Co-Morbidity in Bariatric Patients

A new approach in quantifying the severity

T.P.A. Beishuizen (0791613)
Biomedical Engineering - Computational Biology
Data Engineering - Information Systems
Eindhoven, University of Technology
Email: t.p.a.beishuizen@student.tue.nl

November 2, 2017

Contents

1 Introduction **2**
1.1 Data Sets 2

2 First Proposal **4**

3 Proposal Comparisons **4**

4 Final Proposal **4**

1 Introduction

Worldwide the number of bariatric surgeries is increasing. Although initially thought otherwise, this type of surgery has added benefits on top of losing weight, the primary reason. Among those benefits the remission of metabolic co-morbidities can be named. Due to binary labelling of those co-morbidities, valuable information is lost, while this labelling is not clearly defined either. To obtain more and better results, this binary labelling can be replaced by a continuous severity score. Ruben Deneer conducted a research on trying to achieve a successful replacement.[1]

The available data used for the research stemmed from the Catharina Hospital in Eindhoven. This extensive data set consisted of 41 markers measured pre- and post-surgery for 2367 patients that underwent gastric sleeve or gastric bypass surgery. These 41 markers are divided in several sub-panels that each describe different processes in the body. Conditions for these co-morbidities that were tested for the severity score were type diabetes mellitus (T2DM), hypertension and dyslipidemia. Extensive literature research was done to connect them with 41 markers to find possible relations. [1] The data sets are described in more detail in subsection **??**

To better specify the research on the data, a main goal was created: *to use data mining techniques to develop score that can objectively quantify the severity of co-morbidities present in bariatric patients based on biomarkers, both before and after surgery.* This means that it is not the goal to predict the outcome of the two types of surgery, but it is to quantify the improvement of the co-morbidities before and after the surgery. [1]

Since the data is not available other than within the hospital, a figurative second study was conducted to complement Ruben Deneer's research. This study consists of three parts. The first part was creating a research question, hypothesis and an own analysis approach without the knowledge of Deneer's way. Since no data is available, this can only be made globally. It should be detailed enough, though, to compare it with Deneer's own approach. Secondly the two are compared and remarks are given on the original approach. At last both are combined to create a final proposal for a possible follow-up research.

1.1 Data Sets

Two data sets were used in the study. The first one is called "The Dutch Audit for Treatment of Obesity" (DATO). This data set is a national database that houses all registrations and health statuses of pre- and post treatment bariatric surgery patients in the Netherlands. Before surgery, the co-morbidities were given a binary label of "Yes/No". The co-morbidities after surgery were given one of the following labels:

1. **Cured** No co-morbidity any more
2. **Improved** Less affected by co-morbidity
3. **Same** No change in co-morbidity status
4. **Worse** More affected by co-morbidity
5. **Denovo** Diagnosed co-morbidity while not present before surgery
6. **Not present** No co-morbidity present

The second data set came from a laboratory database, stored in health records. This extensive data set consisted of 3 clinical and 38 blood markers measured pre- and 6, 12 and 24 months post-surgery. The tests pre-surgery had some additional markers on top of the 41 ones. These markers can be divided in the following categories: (Table **??**) Complete blood count, liver function, kidney function, inflammation, lipid spectrum, coagulation, glucose metabolism, thyroid function and at last minerals and vitamins. The data sets of the patients that underwent bariatric surgery can be extracted from these.

Table 1: The markers present in the bariatric laboratory data set [1]

	Before Surgery/Pre-Op/Screening	After Surgery/Post-Op/Follow-up
Complete blood count	hemoglobin	hemoglobin
	hematocrit	hematocrit
	erythrocytes	erythrocytes
	mean corpuscular hemoglobin	mean corpuscular hemoglobin
	mean corpuscular volume	mean corpuscular volume
	thrombocytes	thrombocytes
	leukocytes	leukocytes
Liver function	bilirubin	bilirubin
	aspartate aminotransferase	aspartate aminotransferase
	alanine aminotransferase	alanine aminotransferase
	lactate dehydrogenase	lactate dehydrogenase
	alkaline phosphatase	alkaline phosphatase
	gamma-glutamyltransferase	gamma-glutamyltransferase
Kidney function	urea	urea
	creatinine	creatinine
	potassium	potassium
	sodium	sodium
	calcium	calcium
	phosphate	phosphate
	albumin	albumin
Inflammation	C-reactive protein	C-reactive protein
Lipid spectrum	cholesterol	total cholesterol
	high-density lipoprotein-cholesterol	high-density lipoprotein-cholesterol
	total total/high-density cholesterol ratio	total/high-density cholesterol ratio
	low-density lipoprotein-cholesterol	low-density lipoprotein-cholesterol
	triglycerides	triglycerides
Coagulation	prothrombin time	prothrombin time
Glucose metabolism	hemoglobin A1c (IFCC)	hemoglobin A1c (IFCC)
	glucose	glucose
	insulin	-
	C-peptide	-
Thyroid function	parathyroid hormone	parathyroid hormone
	thyroid-stimulating hormone	-
	free T4	-
	cortisol	-
Minerals and vitamins	iron	iron
	ferritin	ferritin
	folic acid	folic acid
	zinc	-
	magnesium	-
	vitamin A	-
	vitamin B1	vitamin B1
	vitamin B6	vitamin B6
	25-OH vitamin D	25-OH vitamin D
	vitamin B12	vitamin B12

A smaller part of the research is to combine these two data sets. Some challenges arise when doing so. Such a challenge is obviously to find the right connection between them, using the survey and lab data of the same patient. What can the markers say about the severity of co-morbidities?

A second challenge would be to define what to do with non-matching data. In the pre-treatment for example more markers were used than in the post-treatment. These missing ones might be more useful for scoring the co-morbidity severity than the known markers. There also might be measurements that were missing or corrupted, whereas other markers might still be useful enough for a result. How to tackle this missing data challenge should be defined properly.

2 First Proposal

3 Proposal Comparisons

4 Final Proposal

References

[1] R. Deneer, "Scoring co-morbidity severity in bariatric patients based on biomarkers: a data mining approach," 2017.