MASTER

An empirical study of the impact of self-management and digital communication on diabetes patient’s health

Bos, M.D.W.

Award date:
2018

Link to publication
An empirical study of the impact of self-management and digital communication on diabetes patient’s health

Master Thesis

M.D.W. (Matthias) Bos BSc.

Supervisors:
- dr. M. (Murat) Firat – TU/e
- dr. P.M.E. (Pieter) Van Gorp – TU/e
- dr.ir. H. (Rik) Eshuis – TU/e
- H.R. (Hero) Torenbeek – VitalHealth Software

Final Version

Eindhoven, May 22nd, 2018
TUE. School of Industrial Engineering.
Series Master Thesis Operations Management and Logistics

Keywords: eHealth, Self-Management, Self-Management Platform, Patient-HCP Interaction, Digital Communication, Diabetes Mellitus Type 2, DM2
Abstract

Diabetes is a chronic disease and a risk factor for other diseases. The number of diabetic patients is increasing over the years. More than 400 million people had diabetes in 2014 already. In 2016, five percent of the Dutch citizens had diabetes.

Within diabetes, a distinction can be made in several types, from which only Type 2 Diabetes (DM2) is considered in this study. Some of the risk factors for getting DM2 are a low physical activity and high-calorie diets. Symptoms that occur by patients having DM2 are frequent urination, thirst, blurred vision, and more frequent or slower-healing infections. Many patients do not need medication, since a change in lifestyle is an important factor to heal from DM2. Although after a while, people start using medication pills and insulin.

HbA1c is one of the blood pigments and supports the oxygen transfer in the blood. It binds glucose and indicates the average blood glucose of the past months and is normally measured each quarter by analyzing blood samples in a laboratory. A value below 53 mmol/mol is recommended (for patients with an age below 70), although this value is adapted when the patient gets older. In Dutch general practices a physician assistant (HCP) is working who supports the diabetic patients, among others. Patients visit most of the four general practitioner visits a year this HCP. In between those visits, the patient has to monitor his blood glucose level by himself. Monitoring is important, because a too high blood glucose level will harm patient’s health.

Therefore, it is important that patients reach a certain level of self-management. This level indicates to what extent the patient is able to participate in his treatment and several measures exist for this level. Moreover, several self-management platforms exist to support the patient to interact with his HCP, perform assigned tasks, and to learn how to manage his disease.

These interactions have an impact on patients’ health; so besides health expense savings, also people’s quality of life will improve when one becomes healthier. A few studies showed that self-management platforms can have a positive impact on patient’s health. In these studies patients knew they were participating in a research which could have influenced the results. In addition, these studies were performed with a limited number of participants, both patients and HCPs, during a relatively short period of time.

This study includes more participants and data from a longer time period. These participants did not know they were participating in a research (although they gave consent to analyze their anonymized data). This study followed the standard CRISP-DM process for the data mining part to answer the research questions, which were defined as follows:
RQ1 What measure can be used for the level of self-management?

RQ2 In which way is the interaction between the HCP and patient related to the level of self-management and how could it be improved?

RQ3 How does the level of self-management influence patient’s health level?

RQ4 How does the patient-HCP interaction influence patient’s health level?

For the estimation of patient’s self-management level an own measure, the quantifier SML, was derived with the normalized values of patient’s login frequency and percentage of finished tasks. The details of this derivation are described in section 2.2.3.

A data set with 533 patients and 59 HCPs was exported. For not all patients having (enough) HbA1c measurement values, a sub set of 105 patients that had at least one measurement before using the self-management platform, at least two measurements after starting with the self-management platform, an age below 70, and that did not use insulin was used. This sub set was split into a training (60%) and a test set (40%) for the creation of prediction models.

The chat messages were split into chat statements which were classified into nine chat statement categories. One of these categories consisted of emotional social support (ESS) statements, e.g. “Call me if you need anything”. Patients have used the platform on average for more than one and a half year and at maximum almost two and a half year. They had on average an age of 58.5 and logged in on average approximately once per four weeks.

Patients had been assigned to multiple tasks and they did not finish all of them. Some patients had been assigned to more than 30 tasks, while many others had been assigned to less than ten tasks. HCPs logged in on average approximately once each six days and had on average 9.2 patients with DM2 and used the self-management platform.

Patient-HCP interaction was quantified as the frequency of receiving chats, (HCP) login frequency, (HCP) sent-received chat ratio, and the chat content. It was found that patients with more patient-HCP interaction have a higher self-management level.

Multiple methods for building a prediction model exist. Because the dependent variables in this study – the change in HbA1c level and the SML – are continuous variables (and the other variables are numerical), the prediction model can be built with regression. Methods that only predict categorical dependent variables were not applicable in this study. For the case of estimating a continuous variable is chosen to use simple linear regression, multiple linear regression, and decision tree regression.

With the use of HCP grouping and linear regression (using a training set for building the model and a test set for validation) was found that the patients who receive more than 0.005 chat messages per day have on average a significantly higher self-management level.

It was found that patients have on average a decrease in HbA1c level during the first, approximately, fifteen months after being diagnosed with diabetes. However, after these fifteen months their HbA1c level starts increasing. Therefore, the 105 patients of the data set were grouped into three groups based on patients’ self-management level. It was found that during the first (approximately) six months the average percentual change in HbA1c level of the three groups was significantly different. In addition, the percentual
change during the entire period that patients used the self-management platform was
different over the three groups. Although only the medians were significantly different,
means of +12.0%, +1.6%, and -5.4% were found for respectively group low, medium, and
high self-management level. With the use of linear regression was found that a higher self-
management level results in an improvement in health, which is quantified as a decreasing
HbA1c level.

Next, the relation between patient-HCP interaction and change in HbA1c level was in-
vestigated by creating prediction models. Simple constant models, linear models, and
regression trees were built with the training set. With the test set these models were
validated based on the mean absolute and mean squared error (MAE and MSE).

In conclusion can be said that the more frequent an HCP sends a chat message (and
the patient receives more often a chat message), the higher patient’s self-management
level becomes. Secondly, the more chats with social integration support statements a
patient receives, the higher patient’s self-management level. Thirdly, the higher patient’s
self-management level, the more decrease in HbA1c level. Fourthly, the more chats with
esteem social support statements the patient receives, the lower his HbA1c level becomes.
The more decrease in HbA1c level, the more improvement in patient’s health.

This study was able to show that increased used of eHealth was related with a lower
HbA1c level based on data of 105 patients who have used a self-management platform
for a period of more than one year.

It is suggested to extend this research with more data of the patients and care groups,
such as social-demographic details, implementation, lifestyle, and vital parameters such
as BMI and blood pressure, to improve the prediction models. In addition should be
examined whether patient’s health improvement remains during a longer period of time
(of at least two years with patients having at least five measurements during the use of
the self-management platform) and for that study it would be useful to be able to contact
the participants.

Finally, with these results cannot be concluded that using a self-management platform
results in getting a better health – because patient’s health cannot be defined in an
easy way and changing lifestyle is not only done by making a self-management platform
available to the patient – although these platforms seem to help patients to learn about
improving lifestyle and becoming dedicated to work on it by having a decrease in HbA1c
level and an improving health.
Preface

At the moment of writing this chapter, I almost finished my studies at the Eindhoven University of Technology. This report is the result of a Master Thesis project, performed to receive a Master of Science degree in Operations Management and Logistics at the Eindhoven University of Technology and is conducted at VitalHealth Software.

During the period of my studies I learned a lot by taking courses, working in committees of study and student associations, performing two board years, going on exchange, and by a lot more. With this thesis, I finish my student life, but before that ends I would like to thank several people.

Firstly, I would like to thank my supervisors from university. Murat Firat, you were my first supervisor during the final phase of my studies. We had a lot of discussions about the purpose of my research and you were not afraid to discuss details too. Thank you for guiding me, your concerns, and dedication during this final phase of my studies. Pieter Van Gorp, my second supervisor and former mentor, thank you for guiding me during my studies, providing feedback and for mentioning and getting me in contact with VitalHealth Software. Thanks to Rik Eshuis for completing the Assessment Committee.

Secondly, I give thanks to my company supervisor Hero Torenbeek. It was a great pleasure having you as supervisor. You gave great feedback regarding the meaning of the results and kept me focusing on what found results mean for patients. You were enthusiastic and always made time free for me to answer my questions. Thank you!

Thirdly, thanks to my colleagues from VitalHealth Software. In particular thanks to Michiel for your advice, ideas, opportunities, and for getting me into contact with interesting people. I would like to thank all my other colleagues as well for their accompany, laughing, patience and their willingness to support me where and whenever needed.

Moreover, I am really grateful to the (former) doctors I have spoken with, the care group boards that gave access to data, and the HCPs and patients that allowed me to analyze their – anonymized – data.

Lastly, friends (e.g. LunchDate) and family, thank you for your support and all the things we did over the years.

Matthias Bos
Contents

Contents ix

List of Figures xiii

List of Tables xv

1 Introduction 1
  1.1 Diabetes ......................................................... 1
  1.2 Self-Management and Health .................................. 2
  1.3 Self-Management Platform .................................... 4
    1.3.1 Example Self-Management Platform ..................... 4
    1.3.2 Patient-HCP Communication and Interaction .......... 7
    1.3.3 Impact on Patient’s Health ............................ 7
  1.4 Research Questions ........................................... 9

2 Methodology 11
  2.1 Process for Data Mining ....................................... 11
  2.2 Data Preparation ............................................... 12
    2.2.1 Patient Attributes ....................................... 13
    2.2.2 HCP Attributes ........................................... 14
    2.2.3 Quantify Self-Management Level with Aggregation .... 14
    2.2.4 Chat Statements .......................................... 17
    2.2.5 HCP Grouping ............................................. 18
    2.2.6 Patient Grouping ......................................... 18
  2.3 Performed Analyses ............................................ 19
  2.4 Evaluation of Prediction Models ............................. 20
  2.5 Questionnaire and Interview .................................. 21
Appendices

A Sample Sizes Previous Research

B Questionnaire

B.1 Questions
B.2 Results
B.3 Questionnaire Patient
B.4 Questionnaire HCP

C Data Description

D Extra Details Results
List of Figures

1.1 Patient view e-Vita ................................................. 6
1.2 Patient view measurements e-Vita ................................. 6
1.3 Model with expected (positive) relations ........................ 9
2.1 CRISP-DM .......................................................... 12
2.2 Patient’s measurement values over time .......................... 14
3.1 Histogram age of patients ........................................... 24
3.2 Login frequency of patients ......................................... 24
3.3 Patient usage days of e-Vita per care group ....................... 25
3.4 Days between first login and $A_m$ ................................. 25
3.5 Distribution number of open tasks ................................. 25
3.6 Scatter plot SML and fraction of chats containing SIS statements .......................... 26
3.7 Scatter plot SML and fraction of chats containing SD statements .......................... 26
3.8 Scatter plot HbA1c levels $B_n$ .................................... 26
3.9 Scatter plot HbA1c levels $A_2$ .................................... 26
3.10 Scatter plot HbA1c levels $A_m$ ................................... 26
3.11 Scatter plot percentual change HbA1c and received chat frequency .................. 27
3.12 Scatter plot percentual change HbA1c and HCP login frequency .................. 27
3.13 Scatter plot percentual change HbA1c and number of days between first login and $A_m$ .................................................. 27
3.14 Scatter plot percentual change HbA1c and SML .................. 27
3.15 Scatter plot percentual change HbA1c and ISS ChatPerc ........ 28
3.16 Scatter plot percentual change HbA1c and ESSup ChatPerc ....... 28
3.17 Scatter plot percentual change HbA1c and HIR ChatPerc ....... 28
3.18 Scatter plot percentual change HbA1c and SD ChatPerc .......... 28
LIST OF FIGURES

3.19 HCP usage days of e-Vita per care group ........................................ 29
3.20 HCP login and sent chat frequency .................................................. 29
4.1 Boxplot quantifiers of self-management level ....................................... 32
4.2 Boxplot quantifiers of self-management level - zoomed in ................... 32
4.3 Boxplot SML; grouped on HCP’s sent chat frequency .......................... 33
4.4 Boxplot SML; grouped on HCP’s sent chat frequency, zoomed in .......... 33
4.5 Boxplot SML; grouped on received chat frequency .............................. 34
4.6 Boxplot SML; grouped on received chat frequency, zoomed in ............ 34
4.7 HbA1c levels $B_1 - B_{16}$ ................................................................. 37
4.8 HbA1c levels $B_n - B_n$ ............................................................... 37
4.9 Boxplot with percentual HbA1c change on $B_n - A_m$ per patient group .. 38
4.10 Boxplot with percentual HbA1c change on $B_n - A_2$ per patient group .. 38
4.11 Average HbA1c level on $B_n - A_m$ per patient group ......................... 38
4.12 Average HbA1c level on $B_n - A_2$ per patient group ......................... 38
4.13 Boxplot with login frequency per patient group, based on SML .......... 39
4.14 Boxplot with percentage of tasks finished per patient group, based on SML 39
4.15 Boxplot with sent chat frequency per patient group, based on SML ...... 39
4.16 Boxplot with received chat frequency per patient group, based on SML .. 39
4.17 Boxplot with number of active days per patient group, based on SML .................. 40
4.18 Boxplot with number of inactive days per patient group, based on SML .. 40
4.19 Regression tree - no chat content .................................................... 43
4.20 Regression tree ............................................................................. 43
D.1 Age and gender of patients with their login frequency ....................... 69
List of Tables

2.1 Patient attributes .................................................. 13
2.2 HCP attributes ..................................................... 14
3.1 Descriptive data from HCPs ......................................... 30
4.1 Comparison of models ................................................ 35
4.2 Linear model to predict HbA1c change with SML ............... 40
4.3 Comparison of models ................................................ 44
A.1 Summary literature self-management platforms and health .... 60
D.1 Descriptive data from patients ..................................... 70
D.2 Descriptive data chat statements .................................. 71
D.3 Chat statement coding categories ................................. 72
D.4 Linear model to predict SML with patient’s received chat content .... 72
D.5 Linear model to predict SML with patient’s received chat content .... 72
D.6 Linear model to predict SML, forward approach ................ 72
D.7 Linear model to predict SML, backward approach .............. 73
D.8 Linear model to predict HbA1c change with DaysFLAm .......... 73
D.9 Linear model to predict HbA1c change with HCPLloginFreq + ESSup ChatPerc ............................................. 73
D.10 Linear model to predict HbA1c change with HCPLloginFreq + \{TSS + TS + ESSup\} ChatPerc + SML ............................ 73
Chapter 1

Introduction

Both diabetes and self-management are key topics of this study and these topics are introduced in this chapter. This is followed by a description of self-management platforms with an example of a platform and the impact on patients’ health. This chapter is concluded with the current gap in the literature, from which the research questions are drawn.

1.1 Diabetes

Diabetes is a chronic disease and is a risk factor for eye and kidney diseases, along with having diabetes is a risk factor for getting other bad health outcomes as well (Gerstein & Haynes, 2001, p. 62). In 2011, 2015, and 2016 respectively 2.8%, 4.8%, and 5.0% of the Dutch people had diabetes (CBS, 2017, 2016). Compared with other European countries these percentages are relatively low, although the Dutch expenses for diabetic care were 1.7 billion euros in 2011 already (CBS, 2016; Volksgezondheidenzorg.info, n.d.). In the USA, 8% of the people older than 20 years had diabetes in 2011 (American Diabetes Association, 2011, p. 6). It is estimated by the International Diabetes Federation in 2010 that 430 million people will have diabetes by 2030, but in 2014 there were already 422 million people with diabetes (World Health Organization, 2017; American Diabetes Association, 2011, p. 10). Therefore, the number of diabetic patients seems increasing even more than expected.

A distinction can be made in Type 1 (DM1), Type 2 Diabetes (DM2), gestational diabetes, and other types of diabetes (Ahmad, 2013, p. 12-13). People with DM1 are insulin-dependent as a result of genetic triggers that destruct the beta-cells (Codario, 2011, p. 13). DM2 is also called non-insulin-dependent diabetes and more than 90% of the diabetic patients has this type (Codario, 2011, p. 13). Most people with DM2 are diagnosed after the age of 30. A low physical activity, and high-calorie diets (overweight) are some of the risk factors for getting DM2 (Codario, 2011, p. 16).

Many patients with DM2 do not need medication, but they have to change lifestyle by becoming more physically active and by eating healthier. Common symptoms are frequent urination, increased thirst, fatigue, blurred vision, and more frequent or slower-healing infections (American Diabetes Association, 2011, p. 34-43).
HbA1c is one of the blood pigments and supports the oxygen transfer in the red blood cells. HbA1c binds glucose and the level of HbA1c indicates the average blood glucose of the past two to three months. This parameter is normally measured each quarter by analyzing blood in a laboratory (Stöppler, Melissa C, 2018; Diaboss, n.d.; Iqbal, Morgan, Maksoud & Idris, 2008; Saudek & Brick, 2009). This value is recommended being below 53 mmol/mol (7%) by the American Diabetes Association (ADA), the American Heart Association (AHA), the Dutch College of General Practitioners, and Diabetes Fonds, among others (NHG, 2013; Stöppler, Melissa C, 2018; Diabetes Fonds, n.d.; Ahmad, 2013, p. 13).

In most general practices, a physician assistant (Dutch: “PraktijkOndersteuner Huisartsenzorg” (POH)) with a medical background is working, and will be referred to as HCP in this thesis. Patients with Diabetes are visiting the general practice (on average) four times a year for a quarterly check. Most of the visits done by a patient are with the HCP.

In between those visits, patients have to monitor themselves. The frequency should be aligned with the HCP and may variate between once a week till eight times a day depending on patients’ health, goals, and willingness. Monitoring the glucose levels helps the patient and the HCP to evaluate the progress and to adapt the treatment (American Diabetes Association, 2011, p. 56-65). This monitoring can be done with lancets, test strips and blood glucose meters. With this, logbooks can be made to understand the blood glucose levels. There exist continuous glucose monitors that can be placed below the skin. By continuous measuring, patients can set alarms for bad values and this may be beneficial for them (American Diabetes Association, 2011, p. 66-80).

In the remainder of this thesis only patients with DM2 are considered.

1.2 Self-Management and Health

Health is defined and not amended since 1946 as “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity” and therefore the health level of a patient cannot be described with only vital parameter values (WHO, 2018; WHO, 1946, p. 100).

Self-management is the active participation of a patient in his treatment (Lorig & Holman, 2003). So, the self-management level indicates to what extent the patient is able to participate and participates in his treatment. Patient’s confidence (self-efficacy) to be able to reach a goal is important in self-management (Bodenheimer, Lorig, Holman & Grumbach, 2002).

More and more mobile apps are developed to support patient’s self-management resulting in some clinical improvements (Hood et al., 2016). In some of these apps gamification is used to increase patient’s number of measurements (Cafazzo, Casselman, Hamming, Katzman & Palmert, 2012). In addition, self-management is supported by Internet delivered education in which online training and information is given to improve patient’s knowledge and glycemic control (Barlow, Wright, Sheasby, Turner & Hainsworth, 2002; Pereira, Phillips, Johnson & Vorderstrasse, 2015).
Self-management becomes more important to manage patients' health in a time where health care costs are increasing (CBS, 2016). Self-management programs can improve health of patients and reduce hospitalizations and therefore decrease health care costs (Lorig et al., 1999, 2001). Therefore, it is important to make good use of the currently existing possibilities of self-management platforms or even improve those platforms. By monitoring the vital values over time, exacerbations of chronic ill patients could be prevented. Moreover, a decreasing disease burden may support patients and their environment when exacerbations are prevented. These exacerbations are expensive, so prevention results in cost savings and the patient does not have to go to the hospital (Miravitlles et al., 2013; Wouters, 2003).

Literature review and meta-analysis was done by Broadbent et al. (2015) and they found that worse blood glucose levels are associated with low health status and expected bad consequences. The illness perception in the patient sample of 200 patients of Lee, Teh, Malar, Ong and James (2017) is significantly correlated with the HbA1c level (as metabolic control). Better treatment outcomes (lower HbA1c levels) will follow by better communication between patients and HCPs (Lee et al., 2017). Perceived control over patients’ disease was investigated by Lange and Piette (2005) in a sample of more than 600 patients. They found that perceived control is associated with the HbA1c level, where the higher the HbA1c level, the lower the perceived control. On top of that, comorbidities and acute symptoms were correlated as well.

There are models and questionnaires that describe and/or measure the level of health or self-management. These are e.g. Schermers’ categories and the Patient Activation Measure (PAM). According to Elissen et al. (2013) the level of self-management refers to the active participation of patients. In the next paragraphs two measures are described, although in this study an own measure will be defined and used for not having the data and not being able to contact the patients.

**Schermer**

Schermer (2009) defined the level of self-management in three levels. In the first level the patient is able to perform some practical tasks to support the HCP. These tasks are especially measurement tasks. At this level, the patient does not make his own decisions regarding health management. At the second level the patient is able to perform the tasks as in the first level, but now he understands the data from the measurements and can decide what the best is for him in medical perspective, based on HCP’s input. Once a person reaches the third level, he can find his own way of living with the condition. Choices that are made are not always the best in medical perspective, but it might increase patient’s health (Schermer, 2009).

**Patient Activation Measure**

Another measure for self-management level is the Patient Activation Measure (PAM). From this measure the stage of activation will be the output. These stages are “Believes Active Role Important, Confidence and Knowledge to Take Action, and Staying the Course under Stress” (Hibbard, Stockard, Mahoney & Tusler, 2004; Hibbard, Mahoney, Stockard & Tusler, 2005). The higher the level, the more the patient understands his disease and the more he is being active on working on his health. The statements in the PAM questionnaire are related to the knowledge a patient has (e.g. “I know what each of my prescribed medications do”), the confidence of being able to handle the disease (e.g. “I am confident I can figure out solutions when new situations or problems arise with my
health condition”), and to be active in maintaining lifestyle changes (e.g. “I have been able to maintain the lifestyle changes for my health that I have made”) (Hibbard et al., 2005). In the third level, the patient is taking action and is further developing his self-management skills (Insignia Health, n.d.). The activation levels of PAM are significantly related to the self-management level of patients (Hibbard et al., 2004). Patients with a higher PAM level are more likely to keep track of a glucose logbook (Hibbard et al., 2004) and therefore, they are willing to perform more activities that may result in a better health.

1.3 Self-Management Platform

A self-management platform is a platform that supports patient’s self-management. Platforms exist for e.g. enhancing medication adherence, communicating with HCPs and monitoring the disease (Lakshminarayana et al., 2017; McMahon et al., 2005). Results of the use of such platforms in terms of health outcomes are described in section 1.3.3. In section 1.3.1 an example of such a self-management platform is described. Moreover, the data used in this thesis is extracted (after users’ consent) from this platform.

1.3.1 Example Self-Management Platform

In this section an example of a self-management platform is given. First, information is given about the company that made the platform. Secondly, the self-management platform e-Vita is described.

VitalHealth Software

VitalHealth Software (VHS) is founded by Mayo Clinic and Noaber Foundation in 2006. The company is located in India, Germany, the Netherlands, and the United States of America. VHS has the mission to improve “the health of millions of people through eHealth solutions” (VitalHealth Software, n.d.). Their focus is on patients with chronic diseases, such as diabetes, COPD, heart failures, depression, cancer, and Alzheimer’s disease. The company is growing over the past years in number of employees and locations and is acquired by Philips at the end of 2017 (Philips, 2017; VitalHealth Software, 2017).

For realizing the mission and vision, VitalHealth made (and develops) several products like VitalHealth Collaborative Health Management (CHM), VitalHealth QuestManager (for measuring effects and making diagnoses), VitalHealth e-Vita (for patient engagement), and VitalHealth Insight (for giving insight into the data of the HCP). In the next subsection the self-management platform, e-Vita, is described.
VitalHealth e-Vita

VitalHealth e-Vita is the product of VitalHealth Software that engages patients. Only parts that are relevant for this research are described in this section; it is not a complete description of all features and data of e-Vita. Patients can see their history of lab results, can communicate safely with their HCPs, can enter self-measurements, can/have to perform tasks that are allocated by the HCP, and patients can read in the library to get more knowledge. The tasks that an HCP can assign to a patient are related to reading and learning about their disease, entering measurement values, testing their knowledge by e.g. sending a questionnaire, and by setting goals for example. VitalHealth e-Vita is used as addition to VitalHealth CHM where CHM is an addition to the general practitioner information system (Dutch: huisartsinformatiesysteem (HIS)). Information from CHM is synchronized to e-Vita and vice versa. A patient is first included in CHM and when the HCP decides to include the patient in e-Vita the patient can access e-Vita. From that point in time, the patient has all the benefits of having a self-management platform.

There are two ways in which the HCP can enter the platform of e-Vita. The first option is via logging in directly into e-Vita. From that point, one sees the patient list with an overview of one’s patients, when they logged in for the last time, how many tasks are still open, and whether there are new alerts and measurements. The second more often used option is a login via another system (CHM) one is using. If one enters e-Vita via that way, one directly goes to the overview of a specific patient. In this overview are several components shown. These components are the vitals (such as blood pressure and blood glucose value), tasks, alerts, and messages. Via the menu bar it is possible to go to a specific page where more can be found regarding to tasks, messages, measurements, care plan, medicines, diary, personal details, and alerts. The HCP can assign new tasks, medicines, and measurements (lab values) to the patient. Moreover, it is possible to send (chat) messages to the patient. The HCP can go back to the patient list to see the list as described before.

The patient uses the same platform. One logs in via the web portal and after logging in, an overview in which the vital parameters, tasks, and chats are shown as is drawn in Figure 1.1. Moreover, it is possible to navigate via a menu to other pages with the task list, measurements, chats, medicines, and the library. In this portal the patient can perform assigned tasks, read in the library, and see measurements. The measurements are patient disease specific, but it is possible to configure that one can enter (and see) more measurement instances. In Figure 1.2 the patient view of his measurements is drawn. These measurement instances are for example blood glucose level (HbA1c), fasting blood glucose, blood pressure, and weight. However, for the further analyses only HbA1c values that are entered by the HCP are used (to ensure lab measurement values). Moreover, the patient can use the mobile app for some of the before described functionalities. After a fixed time, the patient is logged out automatically, even in the app.

The HCP can see what the patient records, so by entering self-measurement values that are bad for patient’s health, e.g. a fasting blood glucose level of 30mmol/l (critical value), the HCP could see it (and the patient gets a notification that it might be useful to contact the professional). In the next consult the HCP can see the measurements directly in a graph instead of asking the patient the values. In addition, the HCP can enter the library and send an article as a task to a selected set of patients or the complete population
**Figure 1.1:** Patient view e-Vita; the information shown in the figure is fictional. Via the menu bar on top, the patient can navigate through the portal. The blue box shows the last measurement values for blood pressure, glucose, weight, temperature and steps. Moreover, in this box the trend and a progress indicator are shown. Below this box one of the tasks the patient has to perform and the last chat message are listed. On the right side patient’s HCP is listed with a menu to contact him.

**Figure 1.2:** Patient view measurements e-Vita; the information shown in the figure is fictional. Via the menu bar on top, the patient can navigate through the portal. On the left, one can select the different measurement instances which he want to see and/or add. The graphs show the progress over time. On this screen the patient can add values as well.
with a specific disease. Benefits of this product are remote monitoring, personalized and validated online education (via library and tasks), user friendly interfaces, and continuous automatic monitoring (VitalHealth Software, n.d.).

### 1.3.2 Patient-HCP Communication and Interaction

The patient-HCP communication is essential in patients’ care (Asan, Young, Chewning & Montague, 2015; Street, 2013). For example, the design of the consultation room influences patients’ behavior in terms of communication (Asan et al., 2015; Voran, 2017). Kaplan, Greenfield and Ware (1989) concluded that patient involvement is important for improvements of the blood glucose level. Patient’s adherence will become better if patients’ knowledge is taken into account as well (Yelovich, 2016). Background variables, such as culture, doctor-patient relationship, and disease characteristics influence the communication and patient outcomes. Moreover, the communicative behavior (e.g. instrumental versus affective and verbal versus non-verbal) influences the patient outcomes as well (Ong, de Haes, Hoos & Lammes, 1995). However, not all findings of in real-life patient-HCP conversations/communication are applicable and relevant for digital communication on a self-management platform. Stewart (1995) found that good communication between the HCP (physician) and patient can improve health outcomes. Giving education to the patient resulted in an affection of the blood pressure, blood glucose level, and physical health (Stewart, 1995). The behavior during a consult influences patient’s involvement (Kaplan et al., 1989).

Nowadays, the computer is included in the clinical office, patient engagement is more important, and showing the digital data during consults is important for patients (Voran, 2017). Therefore is assumed that when one communicates in a digital way, seeing the results and patient’s health values is important as well. The digital communication is possible on a self-management platform, so this can be tested there. However, not all general practitioners like to communicate digitally because they are concerned about privacy issues and they think that digital communication could disadvantage people that do not use technologies (Huxley, Atherton, Watkins & Griffiths, 2015). When these privacy issues are solved, the digital communication can still improve patient’s health. The (digital) communication remains important for managing patient’s health, because it is highly correlated to patient adherence to the treatment (Haskard-Zolnierek & DiMatteo, 2009). Furthermore, patients are interested in exchanging e-mails with their physicians (Ferguson, 1998). Therefore, digital and secure communication is relevant. Moreover, there has taken place a digital transformation in the industry (Karpeh & Bryczkowski, 2017).

### 1.3.3 Impact on Patient’s Health

Web-based care management is tested for patients with poorly controlled diabetes. Researchers used the web portal of MyCareTeam and found a significant reduction of the HbA1c level for the patients that used the platform (McMahon et al., 2005; Robinson, Turner, Levine & Tian, 2010; Smith et al., 2004). The used sample sizes were 52 and 16 patients. Moreover, patients that used the platform more often were associated with larger decreases in the HbA1c levels than patients that used the platform less often.
The self-management platform e-Vita can be compared with the platform MyCareTeam in terms of functionalities. Sending text-messages as reminder or as motivation support results in decreasing HbA1c levels (Krishna, Austin Boren & Balas, 2009; Rami, Popow, Horn, Waldhoer & Schober, 2006). During a period less than six months was found that patients who communicate (via the web portal) more with their HCPs have a bigger reduction in HbA1c level (Levine, Turner, Robinson, Angelus & Hu, 2009; Robinson, Turner, Levine & Tian, 2011; Turner et al., 2013). These messages support patients for treatment adherence to monitor their blood values (Franklin, Waller, Pagliari & Greene, 2006; Rami et al., 2006).

If the HCP sends (motivational) messages to the patient, this results in more patient logins on a self-management platform and in more entered measurements. Person-centered messages – messages that contain personal information of the patient – seem the most important ones to decrease HbA1c levels (Levine et al., 2009; Robinson et al., 2011; Turner et al., 2013). Therefore is expected that when the HCP sends more (motivational) messages to the patient, the patient will be more active on the self-management platform by logging in more frequently and by entering more measurements. It might be the case that the more active the HCP, the lower HbA1c level of the patient. In the study of Willaing, Rogvi, Bøgelund, Almdal and Schiøtz (2013) 993 patients with diabetes reported their HbA1c level. These levels were combined with data about their self-management behavior, activation and perception of care, using the PAM. They found that a low level of patient activation was significantly associated with a high HbA1c level. The same holds for having too less knowledge about the treatment of diabetics, e.g. target HbA1c level (Willaing et al., 2013). Insignia Health (n.d.) claims that each point someone increases in PAM score results in two percent less hospitalization and an increase of two percent in adherence to medication. In addition, it helps patients and HCPs to get patients engaged. Increasing patient activation may help the reduction of costs and increase health outcomes (Hibbard, Greene, Shi, Mittler & Scanlon, 2015). An overview of the sample sizes, main findings, and period lengths of previous literature is listed in Table A.1.
1.4 Research Questions

In this section the current gap in literature and the research questions of this study are described.

As follows from section 1.1 monitoring vital parameters and changing lifestyle is important for people with DM2. For active participation in patient’s treatment multiple self-management tools exist. An example of such a tool is a self-management platform.

From sections 1.3.2 and 1.3.3 follows that several studies showed that self-management platforms can have a positive impact on diabetes patient’s health. In addition, communication and cooperation between patient and HCP is important. However, in previous research patients and HCPs were volunteers and knew that they participated in a study, which could have influenced the results. Moreover, the e-mails from only 13 HCPs were analyzed (Levine et al., 2009; Robinson et al., 2011; Turner et al., 2013). Smith et al. (2004) found a reduction in HbA1c level in a time period of six months in a sample size of only 16 patients. So, the samples were limited, both in terms of number of patients and number of HCPs. Moreover, it is still unknown how patient’s health improvement (or change) remains on the long-term.

These results, time period, and sample sizes make it interesting to perform research to see whether similar results are found when data of a longer time period and more participants are analyzed.

Therefore, in this study, a data set with more participants (HCPs and patients) and from a longer time window will be analyzed. Moreover, the data set that will be analyzed has data from previous year(s), so participants do not know specifically (although they gave consent for analyzing entered data before using the self-management platform) that they are participating in a research. With these extensions, the results will be more generalizable than previous results.

Based on this chapter, research questions with their hypotheses are formulated which are listed in this section. In Figure 1.3 is visually shown which relations are expected.

![Diagram](image-url)

**Figure 1.3:** Model with expected (positive) relations
CHAPTER 1. INTRODUCTION

RQ1 What measure can be used for the level of self-management?
   In section 1.2 some measures are explained already. If measures cannot be used (or
   the data are not available) an own quantifier for the level of self-management will
   be created and tested.

RQ2 In which way is the interaction between the HCP and patient related to the level
   of self-management and how could it be improved?

   SQ2.1 What is the relation between the number of logins of the HCP and the
   number of logins of the patient per time unit?

   SQ2.2 What is the relation between the number of activities performed by a patient
   and the number of activities of a HCP per time unit?

   SQ2.3 What is the relation between the number of logins of a HCP and the number
   of activities performed by a patient per time unit?

   SQ2.4 What is the relation between the communication style (via chat) between
   the HCP and the patient and patient’s self-management level?

   Based on what is found in the literature, it is expected that the more chat messages
   a patient receives from his HCP, the more he logs in. Furthermore, other forms of
   patient-HCP interaction may improve patient’s self-management level as well. The
   answer on this research question explains arrow A of Figure 1.3.

RQ3 How does the level of self-management influence patient’s health level?
   It is expected that the higher the level of self-management, the higher patient’s
   (perceived) health will be. The answer on this research question explains arrow B
   of Figure 1.3.

RQ4 How does the patient-HCP interaction influence patient’s health level?
   It is expected that the patient-HCP interaction influences patient’s health. More
   specific, receiving more emotional social support chat statements is expected to
decrease the HbA1c level. The answer on this research question explains arrow C
   of Figure 1.3.

In chapter 2 is explained how the analyses were performed. Chapter 3 consists of the data
description. In chapter 4 the results following from the analyses as explained in chapter
2 are described. These chapters are followed by chapter 5 in which the research questions
are answered by concluding the results. Finally, in chapter 6 the results are discussed.
Chapter 2

Methodology

In this chapter the used methodology is described. First is explained in section 2.1 what the Cross Industry Standard Process for Data Mining (CRISP-DM) is and how this structure is applied in this study. Secondly, in section 2.2 is explained how the data is prepared to be able to analyze the data. Thirdly, in section 2.3 the performed analyses are explained and motivated. Section 2.4 covers the description of the evaluation of the created prediction models. Lastly, this chapter ends with section 2.5 in which the plan for the questionnaire and interviews is described (although these were not performed as is described in section 4.5).

2.1 Process for Data Mining

This section describes the process for data mining that is followed in this study.

Data analyses can be performed in a structured manner. A standardized way is the Cross Industry Standard Process for Data Mining (CRISP-DM) of Chapman et al. (2000), which is shown in Figure 2.1. The complete process is split into six phases.

The first phase of the CRISP-DM is Business Understanding. The context of the patient with diabetes who uses the self-management platform e-Vita of VitalHealth Software is described in section 1.3.1. Once the data were collected, it can be understood by describing and exploring the data. For understanding the data, the data had to be collected. Firstly, four care group boards were asked for permission to analyze the data from their care group. After written approval, the data set with patients and HCPs who gave an informed consent for analyzing their data, was exported. After the collection, the data was explored with R and Microsoft Excel and this is described in chapter 3. A more detailed overview of the data attributes is added in Appendix C. In the third phase, the data were selected and cleaned to provide a data set. After the data preparation, phase four, the modeling phase starts. During the modeling phase techniques are chosen with which models are created assessed. During the fifth phase the results are evaluated and after evaluation, phase six starts. The sixth phase is the deployment phase in which the results are translated for interpretation in a broader context (Chapman et al., 2000).

This process and these phases are used in this study. However, the structure of this document does not strictly follow the described steps.
2.2 Data Preparation

This section covers the description how the data was prepared before it was analyzed.

All the patient and HCP data from e-Vita – from them who have given consent for analyzing their data – were exported to a protected server after written approval of the care group board. All the test accounts were removed from the data set and the measurement values were filtered based on ranges (upper and lower abnormal limit) of the thresholds that are used within VHS. Moreover, the data set has been anonymized to ensure privacy. All patients in the data set were diagnosed with diabetes type 2, were treated in the Netherlands, and got protocol based primary care. For there is a standard protocol for diabetic care in the Netherlands, it was assumed that the treatment would not be influenced by the fact of being cared in another care group (another area in the country) and could be analyzed simultaneously (NHG, 2013).

For being able to compare the change in HbA1c level, patients with an age below 70 and at least one measurement value before and two after starting with e-Vita and who did not use insulin were selected. Above 70 years another target value is used and insulin influences the HbA1c level by itself (NHG, 2013). Patients with too low HbA1c measurement values (below 10) were dropped from the data set, because these values were probably representing percentages instead of mmol/mol. This sub set is used in the analyses of this study.

The results of the data preparation phase is described in chapter 3.
2.2.1 Patient Attributes

This sub section covers a list with patient attributes that were existing in the data set or were calculated for the analyses.

For analyzing the data, several variables were calculated and determined. These variables are listed in Table 2.1.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Login frequency</td>
<td>Total number of logins divided by the number of days since the first login</td>
</tr>
<tr>
<td>Percentage of finished tasks</td>
<td>Number of tasks finished divided by total number of tasks assigned</td>
</tr>
<tr>
<td>Self-measurement frequency</td>
<td>Total number of self-measurements divided by the number of days since the first login</td>
</tr>
<tr>
<td>Received chat frequency</td>
<td>Total number of received chats divided by the number of days since the first login</td>
</tr>
<tr>
<td>The percentage of received (for patients) chats containing statements from a certain category</td>
<td>Per patient is calculated which percentage of chats contains certain statements from categories as described in section 2.2.4 (e.g. ISSChatPerc).</td>
</tr>
<tr>
<td>Percentual changes in HbA1c levels</td>
<td>(Newest HbA1c value minus older HbA1c value) divided by the older HbA1c value: ( \text{PercChange}_{\text{HbA1c}} = \frac{\text{newvalue} - \text{oldvalue}}{\text{oldvalue}} \times 100% )</td>
</tr>
<tr>
<td>DaysFLAm</td>
<td>Number of days between patient’s first login and Am, the last measurement date.</td>
</tr>
</tbody>
</table>

Patient’s HbA1c measurement values were sorted based on the measurement date. The first measurement value was named \( B_1 \), the second one \( B_2 \), until the last, \( n^{th} \), one with a measurement date before the patient started with the self-management platform as \( B_n \). The first measurement value with a measurement date after starting with the self-management platform was named \( A_1 \), the second one \( A_2 \), until the last, \( m^{th} \), one as \( A_m \). This is visually shown in Figure 2.2. The percentual change in HbA1c level is used instead of the absolute change in HbA1c level to be able to compare the relative changes with each other. The data is described in chapter 3.

As described in section 2.2.3 patient’s login frequency and percentage of finished tasks are used in an aggregation to quantify patient’s self-management level. By means of chat frequency and content the patient-HCP interaction will be explored.
2.2.2 HCP Attributes

This sub section covers HCP’s attributes. More details are listed in Appendix C.

For analyzing the data, several variables were calculated and determined. These variables are listed in Table 2.2. All chats sent by HCPs are received by patients and in this study these attributes are covered with patient attributes.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Login frequency</td>
<td>Total number of logins divided by the number of days since the first login</td>
</tr>
<tr>
<td>Sent chat frequency</td>
<td>Total number of received chats divided by the number of days since the first login</td>
</tr>
<tr>
<td>Sent-received chat ratio</td>
<td>Total number of sent chats divided by the total number of received chats</td>
</tr>
<tr>
<td>Number of usage days</td>
<td>Number of days between last login and first login</td>
</tr>
<tr>
<td>Number of patients</td>
<td>Total number of patients that is cared by the HCP, is diagnosed with DM2, has started to use the self-management platform, and did not withdraw from the self-management platform</td>
</tr>
</tbody>
</table>

2.2.3 Quantify Self-Management Level with Aggregation

This sub section covers the description why and how patient’s self-management level is quantified.

No self-management level data were included in the data set and by not being allowed to contact the patients during this research an own measure for the level of self-management had to be defined. For this, a literature review was performed to see which other measures
CHAPTER 2. METHODOLOGY

existed already and which ones could be used. In section 1.2 these measures are described. Moreover, with this gathered information an own quantifier was developed based on the literature and the existing data in the data set. According to Hibbard et al. (2004) patients with a higher self-management level are performing more activities that may result in a better health, like keeping track of a glucose logbook. Also, the activation levels of PAM are significantly related to the self-management level of patients (Hibbard et al., 2004; Elissen et al., 2013) and the levels of self-management as defined by Schermer (2009) include performing assigned tasks, including self-measuring tasks. So, the more activities a patient performs on a self-management platform, the more active he is (and the higher the PAM).

A quantifier for the level of self-management was derived with an aggregation in which multiple variables were used. For there was no real value of the self-management level, the derived quantifier was not validated.

As can be derived from the list with patient attributes in section 2.2.1, description in Appendix C and the data description in chapter 3, from all patients is known their login frequency, sent chat frequency, received chat frequency, entered measurement frequency, and the percentage of finished tasks, among other things.

As written in section 1.3.1, patients log in to see lab results, read in the library, send chat messages to their HCP, enter lab results, or perform tasks by themselves. In addition, they can have received an e-mail with the notification that for example a new lab result is added, or that they have received a chat message from their HCP. So, the login behavior really depends on the patient and on the agreements between the patient and HCP.

Each HCP decides per patient which tasks they assign in addition to the automatically – based on the diagnosis – assigned tasks. The percentage of finished tasks per patient indicates what percentage of the tasks are finished by the patient. The higher this number, the more the patient has done; related to the number of tasks he had to do.

Patients enter self-measurements if they want to have a good and nice overview of their values over time. In addition, sometimes an HCP asks specifically for entering e.g. blood values before and after meals during a certain week to see how the blood values behave over time.Patients can enter as many measurements as they want and no rule exists how often a patient should enter a measurement. In addition, no agreements between the HCP and patient on the number of self-measurements were stored. So, many patients could have entered zero self-measurements, resulting in a frequency of zero and an estimated self-management level of zero (in case of multiplication). Therefore, using the entered self-measurement frequency for the estimation of the self-management level seemed not be the best option.

Sending chats to the HCP can be done as an answer to a chat message from an HCP or when the patient has a question for his HCP. It can occur that a patient knows how to manage his health without sending any chat message. So, many patients could have sent zero chats, resulting in a frequency of zero and an estimated self-management level of zero (in case of multiplication). Therefore, using the sent chat frequency in the estimation of patients self-management level seemed not be the best option.

From these variables, the login frequency and the percentage of finished tasks were potential variables to include in the quantifier.
First, the login frequency and the percentage of finished tasks for each patient were determined. The values per patient were normalized via the formula as shown in Equation 2.2, where $V_{i,p}$ denotes the property value of patient $p \in P$, and $i$ the set as defined in Equation 2.1.

$$i \in \{\text{LoginFrequency, PercentageFinishedTasks}\} \quad (2.1)$$

$$\mu_{i,p} = \frac{V_{i,p} - \min_p V_{i,p}}{\max_p V_{i,p} - \min_p V_{i,p}} \in [0,1] \quad (2.2)$$

Although logging in on the self-management platform is an indicator of interest in self-management, it is a start before a patient can do anything on the self-management platform; in other words: the percentage of tasks finished is dependent on the login frequency. Therefore, it is necessary to check the correlation between the percentage of finished tasks and the login frequency of the patients. If for patients with a high percentage of finished tasks the login frequency is always high, and the login frequency is low for patients with a low percentage of finished tasks (if the correlation between these two variables is high), it is not needed to include both variables in the estimation. If both variables are included in the estimation, a weight is given to the variables in the aggregation to indicate the importance.

The multiplication of the normalized values to find one value for a composite measure is called aggregation. The percentage of finished tasks was given a relative weight $w$ of 2 and the login frequency a weight $w$ of 1 in the aggregation. As a T norm operator multiplication was chosen and therefore, the quantifier for the self-management level was defined as in equation 2.3.

$$SML_p = \prod_i (\mu_{i,p})^{1/w_i} \quad (2.3)$$

It was expected that the higher this quantifier, the lower the HbA1c level and the higher $SML_p$ the more decrease in HbA1c level.

Two quantifiers were derived for the self-management level. These quantifiers are given in Equations 2.4 and 2.5:

$$SML_{1p} = \mu_{\text{LoginFrequency},p} \cdot \sqrt{\mu_{\text{PercentageFinishedTasks},p}} \quad (2.4)$$

$$SML_{2p} = \mu_{\text{PercentageFinishedTasks},p} \quad (2.5)$$

In section 4.1 is described what the correlation between the percentage of finished tasks and login frequency is and which quantifier was chosen. The SML values are used in chapter 4 for prediction models to predict the change in patient’s health. In addition, prediction models to estimate patient’s SML are included in chapter 4 as well.
2.2.4 Chat Statements

This sub section describes how and why chat messages were split and coded into chat statements. The chat messages were anonymized by replacing names, contact details, and other personal identifiers in the message text with a code (such as \{NAME\} and \{PHONE NUMBER\}). After that, the messages sent by HCPs were classified in chat statement coding categories that were used by Robinson et al. (2011). Besides their eight categories, the category ‘Other’ was added for the statements that did not fit in the other categories. This classification was done by dividing the chat messages sent by HCPs in multiple chat statements and assigning one category per statement. A description of these nine categories is given in the following list.

1. **Informational Social Support (ISS)** These statements provide information about the illness or treatment. Examples are “The values are lower than they should be”, “Please change the medicine dosage to three units”, and “Those side effects are normal”.

2. **Tangible Social Support (TSS)** These statements directly help the patient with doing his tasks. Examples are “I changed the settings for you”, “I made an appointment for next Tuesday”, and “I sent you a form for the lab”.

3. **Social Integration Support (SIS)** This category include statements that show involvement by making contact. Statements such as “Thank you for performing that task”, “I tried to call you three times”, “See you next week”, “Hello Mr. Jansen”, and “Regards, Noah - HCP” are covered in this category.

4. **Emotional Social Support (ESS)** HCPs who use these statements indicate their care, concern, and general willingness to help. Examples are “Call me if you need anything”, “I am concerned”, and “It is hard to get a stable blood glucose”.

5. **Esteem Social Support (ESSup)** With these statements an HCP compliments the patient. Examples are “You are doing great!”, “You are my best patient!”, “Those values are really good.”, and “Nice, keep going that way!”.

6. **Health Information Requests (HIR)** This category contains questions regarding patient’s health, like “Can you send a day curve?”, “Can you send you current blood pressure?”, “Can you measure your glucose six times a day next week?”.

7. **Self-Disclosure (SD)** With these statements an HCP writes something personally, such as “I’ll go on holiday next week”, “Next month, I’ll get a baby”, and “I was sick last week”.

8. **Technical Support (TS)** This category contains statements that provide information on using the self-management platform. Examples are “Click on the button at the right side of your screen” and “You can edit it on your profile page”.

9. **Other** This category includes statements that did not fit in the other categories. An example is “Okay.”.

In chapter 4 prediction models are developed to estimate patient’s health (improvement) to explore the influence of patient-HCP interaction by means of chat content.
2.2.5 HCP Grouping

This sub section describes how patients were grouped based on their HCP’s activity. The groups are used in section 4.2.1 to compare patient’s self-management level within the groups.

When HCPs are logged in on the self-management platform, they can assign tasks, measurements and medicines and can send chat messages to the patients, among others. Because most of the tasks are assigned automatically, only HCP’s login frequency and sent chat frequency could be interesting for grouping HCPs based on their activity. While the login frequency does not give any information of how many patients were served that time, using the login frequency is not the best idea. Therefore, HCPs were grouped based on the frequency of sending chat messages. In addition, the frequency with which an HCP sends chats does not reflect directly how often a patient receives a chat messages, because HCPs (can) have multiple patients. Therefore, a second grouping is done based on patients received chat frequency. Because all chats received by patients are sent by HCPs this is still a grouping based on HCP’s activity.

Comparing the groups was done with a Mann-Whitney (Wilcoxon rank-sum) test because it could not be assumed that the group samples were normally distributed. This is described in section 4.2.

2.2.6 Patient Grouping

This sub section describes how patients were grouped based on their self-management level. The groups are used in section 4.3.1 to compare the HbA1c changes within the groups.

Patients were grouped in three groups based on their self-management level, SML. Patients in the group ‘low’ had an SML of 0. The twenty patients with the highest SML were grouped as ‘high’ and the others as ‘medium’. With this grouping was checked whether the top-twenty patients in terms of self-management level had a different percentual change in HbA1c level than patients with a self-management level of zero or lower than the top-twenty. So, the twenty patients with the highest self-management level were compared with the others.

With a Kruskal-Wallis test was tested whether the distributions of the values in the three groups were different. When a p-value smaller than 0.05 was found, it was concluded that the medians were not equal and that the distributions were not similar. After that test, with independent two-group Mann-Whitney tests was tested which groups were significantly different from each other. These tests were one-sided, because the expectation was that in higher groups the percentual changes were smaller than in the others. The results of this patient grouping are described in section 4.3.
2.3 Performed Analyses

This section describes the analyses that followed after the data preparation as denoted in section 2.2. The results following from these analyses are given in chapter 4.

Once the data set was understood and prepared, the modeling phase was started. The data set was randomly split into a training and a test set. The training set consisted of 60% of the patients and the test set of 40%.

Multiple methods for building a prediction model exist. Because the dependent variables in this study – the change in HbA1c level and the SML – are continuous variables (and the other variables are numerical), the prediction model can be built with regression (Bramer, 2016, p. 5). SML values were found by using two attributes of the data set (login frequency and percentage of finished tasks) as explained in section 2.2.3.

There exist several regression types. For the case of estimating a continuous variable is chosen to use simple linear regression, multiple linear regression, and decision tree regression. Other methods that only predict categorical dependent variables are not applicable in this study. For the multiple linear regression stepwise approaches (both backward elimination and forward selection) are used. In a forward selection approach is started with a model without any predictors. Then by iteration the most contributive predictor is added until no significant improvements can be made. In case of backward elimination is started with all predictors and by each iteration the least contributive predictor is removed until a model is found that is significant (Sutter & Kalivas, 1993; Weisberg, 2005, p. 221-222).

A regression tree splits the total space into disjoint regions. For each region (leaf) a prediction value is given, which is the expected value of all items in that region (Ruczinski, n.d., p. 258). First is started with one node containing all points. At each iteration the best split is chosen where the best split is the split that results in gaining most purity (entropy) (Ruczinski, n.d., p. 260). So for each split the sum of squared errors between the observation and the prediction value of the leaf is minimized (Moisen, 2008).

For determining the relation between the patient-HCP interaction and the self-management level (relation A in Figure 1.3) the patients were grouped based on the activity of their HCP (see section 2.2). In addition, a linear regression model was built with the training set and tested with the test set. In this study patient-HCP interaction is quantified with patient’s and HCP’s login frequency, received chat frequency, and the chat content categories (as defined in section 2.2.4).

The relation between SML and the change in HbA1c level, patient’s health, (arrow B of the model in Figure 1.3) was found with patient grouping and a regression model. Regression models can be used for predicting numeric values based on (numerical attributes). The SML is a numerical value and for predicting the SML based on other attributes a regression model can build.

Patients were grouped based on their SML. If the SML was equal to zero, patients were included in the group low SML. When patients had an SML of more than the average, they were included in the low SML group. All other patients were included in the medium SML group. The three groups were compared with each other to find the differences in change in HbA1c level between the groups. In addition, with linear regression a model
was built and tested with respectively the training and test set.

For determining the relation between the patient-HCP interaction and the change in HbA1c level (relation C in Figure 1.3) the variables regarding chat statement categories and HCP login frequency were used. These included e.g. the percentage of chats containing emotional support chat statements and the percentage of chat statements that were of the social integration support.

From Turner et al. (2013) was expected that the percentage of received chats containing emotional support statements resulted in a decreasing HbA1c level. Therefore, the model

\[
\text{Change} = \beta \cdot \text{ESSPercentage} + c
\]

with Change as the percentual difference in HbA1c level between \( B_n \) and \( A_m \) was tested. In addition, with a forward approach other linear regression models were evaluated as well.

A regression tree can be used for predicting values for a certain numerical variable \( Y \), in this study the percentual change in HbA1c level. Each node in the tree predicts the sample mean of \( Y \). By building a regression tree, a set of decision rules is created. By iteration, each rule is set up by choosing the split that maximizes the homogenity of the newly created groups (Prasad, Iverson & Liaw, 2006). These models are relatively easy to interpret, but the prediction accuracy is not always the best (Loh, 2011). With these regression trees, prediction models were build.

### 2.4 Evaluation of Prediction Models

In this section is explained how found results were evaluated. Once the prediction models were created, the results were analyzed based on their meaning and two error measures; namely the mean squared error (MSE) and the mean absolute error (MAE).

#### MSE

The MSE is always non-negative and the smaller it is, the better the prediction is. This measure is calculated as follows:

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2
\]

where \( \hat{y}_i \) is the predicted value and \( y_i \) the actual value as part of the test set.
**MAE**

The MAE is also always non-negative and the smaller this value, the better the model. The MAE is calculated via the following formula:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_i - y_i|$$

in which $\hat{y}_i$ is the predicted value and $y_i$ the actual value as part of the test set.

### 2.5 Questionnaire and Interview

This section covers the description of the purpose and goal of the planned questionnaire and interview. It was planned to ask at least five to ten patients and at least five to ten HCPs to undergo a semi-structured interview or to fill out a questionnaire to be able to detect at least 80% of the possible answers (Davids, Harvey, Halperin & Chikte, 2015). Patients and HCPs were asked how they perceive their use of the self-management platform and how they think patients would become more motivated to enter measurements and to use the platform more frequently. Moreover, patients were asked what kind of communication/feedback they would like to receive from their HCP and in which way they would like to get rewards from the system. The collected information would have been used as input for an intervention pilot if there were enough responses. More details about the questions is given in Appendix B.
Chapter 3

Data Description

This chapter describes the data in the data set. First the attributes of the patient are described in section 3.1 and secondly those of the HCPs are described in section 3.2. Only the numbers and figures that are relevant for this study are given and an overview of the attributes of the data instances is given in Appendix C.

3.1 Descriptive Statistics of Patients

This section covers the descriptive statistics of the patient data instance. In following sub sections specific attributes of the patient data instance are described. All patients in the data set were diagnosed with diabetes type 2, were treated in the Netherlands, and got protocol based primary care. Because there is a standard protocol for diabetic care in the Netherlands, it is assumed that the treatment would not be influenced by the fact of being cared in another care group (another area in the country) and could be analyzed simultaneously (NHG, 2013).

A total number of 533 active patients that have accepted the disclaimer and having diabetes type 2 was in the system and was exported. From this data set a sub set was created in which all patients did not use insulin, had an age below 70, had at least one HbA1c measurement before and at least two measurements during the use of the self-management platform as described in section 2.2. This sub set consisted of 105 patients to which is referred to as the data set in the remainder of this study.

3.1.1 Age and Gender

From these patients 66.7% was male and the average age was 58.5 years. From Figure 3.1 can be seen that the range of ages is [34-69]. In addition, the mean age is equal to 59.4, 56.7, and 58.5 years for males, females, and all patients respectively. These numbers (and other detailed numbers from this section) are listed in Table D.1. Moreover, a detailed plot with the age and gender of the patients is drawn in Figure D.1.
3.1.2 Login Frequency

The login frequency variated between 0.0011 to 0.387 times a day and was on average 0.03905 times a day. This corresponds with one login per 25.6 days. Four patients logged in more than once a week and the top ten most frequently logging in patients logged in on average 0.1663 times a day, which corresponds with approximately one login each six days. A histogram with patient’s login frequency is drawn in Figure 3.2. Patients with a high SML log in on average more frequently than the others (see Table D.1).

3.1.3 Active days and FLAm

The number of active days is for patients on average and maximal 251.9 and 910 days respectively. The boxplot of the usage days per care group is drawn in Figure 3.3. In addition, the boxplot with the number of days between First login and $A_m$ (FLAm) is drawn in 3.4. The number of days is on average and maximal 415.5 and 871 respectively. From these two figures can be seen that the usage days and days between first login and $A_m$ differ per care group.

3.1.4 Chat messages

In total 246 chat messages in the data set were received by the patients (only sent by HCPs) and these chat messages were divided into 1295 chat statements and coded as described before. The numbers per category are listed in Tables D.2 and D.3. The percentages per SML group are quite different from each other.
3.1.5 Tasks

In total, there were 1032 tasks assigned to patients, from which 383 were assigned by an HCP and patients have finished 844 of them. At the moment of exporting the data, there were 183 open tasks (from which the maximum per patient was 8 tasks). The number of open tasks per patient is drawn in Figure 3.5. On average, patients do not finish the task before the deadline in 51.58% of the tasks. From Table D.1 can be seen that patients in the high SML group have been assigned to more tasks and they have finished a higher percentage than patients of the other groups.
CHAPTER 3. DATA DESCRIPTION

3.1.6 Self-Management Level and Chat Statements

The scatter plots of patient’s SML with the fraction of chats containing SIS and SD statements are given in Figures 3.6 and 3.7 for having the biggest and smallest (absolute) correlation. The figures of the other chat statement categories are similar. The detailed numbers are listed in Table D.2.

![Scatter plot SML and fraction of chats containing SIS statements](image1)

**Figure 3.6:** Scatter plot SML and fraction of chats containing SIS statements

![Scatter plot SML and fraction of chats containing SD statements](image2)

**Figure 3.7:** Scatter plot SML and fraction of chats containing SD statements

3.1.7 HbA1c Levels

In Figures 3.8, 3.9, and 3.10 the scatter plots of the HbA1c values drawn respectively for time $B_n$, $A_2$, and $A_m$. The mean values are 55.2, 53.8, and 54.9 mmol/mol respectively. More details are listed in Table D.1.

![Scatter plot HbA1c levels $B_n$](image3)

**Figure 3.8:** Scatter plot HbA1c levels $B_n$

![Scatter plot HbA1c levels $A_2$](image4)

**Figure 3.9:** Scatter plot HbA1c levels $A_2$

![Scatter plot HbA1c levels $A_m$](image5)

**Figure 3.10:** Scatter plot HbA1c levels $A_m$

In Figures 3.11 to 3.14 the scatter plots of the percentual change in HbA1c level with the received chat frequency, HCP login frequency, days between first login and $A_m$, and SML are drawn.
CHAPTER 3. DATA DESCRIPTION

From Figure 3.11 can be seen that there is more decrease (and less increase) in HbA1c level when the patient receives more frequently a chat message. The higher patient’s HCP login frequency, the less increase (and more decrease) in HbA1c level results from Figure 3.12.

When the number of days between patient’s first login and $A_m$ are plotted against the percentual change in HbA1c level, see Figure 3.13, one can see that the HbA1c level decreases less when the number of days increases. However, this is mainly due to the patients with a low and medium SML.

From Figure 3.14 follows that the higher patient’s self-management level, the smaller (and more negative) the change in HbA1c level.

![Figure 3.11: Scatter plot percentual change HbA1c and received chat frequency. The red, yellow, and green dots indicate respectively the patients with a low, medium, and high SML.](image1)

![Figure 3.12: Scatter plot percentual change HbA1c and HCP login frequency. The red, yellow, and green dots indicate respectively the patients with a low, medium, and high SML.](image2)

![Figure 3.13: Scatter plot percentual change HbA1c and number of days between first login and $A_m$. The red, yellow, and green dots indicate respectively the patients with a low, medium, and high SML.](image3)

![Figure 3.14: Scatter plot percentual change HbA1c and SML. The red, yellow, and green dots indicate respectively the patients with a low, medium, and high SML.](image4)
CHAPTER 3. DATA DESCRIPTION

In addition, scatter plots in which the percentual change in HbA1c level is plotted against the fraction of chats that contain certain statements are drawn. In Figures 3.15, 3.16, 3.17, and 3.18 these scatter plots are included for respectively ISS, ESSup, HIR, and SD statements. The figures with the other chat statement categories look similar to the shown figures. In Figures 3.15 and 3.16 a clear (positive and negative) linear trend can be seen. In both Figure 3.17 and Figure 3.18 the data points are scattered and no (direct) relation can be seen.

![Figure 3.15](image1.png) **Figure 3.15:** Scatter plot percentual change HbA1c and fraction of chats that contain ISS statements. The red, yellow, and green dots indicate respectively the patients with a low, medium, and high SML.

![Figure 3.16](image2.png) **Figure 3.16:** Scatter plot percentual change HbA1c and fraction of chats that contain ESSup statements. The red, yellow, and green dots indicate respectively the patients with a low, medium, and high SML.

![Figure 3.17](image3.png) **Figure 3.17:** Scatter plot percentual change HbA1c and fraction of chats that contain HIR statements. The red, yellow, and green dots indicate respectively the patients with a low, medium, and high SML.

![Figure 3.18](image4.png) **Figure 3.18:** Scatter plot percentual change HbA1c and fraction of chats that contain SD statements. The red, yellow, and green dots indicate respectively the patients with a low, medium, and high SML.
CHAPTER 3. DATA DESCRIPTION

3.2 Descriptive Statistics of HCPs

This section covers the descriptive statistics of the HCP data instance. The data set with 59 HCPs, who all have patients on the self-management platform, was exported. For all of them was checked whether there were differences in attributes per HCP between the care groups. With Bartlett tests and ANOVAs (because the number of HCPs was 41, 4, 9, and 5 for respectively care group 1, 2, 3, and 4) was found that there were no significant differences between the care groups over the attributes number of logins, number of sent chat messages, number of received chat messages, number of patients, login frequency, sent chat frequency, sent-received chat ratio, and number of sent chat statements. Although a significant difference in the usage period of e-Vita was found, it is concluded that all HCPs could be analyzed as one group. The boxplot of the usage days per care group is drawn in Figure 3.19. It can be easily seen that the HCPs of care group 4 and 2 use the self-management platform for a longer time than the other two.

![Figure 3.19: HCP usage days of e-Vita per care group](image)

The login frequency varied between 0.0024 and 0.7725 times a day and was on average 0.1655 times a day. The average login frequency is similar to one login per 6.04 days which is about one time a working week. The average login frequency of the ten HCPs with the highest login frequency was 0.4734 times a day. These numbers are listed in Table 3.1 too. In Figure 3.20 the sent chat frequency and the login frequency is plotted per HCP. From this figure can be derived that the HCPs with a high login frequency have a high sent chat frequency as well.

![Figure 3.20: HCP login and sent chat frequency; sorted on login frequency. The (open) circles indicate the login frequency while the (black) diamonds indicate the sent chat frequency.](image)

On average, the HCPs have sent 0.0518 chat messages per day. The sent-received chat ratio was on average 1.033 and varied between 0 and 5. The HCPs had at least one, on average 9.15, and at maximum 46 patients. A remark for the number of patients is that the HCPs have more patients that are active on the self-management platform than are included in the sub set. Moreover, they have much more patients in their practices.
### Table 3.1: Descriptive data from HCPs

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Login frequency (per day)</td>
<td>0.0024</td>
<td>0.7725</td>
<td>0.1655</td>
</tr>
<tr>
<td>Sent chat frequency (per day)</td>
<td>0</td>
<td>0.5518</td>
<td>0.0518</td>
</tr>
<tr>
<td>Sent-received chat ratio</td>
<td>0</td>
<td>5</td>
<td>1.033</td>
</tr>
<tr>
<td>Number of patients</td>
<td>1</td>
<td>46</td>
<td>9.15</td>
</tr>
</tbody>
</table>
Chapter 4

Results

This chapter describes the results of the data preparation and modeling phase by performing the analyses are explained in section 2.3. First the results of the quantification of the self-management level are described in section 4.1. Then in sections 4.2, 4.3, and 4.4 the relations between patient-HCP interaction and self-management level, self-management and health improvement, and patient-HCP interaction and health improvement are examined (representing the arrows of the research model as displayed in Figure 1.3). Section 2.5 describes the results of the questionnaire and interview. This chapter ends with section 4.6 in which a summary of the results is given. The conclusion and discussion of these results can be found in chapters 5 and 6.

4.1 Measure Self-Management Level

Based on writings of Hibbard et al. (2004) and Elissen et al. (2013), the existing data of all 105 patients with number of logins, total days that a patient is in the system, and the percentage of finished tasks per patient, quantifiers were defined for the self-management level as discussed in section 2.2.3 (see Equations 2.4 and 2.5). This section covers the results of the analyses.

In Figure 4.1 the boxplots of the quantifiers for the self-management level are drawn. In Figure 4.2 the same boxplots are drawn where the self-management level is limited to 0.05 to make the boxplot clearer.

The values for SML1 varies between 0 and 0.967 and is on average equal to 0.0897. For SML2 these values are respectively 0, 1, and 0.677.

The correlation between the (normalized) percentage of finished tasks and (normalized) login frequency was found equal to 0.403. This correlation value is weak to moderate, and therefore was concluded that both login frequency and percentage of finished tasks had to be included in the quantifier. This results in selecting $SML_{1,p}$ (as displayed in Equation 2.4) as the quantifier for the self-management level in this research, resulting in Equation 4.1.

$$SML_p = \mu_{LoginFrequency,p} \cdot \sqrt{\mu_{PercFinishedTasks,p}}$$ (4.1)
CHAPTER 4. RESULTS

4.2 Patient-HCP Interaction and Self-Management Level

In this section is the relation between patient-HCP interaction and patient’s self-management level (SML) examined. This section is split into multiple sub sections. First, the patients are grouped based on HCP’s activity level, see section 4.2.1. Secondly, regression models are built to predict patient’s SML, which is described in section 4.2.2. Lastly, in section 4.2.3 the found models of section 4.2.2 are compared.

4.2.1 Patients Grouped based on HCP’s Activity Level

The patients were grouped based on HCP’s activity level. As described in section 2.2.5, this is done based on HCP’s sent chat frequency.

When the HCPs were grouped based on their sent chat frequency and the ones sending more than average (0.0518 chats per day) was determined as ‘high frequent’ and the rest as ‘low frequent’, two groups of respectively 61 and 44 patients were made. For comparing means of two groups with a t-test, the equality of variances should be known. These variances were tested with a Levene’s test and the result was that the variances could be assumed as equal (p=0.622). With an independent t-test was found that patients’ SML is not significant different in both groups (p=0.874). See Figure 4.3 for the boxplot with the SML per HCP group and Figure 4.4 for a zoomed in version of the boxplot. The means in SML were 0.0878 and 0.0922 for respectively patients with a high and low sent chat frequency.

This sent chat frequency indicates only how often an HCP sends a chat message and not how often a patient receives a chat message. Therefore, patients were grouped based on
their received chat frequency. Because all chats received by patients are sent by HCPs this is still a grouping based on HCP’s activity.

When the patients were grouped based on their received chat frequency and the patients with a received chat frequency above average (0.00496 chats a day) were determined as ‘high frequent’ and the rest as ‘low frequent’, two groups of respectively 31 and 74 patients were made. The variances and means were unequal (respectively $p = 0.0009$ and $p = 0.0013$). It was found with an independent 2-group Mann-Whitney test that the median of the SML significantly different in both groups ($p < 0.00001$). See Figure 4.5 for the boxplot with the SML per HCP group and Figure 4.6 for a zoomed in version of the boxplot.

![Figure 4.3: Boxplot SML; the HCPs that sent chats more frequently than average are called high active](image1)

![Figure 4.4: Boxplot SML, limited y-axis; the HCPs that sent chats more frequently than average are called high active](image2)

From these numbers and figures follows clearly that patients who receive more often a chat messages, have on average a higher self-management level.
CHAPTER 4. RESULTS

4.2.2 Linear Regression Models

Regression models can be used for predicting numeric values based on numerical attributes. First, patient-HCP interaction is quantified with only the chat content. Using the percentage of chats that contain a certain statement category for the first eight categories resulted in the two models as displayed in Equations 4.2 and 4.3. In the first case is forced to include all categories while in the second case via a stepwise approach the best model was found. In Tables D.4 and D.5 the details are shown. The non-significant coefficients are marked with an asterisk and the prediction errors are shown in Table 4.1.

\[
\hat{SML} = 0.016951^* + 0.012427 \cdot ISSChatPerc^* + 0.002788 \cdot TSSChatPerc^* \\
+ 0.127521 \cdot SISChatPerc - 0.035639 \cdot ESSChatPerc^* \\
+ 0.200757 \cdot ESSupChatPerc - 0.065080 \cdot HIRChatPerc^* \\
- 0.320892 \cdot SDChatPerc^* - 0.172662 \cdot TSChatPerc^* \\
\] (4.2)

Secondly, with a stepwise approach the model (with only chat content) is found as displayed in Equation 4.3. The details of this regression are listed in Tables 4.1 and D.5.

\[
\hat{SML} = 0.0184^* + 0.0743 \cdot SISChatPerc + 0.1821 \cdot ESSupChatPerc \\
\] (4.3)

From these regressions can be seen that the percentage of chats that contain Social Integration Support and Esteem Social Support statements influences the self-management level more than the other statement categories, which could also have been derived from the first model (Eq. 4.2).

Thirdly, a simple regression model is built with chat content and all other attributes (received chat frequency, sent-received chat ratio, HCP’s login frequency, HCP’s sent-
received chat ratio, and HCP’s sent chat frequency). Before this was done the correlation between all attributes was determined. It was found that all correlations are weak to moderate, but below 0.60 except for the correlation between SentChatFreq and RecChatFreq and the correlation between SentRecChatRatio and SentChatFreq. With both a forward and backward approach models were found. First, with a forward approach the model as shown in Equation 4.4 is found. The details are listed in Table D.6 and the prediction errors are 0.010 and 0.068 for respectively MSE and MAE.

\[
\hat{\text{SML}} = 0.0199^* + 0.1404 \cdot \text{SentRecChatRatio} + 4.754 \cdot \text{RecChatFreq}^* \tag{4.4}
\]

With a backward approach also a prediction model is created. The result of this approach is denoted in Equation 4.5. The details are listed in Table D.7 and the prediction errors are 0.0099 and 0.068 for respectively MSE and MAE.

\[
\hat{\text{SML}} = 0.0161^* + 0.0598 \cdot \text{SISChatPerc}^* - 0.0802 \cdot \text{HIRChatPerc}^*
- 0.3557 \cdot \text{SDChatPerc}^* - 0.1587 \cdot \text{TSChatPerc}^*
+ 4.9722 \cdot \text{RecChatFreq}^* + 0.1379 \cdot \text{SentRecChatRatio} \tag{4.5}
\]

In the next section these found models are compared.

### 4.2.3 Model Comparison

In this section are the models of section 4.2.2 compared. In Table 4.1 the prediction errors are listed per model.

As can be seen from Table 4.1 the model as displayed in Equation 4.5 fits the best on the training data. Moreover, this model results in the smallest MSE. The model as shown in Equation 4.4 has a slightly higher MSE and slightly smaller MAE and adjusted R-squared. For the difference between the squared root of the MSE is bigger than the difference between the MAEs is concluded that the prediction model as displayed in Equation 4.5 fits the best for predicting patient’s SML.

<table>
<thead>
<tr>
<th>Model</th>
<th>Included attributes</th>
<th>(R^2_{adj})</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eq. 4.2</td>
<td>All ChatPerc statements (n.s.)</td>
<td>0.1039</td>
<td>0.01626</td>
<td>0.08238</td>
</tr>
<tr>
<td>Eq. 4.3</td>
<td>{ISS + ESSup} ChatPerc (stepwise)</td>
<td>0.1277</td>
<td>0.01673</td>
<td>0.08237</td>
</tr>
<tr>
<td>Eq. 4.4</td>
<td>SentRecChatRatio + RecChatFreq (forward)</td>
<td>0.3554</td>
<td>0.01039</td>
<td>0.06786</td>
</tr>
<tr>
<td>Eq. 4.5</td>
<td>{SIS + HIR + SD + TS} ChatPerc + RecChatFreq + SentRecChatRatio (backward)</td>
<td>0.3732</td>
<td>0.00986</td>
<td>0.06826</td>
</tr>
</tbody>
</table>
It can be said that the more frequent a patient receives a chat message, the higher his self-management level becomes. In addition, the more chats contain statements of certain categories the higher (or lower) patient’s self-management level becomes as well.

4.3 Self-Management Level and Change in HbA1c Level

This section describes the results of analyses performed to find a relation between patient’s self-management level and the change in HbA1c level.

In Figures 4.7 and 4.8 the mean HbA1c levels are plotted for respectively the periods $B_1 - B_{16}$ and $B_{n-8} - B_n$. It can be seen that the average HbA1c decreases until $B_5$ and then starts increasing.

The overall mean change in HbA1c level on the period $B_n - A_2$ of the data set with 105 patients was found as being $-0.43\%$. Over the period $B_n - A_m$ the average percentual change was found as $+1.97\%$ and therefore the patients were grouped to see whether the self-management level was a factor that influenced the HbA1c level.

4.3.1 Patients Grouped, based on Self-Management Level

Grouping the patients in three groups, based on their SML, as described in section 2.2.6, resulted in groups of size 17, 68, and 20 for patients with an SML of respectively low, medium and high level.

The percentual changes in HbA1c level on $B_n - A_m$ of the three patient groups were tested with a Kruskal-Wallis test. This test resulted in the conclusion that not all medians were equal and that the distributions of the values in the three groups were not equal ($\chi^2 = 6.81, p = 0.033$). With Mann-Whitney tests was found that the median of the group high was significantly different from medium ($p = 0.0428$) and low ($p = 0.0065$). The medians in the groups medium and low were not found significantly different ($p = 0.055$).

In Figure 4.9 are the boxplots drawn with percentual changes per group.

In Table D.1 the differences in means of several variables between the three patient groups are listed. It can be seen that the patients with a high SML log in on average 5.2 and 11.3 times more than patients with respectively a medium and low SML. In addition, the percentage of finished tasks is higher. These numbers are also drawn in boxplots in Figure 4.13 till 4.18. The average change in HbA1c level on $B_n - A_m$ is a decrease of 5.4% for patients with a high SML versus an increase of 1.6% and 12.0% for patients with respectively a medium and low SML. Furthermore can be seen from Table D.1 that patients with a high SML use the self-management platform for a longer time period (number of active days) than the other patients and their number of inactive days is lower than in both other groups. So, based on these numbers can be said that patients with a high SML are more committed to use the self-management platform and their change in HbA1c level is better for their health. The total number of days (number of active days + number of inactive days) of the three groups is tested on difference with a Kruskal-Wallis test. With that test was not found that the medians were different.
CHAPTER 4. RESULTS

Figure 4.7: Average HbA1c levels on the period $B_1 - B_{16}$ where $B_1$ indicates the first HbA1c measurement of the patient and $B_i$ the $i^{th}$ measurement. All shown measurements are from the period in which patients did not use the self-management platform. The more to the right in the figure, the less patients are included because not each patient has multiple measurement values before using the self-management platform.

Figure 4.8: Average HbA1c levels on the period $B_{n-8} - B_n$ where $B_n$ is the last measurement value before starting with the self-management platform. The dashed line indicates the moment on which the patient starts using the self-management platform.

$(\chi^2 = 2.95, p = 0.23)$, so is assumed that these number of days over the groups is not different.

In addition, the percentual change on the period $B_n - A_2$ was calculated for the three groups. These numbers are listed in Table D.1 and the boxplots are drawn in Figure 4.10. With a Kruskal-Wallis test was found that the medians of the three groups are not all equal ($\chi^2 = 9.42, p = 0.0090$). Therefore, was tested with Mann-Whitney tests for differences between low and medium, medium and high, and high and low. Respectively $p$-values of 0.0286, 0.0021, and 0.0205 were found, so all medians are different from each other. Moreover, with a Shapiro’s test was found that the distributions of the percentual changes on $B_n - A_2$ could be assumed as normal distributed ($p > 0.05$). With an ANOVA test was found that the means of the percentual changes within the groups were different ($p = 0.0424$). In Figures 4.12 and 4.11 these averages per group are drawn for respectively the periods $B_n - A_m$ and $B_n - A_2$. These figures clearly show that the average HbA1c level of patients with a high SML decreases to almost below the average of the patients with a medium and low SML at $A_2$. On the longer term ($A_m$) these lines do intersect as well. In the remainder of this section the difference in HbA1c level on the period $B_n - A_m$ is considered.
CHAPTER 4. RESULTS

Figure 4.9: Boxplot with percentual HbA1c change on $B_n - A_m$ per patient group

Figure 4.10: Boxplot with percentual HbA1c change on $B_n - A_2$ per patient group

Figure 4.11: Average HbA1c level on $B_n - A_m$ per patient group

Figure 4.12: Average HbA1c level on $B_n - A_2$ per patient group
CHAPTER 4. RESULTS

Figure 4.13: Boxplot with login frequency per patient group, based on SML.

Figure 4.14: Boxplot with percentage of tasks finished per patient group, based on SML.

Figure 4.15: Boxplot with sent chat frequency per patient group, based on SML.

Figure 4.16: Boxplot with received chat frequency per patient group, based on SML.
4.3.2 Linear Regression Model

It was tested how patient’s self-management level and the percentual change in HbA1c level were related to each other. This sub section covers the analyses with linear regression.

First, a correlation of -0.170 was found and with linear regression a relation between these variables was found as displayed in Equation 4.6 and Table 4.2.

\[
PercDiff = 0.075 - 0.299 \cdot SML^* \tag{4.6}
\]

|          | Estimate | Std. error | t-value | \(P(>|t|)\) |
|----------|----------|------------|---------|-------------|
| Intercept| 0.07533  | 0.03503    | 2.150   | 0.0355      |
| SML      | -0.29902 | 0.21266    | -1.406  | 0.1648      |

With the test set a mean squared error of 0.0339 and a mean absolute error of 0.1842 were found; these errors are relatively high. From Table 4.2 follows that the (non-significant) coefficient of SML is equal to \(\beta = -0.30\) and this means that the higher the SML, the more decrease in HbA1c level is expected. This result is consistent with the found differences between the three patient groups.
4.4 Patient-HCP Interaction and Change in HbA1c Level

This section describes the relation between the patient-HCP interaction and the change in HbA1c level. As discussed in section 2.3, linear regression and regression trees are used for prediction models. The analyses were started with using constant models in section 4.4.1 as prediction model, for having a reference point in prediction errors. Secondly, the results of the linear regression models are given in section 4.4.2. Lastly, the results of the regression trees are described in section 4.4.3. In addition, the comparison of the found models and a conclusion which model fits the best is done in section 4.4.4. PercDiff refers to the percentual change in HbA1c level during the period $B_n - A_m$.

4.4.1 Constant Models

The first, simple, prediction models are using the minimum, average, and maximum of the percentual change in HbA1c level of the training set as prediction value. This results in the formulas as given in Equations 4.7 to 4.9 and the prediction errors as listed in Table 4.3. $P_{tr} \subseteq P$ is the training set of the patient set.

$$\hat{\text{PercDiff}} = \min_{p \in P_{tr}} \frac{A^p_m - B^p_n}{B^p_n} = +5.28\%$$  

$$\hat{\text{PercDiff}} = \text{mean}_{p \in P_{tr}} \frac{A^p_m - B^p_n}{B^p_n} = -47.73\%$$  

$$\hat{\text{PercDiff}} = \max_{p \in P_{tr}} \frac{A^p_m - B^p_n}{B^p_n} = +111.5\%$$

4.4.2 Linear Regression Models

In this sub section the with linear regression built prediction models are described.

First, the patient-HCP interaction is quantified with all variables except the chat content itself (SML, HCP login frequency, HCP sent-received chat ratio, sent-received chat ratio, received chat frequency, and days between first login and $A_m$). With both a forward and a backward approach the model as displayed in Equation 4.10 is found. The prediction errors are listed in Table 4.3 and more details are given in Table D.8.

$$\hat{\text{PercDiff}} = -0.113 + 0.00038 \cdot \text{DaysFLAm}$$  

From this model can be concluded that the number of days between patient’s first login and $A_m$ is an important factor. The longer a patient uses the self-management platform, the less decrease (or more increase) in HbA1c level is expected. This is not completely in line with the expectations after the results as showed in section 4.3.1.
Secondly, the chat content was included in the used variable set as well. With all these variables three different approaches were used, namely forcing to include all variables, a forward, and a backward approach. This resulted respectively in a non-significant model, and two times the model as shown in Equation 4.10. Therefore the both forward and backward approaches were used for the complete variable set in which DaysFLAm was neglected. This resulted in the prediction models as shown in respectively Equation 4.11, and 4.12. The prediction errors are listed in Table 4.3 and more details are given in Tables D.9 and D.10.

\[ \text{PercDiff} = 0.176 - 0.351 \cdot HCPLoginFreq^* - 0.231 \cdot ESSupChatPerc^* \] \hspace{1cm} (4.11)

\[ \text{PercDiff} = 0.204 + 0.169 \cdot TSSChatPerc^* - 0.221 \cdot ESSupChatPerc^* - 0.277 \cdot TSChatPerc^* - 0.332 \cdot SML^* - 0.445 \cdot HCPLoginFreq^* \] \hspace{1cm} (4.12)

From the model in Equation 4.11 follows that the more often patient’s HCP logs in and the more chats that contain esteem social support statements a patient receives, the lower his HbA1c level becomes. The model of Equation 4.12 adds that the more chats contain technical support and the higher patient’s self-management level, the lower the HbA1c level becomes as well.

Lastly, based on the literature as discussed in section 1.3.3 was examined what the relation between the chat statement percentages and the change in HbA1c level were. All fractions of chat statement categories (ChatPerc) were separately set as independent variable to predict the dependent variable PercDiff. No significant models were found in this way. From the scatter plots in which these categories are plotted against the percentual change in HbA1c level (as shown in section 3.1.7) follow that a correlation between fractions of chats that contain certain statements and the percentual change in HbA1c level exist. In case of esteem social support statements, the correlation coefficient is equal to -0.139. So, the more chats contain esteem social support statements, the lower patient’s HbA1c level.

### 4.4.3 Regression Trees

In this sub section the found regression trees to predict the percentual HbA1c change are drawn. Moreover, all the prediction errors (tested with the test set) are listed in Table 4.3.

In Figure 4.19 a regression tree based on all variables except the chat content is drawn. The mean squared error and mean absolute error are respectively 0.063 and 0.180. As can be seen from this tree, patients with an HCP who logs in more than 0.176 and less than 0.379 times a day and having less than 444.5 days between patient’s first login and Am have a predicted decreasing HbA1c level. When patient’s HCP logs in more frequently than 0.397 times a days and the patient has less than 366 days between his first login and Am his HbA1c level is predicted decreasing as well.
Chapter 4. Results

Figure 4.19: Regression tree - no chat content; RecChatFreq indicates the received chat frequency by the patient, HCPLoginFreq indicates the login frequency of patient’s HCP, DaysFLAm indicates the number of days between patient’s first login and $A_m$, SML indicates patient’s self-management level, and HCPSentRecChatRatio indicates the ratio of sent-received chats by patient’s HCP. The numbers at the leaves indicate the predicted percentual change in HbA1c level.

In Figure 4.20 the regression tree based on all variables (including chat content) is drawn. The mean squared error and mean absolute error are respectively 0.067 and 0.197. As can be seen from this tree, patients will have a predicted decreasing HbA1c level when their HCP logs in more than 0.379 times a day and have less than 366 days between their first login and $A_m$. If patient’s HCP logs in less than 0.379 but more than 0.176 times a day and less than 29.17% of the chats received by the patient contain tangible social support statements, the patient will have a predicted decreasing HbA1c level as well.

Figure 4.20: Regression tree; RecChatFreq indicates the received chat frequency by the patient, HCPLoginFreq indicates the login frequency of patient’s HCP, DaysFLAm indicates the number of days between patient’s first login and $A_m$, SML indicates patient’s self-management level, TSS ChatPer indicates the fraction of chats that contain tangible social support statements, and HCPSentRecChatRatio indicates the ratio of sent-received chats by patient’s HCP. The numbers at the leaves indicate the predicted percentual change in HbA1c level.
4.4.4 Comparison of the Models

This sub section includes and compares the prediction errors of the found prediction models. In Table 4.3 are these prediction errors listed. From the constant models it can be easily seen that the model with only the average HbA1c level of the training set as prediction value scores the best, according to the error terms. Comparing the linear models with the regression trees results in concluding that the prediction errors of the trees are bigger than those of the linear models. The smallest prediction errors are found with the model that predicts the percentual change in HbA1c level with only the number of days between patient’s first login and his last measurement (Equation 4.10). The model that predicts the best after this model, is the model with patient’s HCP’s login frequency and the fraction of chats that contain esteem social support statements (Equation 4.11). Including more variables into the models did not always improve prediction results. All prediction errors are relatively big, so no perfect model for the prediction is found.

Table 4.3: Comparison of models with their mean squared error (MSE) and mean absolute error (MAE) of the test set. $R^2_{adj}$ is the percentage of variance of the training set that is explained by the model. C, L, and T mean respectively constant model, linear model, and regression tree model. For the regression model marked with $^1$ only the patients with a high SML are used to build the model.

<table>
<thead>
<tr>
<th>Model</th>
<th>Type</th>
<th>Included attributes</th>
<th>$R^2_{adj}$</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eq. 4.7</td>
<td>C</td>
<td>Minimum</td>
<td>-</td>
<td>0.22846</td>
<td>0.44728</td>
</tr>
<tr>
<td>Eq. 4.8</td>
<td>C</td>
<td>Average</td>
<td>-</td>
<td>0.03526</td>
<td>0.14405</td>
</tr>
<tr>
<td>Eq. 4.9</td>
<td>C</td>
<td>Maximum</td>
<td>-</td>
<td>1.34031</td>
<td>1.14538</td>
</tr>
<tr>
<td>Eq. 4.10</td>
<td>L</td>
<td>DaysFLAm</td>
<td>0.07972</td>
<td>0.03437</td>
<td>0.14239</td>
</tr>
<tr>
<td>Eq. 4.11</td>
<td>L</td>
<td>HCPLoginFreq + ESSup ChatPerc</td>
<td>0.05721</td>
<td>0.04022</td>
<td>0.16129</td>
</tr>
<tr>
<td>Eq. 4.12</td>
<td>L</td>
<td>{TSS + ESSup + TS} ChatPerc + SML + HCPLoginFreq</td>
<td>0.08345</td>
<td>0.04874</td>
<td>0.17409</td>
</tr>
<tr>
<td>Fig. 4.19</td>
<td>T</td>
<td>HCPLoginFreq + SML + RecChatFreq + DaysFLAm + HCPSentRecChatRatio</td>
<td>-</td>
<td>0.06278</td>
<td>0.18024</td>
</tr>
<tr>
<td>Fig. 4.20</td>
<td>T</td>
<td>HCPLoginFreq + SML + RecChatFreq + DaysFLAm + TSS ChatPerc + HCPSentRecChatRatio</td>
<td>-</td>
<td>0.06667</td>
<td>0.19684</td>
</tr>
</tbody>
</table>

4.5 Questionnaire and Interview

This section covers the results of the planned and performed questionnaire and interviews.

One of the two asked care group boards gave permission to approach their HCPs and none of them gave permission to approach their patients. Therefore, patients from a patient panel were asked for an interview. Three out of four HCPs and two out of three patients responded. These numbers are much too low to perform statistical analyses on the results. Therefore, this input was not used in the further research and therefore was only focused on the data analysis. More details can be found in Appendix B.
4.6 Summary of Results

This section summarizes the results that are described in this chapter. These results are summarized with the statements as listed below. More details regarding these statements can be found in this chapter.

1. An own estimator for the self-management level is found as the product of the normalized login frequency and the square root of the percentage of tasks that are finished.

2. Patients who receive more than 0.005 chat messages on average per day have a significantly higher self-management level \((p < 0.00001)\) than patients who receive less chat messages.

3. The more frequent a patient receives a chat message, the higher patient’s self-management level \((\beta = 4.754, p = 0.0907)\).

4. The more frequent a patient receives a chat message with social integration support, the higher patient’s self-management level \((\beta = 0.0598, p = 0.179)\).

5. Patients with a high self-management level have a significantly different change (a decrease) in HbA1c level than patients with a low or medium self-management level. In addition, their login frequency, percentage of tasks that is finished, sent and received chat frequency, and active days are higher.

6. The higher patient’s self-management level, the more decrease in HbA1c level (and the more health improvement) \((\beta = -0.30, p = 0.16)\).

7. The more days between patients first login and last measurement \((A_m)\), the lower his decrease (or the higher his increase) in HbA1c level.

8. The more chats sent with esteem social support chat statements, the more decrease in HbA1c level and the more health improvement \((\beta = -0.231, p = 0.11)\).

In chapter 5 the conclusions are drawn and in chapter 6 these findings are discussed.
Chapter 5

Conclusion

In this chapter answers on the research questions are given based on the results following from the methodology and results as described in respectively chapters 2 and 4.

Self-management level

From the literature followed that the level of self-management could be determined by for example the Patient Activation Measure (PAM) and the categories of Schermer (2009). Moreover, the PAM levels were related to the number of activities patients perform on a platform. As no PAM data were included in the data set – and no questionnaire could be sent for licensing and contacting reasons – an own quantifier was developed based on the performed actions on the self-management platform. The quantifier is defined in Equation 4.1 in section 4.1 as follows:

\[
SML_p = \mu \text{LoginFrequency}_p \cdot \sqrt{\mu \text{PercFinishedTasks}_p}
\]

and is based on patient’s normalized login frequency and the percentage of tasks that were finished. With simple linear regression was found that the higher this SML, the more decrease in HbA1c level. This quantifier has the same direction as other measures, like the PAM (Insignia Health, n.d.; Willaing et al., 2013).

Patient-HCP interaction and self-management level

It was found that patients who receive on average more than 0.005 chat messages per day have a significantly higher self-management level than patients who receive less chat messages \((p < 0.00001)\). Receiving chat messages is a result of the fact that HCPs send chat messages. Therefore can be concluded that the more often HCPs send chat messages (to patients) the higher their self-management level becomes \((\beta = 4.754, p = 0.09)\). Furthermore, patients who receive (relatively) more chat messages that contain social integration support statements get a higher SML too \((\beta = 0.060, p = 0.179)\). So, just HCP activity as login frequency is not important to patients, but the content of their interaction is important. Patient’s self-management level can be improved by receiving more frequently chat messages (with statements of certain categories).
CHAPTER 5. CONCLUSION

Level of self-management and patient’s health

Patients were divided into three groups based on their self-management level (SML). Patients with a high SML use the self-management platform more than the other patients. The average percentual changes per group in the period $B_n - A_2$ (approximately half a year) were significantly different ($p = 0.0424$). Patients with a high SML had a decreasing HbA1c level of 7.9%, which means an improving (perceived) health, as followed from Broadbent et al. (2015). In addition was found that the patients from the group ‘high SML’ had a significantly different median percentual change in HbA1c level during the period $B_n - A_m$ with respect to the groups ‘low SML’ ($p = 0.0065$) and ‘medium SML’ ($p = 0.0428$). Although the HbA1c level at the baseline $B_n$ was higher than that one of the other groups, having a higher self-management level is related with a better health. Patients of the group with low SML have an increasing HbA1c level that transcends the target value of 53. Moreover, on the longer term ($A_m$) the averages converge, in the disadvantage of patients with a medium and low SML.

For the patients with an age below 70, at least one measurement before the start and at least two after the start with the self-management platform, and using no insulin a relation between the SML and percentual HbA1c change was found ($\beta = -0.30, p = 0.16$). An increase of 0.01 in SML results in an additional decrease of 0.30 percent points in percentual HbA1c change. So, it is important to support the patient in a way his self-management level increases to improve patient’s health.

Patient-HCP interaction and patient’s health

For the relation between patient’s health and the patient-HCP interaction multiple models were derived, based on the data set. The model with the smallest prediction errors was based on just the number of days between patient’s first login and patient’s last HbA1c measurement date, $A_m$. It was found – as expected – that receiving a higher percentage of chats containing esteem social support statements had an effect on the decrease in HbA1c level. In addition, the more often patient’s HCP logs in, the lower (more negative) his predicted change in HbA1c level ($\beta = -0.351, p = 0.09$). So, the more frequent patient’s HCP logs in and the more frequent a patient gets a chat message from his HCP (the better the online patient-HCP interaction), the better patient’s HbA1c value becomes. In addition, esteem social support chat statements seem to make the percentual change in HbA1c level even stronger.

Overall conclusion

Based on the performed data analyses is found that patients with a higher self-management level have a (higher) decrease in HbA1c level compared to patients with a low and medium self-management level. Moreover, patients who receive chat messages more frequently, have a higher self-management level. Also, having a higher self-management level results in getting a lower HbA1c level. In addition, having more interaction (e.g. receiving more chat messages from the HCP) is related with getting a lower HbA1c level as well. It seems that using a self-management platform – in cooperation with the HCP – is related with improving patient’s health (in terms of a decreasing HbA1c level).

48
Chapter 6

Discussion

As was described in section 1.4 previous research showed results of small sample sizes and relatively short time periods by which patients new they were participating in a research. One of the strengths of this study is that the results are based on data from a period of more than one year of 105 patients who did not know that they were participating in a study.

An additional strength is the period length of the research. The number of days during the period $B_n - A_m$ was on average 527 days and was almost one and a half year. So, results on a longer term than was analyzed by e.g. Smith et al. (2004) are seen. Still after more than a year, one sees a decreasing HbA1c level for patients with a high SML, which is a health improvement.

For the self-management level an own quantifier is drawn. The relation between this quantifier and the change in health outcome is similar to the results of the analyses with PAM where was found by Insignia Health (n.d.) and Willaing et al. (2013). They found that higher PAM scores are related with more decreasing HbA1c levels. Patient adherence is often measured as the more use of a self-management platform, the better it is (Sieverink, Kelders & van Gemert-Pijnen, 2017). However, an increasing health level is more important than a patient that uses a self-management platform more frequently. Therefore, eHealth should be tailored to the need of patient (Huygens et al., 2016).

In addition to the strengths, this research has its limitations too. This first limitation is related to the last strength because an own quantifier is used. The results will become more strengthened when a validated measure is used.

Secondly, the usage of the self-management platform differs per patient. Not all included patients were using the self-management platform in the same way, because the adoption of such a platform could be affected by their HCP. According to Talboom-Kamp (2017, p. 197) the usage of self-management platforms is better when the platforms are perfectly integrated and when patients get personal coaching where needed. Once platforms add value to patient’s life, the usage becomes higher. HCPs indicated in the interviews of Sieverink et al. (2014) that a patient platform (patient health record system) has to have advantages for the HCP to ensure they stimulate the self-management of patients. They knew the importance of stimulation and needed training and guidance to use the system in their daily practice. Moreover, when patients are encouraged by their HCP platform usage becomes higher (Talboom-Kamp et al., 2017). This difference could be
caused by the implementation and introduction of the platform by the HCP. This may differ per care group and even per HCP and was not included in this study. Including the background details of the introduction and socio-demographic data of the patients in the research could improve found results. These kind of details may be exported from the general practitioner information system and/or the collaborative health management system. An additional benefit of inclusion of those data is that drug usage and HCP visits could be taken into account too. If patients visit their HCP less often, HCPs will have more time for using the self-management platform. Moreover, inclusion of HCP performance over time (seeing how active (s)he is over time with his/her patients) will give more insight how to interact.

In future studies one could identify the different types of tasks and education instead of only considering the percentage of tasks that were finished to do research how certain types give different results. With this addition the teaching that is needed according to Bodenheimer et al. (2002) can be investigated.

Communication via chat messages was included in the analyses by first dividing the chat messages into chat statements and classifying these chat statements into nine categories. This was done by only one person and in future research the results would become stronger if multiple researchers classify these statements by which it is possible to check the intercoder agreement level.

In the comparison of change in HbA1c level on $B_n - A_m$ of the three patient groups was found that the means in total days were not significantly different ($p = 0.23$). Although these means were not significantly different, also the period $B_n - A_2$ was used to compare the groups. The longer the period, the harder it is to stabilize or decrease the HbA1c level. Therefore is suggested to repeat the analyzes with data of a longer period in which also the number of days do not differ that much between the patient groups. By analyzing data of a longer time period, also the HbA1c change on the long-term could be tested.

In addition, patients with a too high HbA1c level at the start of using a self-management platform (which is the case for patients of the group ‘high SML’) might be more motivated to reach a higher self-management level and to work on a better health. However, it was found that patients who started with (on average) a good HbA1c level (50.4 mmol/ mol for the group ‘low SML’) and a low SML showed a too high HbA1c level. Moreover, the found prediction models for the percentual change in HbA1c level had relatively big prediction errors. This might be explained by the fact that the percentual HbA1c change is not only explained by variables used in the data analysis. According to the research of Keer Diabetes2 Om (2018) other lifestyle and eating patterns are really important for improving health or even healing from DM2 (NOS, 2018). Therefore, research that includes also these factors will improve found results.

Although it was found that patients with a high self-management level had a health improvement during the use of the self-management platform, it is not known how this improvement remains over a longer period. The number of HbA1c measurements during the use of the self-management platform, $m$, ranged from two to seven, where only two patients had more than five measurements. To examine the influence of a self-management platform on patients’ health, a longer usage period and a higher number of measurements are necessary. As is seen from the period before using e-Vita, patients had (on average) a decrease in HbA1c until their fifth measurement. With this research cannot be determ-
ined whether having a self-management platform facilitates a decrease in HbA1c level for a longer period (until target values are reached). In the study of Keer Diabetes2 Om (2018) patients were working on changing their lifestyle. After a training of six months many of them improved their health by changing lifestyle and these outcomes remained after 12 months, so it is possible to change in about one year. Combining this with the seen trend of the first five measurements is concluded that data of at least two years (where each patients has more than five measurements) will improve found results.

In this study is found that not all patients communicate with their HCP; this could be in line with Huygens et al. (2015) that patients are not aware of the existing communication possibilities. Investigating why and how patients use (or do not use) the communication possibilities on a self-management platform is an interesting part of being explored in future research, especially because the results show that patient-HCP interaction improves patient’s health. If patients do not use the communication option, patient and HCP lack knowledge and/or awareness. The HCP should then start a chat and unveil this option to the patient.

With the results of this study, no one-to-one causation between using a self-management platform and an improving health can be concluded. The introduction of a self-management platform to patients could be a factor helping to improve their health. Patients can be educated, but HCPs might be more trained on this, because that is more scaleable than training all the patients. However, knowing as patient to take action himself and that the HCP supports him might be motivating the patient. Moreover, input from relatives or other external factors are not known and may be important as well. Self-management starts with the patient as his own manager in cooperation with his HCP and results in an improved health (Simmons, Baker, Schaefer, Miller & Anders, 2009). The found results indicate a health improvement for patients with a high self-management level and for patients whose HCP communicates frequently with them. A limitation of this study is that patients’ (perceived) health is estimated with only the HbA1c level. Although according to Broadbent et al. (2015) the HbA1c level indicates patients’ perceived health, the results will be stronger if more factors – such as BMI, blood pressure, and glucose levels during the day – are used in estimating patient’s health. To find prove for a one-to-one relation, a longitudinal research in which more of the patients is known, multiple causative factors are included, and the patients can be asked for clarification is needed. However, having and using a self-management platform as patient can still help to learn more about improving lifestyle and to become dedicated to work on it. Moreover, research can be done to check for similar results in other diseases, like cardiovascular diseases, or to check for similar results in other sub sets of patients.

All in all can be said that patients who use the self-management platform and have a high self-management level, although the number of days between patient’s first login and \( A_m \) impacts the percentual change, have an increasing health status and less development of complications because their HbA1c level is decreasing (Ridderstråle et al., 2016). According to Juarez et al. (2013) this will result in a reduction of health care costs as well.

Finally, patient’s health cannot be defined in an easy way and changing lifestyle – as is needed for healing from DM2 – is not only done by making a self-management platform available to the patient. However, such a platform seems to help patients to keep motivated and to get the needed help from the HCP to improve their health.
References


References


References


References


Appendix A

Sample Sizes Previous Research

In Table A.1 a summary, in terms of sample size and period length, of the used literature of section 1.3.3 is given.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Main findings</th>
<th>Type</th>
<th>Number of patients</th>
<th>Period length</th>
</tr>
</thead>
</table>
| Smith et al. (2004)     | Reduction in HbA1c level  
Heavy users have a bigger reduction than light users                                                                                                                                                    | DM1 &  
DM2 | 16                 | 6 months         |
| McMahon et al. (2005)   | Web-based care results in a reduction in HbA1c level  
People with more measurement uploads have a bigger decrease  
People who use the platform more often, have a bigger decrease  
Decline in systolic blood pressure  
Increase in HDL  
Decrease in triglyceride levels                                                                                                    | DM     | 52 + 52 (control)  | 12 months     |
| Franklin et al. (2006)  | Improved self-efficacy, adherence, and self-management  
HbA1c decreased by patients using SweetTalk and intensive insulin therapy  
Daily personalized text messages supported patients                                                                                           | DM1    | 64 + 28            | 12 months     |
| Rami et al. (2006)      | During telemedicine phase the HbA1c levels improved  
Text-messages as reminder support patient’s adherence                                                                                            | DM1    | 36                 | 6 months      |
| Levine et al. (2009)    | Receiving (person-centered) chat messages is related to the frequency of entering self-measurements                                                                                                       | DM1 &  
DM2 | 109                | 6 months         |
| Robinson et al. (2011)  | Person-centered chat messages were the best predictor of patients’ involvement  
More logins and measurements ...                                                                                                               | DM1 &  
DM2 | 109                | 6 months         |
| Turner et al. (2013)    | Login frequency and measurement upload frequency did not predict change in HbA1c  
Significant decrease in HbA1c level during research  
The higher the percentage of chats containing emotional social support, the lower the HbA1c level                                                                                               | DM1 &  
DM2 | 41                 | 6 months         |
| Willaing et al. (2013)  | Low level patient activation is significantly associated with a high HbA1c level  
High HbA1c level is related to unhealth eating behavior, low level of exercise, low educational level, and having too less knowledge                                                                            | DM2    | 993                | 1 measurement |
Appendix B

Questionnaire

B.1 Questions

The questions that were asked to patients are questions about the frequency of sending chat messages, entering measurements, logging in, and performing assigned tasks. Questions about the expected usage of the chat functionality and response time were asked as well. Moreover, one was be asked what could motivate them to enter more measurements and to become more active. On top of that, specific questions regarding rewards (open and closed questions) were stated. At the end, some statements were stated to check the extent to which patients agreed on these statements, and to validate earlier given answers.

The HCPs were asked for some demographic characteristics (age, gender, and education), the digital communication with their patients and expected response time. On top of that, one was be asked how one thinks they would be motivated to become more active with the system. The open questions were analyzed by checking the text that was written and by counting certain statements to create a total overview. The answers to statements and questions with fixed answer options were processed by counting how often certain answers were given. The demographic characteristics would be used to make groups of participants. It was tried to keep the questionnaire short in a way that people could answer it within ten minutes. A translation of the questions is given in sections B.3 and B.4.

B.2 Results

From the interviews with two patients (both male, age 56 and 69, and having at least a Bachelor’s degree) is found that they use the self-management platform for registration of measurements and keeping contact with the HCP. One uses the web portal once a week for all registrations and the other does not use the web portal, but only the app (with which he has problems with logging in). For having contact with the HCP, one is expecting supporting messages and just normal conversation for clarification. Both patients said they will not become more active on the system when they will be rewarded for it; seeing their progress is motivating already. For these two patients can be said that
APPENDIX B. QUESTIONNAIRE

(extra) gamification will not give more motivation to become more active.

The average age of the HCPs that responded is 46 years; they all are female and have a Bachelors degree. They want to use the chat functionality for answering patients’ questions, discussing measurement values, faster communication, and communication with multiple disciplines (other HCPs) at the same time. The expected response time on chats differs between one day and one week.

One HCP wants to receive money each quartile for using the system to become more active, while the others do not expect becoming more active if they will be rewarded. In their opinion, it is more important that patients become more active and seeing that colleagues are active is expected as motivating themselves.

B.3 Questionnaire Patient

This questionnaire was conducted in Dutch. Therefore, only the translation of the questions is given in this section. For questions with similar multiple possible questions, these lists are given at the beginning of their section.

<table>
<thead>
<tr>
<th>Introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Introduction text]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demographic characteristics and general questions e-Vita</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is your age? (text box)</td>
</tr>
<tr>
<td>What is your gender? (male/female)</td>
</tr>
<tr>
<td>What is your highest enrolled education level? (None / secondary school / Bachelor / Master / Doctor / Higher)</td>
</tr>
<tr>
<td>Which diseases do you have? (Asthma / COPD / Diabetics / ...)</td>
</tr>
<tr>
<td>Since when do you use the self-management platform e-Vita? (text box)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1] Answer options: (less than once a month / once a month / more than once a month / once a week / more than once a week / each day / more than once a day)</td>
</tr>
<tr>
<td>How often do you log in on the self-management platform? [1]</td>
</tr>
<tr>
<td>How often do you send a chat message to your HCP? [1]</td>
</tr>
<tr>
<td>How often do you enter new measurements? [1]</td>
</tr>
<tr>
<td>How often do you perform an assigned task? [1]</td>
</tr>
<tr>
<td>Do you want to use e-Vita more often? (yes/no)</td>
</tr>
<tr>
<td>General remarks (text box)</td>
</tr>
</tbody>
</table>

62
## (Digital) communication with HCP

<table>
<thead>
<tr>
<th>Question</th>
<th>Options/Text Box</th>
</tr>
</thead>
<tbody>
<tr>
<td>What kind of feedback/communication do you expect from your HCP via e-Vita?</td>
<td>(text box)</td>
</tr>
<tr>
<td>How fast do you want to receive a reply on your chat message?</td>
<td>(more than a week / within a week / within three days/ within a day / within four hours / within one hour / within ten minutes)</td>
</tr>
<tr>
<td>What do you think of the quality of the answers to your chat messages?</td>
<td>(text box)</td>
</tr>
<tr>
<td>Do you have questions, tips, or comments about the chat functionality?</td>
<td>(text box)</td>
</tr>
<tr>
<td>How often do you have non-digital contact with your HCP? (This can be contact by phone, by post, or by a visit)</td>
<td>(less than once a year / once a year / once a quarter / once a month / more than once a month / once a week / more than once a week / once a day / more than once a day)</td>
</tr>
<tr>
<td>General remarks</td>
<td>(text box)</td>
</tr>
</tbody>
</table>

## Measurements – I

<table>
<thead>
<tr>
<th>Question</th>
<th>Options/Text Box</th>
</tr>
</thead>
<tbody>
<tr>
<td>What would make you entering more measurement values on e-Vita?</td>
<td>(text box)</td>
</tr>
<tr>
<td>What should change to make you (more) active on e-Vita?</td>
<td>(text box)</td>
</tr>
</tbody>
</table>

## Measurements – II

<table>
<thead>
<tr>
<th>Question</th>
<th>Options/Text Box</th>
</tr>
</thead>
<tbody>
<tr>
<td>Would you enter measurements more often if you will receive a reward for this?</td>
<td>(yes/no)</td>
</tr>
<tr>
<td>What kind of (digital) reward would you like to receive?</td>
<td>(text box)</td>
</tr>
<tr>
<td>How often would you like to receive a reward?</td>
<td>(text box)</td>
</tr>
</tbody>
</table>

## Statements


- When I receive more frequently a response from my HCP, I will become more active on e-Vita. [2]
- When I receive points with which I can earn digital badges for my activities on e-Vita, I will become more active on e-Vita. [2]
- It will be fun to see how active other patients are on e-Vita. [2]
- It will be motivating to see how active other patients are on e-Vita. [2]
- I would like to use e-Vita more often. [2]
- I regularly ask questions via the chat functionality to my HCP. [2]
- I regularly receive motivational messages via the chat functionality from my HCP. [2]
- I regularly use the chat functionality. [2]
| General remarks | (text box) |

## Completion

[Word of thank and contact details]
## B.4 Questionnaire HCP

This questionnaire was conducted in Dutch. Therefore, only the translation of the questions is given in this section. For questions with similar multiple possible questions, these lists are given at the beginning of their section.

### Introduction

[Introduction text]

### Demographic characteristics and general questions e-Vita

<table>
<thead>
<tr>
<th>Question</th>
<th>Options/Text Box</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is your age?</td>
<td>(text box)</td>
</tr>
<tr>
<td>What is your gender?</td>
<td>(male/female)</td>
</tr>
<tr>
<td>What is your highest enrolled education level?</td>
<td>(None / secondary school / Bachelor / Master / Doctor / Higher)</td>
</tr>
</tbody>
</table>

### (Digital) communication with patients

<table>
<thead>
<tr>
<th>Question</th>
<th>Options/Text Box</th>
</tr>
</thead>
<tbody>
<tr>
<td>For what purpose do you want to use the chat functionality with your patients?</td>
<td>(text box)</td>
</tr>
<tr>
<td>How fast do you want to receive a reply on your chat message?</td>
<td>(more than a week / within a week / within three days / within a day / within four hours / within one hour / within ten minutes)</td>
</tr>
<tr>
<td>Do you have questions, tips, or comments about the chat functionality?</td>
<td>(text box)</td>
</tr>
<tr>
<td>How often do you have non-digital contact with your patients? (on average per patient)</td>
<td>(less than once a year / once a year / once a quarter / once a month / more than once a month / once a week / more than once a week / once a day / more than once a day)</td>
</tr>
</tbody>
</table>

### General remarks

[text box]

### Activities on e-Vita

<table>
<thead>
<tr>
<th>Question</th>
<th>Options/Text Box</th>
</tr>
</thead>
<tbody>
<tr>
<td>Would you become more active on e-Vita if you will receive a reward for this?</td>
<td>(yes/no)</td>
</tr>
<tr>
<td>What kind of (digital) reward would you like to receive?</td>
<td>(text box)</td>
</tr>
<tr>
<td>How often would you like to receive a reward?</td>
<td>(text box)</td>
</tr>
<tr>
<td>What should change to make you (more) active on e-Vita?</td>
<td>(text box)</td>
</tr>
</tbody>
</table>
**APPENDIX B. QUESTIONNAIRE**

**Statements**


| When I receive more frequently a response from my patients, I will become more active on e-Vita. |
| When I receive points with which I can earn digital badges for my activities on e-Vita, I will become more active on e-Vita. |
| It will be fun to see how active colleagues are on e-Vita. |
| It will be motivating to see how active colleagues are on e-Vita. |
| I would like to use e-Vita more often. |
| I regularly ask questions to patients via the chat functionality. |
| I regularly send motivational messages to patients via the chat functionality. |
| I regularly use the chat functionality. |
| I would recommend e-Vita to other HCPs. |
| e-Vita helps me to monitor patients. |
| I encourage patients to use e-Vita. |
| General remarks (text box) |

**Completion**

[Word of thank and contact details]
Appendix C

Data Description

In this appendix a more detailed version of the data description is given. Data of all performed actions are stored. In this section only the for this research relevant data instances and properties are described.

The first instance that was analyzed is the patient. For each patient instance there exists a row. These patient instances have several attributes, which are listed in this section. First the following sets are defined:

- $DS$ the finite set with all diseases
- $C$ the finite set with all HCPs (e.g. GPs and physician assistants)
- $P$ the finite set with all patients
- $I$ the finite set with all possible measurements
  (e.g. $I = \{BMI, HbA1c, Smoking, bloodpressure, weight, \ldots\}$)

Different measurements are distinguished and are seen as a group of columns for measurement type $i, i \in I$, and this is a finite set. These measurements are disease related and therefore not applicable to each patient. For example, HbA1c and blood glucose level are relevant for patients with diabetes while smoking is relevant for patients with COPD or Asthma.

A tuple is a finite ordered list of elements. For example, the measurement attribute type $i$ consists of a set of tuples $HMS = (v, d)$ where $v$ indicates value, and $d$ the timestamp of the measurement.

The timestamps can be written down as a datetime of as unix timestamp. For the analyses it does not matter in which format these timestamps are in the beginning. It can be formatted later to a useful format. This remark remains for all the mentioned timestamps. The set $I$ of home measurements given before is not complete, but indicates some of the available measurements.

Then the patient instance, $p \in P$, has the following attributes:
- Age (in years) $a$ numerical (ratio) $a \in \mathbb{N}_0$
- Gender $g$ categorical $g \in \{\text{male, female}\}$
- Diseases $DP$ categorical $DP \subseteq DS$
- Insulin $ins$ categorical $ins \in \{\text{yes, no}\}$
- Caregiver $CP$ categorical $CP \subseteq C$
- Creation date $cd$ timestamp on which the account of the patient is made
- Measurements $HM \forall i \in I: \text{set of tuples } hms$ $hms = (v, d)$
  $v$ numerical double
  $d$ timestamp datetime/integer
- Tasks $TP \text{ set of tuples } tp$ $tp = (s, sd, ed, r, df, do)$
  $t$ indicates the task string to identify the task
  $s$ indicates the status e.g. $s \in \{\text{open, finished}\}$
  $sd$ the start datetime datetime/integer
  $ed$ the end datetime datetime/integer
  $r$ the datetime on which a reminder is sent $r$ can be empty, datetime/integer
  $df$ the datetime on which the task is finished datetime/integer
  $do$ the datetime on which the task is opened datetime/integer
- Chats $CP \text{ set of tuples } cp$ $cp = (se, re, cc, tc)$
  $se$ the sender $se \in (p \cup CP)$
  $re$ the receiver $re \in (p \cup CP)$
  $cc$ chat content (Chat message) text string
  $tc$ timestamp when the chat is sent datetime/integer
- Logins $LP \text{ set of tuples } lp$ $lp = (lid, dl, lod)$
  $lid$ timestamp on which the patient logs in datetime/integer
  $dl$ the device with which the patient logs in $dl \in \{\text{web, app}\}$
  $lod$ timestamp on which the patient logs out $lod$ can be empty, datetime/integer
The second instance is the HCP. This instance, \( c \in C \), has the following attributes:

- **Creationdate** \( cd_c \) timestamp on which the account of the HCP is made
- **Profession** \( p_c \) categorical \( p_c \in P_c \) where \( P_c = \{GP, Assistant, ...\} \)
- **Patients** \( P_c \) \( P_c \subseteq P \)
- **Chats** \( CC_c \) set of tuples \( cc \)
  - \( se_c \) the sender \( se_c \in (c \cup P) \)
  - \( re_c \) the receiver \( re_c \in (c \cup P) \)
  - \( cc \) chat content (Chat message) \( \text{text string} \)
  - \( tc_c \) timestamp when the chat is sent \( \text{datetime/integer} \)
- **Logins** \( LC \) set of tuples \( lc \)
  - \( lid_c \) timestamp on which the HCP logs in \( \text{datetime/integer} \)
  - \( dl_c \) the device with which the HCP logs in \( dl_c \in \{web, app\} \)
  - \( lod_c \) timestamp on which the HCP logs out \( lod_c \) can be empty, \( \text{datetime/integer} \)

Some of the attributes of the HCP instance are similar or linked to attributes of the patient instance. For example, the chats are sent between HCPs and patients and therefore the content of the attribute of the patient instance is more or less equal to the content of the attribute of the HCP instance.
Appendix D

Extra Details Results

In this appendix tables and figures with extra details with respect to Chapter 4 are listed. In Chapter 4 is referred to these tables and figures.

**Figure D.1**: Age and gender of patients with (and sorted by) their login frequency. The (orange) dots indicate the female patients and the (black) diamonds indicate the male patients. The left axis represents the age, so the higher a (orange) dot or a (black) diamond, the older the patient is. The small (and red) dots indicate per patient the login frequency from which the scale is drawn on the right side. Moreover, the minimum, average, and maximum ages are drawn with respectively dashed, dotted, and solid lines.
## APPENDIX D. EXTRA DETAILS RESULTS

Table D.1: Descriptive data from patients - Numbers are represented as $m \pm sd$ with $m$ as average and $sd$ the standard deviation. In $N(x)$ represents $N$ the total number and $x$ the percentage. Numbers may not add up to 100% due to rounding. SML groups are defined in section 2.2.6. The frequencies are numbers per day. a: $p = 0.0286$, b: $p = 0.0021$, c: $p = 0.0205$, d: $p = 0.0428$, e: $p = 0.0065$, f: $p = 0.055$, and g: $p = 0.0424$. These p-values are results from section 4.3. Some cells have been intentionally left empty.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sub set</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>105</td>
<td>17</td>
<td>68</td>
<td>20</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>58.5 ± 7.4</td>
<td>56.6 ± 8.4</td>
<td>58.1 ± 7.4</td>
<td>61.2 ± 6.3</td>
</tr>
<tr>
<td>Male</td>
<td>59.4 ± 7.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>56.7 ± 8.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>70 (66.7)</td>
<td>9 (52.9)</td>
<td>45 (66.2)</td>
<td>16 (80.0)</td>
</tr>
<tr>
<td>Female</td>
<td>35 (33.3)</td>
<td>8 (47.1)</td>
<td>23 (33.8)</td>
<td>4 (20.0)</td>
</tr>
<tr>
<td>Login frequency</td>
<td>0.0391 ±</td>
<td>0.0107 ±</td>
<td>0.0222 ±</td>
<td>0.1205 ±</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number active days</td>
<td>252 ± 268</td>
<td>38 ± 58</td>
<td>264 ± 284</td>
<td>393 ± 207</td>
</tr>
<tr>
<td>Number inactive days</td>
<td>289 ± 236</td>
<td>494 ± 203</td>
<td>301 ± 224</td>
<td>76 ± 85</td>
</tr>
<tr>
<td>Tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number assigned</td>
<td>9.83 ± 8.4</td>
<td>3.12 ± 1.73</td>
<td>10.57 ± 8.90</td>
<td>13.00 ± 7.44</td>
</tr>
<tr>
<td>Number open</td>
<td>1.74 ± 1.92</td>
<td>2.94 ± 1.56</td>
<td>1.88 ± 2.00</td>
<td>0.25 ± 0.55</td>
</tr>
<tr>
<td>Finished (%)</td>
<td>67.7 ± 35.8</td>
<td>2.9 ± 12.1</td>
<td>75.1 ± 23.7</td>
<td>97.8 ± 4.9</td>
</tr>
<tr>
<td>Chat frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sent</td>
<td>0.000289 ±</td>
<td>0.00046 ±</td>
<td>0.00126 ±</td>
<td>0.01052 ±</td>
</tr>
<tr>
<td></td>
<td>0.00968</td>
<td>0.0014</td>
<td>0.0034</td>
<td>0.0122</td>
</tr>
<tr>
<td>received</td>
<td>0.00496 ±</td>
<td>0.00155 ±</td>
<td>0.00321 ±</td>
<td>0.01349 ±</td>
</tr>
<tr>
<td></td>
<td>0.00820</td>
<td>0.0032</td>
<td>0.0046</td>
<td>0.0135</td>
</tr>
<tr>
<td>HbA1c levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(mmol/mol)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B_n$</td>
<td>55.2 ± 13.7</td>
<td>50.4 ± 12.5</td>
<td>54.6 ± 11.8</td>
<td>61.0 ± 18.6</td>
</tr>
<tr>
<td>$A_2$</td>
<td>53.8 ± 11.5</td>
<td>53.6 ± 13.5</td>
<td>53.8 ± 11.3</td>
<td>53.7 ± 11.1</td>
</tr>
<tr>
<td>$A_m$</td>
<td>54.9 ± 13.0</td>
<td>56.0 ± 18.2</td>
<td>54.3 ± 11.5</td>
<td>55.8 ± 13.3</td>
</tr>
<tr>
<td>HbA1c change (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median $B_n - A_2$</td>
<td>0 ± 7.6$^{a,c}$</td>
<td>0$^a$</td>
<td>−10.3$^{b,c}$</td>
<td></td>
</tr>
<tr>
<td>Mean $B_n - A_2$</td>
<td>−0.2 ± 18.1</td>
<td>+6.8 ± 10.1$^g$</td>
<td>0.3 ± 18.1$^g$</td>
<td>−7.9 ± 21.9$^g$</td>
</tr>
<tr>
<td>Median $B_n - A_m$</td>
<td>0 ± 6.3$^{e,f}$</td>
<td>0$^d$</td>
<td>−7.9$^{e,d}$</td>
<td></td>
</tr>
<tr>
<td>Mean $B_n - A_m$</td>
<td>+2.0 ± 22.4</td>
<td>+12.0 ± 28.1</td>
<td>+1.6 ± 21.0</td>
<td>−5.4 ± 19.5</td>
</tr>
<tr>
<td>Minimum $B_n - A_m$</td>
<td>−47.7</td>
<td>−15.8</td>
<td>−47.7</td>
<td>−44.4</td>
</tr>
<tr>
<td>Maximum $B_n - A_m$</td>
<td>+111.5</td>
<td>+111.5</td>
<td>+70.0</td>
<td>+38.5</td>
</tr>
<tr>
<td>Number days $FLA_m$</td>
<td>415 ± 201</td>
<td>401 ± 191</td>
<td>436 ± 204</td>
<td>357 ± 195</td>
</tr>
<tr>
<td>m</td>
<td>2.80 ± 1.03</td>
<td>2.24 ± 0.44</td>
<td>2.87 ± 1.02</td>
<td>3.05 ± 1.28</td>
</tr>
</tbody>
</table>
### Table D.2: Descriptive data chat statements - Numbers may not add up to 100% due to rounding. SML groups are defined in section 2.2.6.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chat statements of category</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISS (%)</td>
<td>5.6 ± 10.4</td>
<td>2.50 ± 7.09</td>
<td>4.84 ± 9.79</td>
<td>10.85 ± 12.94</td>
</tr>
<tr>
<td>TSS (%)</td>
<td>5.7 ± 9.8</td>
<td>3.53 ± 8.43</td>
<td>4.68 ± 7.68</td>
<td>10.96 ± 14.71</td>
</tr>
<tr>
<td>SIS (%)</td>
<td>21.8 ± 20.4</td>
<td>13.79 ± 20.72</td>
<td>21.41 ± 21.11</td>
<td>29.81 ± 14.56</td>
</tr>
<tr>
<td>ESS (%)</td>
<td>3.5 ± 6.0</td>
<td>2.14 ± 4.96</td>
<td>3.70 ± 6.53</td>
<td>3.85 ± 5.06</td>
</tr>
<tr>
<td>ESSup (%)</td>
<td>2.4 ± 6.8</td>
<td>1.36 ± 4.24</td>
<td>1.84 ± 6.76</td>
<td>4.93 ± 8.16</td>
</tr>
<tr>
<td>HIR (%)</td>
<td>4.4 ± 7.7</td>
<td>0.19 ± 0.78</td>
<td>4.91 ± 7.93</td>
<td>6.24 ± 9.08</td>
</tr>
<tr>
<td>SD (%)</td>
<td>0.3 ± 1.2</td>
<td>0.38 ± 1.56</td>
<td>0.30 ± 1.26</td>
<td>0.15 ± 0.45</td>
</tr>
<tr>
<td>TS (%)</td>
<td>0.8 ± 3.5</td>
<td>2.65 ± 7.52</td>
<td>0.31 ± 1.31</td>
<td>1.10 ± 2.81</td>
</tr>
<tr>
<td><strong>Chats containing category</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISS (%)</td>
<td>15.2 ± 27.0</td>
<td>9.80 ± 28.30</td>
<td>13.41 ± 26.22</td>
<td>26.02 ± 27.13</td>
</tr>
<tr>
<td>TSS (%)</td>
<td>22.5 ± 33.4</td>
<td>14.71 ± 34.30</td>
<td>19.40 ± 31.18</td>
<td>39.64 ± 35.90</td>
</tr>
<tr>
<td>SIS (%)</td>
<td>56.3 ± 48.1</td>
<td>35.29 ± 49.26</td>
<td>53.43 ± 48.83</td>
<td>83.82 ± 31.56</td>
</tr>
<tr>
<td>ESS (%)</td>
<td>17.3 ± 30.6</td>
<td>12.75 ± 29.77</td>
<td>18.21 ± 32.41</td>
<td>17.89 ± 25.39</td>
</tr>
<tr>
<td>ESSup (%)</td>
<td>8.1 ± 20.5</td>
<td>7.84 ± 25.08</td>
<td>5.15 ± 15.76</td>
<td>18.40 ± 27.33</td>
</tr>
<tr>
<td>HIR (%)</td>
<td>19.0 ± 32.1</td>
<td>1.96 ± 8.08</td>
<td>21.08 ± 33.99</td>
<td>26.29 ± 34.51</td>
</tr>
<tr>
<td>SD (%)</td>
<td>1.4 ± 5.7</td>
<td>1.96 ± 8.08</td>
<td>1.40 ± 5.70</td>
<td>0.77 ± 2.39</td>
</tr>
<tr>
<td>TS (%)</td>
<td>3.9 ± 15.3</td>
<td>11.76 ± 33.21</td>
<td>1.73 ± 7.03</td>
<td>4.48 ± 10.61</td>
</tr>
</tbody>
</table>
Table D.3: Chat statement coding categories; total number received by patients, mean, minimum, and maximum received per patient

<table>
<thead>
<tr>
<th>Category</th>
<th>Total</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informational Social Support</td>
<td>160</td>
<td>1.524</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>Tangible Social Support</td>
<td>99</td>
<td>0.943</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Social Integration Support</td>
<td>432</td>
<td>4.114</td>
<td>0</td>
<td>41</td>
</tr>
<tr>
<td>Emotional Social Support</td>
<td>83</td>
<td>0.790</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Esteem Social Support</td>
<td>57</td>
<td>0.543</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Health Information Requests</td>
<td>93</td>
<td>0.886</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Self-Disclosure</td>
<td>11</td>
<td>0.105</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Technical Support</td>
<td>15</td>
<td>0.142</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Table D.4: Linear model to predict SML based on chat content categories, \( R^2 = 0.2195, R^2_{adj} = 0.07915. \)

| Estimate | Std. error | t-value | \( P(>|t|) \) |
|----------|------------|---------|---------------|
| Intercept| 0.016951   | 0.02694 | 0.629         | 0.5319       |
| ISS ChatPerc | 0.012427 | 0.07662 | 0.162         | 0.8718       |
| TSS ChatPerc | 0.002788 | 0.07742 | 0.036         | 0.9714       |
| SIS ChatPerc | 0.127521 | 0.06087 | 2.095         | 0.0409       |
| ESS ChatPerc | -0.035639 | 0.08371 | -0.426        | 0.6720       |
| ESSup ChatPerc | 0.200757 | 0.09019 | 2.226         | 0.0302       |
| HIR ChatPerc | -0.065080 | 0.07295 | -0.894        | 0.3753       |
| SD ChatPerc  | -0.320892 | 0.27656 | -1.160        | 0.2510       |
| TS ChatPerc  | -0.172662 | 0.11029 | -1.565        | 0.1233       |

Table D.5: Linear model to predict SML based on chat content categories - stepwise approach, \( R^2 = 0.1559, R^2_{adj} = 0.1277. \)

| Estimate | Std. error | t-value | \( P(>|t|) \) |
|----------|------------|---------|---------------|
| Intercept| 0.01835    | 0.02656 | 0.691         | 0.4922       |
| SIS ChatPerc | 0.07248    | 0.03635 | 2.044         | 0.0454       |
| ESSup ChatPerc | 0.18212    | 0.08364 | 2.177         | 0.0334       |

Table D.6: Linear model to predict SML based on all attributes, forward approach, \( R^2 = 0.3762, R^2_{adj} = 0.3554. \)

| Estimate | Std. error | t-value | \( P(>|t|) \) |
|----------|------------|---------|---------------|
| Intercept| 0.01985    | 0.01832 | 1.084         | 0.2828       |
| SentRecChatRatio | 0.1404    | 0.03292 | 4.265         | 0.000072     |
| RecChatFreq | 4.7541    | 2.76487 | 1.719         | 0.0907       |
### Table D.7: Linear model to predict SML based on all attributes, backward approach, $R^2 = 0.4338$, $R^2_{adj} = 0.3732$.

|        | Estimate | Std. error | t-value | $P(>|t|)$ |
|--------|----------|------------|---------|-----------|
| Intercept | 0.01608  | 0.02240    | 0.718   | 0.4759    |
| SIS ChatPerc | 0.05975  | 0.04388    | 1.362   | 0.1787    |
| HIR ChatPerc  | -0.08022 | 0.05535    | -1.449  | 0.1529    |
| SD ChatPerc   | -0.35569 | 0.22452    | -1.584  | 0.1188    |
| TS ChatPerc   | -0.15869 | 0.08619    | -1.841  | 0.0709    |
| RecChatFreq   | 4.97218  | 3.11903    | 1.594   | 0.1165    |
| SentRecChatRatio | 0.13786 | 0.03288    | 4.193   | 0.00099   |

### Table D.8: Linear model to predict percentual change in HbA1c level. DaysFLAm indicate the number of days between patient’s first login and $A_m$, $R^2_{adj} = 0.07972$.

|        | Estimate | Std. error | t-value | $P(>|t|)$ |
|--------|----------|------------|---------|-----------|
| Intercept | -0.11303 | 0.0723     | -1.561  | 0.1229    |
| DaysFLAm | 0.00038  | 0.0002     | 2.524   | 0.0142    |

### Table D.9: Linear model to predict percentual change in HbA1c level. HCPLoginFreq indicates how often the HCP logs in per day and ESSup ChatPerc indicates the fraction of chats that contain esteem social support statements, $R^2_{adj} = 0.05721$.

|        | Estimate | Std. error | t-value | $P(>|t|)$ |
|--------|----------|------------|---------|-----------|
| Intercept | 0.1760   | 0.06808    | 2.585   | 0.0122    |
| HCPLoginFreq | -0.3511  | 0.20451    | -1.717  | 0.0912    |
| ESSup ChatPerc | -0.2313  | 0.14405    | -1.606  | 0.1135    |

### Table D.10: Linear model to predict percentual change in HbA1c level. HCPLoginFreq indicates how often the HCP logs in per day, ESSup ChatPerc indicates the fraction of chats that contain esteem social support statements, TSS ChatPerc the fraction of chats that contain tangible social support statements, TS ChatPerc the fraction of chats that contain technical support statements, and SML patient’s self-management level, $R^2_{adj} = 0.08345$.

|        | Estimate | Std. error | t-value | $P(>|t|)$ |
|--------|----------|------------|---------|-----------|
| Intercept | 0.2035   | 0.0694     | 2.932   | 0.0048    |
| TSS ChatPerc | 0.1692   | 0.1118     | 1.513   | 0.1358    |
| ESSup ChatPerc | -0.2209  | 0.1537     | -1.438  | 0.1560    |
| TS ChatPerc | -0.2768  | 0.1766     | -1.568  | 0.1224    |
| SML     | -0.3322  | 0.2223     | -1.495  | 0.1405    |
| HCPLoginFreq | -0.4449  | 0.2122     | -2.096  | 0.0405    |