOI: https://doi.org/10.1002/bjs.4608. URL: https://bjssjournals.onlinelibrary.wiley.com/doi/abs/10.1002/bjs.4608.

[29] Hossein Masoomi et al. "Risk Factors for Conversion of Laparoscopic Colorectal Surgery to Open Surgery: Does Conversion Worsen Outcome?" In: *World Journal of Surgery* 39.5 (2015), pp. 1240–1247. ISSN: 1432-2323. DOI: 10.1007/s00268-015-2958-z. URL: https://doi.org/10.1007/s00268-015-2958-z.

[30] Ahmed M. Al-Mazrou et al. "Characterization of Readmission by Day of Rehospitalization After Colorectal Surgery". In: *Diseases of the Colon and Rectum* 60.2 (2017), pp. 202–212. ISSN: 0012-3706. DOI: 10.1097/dcr.0000000000000734. URL: https://journals.lww.com/dcrjournal/fulltext/2017/02000/characterization_of_readmission_by_day_of.11.aspx.

[31] *Men and women*. Web Page. URL: https://www.cbs.nl/en-gb/visualisations/dashboard-population/men-and-women.

[32] Katiuscha Merath et al. "Use of Machine Learning for Prediction of Patient Risk of Postoperative Complications After Liver, Pancreatic, and Colorectal Surgery". In: *Journal of Gastrointestinal Surgery* 24.8 (2020), pp. 1843–1851. ISSN: 1873-4626. DOI: 10.1007/s11605-019-04338-2. URL: https://doi.org/10.1007/s11605-019-04338-2.

[33] M. Michailidou and V. N. Nfonsam. "Preoperative anemia and outcomes in patients undergoing surgery for inflammatory bowel disease". In: *The American Journal of Surgery* 215.1 (2018), pp. 78–81. ISSN: 0002-9610. DOI: https://doi.org/10.1016/j.amjsurg.2017.02.016. URL: https://www.sciencedirect.com/science/article/pii/S0002961016309837.

[34] Yuji Miyamoto et al. "Postoperative complications are associated with poor survival outcome after curative resection for colorectal cancer: A propensity-score analysis". In: *Journal of Surgical Oncology* 122.2 (2020), pp. 344–349. ISSN: 0022-4790. DOI: https://doi.org/10.1002/jso.25961. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jso.25961.

[35] Zhobin Moghadamyeghaneh et al. "Outcomes of colon resection in patients with metastatic colon cancer". In: *The American Journal of Surgery* 212.2 (2016), pp. 264–271. ISSN: 0002-9610. DOI: https://doi.org/10.1016/j.amjsurg.2016.01.025. URL: https://www.sciencedirect.com/science/article/pii/S0002961016301234.

[36] Ajinkya More. "Survey of resampling techniques for improving classification performance in unbalanced datasets". In: *arXiv preprint arXiv:1608.06048* (2016).

[37] David Ortiz-López et al. "Utility of a new prognostic score based on the Comprehensive Complication Index (CCI®) in patients operated on for colorectal cancer (S-CRC-PC score)". In: *Surgical Oncology* 42 (2022), p. 101780. ISSN: 0960-7404. DOI: https://doi.org/10.1016/j.suronc.2022.101780. URL: https://www.sciencedirect.com/science/article/pii/S0960740422000731.

[38] M L Peters et al. "Predictors of physical and emotional recovery 6 and 12 months after surgery". In: *British Journal of Surgery* 97.10 (2010), pp. 1518–1527. ISSN: 0007-1323. DOI: 10.1002/bjs.7152. URL: https://doi.org/10.1002/bjs.7152.

[39] Basilio Pirrera et al. "E.R.A.S. pathway in colorectal surgery in elderly: Our experience: A retrospective cohort study". In: *International Journal of Surgery* 43 (2017), pp. 101–106. ISSN: 1743-9191. DOI: https://doi.org/10.1016/j.ijsu.2017.05.013. URL: https://www.sciencedirect.com/science/article/pii/S1743919117303953.

[40] *Postoperatieve misselijkheid en braken (PONV)*. Government Document. 2023.

[41] *Prevalence of anaemia in women of reproductive age (aged 15-49)*. Web Page. URL: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-anaemia-in-women-of-reproductive-age-(-).

[42] Femke Raijmakers. *towards colon surgery in daycare in the mst*. Report. university of twente, 2021.

[43] Suvi Rasilainen et al. "ERAS failure and major complications in elective colon surgery: Common risk factors". In: *Surgery in Practice and Science* 10 (2022), p. 100080. ISSN: 2666-2620. DOI: https://doi.org/10.1016/j.sipas.2022.100080. URL: https://www.sciencedirect.com/science/article/pii/S2666262022000249.

[44] Ahmet Rencuzogullari et al. "Nomogram-Derived Prediction of Postoperative Ileus after Colectomy: An Assessment from Nationwide Procedure-Targeted Cohort". In: *The American Surgeon™* 83.6 (2017), pp. 564–572. DOI: 10.1177/000313481708300620. URL: https://journals.sagepub.com/doi/abs/10.1177/000313481708300620.

[45] Claudia M. Simões et al. "Predictors of major complications after elective abdominal surgery in cancer patients". In: *BMC Anesthesiology* 18.1 (2018), p. 49. ISSN: 1471-2253. DOI: 10.1186/s12871-018-0516-6. URL: https://doi.org/10.1186/s12871-018-0516-6.

[46] F. J. van der Sluis et al. "Predicting postoperative mortality after colorectal surgery: a novel clinical model". In: *Colorectal Disease* 16.8 (2014), pp. 631–639. ISSN: 1462-8910. DOI: https://doi.org/10.1111/codi.12580. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/codi.12580.

[47] Esteban T. D. Souwer et al. "Risk prediction models for postoperative outcomes of colorectal cancer surgery in the older population - a systematic review". In: *Journal of Geriatric Oncology* 11.8 (2020), pp. 1217–1228. ISSN: 1879-4068. DOI: https://doi.org/10.1016/j.jgo.2020.04.006. URL: https://www.sciencedirect.com/science/article/pii/S1879406819304126.

[48] Cloë L. Sparreboom et al. "Different Risk Factors for Early and Late Colorectal Anastomotic Leakage in a Nationwide Audit". In: *Diseases of the Colon and Rectum* 61.11 (2018). ISSN: 0012-3706. URL: https://journals.lww.com/dcrjournal/fulltext/2018/11000/different_risk_factors_for_early_and_late.6.aspx.

[49] Krishna K. Varadhan et al. "The enhanced recovery after surgery (ERAS) pathway for patients undergoing major elective open colorectal surgery: A meta-analysis of randomized controlled trials". In: *Clinical Nutrition* 29.4 (2010), pp. 434–440. ISSN: 0261-5614. DOI: https://doi.org/10.1016/j.clnu.2010.01.004. URL: https://www.sciencedirect.com/science/article/pii/S0261561410000099.

[50] A. K. Warps et al. "Interhospital referral of colorectal cancer patients: a Dutch population-based study". In: *Int J Colorectal Dis* 36.7 (2021). 1432-1262 Warps, A K de Neree Tot Babberich, M P M Dekker, E Wouters, M W J M Dekker, J W T Tollenaar, R A E M Tanis, P J Orcid: 0000-0002-3146-3310 Dutch ColoRectal Audit Journal Article Germany 2021/03/21 Int J Colorectal Dis. 2021 Jul;36(7):1443-1453. doi: 10.1007/s00384-021-03881-2. Epub 2021 Mar 20., pp. 1443–1453. ISSN: 0179-1958 (Print) 0179-1958. DOI: 10.1007/s00384-021-03881-2.

[51] Wenying Yang et al. "Observational studies: going beyond the boundaries of randomized controlled trials". In: *Diabetes Research and Clinical Practice* 88 (2010), S3–S9. ISSN: 0168-8227. DOI: https://doi.org/10.1016/S0168-8227(10)70002-4. URL: https://www.sciencedirect.com/science/article/pii/S0168822710700024.

[52] Xuexue Zhang et al. "Effect of comorbidity assessed by the Charlson Comorbidity Index on the length of stay, costs, and mortality among colorectal cancer patients undergoing colorectal surgery". In: *Current Medical Research and Opinion* 39.2 (2023). doi: 10.1080/03007995.2022.2139053, pp. 187–195. ISSN: 0300-7995. DOI: 10.1080/03007995.2022.2139053. URL: https://doi.org/10.1080/03007995.2022.2139053.

A

Nested data

1. *Preoperative nutritional status assessment*
   Screenings intrument
2. *Smoker*
   Termination of smoking no weeks before surgery
3. *Alcohol usage*
   Standard units
   Termination of alcohol no weeks before surgery
4. *Diabetes mellitus*
   Last HBA1c value mmolmol
5. *Preoperative chemotherapy*
   Days between admission and the last chemotherapy
6. *Was the patient screened for anaemia preoperatively*
   Was iron replacement treatment given
7. *Thrombosis prophylaxis*
   When was the first Anticoagulant prophylaxis done
   What was the duration of anticoagulant prophylaxis
8. *Bowel anastomosis*
   Type of bowel anastomosis
   Anastomotic technique
9. *Deep neuromuscular blockade*
   Ensure full reversal of Neuromuscular block
10. *Postoperative epidural analgesia*
    Other main postoperative analgesia
    Time to termination of epidural analgesia nights
    Successful block
    Strong opioids given within 48 hrs postoperatively
11. *Tijdens_*Complications at all during primary stay
    all columns starting with: *Tijdens_* (92 columns start with *Tijdens_*)
12. *Final diagnosis*
    T primary tumour
    N regional lymph nodes
    M distant metastasis

13. *Na_Complications at all after primary stay*
    all columns starting with: *Na_* (82 columns start with *Na_*)
14. *Stomal Procedure*
    Other free notes
15. *Epidural or spinal aneasthesia*
    Level of insertion
16. *Main procedure name*
    Lumbar supplementary analgesia
17. *Urinary drainage postop*
    Time to termination of urinary drainage nights
18. *Nasogastric tube inserted*
    Nasogastric tube nights

B

Variable names of unpreprocessed data

numberofoperation
ERASImplementation
Age
Gender
Totallengthofstaynights
Preoperativeweightchange
Preoperativenutritionalstatusassessment
ScreeningInstrument
Preoperativenutritionaltreatment
BMI
Smoker
Terminationofsmokingnoofweeksbeforesurgery
TerminationofsmokingnoofweeksbeforesurgeryUnknown
Alcoholusage
Standardunitsperweek
StandardunitsperweekUnknown
Terminationofalcoholnoofweeksbeforesurgery
TerminationofalcoholnoofweeksbeforesurgeryUnknown
Diabetesmellitus
Severeheartdisease
LastHbA1cvaluemmolmol
Severepulmonarydisease
LastHbA1cvalueUnknown
PreoperativeWHOperformancescore
Recentimmunosuppressivetreatment
Preoperativechemotherapy
Daysbetweenadmissionandthelastchemotherapy
Anyradiotherapytooperatingfield
Previoussurgerytosameabdominalregion
StomalProcedure
Preadmissionpatienteducationgiven
Recreationaldruguse
Preadmissionstomacounseling
Wasthepatientscreenedforanaemiapreoperatively
WasIronreplacementtreatmentgiven
Preoperativeoralcarbohydratetreatment

Oralbowelpreparation
Preoperativelongactingsedativemedication
Antibioticprophylaxisbeforeincision
Thrombosisprophylaxis
WhenwasthefirstAnticoagulantprophylaxisdone
Whatwasthedurationofanticoagulantprophylaxis
year_ok
Surgeryschedulingtype
Mainprocedurename
Additionalmajorprocedures
Surgicalapproachgroup
Operationconverted
Newstoma
Colostomyclosure
Ileostomyclosure
Parastomalhernia
StomalProcedure
Otherfreenotes
Bowelanastomosis
Typeofbowelanastomosis
Anastomotictechnique
Lengthofincision
LengthofincisionUnknown
Intraoperativebloodloss
IntraoperativebloodlossUnknown
Skinpreparationused
Lengthofoperation
Resectionsitedrainage
Urinarydrainagepostop
ASAphysicalstatusclass
PreviousPONVormotionsickness
PONVprophylaxisadministered
Generalanaesthesia
Nitrousoxideused
Systemicopioidsgiven
Airwaycontrol
Depthofanaesthesiamonitored
Deepneuromuscularblockade
EnsurefullreversalofNeuromuscularblock
Epiduralorspinalanaesthesia
Levelofinsertion
Lumbarsupplementaryanalgesia
Nerveblocksorlocalanaesthesia
Infusionofvasoactivedrugs
Upperbodyforcedairheatingcoverused
Forcedairheatingcoverused
HeatedIVfluidsused
Corebodytemperatureatendofoperation
CorebodytemperatureatendofoperationUnknown
Minimumcorebodytemperatureduringoperation
MinimumcorebodytemperatureduringoperationUnknown
IVvolumeofcrystalloidsintraoperatively
IVvolumeofcrystalloidsintraoperativelyUnknown
Useof09NaCl
IVvolumeofcolloidsintraoperatively
IVvolumeofcolloidsintraoperativelyUnknown

IVvolumeofbloodproductsintraoperatively
IVvolumeofbloodproductsintraoperativelyUnknown
TotalIVvolumeoffluidsintraoperatively
Fluidadministrationguidance
Nasogastrictubeusedpostoperatively
Ondayofsurgerypostoperatively
OndayofsurgerypostoperativelyUnknown
TotalIVvolumeoffluidsdayzero
DurationofIVfluidinfusionnights
Intravenousfluidinfusionrestarted
Weightchangeday1
Weightchangeday2
Weightchangeday3
OralfluidstotalvolumetakenOndayofsurgerypostoperatively
OralfluidstotalvolumetakenOndayofsurgerypostoperativelyUnknown
OralfluidstotalvolumetakenOnpostoperativeday1
OralfluidstotalvolumetakenOnpostoperativeday1Unknown
OralfluidstotalvolumetakenOnpostoperativeday2
OralfluidstotalvolumetakenOnpostoperativeday2Unknown
OralfluidstotalvolumetakenOnpostoperativeday3
OralfluidstotalvolumetakenOnpostoperativeday3Unknown
OralnutritionalsupplementsenergyintakeOndayofsurgerypostoperatively
OralnutritionalsupplementsenergyintakeOndayofsurgerypostoperativelyUnknown
OralnutritionalsupplementsenergyintakeOnpostoperativeday1
OralnutritionalsupplementsenergyintakeOnpostoperativeday1Unknown
OralnutritionalsupplementsenergyintakeOnpostoperativeday2
OralnutritionalsupplementsenergyintakeOnpostoperativeday2Unknown
OralnutritionalsupplementsenergyintakeOnpostoperativeday3
OralnutritionalsupplementsenergyintakeOnpostoperativeday3Unknown
Stimulationofgutmotility
Timetopassageofflatusnights
Timetopassageofstoolnights
Timetotoleratingsolidfoodnights
Independentlymanaginganewstomanights
Artificialnutrition
Atallondayofsurgery
Onpostoperativeday1
Onpostoperativeday1Unknown
Onpostoperativeday2
Onpostoperativeday2Unknown
Onpostoperativeday3
Onpostoperativeday3Unknown
Timetoterminationofurinarydrainagenights
TimetorecoveryofADLabilitynights
Postoperativeepiduralanalgesia
Othermainpostoperativeanalgesia
Timetoterminationofepiduralanalgesianights
Successfulblock
Strongopioidsgivenwithin48hrspostoperatively
Useofperipheralopioidreceptorantagonist
PostoperativeuseofNSAIDS
Timetopaincontrolwithoralanalgesicsnights
Nasogastrictubereinserted
Nasogastrictubereinsertednights
PatientreportedmaximumpainVASOndayofsurgery
PatientreportedmaximumpainVASOndayofsurgeryUnknown

PatientreportedmaximumpainVASOnpostoperativeday1
PatientreportedmaximumpainVASOnpostoperativeday1Unknown
PatientreportedmaximumpainVASOnpostoperativeday2
PatientreportedmaximumpainVASOnpostoperativeday2Unknown
PatientreportedmaximumpainVASOnpostoperativeday3
PatientreportedmaximumpainVASOnpostoperativeday3Unknown
PatientreportedmaximumnauseaVASOndayofsurgery
PatientreportedmaximumnauseaVASOndayofsurgeryUnknown
PatientreportedmaximumnauseaVASOnpostoperativeday1
PatientreportedmaximumnauseaVASOnpostoperativeday1Unknown
PatientreportedmaximumnauseaVASOnpostoperativeday2
PatientreportedmaximumnauseaVASOnpostoperativeday2Unknown
PatientreportedmaximumnauseaVASOnpostoperativeday3
PatientreportedmaximumnauseaVASOnpostoperativeday3Unknown
ObservednausearetchingandvomitingOndayofsurgery
ObservednausearetchingandvomitingOnpostoperativeday1
ObservednausearetchingandvomitingOnpostoperativeday2
ObservednausearetchingandvomitingOnpostoperativeday3
OpioiduseOndayofsurgery
OpioiduseOnpostoperativeday1
OpioiduseOnpostoperativeday2
OpioiduseOnpostoperativeday3
Dischargedwithin30postopdays
Dischargedto
Lengthofstaynightsinhospitalafterprimaryoperation
Timebetweenoperationanddeathnights.x
Tijdens_Complicationsatallduringprimarystay
Tijdens_Numberofnightsreceivingintensivecare
Tijdens_Reoperations
Tijdens_Gradingofmostseverecomplication
Tijdens_Complicationseveritygrade
Tijdens_Primarycauseofdeath
Tijdens_Respiratorycomplications
Tijdens_Lobaratelectasis
Tijdens_Pneumonia
Tijdens_PneumoniaClavien
Tijdens_Pleuralfluid
Tijdens_PleuralfluidClavien
Tijdens_Respiratoryfailure
Tijdens_RespiratoryfailureClavien
Tijdens_Pneumothorax
Tijdens_PneumothoraxClavien
Tijdens_Otherrespiratorycomplication
Tijdens_OtherrespiratorycomplicationClavien
Tijdens_Infectiouscomplications
Tijdens_Woundinfection
Tijdens_WoundinfectionClavien
Tijdens_Urinarytractinfection
Tijdens_UrinarytractinfectionClavien
Tijdens_Infectedlymphocele
Tijdens_Intraperitonealorretroperitonealabscess
Tijdens_IntraperitonealorretroperitonealabscessClavien
Tijdens_Sepsis
Tijdens_SepsisClavien
Tijdens_Septicshock
Tijdens_SepticshockClavien

Tijdens_Infectedgraftorprosthesis
Tijdens_Otherinfectiouscomplication
Tijdens_OtherinfectiouscomplicationClavien
Tijdens_Cardiovascularcomplications
Tijdens_Heartfailure
Tijdens_HeartfailureClavien
Tijdens_Acutemyocardialinfarction
Tijdens_Deepvenousthrombosis
Tijdens_PortalVeinThrombosis
Tijdens_PortalVeinThrombosisClavien
Tijdens_Pulmonaryembolus
Tijdens_Cerebrovascularlesion
Tijdens_Cardiacarrhythmia
Tijdens_CardiacarrhythmiaClavien
Tijdens_Cardiacarrest
Tijdens_Othercardiovascularcomplication
Tijdens_OthercardiovascularcomplicationClavien
Tijdens_Hypertension
Tijdens_Renalhepaticpancreaticandgastrointestinalcomplications
Tijdens_Renaldysfunction
Tijdens_RenaldysfunctionClavien
Tijdens_Urinaryretention
Tijdens_UrinaryretentionClavien
Tijdens_Hepaticdysfunction
Tijdens_Pancreatitis
Tijdens_Gastrointestinalhaemorrhage
Tijdens_GastrointestinalhaemorrhageClavien
Tijdens_Nauseaorvomiting
Tijdens_NauseaorvomitingClavien
Tijdens_Obstipationordiarrhoea
Tijdens_ObstipationordiarrhoeaClavien
Tijdens_Otherorgandysfunction
Tijdens_OtherorgandysfunctionClavien
Tijdens_Incontinence
Tijdens_Surgicalcomplications
Tijdens_Anastomoticleak
Tijdens_AnastomoticleakClavien
Tijdens_Urinarytractinjury
Tijdens_UrinarytractinjuryClavien
Tijdens_Mechanicalbowelobstruction
Tijdens_MechanicalbowelobstructionClavien
Tijdens_Postoperativeparalyticileus
Tijdens_PostoperativeparalyticileusClavien
Tijdens_Deepwounddehiscence
Tijdens_DeepwounddehiscenceClavien
Tijdens_Intraoperativeexcessivehaemorrhage
Tijdens_IntraoperativeexcessivehaemorrhageClavien
Tijdens_Postoperativeexcessivehaemorrhage
Tijdens_PostoperativeexcessivehaemorrhageClavien
Tijdens_Othersurgicaltechnicalcomplicationorinjury
Tijdens_OthersurgicaltechnicalcomplicationorinjuryClavien
Tijdens_Othercomplicationsofreconstructivesurgery
Tijdens_Hematoma
Tijdens_HematomaClavien
Tijdens_Complicationsrelatedtoepiduralorspinalanaesthesia
Tijdens_Postduralpunctureheadache

Tijdens_Epiduralhematomaorabscess
Tijdens_OtherEDAorspinalrelatedcomplication
Tijdens_Anaestheticcomplications
Tijdens_Aspirationofgastriccontents
Tijdens_Hypotension
Tijdens_Hypoxia
Tijdens_Prolongedpostoperativesedation
Tijdens_Otheranaestheticcomplications
Tijdens_Psychiatriccomplications
Tijdens_Astheniaortiredness
Tijdens_Pain
Tijdens_Injuries
Tijdens_othercomplication
Finaldiagnosis
TPrimaryTumour
NRegionalLymphNodes
MDistantMetastasis
dertigdaysurvival
Timebetweenoperationanddeathnights.y
dertigdayfollowupperformed
Timebetweenoperationandfollowupnights
WHOPerformanceScoreatdertigdayspostoperatively
Na_Complicationsatallafterprimarystay
Na_Numberofnightsreceivingintensivecare
Na_Readmissions
Na_Lengthofstayforreadmissions
Na_Lengthofstayforreadmissionsunknown
Na_Reoperations
Na_Gradingofmostseverecomplication
Na_Complicationseveritygrade
Na_Primarycauseofdeath
Na_Respiratorycomplications
Na_Lobaratelectasis
Na_Pneumonia
Na_PneumoniaClavien
Na_Pleuralfluid
Na_Respiratoryfailure
Na_Pneumothorax
Na_Otherrespiratorycomplication
Na_OtherrespiratorycomplicationClavien
Na_Infectiouscomplications
Na_Woundinfection
Na_WoundinfectionClavien
Na_Urinarytractinfection
Na_Intraperitonealorretroperitonealabscess
Na_IntraperitonealorretroperitonealabscessClavien
Na_Sepsis
Na_SepsisClavien
Na_Septicshock
Na_SepticshockClavien
Na_Infectedgraftorprosthesis
Na_Otherinfectiouscomplication
Na_OtherinfectiouscomplicationClavien
Na_Cardiovascularcomplications
Na_Heartfailure
Na_Acutemyocardialinfarction

Na_Deepvenousthrombosis
Na_PortalVeinThrombosis
Na_Pulmonaryembolus
Na_Cerebrovascularlesion
Na_Cardiacarrhythmia
Na_Cardiacarrest
Na_Othercardiovascularcomplication
Na_OthercardiovascularcomplicationClavien
Na_Hypertension
Na_Renalhepaticpancreaticandgastrointestinalcomplications
Na_Renaldysfunction
Na_RenaldysfunctionClavien
Na_Urinaryretention
Na_UrinaryretentionClavien
Na_Hepaticdysfunction
Na_Pancreatitis
Na_Gastrointestinalhaemorrhage
Na_GastrointestinalhaemorrhageClavien
Na_Nauseaorvomiting
Na_NauseaorvomitingClavien
Na_Obstipationordiarrhoea
Na_ObstipationordiarrhoeaClavien
Na_Otherorgandysfunction
Na_OtherorgandysfunctionClavien
Na_Incontinence
Na_Surgicalcomplications
Na_Anastomoticleak
Na_AnastomoticleakClavien
Na_Urinarytractinjury
Na_Mechanicalbowelobstruction
Na_MechanicalbowelobstructionClavien
Na_Postoperativeparalyticileus
Na_PostoperativeparalyticileusClavien
Na_Deepwounddehiscence
Na_DeepwounddehiscenceClavien
Na_Intraoperativeexcessivehaemorrhage
Na_Postoperativeexcessivehaemorrhage
Na_PostoperativeexcessivehaemorrhageClavien
Na_Othersurgicaltechnicalcomplicationorinjury
Na_OthersurgicaltechnicalcomplicationorinjuryClavien
Na_Hematoma
Na_HematomaClavien
Na_Complicationsrelatedtoepiduralorspinalanaesthesia
Na_Postduralpunctureheadache
Na_Epiduralhematomaorabscess
Na_OtherEDAorspinalrelatedcomplication
Na_Anaestheticcomplications
Na_Aspirationofgastriccontents
Na_Hypotension
Na_Hypoxia
Na_Prolongedpostoperativesedation
Na_Otheranaestheticcomplications
Na_Psychiatriccomplications
Na_Astheniaortiredness
Na_Pain
Na_Injuries

Na_othercomplication
Timebetweenadmissioandprimaryoperation

C

Summary of relations in bewteen variables from literature study by P. Kirchhoff et al.

Variable	Influences the risk of another variable
old age	comorbidities
Gender and low rectal anastomoses	anastomotic leakage
prior abdominal surgery	conversion of surgery, surgical injury, postoperative ileus, reoperation and longer operating times
open surgery, colon or rectal insections	adhesion-related problems.
weight loss > 10% or neurologic comorbidity	complications
A hematocrit < 30%, the use of steroids, albumin <3.5 g/L or creatinine >1.4 mmol/L	morbidity and mortality
cancer, ascites, hypernatremia, do not resuscitate status before surgery, ASA classes III-V or a medical history of congestive heart failure.	death
postoperative coma, cardiac arrest, a pre-existing vascular graft prosthesis failing, renal failure, pulmonary embolism, or progressive renal insufficiency	30-day mortality

Variable	Influences the risk of another variable
obesity	longer operating time, prolonged hospital stay, more intraoperative complications and conversion of surgery
BMI > 25	surgical site infection.
immunonutrition	infection complications, length of hospital stay and mortality
nutritional support	morbidity
bowel cleaning	intraoperative spillage of bowel contents
intraoperative spillage of bowel contents	infections
bowel preparation and anastomosis	operation time
major surgery, blood utilization	anemia
anemia	complication and mortality
endoscopic surgery	visceral injury
BMI, ASA grade, type of resection, abscess, Age, Adgesions, bleeding, fistula inflammatory mass and bowel perforation, prior abdominal surgery	conversion
conversion	morbidity
postoperative normal diet (over fasting)	anastomotic complications, infections, length of stay
enteral nutrition	mortality
BMI, ostomy, transfusion, gender, ASA score, laparascopic, immumonutrition, antibiotic prophylaxis colon surgery and (ostomy closure or antibiotics) rectal surgery and (steroids or preoperative radiation or ostomy creation)	wound infection
anastomotic leakage intra- or extra- peritoneal anastomoses, rectal resections, gender, previous abdominal surgery Crohn's disease, rectal cancer,	mortality
prolonged operating time, anaesthesiologist grade	anastomotic leakage
anastomotic leakage and cancer	mortality, recurrence
ileus	morbidity, length of stay
operating time and bloodloss	ileus

D

Variable meanings

Variable name	meaning
Patient characteristics	
Alcohol usage	Whether the patient normally drinks alcohol any number of alcoholic units a week
BMI	Body Mass Index (kg /m^2)
Diabetes mellitus	Whether the patient has diabetes mellitus, a disease affecting sugar regulation in blood.
Gender	Gender of the patient
If Presidease	Whether one of the comorbidities severe pulmonary disease, severe heart disease or diabetes mellitus are present for the patient
Severe hearth disease	Indicates if the patient has severe heart disease
Severe pulmonary disease	Indicates if the patient has severe lung or respiratory disease
ASA physical status class	Score based on health made by an anaesthesiologist
Preoperative WHO performance score	General health score made by WHO, measures mostly mobility
Smoker	Whether the patient smokes
Continous variables	
Age	Age of the patient at time of operation
Preoperative variables	
Previous surgery to same abdominal region	Did the patient already undergo surgery previously to the same abdominal region
Any nonsurgical preoperative treatment	whether the patient underwent the treatments chemotherapy, immunosuppressive or radiotherapy to the operating field before the surgery.
Preadmission stoma counselling	If counselling before surgery was done on how to handle a stoma
Was Anaemia found	Did the patient test positive for anaemia
Preoperative nutritional treatment	Was the patient prescribed a special diet
Preoperative oral carbohydrate treatment	Carbohydrate should improve the outcome of the surgery by supplying nutrients
Oral bowel preparation	Were laxatives administered before surgery
Preoperative thrombosis prophylaxis	Was the thrombosis prophylaxis done before the surgery
Preoperative variables	
Final diagnosis	Indicates disease the patient suffers from
If colloids	Whether a certain type of fluid is given
If blood	If blood loss was more than 100ml or if any blood products were given
Resectionsite drainage	Was a drain inserted for postoperative drainage of fluids
Interoperative anaesthesia and drugs	
General anaesthesia	Was the anaesthesia of the inhalation type or the intravenous type
Depth of anaesthesia monitored	Was the depth of the anaesthesia monitored
If spinal aneasthesia	Was spinal anaesthesia used
Infusion of vasoactive drugs	Were vasoactive drugs used during the operation, these drugs regulate the blood flow, by affecting heart function or the veins and arteries
Intraoperative thrombosis prophylaxis	Was the medication to avoid blood clots started during surgery
If lidocaie	Indicates the use of a certain painkiller
Intraoperative surgery type	
Main procedure name	Name of the main procedure of the operation
Additional major procedures	Were other major surgical procedures performed during the operation (for example, part of the liver removed, or cancerous growth removed)
Stomal Procedure	If a stoma is created, removed or relocated during the operation.
Bowel anastomosis	Is a bowel anastomosis created
If open surgery or converted	If the surgery ended in open surgery, does not contain information about how it started
Continuous variables	
Total IV volume of fluids intraoperatively	ml of fluids intraoperatively administered through IV
Core body temperature at end of operation	Temperature in degrees celsius of the patient
Length of operation minutes	Length of the time necessary for the surgery

Varaible name	**meaning**
Postoperative variables	
PONV observed in 3 days	Was any vomiting or retching observed in the first 3 post-op days
Mobilised on POD 1 or 2	When was the patient not bedbound any more
Intravenous fluid infusion restarted	whether the patient had to be reconnected to a drip
CD at least 2	Any complications with at least a CD of 2 during primary stay
Time to tolerating solid food nights	When could the patient consume food normally
More than one night with urinary drain	Whether urinary drain is used for at least one night
High cancer stage	Specifies the cancer diagnosis based on the stage of the canser
Duration of iv fluid infusion nights	How many nights the patient was connected to an drip
Postoperative drugs usage	
Opioid use on postoperative day 1	Were any opioids used by the patient on the first day after surgery
Postoperative use of NSAIDs	Whether NSAIDs, a type of pain killer was administered to the patient
Time to pain control with oral analgesics nights	Number of nights passed before only oral pain medication was used for pain control
Continuous vairables	
IV volume post operational	ml of fluids from drip (IV) on the day of surgery after the surgery took place
Max pain VAS in 3 days	Max pain experienced in the first 3 postoperative days as stated by the patient themself on a VAS
Length of stay nights in hospital *after primary operation*	Nights spend in hospital recovering from the surgery

E

Overview of outcome related variables

Data directly related to complications, These can not be used as explanatory variables but could be seen as alternative outcomes.

Varaible name	Collected during primary stay	Type of data	Meaning
Time between operation and death nights.x	only during primary stay	number of nights	Time from operation to death if diseased during primary stay
Time between operation and death nights.y	only after primary stay	number of nights	Time from operation to death if diseased after primary stay
Complications at all during primary stay	only during primary stay	binary,	Indicates if any complications occur during primary stay
Number of nights receiving intensive care	during and after primary stay	number of nights	number of nights receiving continuously monitored care
Reoperations	during and after primary stay	number	number of reoperations, operations for complications
Grading of most severe complication	during and after primary stay	CD scale	Clavien Dindo scale (CD scale) of most severe complication
Complication severity grade	during and after primary stay	binary	Was the CD scale for the most severe complication more than 3
Primary cause of death	during and after primary stay	factor	Indicates primary cause of death
Readmissions	only after primary stay	binary	Was the patient readmitted to the hospital
Length of stay for readmissions	only after primary stay	number of nights	How many nights did the patient stay during the readmission
Dertig day survival		binary	Did the patient survive upto 30 days after the surgery

F

CD of complications complete information

name complication	during primary stay	0	1	2	3	4	5	6	7
Grading of most severe complication	during	455	38	33	2	23	5	1	5
	after	464	20	34	4	19	3	0	0
Pneumonia	during	553	0	6	0	1	1	0	1
	after	544	0	0	0	0	0	0	0
Pleural fluid	during	560	0	1	1	0	0	0	0
	after	544	0	0	0	0	0	0	0
Respitory failure	during	562	0	0	0	0	0	0	0
	after	544	0	0	0	0	0	0	0
Pneumothorax	during	560	0	1	1	0	0	0	0
	after	544	0	0	0	0	0	0	0
Other respiratory comlications	during	558	1	2	0	0	0	0	1
	after	544	0	0	0	0	0	0	0
Wound infection	during	553	8	0	1	0	0	0	0
	after	531	8	3	0	1	0	0	0
Urinarytract infection	during	560	0	2	0	0	0	0	0
	after	544	0	0	0	0	0	0	0
Intraperitoneal or retroperitoneal abscess	during	557	1	1	2	1	0	0	0
	after	532	0	9	0	3	0	0	0
Sepsis	during	560	0	0	0	0	0	0	2
	after	544	0	0	0	0	0	0	0
Septicshock	during	561	0	0	0	0	0	0	1
	after	544	0	0	0	0	0	0	0
Other infectious complication	during	560	0	2	0	0	0	0	0
	after	544	0	0	0	0	0	0	0
Hearth failure	during	560	0	1	0	0	0	0	1
	after	544	0	0	0	0	0	0	0
Protal vein thrombosis	during	562	0	0	0	0	0	0	0
	after	544	0	0	0	0	0	0	0

Cariac arrhythmia	during	561	1	0	0	0	0	0	0
	after	544	0	0	0	0	0	0	0
Other cardio vascular complication	during	562	0	0	0	0	0	0	0
	after	544	0	0	0	0	0	0	0
Renaldysfunction (renal failure)	during	560	0	2	0	0	0	0	0
	after	543	0	0	0	1	0	0	0
Renaldysfunction (oliguria)	during	561	0	1	0	0	0	0	0
	after	544	0	0	0	0	0	0	0
Urinary retention	during	547	9	6	0	0	0	0	0
	after	543	0	1	0	0	0	0	0
Gastrointestinal haemorrage	during	561	0	0	0	0	1	0	0
	after	541	0	1	1	0	1	0	0
Nausea or vomiting	during	556	5	0	0	1	0	0	0
	after	541	1	2	0	0	0	0	0
Obstipation or diarrhoea	during	561	1	0	0	0	0	0	0
	after	534	4	5	1	0	0	0	0
Other organ dysfunction	during	560	1	1	0	0	0	0	0
	after	544	0	0	0	0	0	0	0
Anastomotic (reoperated)	during	553	0	0	0	7	1	0	1
	after	538	0	0	0	6	0	0	0
Anastomotic (radiological diagnosis with no intervention)	during	560	0	2	0	0	0	0	0
	after	542	1	1	0	0	0	0	0
Anastomotic (drained percutaneously)	after	541	0	1	0	1	1	0	0
Uninary tract injury	during	561	0	0	0	0	0	0	1
	after	544	0	0	0	0	0	0	0
Mechanical bowel obstruction	during	561	0	0	0	1	0	0	0
	after	543	0	0	0	1	0	0	0
Postoperative paralytic ileus	during	544	5	8	1	3	1	0	0
	after	535	1	6	0	2	0	0	0
Deep wound dehiscence	during	560	0	1	0	1	0	0	0
	after	541	0	1	0	2	0	0	0
Intra operative excessive haemorrhage	during	560	0	0	0	1	1	0	0
Post operative excessive heamorrage	during	552	2	4	0	3	0	1	0
	after	543	0	0	0	1	0	0	0
Other surgical techinical compliation	during	553	3	3	0	3	0	0	0
	after	540	0	1	1	1	1	0	0
Hematoma	during	561	0	0	0	1	0	0	0
	after	537	1	4	0	2	0	0	0

Table F.1: Clavien scale data, Appendix H

G

Modifications text to numbers

No	→ 0
Yes	→ 1
Yes, alive	→ 1
No, dead	→ 0
Yes, reoperated	→ 1
Yes, radiological diagnosis with no intervention	→ 2
Yes, drained percutaneously	→ 3
I	→ 1
II	→ 2
IIIa	→ 3
IIIb	→ 4
IVa	→ 5
IVb	→ 6
V	→ 7
Serious complications	→ 1
Non-serious complications	→ 0
Yes, renal failure	→ 1
Yes, oliguria	→ 2
intraperitoneale correctie van ileostoma	→ 1
extraperitoneale correctie van colostoma	→ 2
eindstandig sigmoïdostomie	→ 3
Revisie colostoma	→ 4
revisie colostoma	→ 4

H

Additional information of studies included in literature based variable selection

Here we have additional information about studies included in the literature-based variable selection.

no. of study	elective or emergency surg.	data source	location of population	statistical model	ref
1	unknown	literature		literature review	link
2	unknown	literature		literature review	link
3	elective	observational	Finland	logistic regression	link
4	elective	observational	Dutch (MST)	logistic regression	link
5	both	observational	India	bivariate statistical testing	link
6	elective	observational	Brazil	logistic regression	link
7	both	observational	Dutch	logistic regression	link
8	elective	questionnaire	Dutch	regression	link
9	elective	observational	Taiwanese	cox proportional hazards	link
10	both	observational	Dutch	logistic regression	link
11	elective	observational	Spanish	cox proportional hazards	link
12	unknown	observational	Chinese	other regression model	link
13	both	observational	United states	bootstrap classification models	link
14	elective	observational	United states	logistic regression	link
15	elective	observational	Switzerland	logistic regression	link
16	both	observational	Dutch	logistic regression	link
17	unknown	literature		literature review	link
18	both	observational		logistic regression	link
19	both	observational	Switserland	logistic regression	link
20	elective	observational	United states	logistic regression	link
21	elective	observational	Portugal	bivariate statistical testing	link

I

Estimated coefficients of logistic regression models predicting any complications

Scenario 1 with any complications
Scenario 1 predictes postoperative complications before the surgery takes place.

Literature variables, Scenario 1, Any complications	
Variable	**Estimated coeficient**
(Intercept)	-1.216
ASA physical status class '3 or more'	0.427
Smoker 'yes'	0.264
Severe heart disease 'yes'	-0.665
Preoperative WHO performance score '1'	1.465
Age	-0.00146
BMI '(21.5-24.9)'	0.0189
BMI '(25-29.9)'	0.115
BMI '(30 or more)'	0.185
If predisease 'yes'	0.376
Diabetes mellitus 'yes'	-0.681
Severe pulmonary disease 'yes'	0.0907
Gender 'female'	-0.219
Alcohol usage 'yes'	-0.126
Preoperative nutritional treatment 'yes'	-0.211
Preoperative oral carbohydrate treatment 'yes'	0.163
Oral bowel preparation 'yes'	0.353
Was aneamia found 'yes aneamia'	0.123
Any nonsurgical preoperative treatment 'yes'	-0.759
Previous surgery to same abdominal region 'yes'	0.161

Table I.1: Coeficients of logistic regression with literature-based variables predicting any complications in Scenario 1 (preoperative).

Marginal variables, Scenario 1, Any complications	
Variable	**Estimated coeficient**
(Intercept)	-1.15
Severe pulmonary disease 'yes'	0.388
ASA physical status class '3 or more'	0.372
Preoperative WHO performance score '1'	1.279

Table I.2: Coeficients of logistic regression with marginal important variables predicting any complications in Scenario 1 (before surgery).

AIC forward variables, Scenario 1, Any complications	
Variable	**Estimated coeficient**
(Intercept)	-1.38
Preoperative WHO performance score '1'	1.553
Preadmission stoma counseling 'yes'	0.193
Preoperative oral carbohydrate treatment 'yes'	0.391

Table I.3: Coeficients of logistic regression with variables selected with AIC forward predicting any complications in Scenario 1 (before surgery).

Scenario 2 with any complications

Scenario 2 predicts postoperative complications right after the surgery takes place.

Literature variables, Scenario 2, Any complications	
Variable	**Estimated coeficient**
(Intercept)	-9.567
ASA physical status class '3 or 4'	0.483
Smoker 'yes'	0.275
Severe heart disease 'yes'	-0.409
Preoperative WHO performance score '1'	1.395
Age	0.0101
BMI '(21.5-24.9)'	0.00872
BMI '(25-29.9)'	0.0580
BMI '(30 or more)'	0.150
If predisease 'yes'	0.288
Diabetes mellitus 'yes'	-0.921
Severe pulmonary disease 'yes'	0.0322
Gender 'female'	-0.144
Alcohol usage 'yes'	-0.120
Preoperative nutritional treatment 'yes'	-0.330
Preoperative oral carbohydrate treatment 'yes'	-0.203
Oral bowel preparation 'yes'	0.289
Was aneamia found 'yes aneamia'	0.0416
Any nonsurgical preoperative treatment 'yes'	-0.828
Previous surgery to same abdominal region 'yes'	-0.0751
If colloids 'yes'	-0.322
Resection site drainage 'yes'	0.317
Depth of anaesthesia monitored 'yes'	0.186
Infusion of vasoactive drugs 'yes'	-0.567
If lidocaine 'yes'	-0.266
Final diagnosis 'cancer'	0.251
Final diagnosis 'no cancer or functional disorder'	0.440
Stomal procedure 'yes'	-0.274
Length of operation minutes	0.00289
Additional major procedures 'yes'	0.721
If open surgery or converted 'no'	-0.125
Main procedure name 'sigmoid resection'	-0.301
Main procedure name 'stoma procedure	0.779
Main procedure name 'anterior resection of the rectum'	0.277
Main procedure name 'uncommon procedure, no ileocaecal'	0.298
If blood 'yes'	0.203
Total IV volume of fluids intraoperatively	0.000226
Core body temperature at end of operation	0.201

Table I.4: Coeficients of logistic regression with literature-based variables predicting any complications in Scenario 2 (right after surgery).

Marginal variables, Scenario 2, Any complications	
Variable	**Estimated coeficient**
(Intercept)	-1.696
Severe pulmonary disease 'yes'	0.213
ASA physical status class '3 or 4'	0.471
Preoperative WHO performance score '1'	1.095
Additional major procedures 'yes'	0.768
If open surgery or converted 'no'	-0.213
If blood 'yes'	0.0327
Resection site drainage 'yes'	0.285
Total IV volume of fluids intraoperatively	0.000132
Length of operation minutes	0.00252

Table I.5: Coeficients of logistic regression with marginal important variables predicting any complications in Scenario 2 (right after surgery).

AIC forward variables, Scenario 2, Any complications	
Variable	**Estimated coeficient**
(Intercept)	-1.109
Total IV volume of fluids intraoperatively	0.000377
If blood 'yes'	0.486
Infusion of vasoactive drugs 'yes'	-0.331
Intraoperative thrombosis prophylaxis 'yes'	-0.340
BMI '(21.5-24.9)'	-0.0448
BMI '(25-29.9)'	0.0429
BMI '(30 or more)'	0.0619

Table I.6: Coeficients of logistic regression with variables selected with AIC forward predicting any complications in Scenario 2 (right after surgery).

Scenario 3 with any complications

Scenario 3 predictes complications after primary stay when the patient is released from the hospital.

Literature variables, Scenario 3, Any complications	
Variable	**Estimated coeficient**
(Intercept)	-6.344
ASA physical status class '3 or 4'	0.115
Smoker 'yes'	-0.129
Severe heart disease 'yes'	0.293
Preoperative WHO performance score '1'	1.124
Age	0.000594
BMI '(21.5-24.9)'	-0.351
BMI '(25-29.9)'	0.0197
BMI '(30 or more)'	0.215
If predisease 'yes'	0.0883
Diabetes mellitus 'yes'	-0.961
Severe pulmonary disease 'yes'	-0.123
Gender 'female'	-0.178
Alcohol usage 'yes'	-0.198
Preoperative nutritional treatment 'yes'	-0.253
Preoperative oral carbohydrate treatment 'yes'	-1.043
Oral bowel preparation 'yes'	0.0916
Was aneamia found 'yes aneamia'	-0.343
Any nonsurgical preoperative treatment 'yes'	0.310
Previous surgery to same abdominal region 'yes'	-0.134
If colloids 'yes'	-0.653
Resectionsite drainage 'yes'	0.483
Depth of anaesthesia monitored 'yes'	-0.289
Infusion of vasoactive drugs 'yes'	-0.273
If lidocaine 'yes'	-0.545
Final diagnosis'cancer'	0.536
Final diagnosis 'no cancer or functional disorder'	0.741
Stomal procedure 'yes'	-0.727
Length of operation minutes	0.00195
Additional major procedures 'yes'	0.750
If open surgery or converted 'no'	-0.307
Main procedure name 'sigmoid resection'	0.342
Main procedure name 'stoma procedure'	1.548
Main procedure name 'anterior resection of the rectum'	0.558
Main procedure name 'uncommon procedure, no ileocaecal'	0.650
If blood 'yes'	-0.527
Total IV volume of fluids intraoperatively	0.00015
Core body temperature at end of operation	0.113
Mobilised on POD 1 or 2 'on POD 2'	0.272
Mobilised on POD 1 or 2 'after POD 2'	-0.132
Time to tolerating solid food nights '1 or more nights needed'	-0.00831
More than one night with urinary drain 'yes'	0.103
Time to pain control with oral analgesics nights '2 or more nights'	-0.227
IV volume postoperational	-0.0000535
Opioid use on postoperative day 1 'yes'	0.695
Max pain VAS in 3 days	0.157
CD at least 2 'yes'	0.492
Length of stay nights in hospital after primary operation	0.00314
High cancer stage 'yes'	-0.00533

Table I.7: Coeficients of logistic regression with literature-based variables predicting any complications in Scenario 3 (right after primary stay).

Marginal variables, Scenario 3, Any complications	
Variable	**Estimated coeficient**
(Intercept)	-2.861
Preoperative WHO performance score '1'	0.914
Additional major procedures 'yes'	0.603
If open surgery or converted 'no'	-0.140
Resection site drainage 'yes'	0.452
Opioid use on postoperative day 1 'yes'	0.621
PONV observed in 3 days 'yes'	0.498
Duration of IV fluid infusion nights 'removed on POD 1'	0.503
Duration of IV fluid infusion nights 'removed on or after POD 2'	0.0740
Total IV volume of fluids intraoperatively	-0.0000202

Table I.8: Coeficients of logistic regression with marginal important variables predicting any complications in Scenario 3 (right after primary stay).

AIC forward variables, Scenario 3, Any complications	
Variable	**Estimated coeficient**
(Intercept)	-3.129
PONV observed In 3 days 'yes'	0.634
Preadmission stoma counseling 'yes'	0.403
Opioid use on postoperative day 1 'yes'	0.731
Intraoperative thrombosis prophylaxis 'yes'	-0.13
Time to tolerating solid food nights '1 or more nights needed'	0.135
IV volume postoperational	-0.000183
Time to pain control with oral analgesics nights '2 or more nights'	-0.0809
Mobilised on POD 1 or 2 'on POD 2'	0.373
Mobilised on POD 1 or 2 'after POD 2'	0.00159
If colloids 'yes'	-0.789
Total IVvolume of fluids intraoperatively	0.000316

Table I.9: Coeficients of logistic regression with variables selected with AIC forward predicting any complications in Scenario 3 (after primary stay).

J

Estimated coefficients of weighted logistic regression models using literature based variables

Literature-based selection, Scenario 1, severe complications			
Variable	**Estimated coeficient**		
	equal (or non) weights	**fixed weights**	**balanced weights**
weight	1	2	8.92
(Intercept)	-3.4	-2.75	-1.48
ASA physical status class '3 or 4'	-0.0103	0.0284	0.146
Smoker 'yes'	-0.124	-0.0781	0.116
Severe heart disease 'yes'	-0.436	-0.401	-0.285
Preoperative WHO performance score '1'	1.08	1.04	0.992
Age	-0.00222	-0.00194	-0.00157
BMI '(21.5-24.9)'	-1.12	-1.11	-1.09
BMI '(25-29.9)'	-0.944	-0.916	-0.864
BMI '(30 or more)'	-0.618	-0.591	-0.52
If predisease 'yes'	0.581	0.539	0.401
Diabetes mellitus 'yes'	0.574	0.635	0.844
Severe pulmonary disease 'yes'	0.0118	0.00951	0.0574
Gender 'female'	-0.12	-0.125	-0.135
Alcohol usage 'yes'	0.273	0.258	0.238
Preoperative nutritional treatment 'yes'	-0.119	-0.094	-0.0474
Preoperative oral carbohydrate treatment 'yes'	1.65	1.64	1.75
Oralbowel preparation 'yes'	0.335	0.363	0.433
Was aneamia found 'yes anemia'	-0.125	-0.141	-0.177
Any nonsurgical preoperative treatment 'yes'	-0.114	-0.13	-0.214
Previous surgery to same abdominal region 'yes'	0.318	0.304	0.284

Table J.1: Estimated coefficients of logistic regression (right), weighted logistic regression with fixed weight (middle) and balanced weighted logistic regression (right) on literature-based variables predicting severe complications in Scenario 1 (before surgery).

Literature-based selection, Scenario 2, severe complications			
Variable	**Estimated coeficient**		
	equal (or no) weights	**fixed weights**	**balanced weights**
weight	1	2	8.92
(Intercept)	-0.144	-0.204	-2.39
ASA physical status class '3 or 4'	0.0746	0.138	0.26
Smoker 'yes'	0.00088	0.00637	0.0702
Severe heart disease 'yes'	-0.447	-0.365	-0.0646
Preoperative WHO performance score '1'	1.03	0.897	0.705
Age	0.00782	0.00874	0.0155
BMI '(21.5-24.9)'	-1.23	-1.27	-1.58
BMI '(25-29.9)'	-1.22	-1.21	-1.35
BMI '(30 or more)'	-0.645	-0.672	-0.887
If predisease 'yes'	0.68	0.615	0.554
Diabetes mellitus 'yes'	0.614	0.483	0.16
Severe pulmonary disease 'yes'	-0.327	-0.252	-0.0626
Gender 'female'	-0.0793	-0.103	-0.207
Alcohol usage 'yes'	0.319	0.272	0.197
Preoperative nutritional treatment 'yes'	-0.478	-0.437	-0.42
Preoperative oral carbohydrate treatment 'yes'	1.74	1.62	1.47
Oral bowel preparation 'yes'	0.069	0.109	0.214
Was aneamia found 'yes aneamia'	-0.237	-0.245	-0.178
Any nonsurgical preoperative treatment 'yes'	-0.28	-0.312	-0.501
Previous surgery to same abdominal region 'yes'	-0.136	-0.079	-0.0641
If colloids 'yes'	0.291	0.263	0.196
Resection site drainage 'yes'	0.55	0.459	0.364
Depth of anaesthesia monitored 'yes'	0.0864	0.107	0.102
Infusion of vasoactive drugs 'yes'	-1.08	-1.04	-1.09
If lidocaine 'yes'	-0.664	-0.703	-0.882
Final diagnosis 'cancer'	0.155	0.309	0.565
Final diagnosis 'no cancer or functional disorder'	0.396	0.535	0.849
Stomal procedure 'yes'	-0.944	-0.914	-0.859
Length of operation minutes	-7.15e-05	0.00112	0.0047
Additional major procedures 'yes'	0.394	0.369	0.348
If open surgery or converted 'no'	-0.0709	0.00858	0.258
Main procedure name 'sigmoid resection'	-0.375	-0.359	-0.31
Main procedure name 'stoma procedure'	1.7	1.81	2.19
Main procedure name 'anterior resection of the recturm'	1.32	1.27	1.26
Main procedure name 'uncommon procedure, no ileocaecal'	0.553	0.604	0.694
If blood 'yes'	0.905	0.786	0.466
Total IV volume of fluids intraoperatively	0.000314	0.000324	0.000411
Corebody temperature at end of operation	-0.118	-0.105	-0.0313

Table J.2: Estimated coefficients of logistic regression (right), weighted logistic regression with fixed weight (middle) and balanced weighted logistic regression (right) on literature-based variables predicting severe complications in Scenario 2 (right after surgery).

Literature-based selection, Scenario 3, severe complications			
Variable	**Estimated coeficient**		
	equal (or no) weights	**fixed weights**	**balanced weights**
weight	1	2	22.6
(Intercept)	-12.5	-12.1	-21.5
ASA physical status class '3 or 4'	0.399	0.216	-0.783
Smoker 'yes'	0.297	0.285	0.586
Severe heart disease 'yes'	-0.266	-0.358	-1.74
Preoperative WHO performance score '1'	-0.293	-0.104	0.536
Age	-0.00749	-0.00648	-0.00166
BMI '(21.5-24.9)'	-1.85	-1.91	-3.05
BMI '(25-29.9)'	-1.8	-1.64	-1.26
BMI '(30 or more)'	-0.641	-0.445	0.609
If predisease 'yes'	0.913	1.04	2.82
Diabetes mellitus 'yes'	1.39	1.32	1.33
Severe pulmonary disease 'yes'	-0.0866	-0.216	-1.53
Gender 'female'	0.291	0.33	0.778
Alcohol usage 'yes'	0.121	0.136	0.485
Preoperative nutritional treatment 'yes'	-0.0931	-0.0265	0.000663
Preoperative oral carbohydrate treatment 'yes'	1.89	2	3.92
Oral bowel preparation 'yes'	-1.5	-1.59	-3.2
Was aneamia found 'yes' aneamia	0.395	0.414	0.808
Any nonsurgical preoperative treatment 'yes'	1.02	0.952	1.68
Previous surgery to same abdominal region 'yes'	-0.278	-0.221	-0.0417
If colloids 'yes'	-1.54	-1.62	-3.25
Resectionsite drainage 'yes'	0.365	0.426	1.02
Depth of anaesthesia monitored 'yes'	0.108	0.0397	-0.197
Infusion of vasoactive drugs 'yes'	-0.581	-0.611	-1.05
If lidocaine 'yes'	-2.2	-2.26	-4.05
Final diagnosis 'cancer	-0.59	-0.465	0.109
Final diagnosis 'no cancer or functional disorder'	-0.449	-0.367	0.824
Stomal procedure 'yes'	-2.04	-1.84	-2.13
Length of operation minutes	-0.0049	-0.00547	-0.0167
Additional major procedures 'yes'	0.295	0.273	-0.0278
If open surgery or converted 'no'	1.4	1.51	3.33
Main procedure name 'sigmoid resection'	1.32	1.31	2.2
Main procedure name 'stoma procedure'	3.18	3	3.83
Main procedure name 'anterior resection of the rectum'	2.19	2.25	3.68
Main procedure name 'uncommon procedure, no ileocaecal'	1.09	1.03	1.34
If Blood 'yes'	0.0869	0.0207	-0.132
Total IV volume of fluids intraoperatively	0.000703	0.000733	0.00168
Core body temperature at end of operation	0.125	0.126	0.298
Mobilised on POD 1 or 2 'on POD 2'	0.597	0.581	0.836
Mobilised on POD 1 or 2 'after POD 2'	2.11	2.05	3.13
Time to tolerating solid food nights 'yes'	1.24	1.28	2.34
More than one night with urinary drain 'yes'	0.985	1	1.67
Time to pain control with oral analgesics nights '2 or more nights'	-2.04	-1.92	-2.54
IV volume postoperational	-0.00149	-0.00147	-0.0021
Opioid use on postoperative day 1 'yes'	2.89	2.8	3.38
Max Pain VAS in 3 days	-0.075	-0.0675	-0.13
CD at least 2 'yes'	-1.3	-1.21	-1.27
Length of stay nights in hospital after primary operation	0.0299	0.0292	0.0364
High cancer stage 'yes'	-0.119	-0.134	-0.191

Table J.3: Estimated coefficients of logistic regression (right), weighted logistic regression with fixed weight (middle) and balanced weighted logistic regression (right) on literature-based variables predicting severe complications in Scenario 3 (after primary stay).