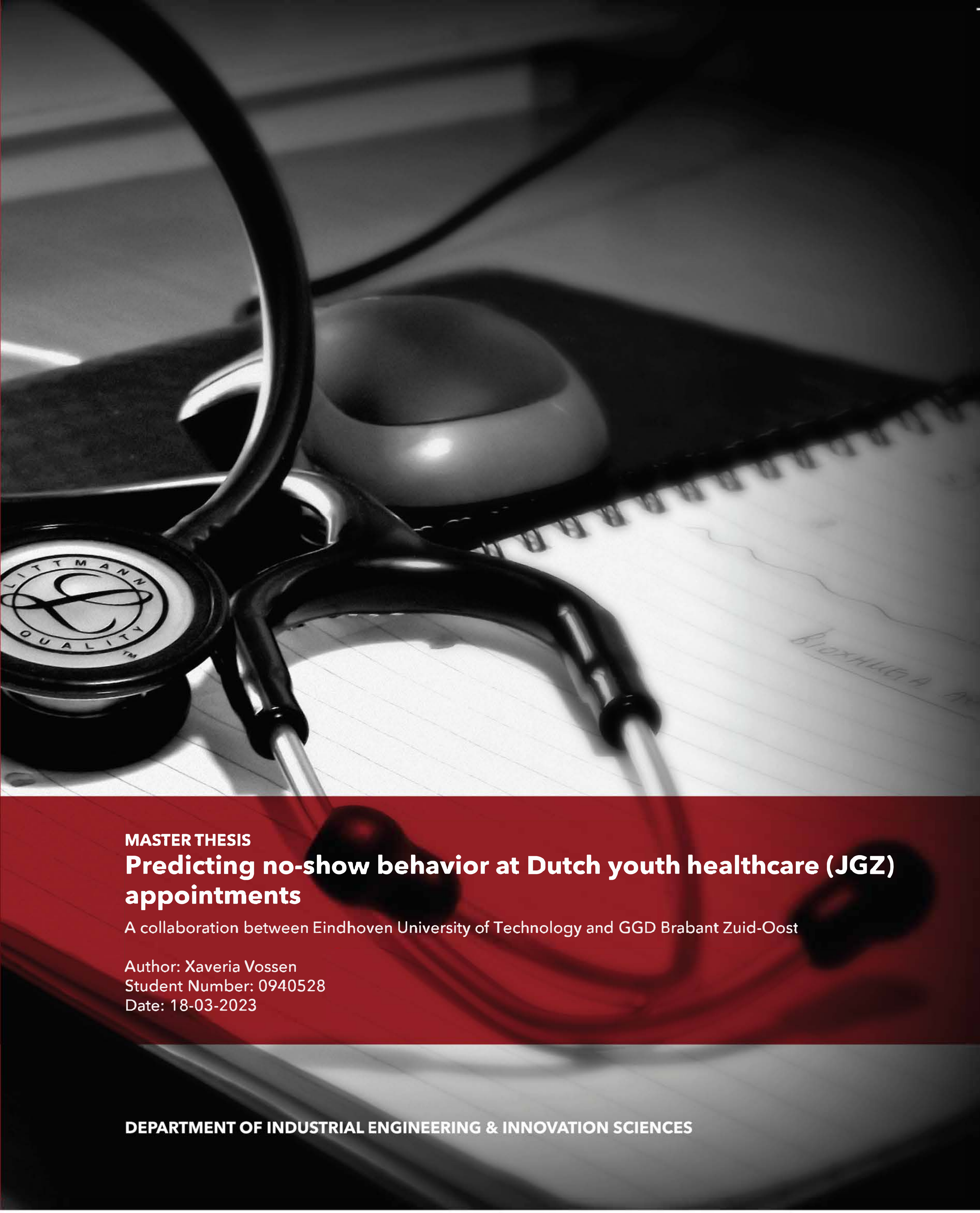
MASTER

Predicting no-show behavior at Dutch youth healthcare (JGZ) appointments

Vossen, Xaveria L.A.

Award date:
2023



MASTER THESIS

Predicting no-show behavior at Dutch youth healthcare (JGZ) appointments

A collaboration between Eindhoven University of Technology and GGD Brabant Zuid-Oost

Author: Xaveria Vossen
Student Number: 0940528
Date: 18-03-2023

DEPARTMENT OF INDUSTRIAL ENGINEERING & INNOVATION SCIENCES

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Eindhoven University of Technology
Human Technology Interaction

Master Thesis Project

Predicting no-show behavior at Dutch youth healthcare (JGZ) appointments

By Xaveria Vossen
x.l.a.vossen@student.tue.nl - 0940528

in partial fulfillment of the requirements for the degree of

Master of Science
in Human Technology Interaction

First Assessor: Rianne Conijn - m.a.conijn@tue.nl
Second Assessor: Chris Snijders - c.c.p.snijders@tue.nl
Company Mentor: Arne Meeldijk (GGD) - a.meeldijk@ggdbzo.nl

Eindhoven, Saturday 18th March, 2023

Table of Content

Preface

The thesis in front of you is the thesis 'Predicting presence at Dutch youth healthcare (JGZ) appointments.' This study was carried out in a collaboration between the Eindhoven University of Technology and GGD Brabant Zuid-Oost. This thesis completes my graduation from the Human Technology Interaction department. From September 2022 until January 2023, I have been working on this thesis. Throughout this process, I was supervised by dr. ir. Rianne Conijn and prof. dr. Chris Snijders within the Human Technology Interaction department, and ir. Arne Meeldijk within the GGD Brabant Zuid-Oost Research department.

The past period has been a period with many ups and downs, in the most moving period of my life. Passing my thesis project over time got linked to overcoming personal setbacks, which drove my motivation throughout this project. Despite all challenges, I completed this project, of which I am really proud.

As a result, I want to thank my supervisor, dr. ir Rianne Conijn, for our weekly meetings, you guided me in the right direction and supported me. I also want to thank my second supervisor Chris Snijders for his help and suggestions throughout the project. Also, I want to thank the GGD Brabant Zuid-Oost, especially ir. Arne Meeldijk for enabling me to graduate within the GGD Brabant Zuid-Oost, and contributing to the first steps of using data-driven approaches in Dutch youth healthcare (JGZ). Last but not least, I will thank my family, friends, and lovely boyfriend, Nic, for providing me with the support, delicious cheese, and love I needed to preserve during my thesis project, especially during the last stretch. It's time for a new chapter!

Enjoy reading this final master thesis

Xaveria Vossen

Abstract

This thesis project studies no-show behavior in the context of Dutch youth healthcare (JGZ) appointments, which are part of GGD Brabant Zuid-Oost. The study aims to increase healthcare delivery efficiency by addressing the primary research question: How can no-show behavior at Dutch youth healthcare (JGZ) appointments be predicted? To answer this question, the study is divided into three sections: an exploratory model that examines the behavioral and environmental factors that determine a person's probability of missing appointments, the most accurate classification method that classifies those who are most likely to exhibit no-show behavior, and machine learning and decision-making, in which the model is evaluated for its capacity to predict independently of data subgroups and to indicate whether domain experts would be likely to accept the assistance of machine learning methods in the decision-making process.

The study shows that demographic factors, such as where a patient lives and socioeconomic status, can predict whether a patient misses an appointment. This may be due to the effort required to attend the appointment. Additionally, past appointment behavior can predict no-show behavior, possibly because people tend to behave consistently over time. Furthermore, characteristics such as the number and type of appointments can also influence whether someone will attend an appointment, possibly due to the perceived importance of the appointment. This study concludes the predictors of no-show behavior found within the context of Dutch youth healthcare (JGZ) appointments are generalizable. In contrast, the psychological constructs causing the effects are understudied and suggest potential reasons why the effect in this study might be found.

Next, the study also found that the differences in accuracy between the random forest and logistic regression methods were rather small. While the random forest method based on accuracy was recommended for predicting no-show behavior, this study suggests using logistic regression in the specific context of the GGD Brabant Zuid-Oost because it is easily interpretable and more useful in practice.
Moreover, the study found that there were no significant differences between subgroups of data, indicating that the prediction model performs equally well independently of data subgroups. However, the domain experts' understanding of the problem did not fully align with how the model interpreted the data, specifically regarding the impact of nationality on appointment non-attendance. Despite this deviation, the study recommends using machine learning algorithms to assist healthcare decision-making.

In terms of practical implications, the study recommends sending reminders to patients and educating them about the importance of attending appointments to improve attendance rates. Using this study's results enables classifying different patient groups that should be targeted with different strategies. The study also suggests using an "open access" scheduling system to minimize the operational impact of no-shows. The ultimate goal is to improve the effectiveness of healthcare in the context of Dutch youth health care (JGZ) appointments.

Chapter 1

Introduction

This study examines appointment behavior in the context of healthcare. Appointment behavior in healthcare refers to patterns and behaviors patients exhibit when scheduling and attending appointments. This study focuses on so-called no-show behavior, where patients fail to attend appointments. Low attendance rates for healthcare appointments have been associated with poor health outcomes and efficiency problems for healthcare providers (Ferro, Brailsford, Bravo, & Smith, 2020). On the one hand, there is a relationship between low attendance rates and adverse health outcomes such as delays in diagnosis and start of treatment, increased use of emergency services, and higher premature mortality rates (Zebina et al., 2019; McQueenie, Ellis, McConnachie, Wilson, & Williamson, 2019; Wallace et al., 2018). On the other hand, high no-show rates decrease the efficiency of healthcare providers. When a patient misses an appointment, another patient may have used the vacated time slot, which raises the cost of care (Mikhaeil et al., 2019; Weaver, Talley, Mullins, & Selleck, 2019).

There are various approaches for dealing with no-show behavior in healthcare. Ferro et al. (2020) illustrate an approach to no-show behavior that improves attendance rates by changing patients' behavior. Another approach is minimizing the operational impact of no-show behavior by adapting the decision-making process regarding appointment allocation, and scheduling (Millhiser & Veral, 2019). Due to patient unpredictability, obtaining probabilities of no-show behavior could improve both approaches' outcomes in supporting the decision-making process behind it (Schwebel & Larimer, 2018; Ahmadi-Javid, Jalali, & Klassen, 2017).

Attendance and non-attendance should be mapped to predict no-show behavior. This study focuses on absenteeism. Several factors come into play when analyzing people's non-attendance for scheduled appointments to predict no-show behavior. Firstly, demographics play a role in no-show behavior, such as age and gender, and environmental factors, such as accessibility of essential facilities, for example, medical facilities. Faisal-Cury, Quayle, Marques, Menezes, and Matijasevich (2015) found missing at least one appointment was associated with being a young adult, female, and living in a deprived neighborhood. Secondly, appointment behavior plays a role in no-show behavior. This refers to the behavior of individuals that predicts the likelihood of a no-show. Chariatte, Berchtold, Akré, Michaud, and Suris (2008) demonstrated that missing a prior visit and a greater delay between appointments increased the likelihood of missing a subsequent appointment. Lastly, appointment characteristics also play a role when looking at no-show behavior. Goffman et al. (2017) demonstrate that the number of prior appointments and the purpose of appointments is key elements in predicting presence at appointments, i.e., a lesser number of previous appointments and an unclear purpose increase the likelihood of missing an appointment. Personal, behavioral, and appointment characteristics play a role in no-show behavior.

However, research reveals a particular context, including outpatient gastroenterology clinics (Neal et al., 2001), university hospitals (Chariatte et al., 2008; Zebina et al., 2019), gynecological clinics (Faisal-Cury et al., 2015), and veteran healthcare centers (Goffman et al., 2017), namely small diagnostic imaging clinics with less than 500 patients. This context indicates direct health implications associated with appointment non-attendance. These studies have in common that all patients are currently being treated for (symptoms of) diseases and exposed to direct health consequences accompanied by missing such an appointment. Also, the patient and the healthcare provider are involved in the costs of care, also referred to as the financial consequences of missing appointments due to the healthcare system in which these appointments occur. This research gap indicates that studying no-show behavior in the context of preventative healthcare, where individuals are not responsible for care costs, would contribute to the existing scientific literature.

Therefore, this thesis project studies no-show behavior in the context of Dutch youth healthcare (JGZ) part of GGD Brabant Zuid-Oost. This government initiative monitors children from age 0 to 19 years old for 21 municipalities. JGZs goal is to promote, protect and secure the health and physical and mental development at individual and population levels (Dunnink & Lijs-Spek, 2008). The GGD operates at a municipal level by contacting every newborn via appointments with the JGZ health clinic. The national government sets healthcare priorities in the Dutch healthcare system and monitors access, quality, and costs (Wammes, Jeurissen, Westert, & Tanke, 2020). In this context, the healthcare provider, GGD Brabant Zuid-Oost, is the only one involved in the financial consequences of JGZ appointments. Besides that, JGZ appointments include preventative healthcare, so patients visiting JGZ appointments are only indirectly exposed to the health consequences of missing JGZ appointments. This addresses the research gap as described in the previous paragraph.

Yearly there are over 350.000 appointments, of which approximately 22% result in an absence. Over time these patients stop attending appointments at all, which is problematic. At the municipal level, this means wasting a lot of money and effort; staff shortages immediately result in long-term health consequences for other patients: reducing no-shows would be beneficial. This study aims to improve the efficiency of healthcare delivery in the context of the GGD Brabant Zuid-Oost by supporting the decision-making process of improving appointment attendance and minimizing the operational impact of no-show behavior by answering the following main research question:
RQ: How can no-show behavior at Dutch youth healthcare (JGZ) appointments be predicted?

On an individual level, this will be utilized to forecast the probability of no-show behavior for each individual. First, this study aims to understand the factors or characteristics behind the behavior to influence the likelihood of individuals attending their appointments. This illustrates to what extent no-show behavior can be influenced to improve attendance. This leads to the first sub-question:
SQ1: What are the behavioral and contextual factors influencing the likelihood of individuals not attending their appointments?

The GGD aims to categorize those with the highest probability of no-show accurately. This provides a foundation for altering appointment allocation and scheduling to mitigate the operational impact of no-show behavior. To make this classification feasible, the GGD uses different approaches. So-called 'black box' approaches, sophisticated analytical machine learning methods incomprehensible to most healthcare managers, and more easily interpretable but less accurate approaches such as regression models are used (Topuz, Uner, Oztekin, & Yildirim, 2018). The GGD's primary objective is to classify the individuals with the highest probability of missing appointments as accurately as possible. This leads to the second sub-question:
SQ2: Which method most accurately predicts no-show behavior for Dutch youth healthcare (JGZ) appointments?

The GGD data is collected by health professionals who administer a system. The other steps of the machine learning process: data pre-processing, the definition of training data, algorithms selection, training, evaluation with test data, parameter tuning, and classification are executed by domain experts, i.e., research and data employees of the GGD (Osisanwo et al., 2017). Using algorithms is only successful when they can be expected to be fair, reliable, trustworthy, and equally accurate for all data subgroups (Kaur, Uslu, Rittichier, & Durresi, 2022). This leads to the third sub-question:
SQ3a: Does the prediction model perform equally well independently of the data subgroups?

On the other hand, only health professionals presently influence decisions on no-show behavior

based on their perspectives. Domain experts can be involved in decision-making by incorporating data-driven approaches. They may support the management team by performing machine learning methods and evaluating the result to provide guidance, contributing to the reasoning of the decision-making process. Alahmari et al. (2018) demonstrate that data producing and pre-processing in health application fields require domain expertise and cannot be performed by ordinary people and crowd workers due to the quality of the data samples and labels. This implies that the beliefs of the domain experts need to align with the predictive model's factors to accept incorporating machine learning methods in the decision-making (Maadi, Akbarzadeh Khorshidi, & Aickelin, 2021). Also referred to as the interaction between domain expertise and machine learning methods. This section intends to demonstrate whether domain experts are likely to accept the assistance of machine learning methods in the decision-making process. This leads to the fourth sub-question:
SQ3b: Is there a deviation between the beliefs of domain experts and the prediction model's factors?

This study potentially improves healthcare delivery efficiency within the scope of the GGD Brabant Zuid-Oost by mapping no-show behavior at Dutch youth healthcare (JGZ) appointments and supporting the decision-making process of improving appointment attendance, and minimizing the operational impact of no-show behavior. This thesis opens with the construction of literature to map potential factors, identify approaches to support the decision-making process, and study the interaction between domain expertise and machine learning methods. The approach of this study will be outlined in the method section. In the results section, every result will be discussed. All sub-questions will be answered in the conclusion section to provide a complete response to the main research question. This thesis finishes with a discussion section considering the practical implications.

Chapter 2

Literature

This study examines how patients behave when scheduling and attending appointments in the healthcare context, focusing on no-show behavior. This chapter first explains different factors predicting no-show behavior and their consequences in different contexts. Then different approaches for dealing with no-show behavior will be discussed. Afterward, the role of domain experts in decision-making and the interaction with machine-learning approaches will be explained.

2.1 Factors of no-show behavior

Categorizing factors of no-show behavior into different characteristics allows dividing them according to their influenceability. A division is made between demographics, i.e., factors that cannot (easily) be changed by individuals, and behavioral characteristics, i.e., factors that can be influenced.

2.1.1 Demographics

The following paragraphs explain one of the analyzed demographics in a situation that gives insight into no-show behavior.

First, **travel distance** between the home address and the location of the appointment has been identified as a significant predictor of appointment attendance behavior. This was demonstrated in a study conducted on patients at a neurology clinic in Marina Alta. Travel distance from patients' homes to the clinic was the primary factor influencing no-show behavior (Morera-Guitart, Mas-Server, & Mas-Sese, 2002). The study found that as the distance between a patient's home and the clinic increased, there was a corresponding decrease in the likelihood of appointment attendance. Morera-Guitart et al. (2002) suggest that accessible travel modes, i.e., car possession and accessibility of public transport, influence the effort patients must undertake for visiting appointments. Increased travel distance increases patients' effort, decreasing the likelihood of attending appointments. Jennings, Boyle, Mahawar, Balupuri, and Small (2013) supports these findings by showing that increased distance between the patient's home and the bariatric center was related to lower appointment attendance. They suggested that compared to owning a car, using public transportation lengthens travel times and increases the effort required, which increases the probability of missing appointments.

The findings of this study are consistent with previous research on the relationship between travel distance and appointment attendance in hospital settings (Daggy et al., 2010; Dantas, Hamacher, Cyrino Oliveira, Barbosa, & Viegas, 2019). Specifically, Daggy et al. (2010) found that an increased travel distance to a Veteran Affairs medical clinic was associated with a higher probability of no-show behavior. Similarly, Dantas et al. (2019) reported a positive correlation between travel distance and no-show behavior in bariatric clinics. Seen in both VA and bariatric clinics, both studies suggested that time investments and accompanying effort explained the decrease in appointment-missing probability. Due to greater travel lengths than the duration of the session, patients are more likely to miss their scheduled appointments. This indicates that the association between travel distance and no-show behavior is generalizable in various hospital contexts (Fan, Deng, Ye, & Wang, 2021). As such, a patient's home address, as a proxy for travel distance, may serve as a valuable predictor of appointment attendance in hospital contexts.

Another demographic influencing appointment attendance behavior is **ethnicity**. A study conducted among patients at a Gastroenterology Clinic, in an academic setting found that ethnicity is a significant predictor of no-show behavior (Shrestha, Hu, & Taleban, 2017). The study sourced data from the electronic health records of patients who had scheduled appointments at the Gastroenterology Clinic at the Banner University Medical Center. The study found that patients from ethnicities other than white were more likely to miss clinic appointments. Shrestha et al. (2017) attribute these differences to familiarity with the system that may affect patients from different ethnic backgrounds. The patient's understanding of the healthcare system and the effort required to engage in appointments might explain why other ethnicities than white were more likely to miss appointments. Ethnicities other than white would be less likely to be familiar with the healthcare system, requiring more effort to obtain the required information about engaging in appointments. The increased effort and lack of understanding would decrease the likelihood of attending appointments. Previous research on cultural disparities in healthcare provides an additional explanation. Namely, Fiscella, Franks, Doescher, and Saver (2002) conducted a cross-sectional analysis of a community tracking telephone survey, which included 60,446 participants, to investigate cultural differences in healthcare. Cultural differences were explained in terms of native language and understanding of the healthcare system. Fiscella et al. (2002) found that native language significantly predicts the likelihood of missing appointments: English-speaking Hispanics are more likely to attend appointments than Spanish-speaking Hispanics. First, this could be because Spanish-speaking Hispanics would not correctly estimate the added value of the appointment or do not understand the concept of having an appointment. The lack of understanding would result in a higher probability of missing appointments. The context of this study was healthcare, where appointments such as physical visits, mental visits, or influenza vaccinations are examples of appointments that participants should attend.

On the extreme of the spectrum, looking at cultural differences, understanding the healthcare system in itself would completely influence the likelihood of attending appointments (Pell et al., 2013). A field study, including in-depth individual interviews with pregnant women, conducted in Ghana, Kenia, and Malawi showed that women only had a vague understanding of the healthcare system. This meant that women did not attend antenatal care appointments at all, which are used to improve maternal and infant health. This highlights the importance of considering patients' cultural backgrounds and language proficiency when designing healthcare systems and interventions to improve appointment attendance.

Marital status also significantly predicts no-show behavior in various healthcare settings. A case-control study was conducted in an academic environment (Shrestha et al., 2017). Shrestha et al. (2017) found that single patients were likelier than married patients to miss their appointments in Gastroenterology Clinics. The mechanism behind this is explained by a so-called 'spousal effect': marriage influences health status through the support and protection it provides and through a more efficient pattern of healthcare utilization (Pandey et al., 2019). Pandey et al. (2019) explain that marriage increases access to resources. First, increasing resources, such as higher total income, enable married couples to take private insurance. Second, greater health-enhancing home support and assistance that improves adherence to the outpatient clinics. Small private outpatient clinics provide higher quality healthcare, and married patients can afford this type of healthcare due to private insurance. Where higher quality of healthcare makes patients more eager to visit appointments due to a higher perceived added value (Goldsmith, 1980). All in all, these combined resources increase the likelihood of attending appointments for married patients. Single patients are less likely to attend appointments compared to married patients due to these lower resources and the fact outpatient clinics are less accessible to them.

Similar findings were observed in a study of gynecology clinic patients, where younger single women were at a higher risk of missing appointments (Pillai, Bhangu, Narayanan, & Yoong, 2012). This study included a retrospective analysis of trends of non-attendance based on a computerized database of all gynecology appointments over 12 months and a prospective case-control study.

The findings are attributed to a less social support network of singles, and single pregnant women experience less stability (Pillai et al., 2012). Ultimately, this illustrates that single patients were likelier than married patients to miss their appointments.

Another demographic that plays a role in appointment behavior are **gender** and **age**. Pillai et al. (2012) found that younger single patients were most likely to miss appointments in mental health clinics. Their age was a significant predictor, where patients who missed their appointments were, on average, eight years younger than patients that attended their appointments. Regarding age, younger people may be more likely to miss appointments due to various factors. First, they may be more forgetful or have busier schedules, making it harder to keep track of appointments (Ofei-Dodoo, Kellerman, Hartpence, Mills, & Manlove, 2019). Additionally, younger people may have less experience with the importance of medical or other types of appointments and may not fully appreciate the consequences of missing them (A. J. Mitchell & Selmes, 2007). As for gender, women are generally more likely than men to attend medical appointments and less likely to miss them. This may be because women tend to place a higher value on preventive healthcare and are more likely to seek medical care (Vallée, Cadot, Grillo, Parizot, & Chauvin, 2010). Regarding males, their attitude of self-reliance and a reluctance to seek help are traits frequently observed in male patients in regional and rural healthcare settings and explain why males are less likely to attend medical appointments (Begg, 2007).

The last demographic that plays a role in appointment behavior is **socioeconomic status**. This refers to an individual's economic and social position within society, including income, education level, occupation, and access to resources (Adler & Ostrove, 1999). For gynecological appointments in low-middle versus high-income countries, it was found that women in low-middle income countries showed lower attendance than high-income countries (Faisal-Cury et al., 2015). This is since people with low socioeconomic status experience financial barriers. For example, those people cannot afford transportation costs or payments associated with medical appointments. Also, people living in neighborhoods with a low socioeconomic status have limited or no access to healthcare facilities, making it more challenging to schedule and attend appointments (Adler & Ostrove, 1999). Last, people with a low socioeconomic status also have limited health literacy, which makes it more challenging to understand medical instructions, follow treatment plans, and recognize the importance of attending appointments (Stormacq, Van den Broucke, & Wosinski, 2019).

C. Kelly, Hulme, Farragher, and Clarke (2016) support these findings in the context of the Irish national diabetic retinopathy screening program, RetinaScreen. Combining this program with publicly available meteorological and geospatial data, they found that social-economic deprivation, the so-called low socioeconomic status, was the main factor for missing appointments. Social-economic deprived areas have fewer necessities of life facilities, such as medical facilities and public transport (Adler & Ostrove, 1999). These facilities, as explained earlier, contribute to attending appointments by influencing the effort required to attend healthcare appointments. S. R. Kelly et al. (2021) also use language and cultural barriers to explain the differences between people having low versus high socioeconomic status. People with lower socioeconomic status are more likely to have limited language skills in the language spoken by healthcare providers or different cultural norms that lower their willingness or ability to seek healthcare (Meléndez-Ackerman et al., 2014). As earlier explained by the study of Fiscella et al. (2002), limited language skills increase the effort required to understand healthcare providers and decrease their health literacy. Therefore, the likelihood of attending appointments is lower for people with a low socioeconomic status than for those with a higher socioeconomic status.

Literature supports demographics such as home address and travel distance, ethnicity, marital status, gender, age, and socioeconomic status to influence no-show behavior. These factors were found due to the following overarching concepts: an individual's availability of resources, required effort to attend appointments, understanding and familiarity with the healthcare system, and perceived importance of the appointment.

2.1.2 Behavioral characteristics

In appointment behavior, behavioral characteristics also play a role. Behavioral characteristics depend on an individual's behavior and accompanying psychological constructs. Psychological constructs are the different ways people understand and interpret things, and these constructs can activate coping mechanisms in response to different situations (Skinner, Edge, Altman, & Sherwood, 2003). Coping mechanisms are strategies people use to deal with situations and can be adaptive or maladaptive. In the context of showing no-show behavior, coping mechanisms can result in either showing up, meaning using an adaptive coping mechanism or not showing up, meaning using a maladaptive coping mechanism, for an appointment. This depends on the individual's psychological construct and the effectiveness of their coping strategy. Several psychological constructs associated with characteristics that predict no-show behavior are explained in the following paragraphs to illustrate how psychological constructs affect individuals' behavior.

The first behavioral characteristic that influences no-show behavior is the **amount of rescheduling an appointment** (Goffman et al., 2017). Goffman et al. (2017) used demographic information, appointment characteristics, and attendance history from four Veterans Affairs healthcare facilities within six separate service areas to predict appointment attendance. The study developed 24 unique predictive models, classifying 1,754 high-risk patients in a pilot study whose probability of missing an appointment was predicted to be at least 80%. The study found that self-efficacy explains the role of rescheduling an appointment in no-show behavior. If a person strongly believes that he or she can attend an appointment, then he or she applies adaptive coping mechanisms and attends the appointment. They argue that whenever a person repeatedly reschedules an appointment, self-efficacy declines, hence diminishing the efficacy of the adaptive coping mechanism. Margolis and McCabe (2006) explain that repeated rescheduling of tasks undermines a person's belief in their ability to attend the appointment as planned. Leading to maladaptive coping mechanisms that result in missing the appointment. Since self-efficacy is the degree to which an individual believes in his or her capacity, both studies explain self-efficacy as a psychological construct that explains the cause of (mal)adaptive coping mechanisms resulting in no-show behavior.

McComb et al. (2017) found that **time between two consecutive appointments**, also known as the age of an appointment, is related to appointment attendance. McComb et al. (2017) studied canceled primary health care appointments of diabetic patients in a cohort study using data of 46,710 appointments. It was found that an increase in the time between two consecutive appointments is associated with a higher probability of not appearing for appointments. A longer time gap increases the risk of no-show behavior because the patient may forget the appointment or become less motivated to attend. McComb et al. (2017) specifically focuses on motivation as the psychological construct causing this effect, where the urgency of having an appointment would specifically drive motivation in this context. They conclude that a loss in motivation directly caused by a lack of perceived urgency affects appointment attendance and the related coping mechanisms; a patient is less effective at employing adaptive coping mechanisms and fails to attend the appointment. Also, a shorter time gap may increase motivation to attend the appointment, illustrating a reciprocal relationship. Rabideau (2005) explains the effect of motivation on behavior in general. This study underlines that a motivated person will likely use adaptive coping mechanisms since people generally have a positive outlook and are more likely to take proactive steps to manage their stressors. Arguing that patients will attend the appointments in a healthcare context.

Another study in the healthcare context shows that these findings are consistent. Shrestha et al. (2017) also found that the time between two consecutive appointments significantly predicts no-show behavior. Specifically, the study found that a longer waiting time was correlated with a higher probability of not showing up for appointments, where motivation was again used as an explanation. Shrestha et al. (2017) studied no-show behavior in the Gastroenterology Clinic, as explained before. The study argues that motivation is driven by the understanding of the healthcare system, where a lack of understanding drives the extent to which a patient is motivated.

Again, a decrease in motivation causes a less effective way of using adaptive coping mechanisms, and the patient exhibits no-show behavior. Both studies show that the psychological construct of motivation is directly related to no-show behavior.

Next, Goffman et al. (2017) found that **appointment type** influences the probability of attendance in several ways. For instance, individuals may be more likely to attend routine or follow-up appointments as they recognize the importance of regular check-ups and health monitoring. The psychological construct influencing this behavior is perceived importance. Conversely, patients may be less likely to attend appointments related to procedures that make them anxious or hesitant. C. Kelly et al. (2016) suggest that when patients feel the health situation is more serious, people are more eager to come, i.e., the perceived importance is high. This study was conducted among the Irish national diabetic retinopathy screening program, RetinaScreen, as earlier explained. Overall, perceived importance may explain the role of appointment type in no-show behavior. It can influence a person's motivation to attend their appointment, sense of responsibility for their health, and use of adaptive coping strategies to ensure attendance. Both studies explain that the psychological construct of perceived importance directly influences no-show behavior.

Having multiple appointments at one day also decreases the probability of missing appointments (Hu, Xu, Li, & Che, 2020). This study analyzed data from a Chinese tertiary care public hospital's appointment system and conducted surveys among hospital patients. Hu et al. (2020) showed that perceived control is the psychological construct that affects no-show behavior by creating a sense of time pressure and reducing an individual's sense of control over their schedule. Moreover, Hu et al. (2020) suggest that whenever patients perceive a lack of control over their schedule, this will decrease their motivation to attend appointments. As explained earlier, demotivated patients are more likely to engage in maladaptive coping strategies, such as ignoring or avoiding the situation, rather than taking proactive steps to manage their schedule and attend appointments. Skinner, Zimmer-Gembeck, Connell, Eccles, and Wellborn (1998) elaborate on the relationship between perceived control and motivation in a general context studied in a longitudinal study on 1,600 individuals. The study shows increased performance influences an individual's sense of control, meaning increased perceived control. Perceived control was in this study also linked to the extent to which an individual is motivated to exhibit desirable behavior. More specifically, engaging in adaptive coping mechanisms towards behavior. The study highlights that perceived control indirectly influences coping mechanisms through motivation. In the end, both studies show, however, using different reasons, that perceived control functions as an indirect psychological construct for no-show behavior.

Lastly, Chariatte et al. (2008) found that **past appointment behavior** greatly influenced the risk of missing an appointment. Specifically, patients who had already missed appointments had a higher risk of missing appointments again. Chariatte et al. (2008) conducted a study in an outpatient clinic for adolescent patients based on their past behavior and other characteristics. The study included 2,193 patients aged 12 to 20 who had scheduled at least four appointments. This study argues that behavioral consistency is why past behavior is repeated. Behavioral consistency refers to the idea that people repeat their past behaviors in similar situations, even if they are not necessarily beneficial or desirable (Ajzen, 1991). If someone has a history of missing appointments or showing up late, for example, they will likely continue to exhibit similar behaviors. This could be due to various factors, such as a lack of motivation, poor time management skills, or a tendency to prioritize other activities over appointments. McComb et al. (2017) also saw that behavioral consistency served as an argument why past appointment behavior was a significant predictor for no-show behavior. This study argues that a lack of motivation mainly drives no-show behavior. Both studies underline that behavioral consistency explains why past appointment behavior plays a role in no-show behavior.

The studies above illustrate that self-efficacy, motivation, and perceived importance play a role in no-show behavior. However, these studies do not fully explain the relationship between self-efficacy, motivation, and perceived importance. Two theories can be used to explain the relationship between those psychological constructs: operant conditioning and compliance theory. First, operant conditioning explains that behavior is influenced by its consequences, and the relationship between self-efficacy, motivation, and no-show behavior can be explained through this theory (Staddon & Cerutti, 2003). Individuals with low self-efficacy may believe they cannot attend appointments or meetings, decreasing their motivation. If these patients do not attend the appointment, they avoid the negative consequences of attending, such as embarrassment or failure. In this case, explained through operation conditioning, the behavior of not attending the appointment is negatively reinforced, which means the behavior is more likely to occur. Conversely, individuals with high self-efficacy and motivation attend appointments, which positively reinforces the behavior of attending. In summary, operant conditioning explains how self-efficacy and motivation can affect no-show behavior by reinforcing or punishing the behavior. Secondly, compliance theory explains that people are motivated to comply with requests from credible sources or when they believe that compliance will lead to positive outcomes (R. B. Mitchell, 2014). This theory explains the relationship between perceived importance and motivation. Perceived importance is a factor that motivates individuals to attend appointments or meetings. When an appointment or meeting is perceived as important, individuals feel more obligated to attend and more motivated by the potential positive outcomes, namely health benefits.

In summary, the literature supports behavioral characteristics such as past appointment behavior, the amount of rescheduling an appointment, and having multiple appointments scheduled on a day. These are all predictors dependent on an individual's behavior. These factors are influenced by psychological constructs such as self-efficacy, motivation, perceived importance, and perceived control. The literature framework in figure 2.1 visually composed the literature above.

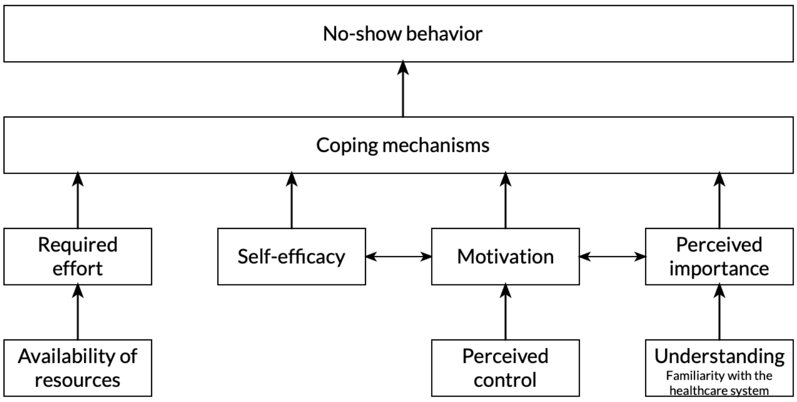


Figure 2.1: Literature Framework: psychological constructs of no-show behavior

As explained in the introduction of this study, current literature reveals a particular context of either small outpatient clinics with less than 500 patients or diagnostic imaging clinics that indicate direct health implications associated with appointment non-attendance. These studies have in common that all patients are currently being treated for (symptoms of) diseases and exposed to direct health consequences accompanied by missing such an appointment. For example, a patient of a diabetic clinic is already experiencing the health effects of being a diabetic patient. Therefore, the perceived importance of an appointment differs from a patient who needs to attend an appointment for a regular health check at the general practitioner. Also, most studies are conducted in countries with the same healthcare system where the patient and the healthcare provider are involved in care costs, also referred to as the financial consequences of missing appointments. Therefore, the research gap in the context of no-show behavior is two-sided.
On the one hand, current literature focuses on direct health consequences for the patient, where

the context of preventative healthcare is understudied. On the other hand, the current literature shows shared financial consequences of missing appointments, where contexts in which there are no financial consequences for the patient are understudied. As the introduction explains, this study's research includes preventative healthcare appointments within Dutch youth healthcare (JGZ) appointments. Due to the Dutch healthcare system, the patients will not be exposed to the financial consequences of missing these appointments.

This study aims to map no-show behavior in Dutch youth healthcare (JGZ) appointments and support the decision-making process of improving appointment attendance and minimizing the operational impact of no-show behavior. The approaches that can be used to reduce no-show behavior can be found in the next section.

2.2 Approaches for reducing no-show behavior

As the introduction explains, Ferro et al. (2020) and Millhiser and Veral (2019) illustrate two different approaches for dealing with no-show behavior from the perspective of the healthcare provider. Ferro et al. (2020) suggests that **improving attendance rates** by changing patients' behavior is an approach for dealing with no-show behavior. Strategies such as phone reminders and education programs are proposed to improve appointment attendance.

Wu et al. (2019) studied the effect of short message service interventions such as phone reminders on no-show behavior among a preliminary colorectal cancer screening program. The study found that both mailed reminder letters and phone reminders effectively reduced non-attendance among patients. Both mailed reminder letters and phone reminders are most effective seven days before the appointment date since the appointment can be canceled in time, enabling another patient to use the available timeslot instead. Wu et al. (2019) suggest that the interventions lead to an increased belief of urgency or importance of attending the appointment. These results are supported by a study that tested the effectiveness of short message reminders on appointment attendance in a pediatric cataract treatment clinic (Lin et al., 2012). They found that short message reminders (SMS) and phone calls reduced the number of no-shows, whereby the probability that the no-show is caused by forgetting the appointment is sidelined. The phone calls and short message reminders (SMS) were found to be effective seven days before the appointment date.

Another strategy used to improve appointment attendance is offering educational programs. People might experience different language, financial, social, educational, and psychological barriers that prevent them from attending appointments (Percac-Lima et al., 2015). Weaver et al. (2019) shows that educating people and guiding patients through digital healthcare systems decreases no-show behavior. The study was conducted among diabetes patients in the PATH Clinic from UAB Hospital. Two interventions were tested: the added value of educating people and the added value of guiding patients through (digital) healthcare systems. Both interventions decreased no-show behavior. The findings are explained as educating people increases the understanding of the healthcare system and understanding the importance of appointments; this increases the perceived importance of appointments. As explained before, this directly decreases the probability of exhibiting no-show behavior. Weaver et al. (2019) explain that guiding patients through (digital) healthcare systems decreases patients' effort to (re)schedule appointments. The less effort a patient requires to undertake action, to higher the likelihood of this patient to proactively act towards attending the healthcare appointment (Morera-Guitart et al., 2002).

Both strategies affect no-show behavior in a different healthcare context. However, Zebina et al. (2019) suggest that a key element in this approach is to be able to correctly identify the patients that should be targeted with each strategy. For example, younger people are likelier to read short message reminders (SMS), whereas older people are more easily approachable through phone calls. Also, not everyone will experience technical difficulties in using the digital healthcare system and needs help. Zebina et al. (2019) illustrates the importance of predefined groups; this problem is tackled in scientific literature as all studies above work within predefined groups in their study's 'design and sample' section. For practical implications, it is important to consider using correctly identified patient groups, each targeted with an associated strategy.

Millhiser and Veral (2019) suggests that **minimizing the operational impact of no-shows** is another approach for dealing with no-show behavior. The operational impact of no-shows can be reduced by more effectively scheduling appointments by adjusting the decision-making process and using different appointment scheduling methods.

The first way to minimize the no-shows' operational impact is more effectively scheduling appointments by adjusting the decision-making process (Brailsford, Harper, & Sykes, 2012). This study shows that specifying patient groups in the context of cancer treatment allows the healthcare provider to adjust the decision-making process regarding resource allocation and scheduling. Healthcare providers can allocate resources more effectively by identifying patients most likely to miss appointments. For example, healthcare providers schedule high-risk patients at different times of the day or week to reduce their likelihood of missing appointments. Additionally, healthcare providers may allocate additional resources to these groups, such as offering transportation assistance, to reduce the likelihood of non-attendance.

Another way to minimize the operational impact of no-shows is using different appointment scheduling methods (Rose, Ross, & Horwitz, 2011). This study illustrates that a so-called 'open access' scheduling strategy decreases the operational impact of no-show behavior by decreasing staffing. 'Open access' scheduling means that all patients can schedule and cancel appointments themselves, which improves the accessibility of healthcare to patients using this strategy. Daggy et al. (2010) explains that the staff, using this strategy, can be more effectively used to implement strategies that improve appointment attendance, such as giving phone call reminders or sending email reminders.

2.3 Machine learning and decision-making

Knowing the factors that predict no-show behavior, different machine learning methods can be used to predict no-show behavior using demographic information, behavioral characteristics, and appointment characteristics.

2.3.1 Machine Learning Methods

Osisanwo et al. (2017) explains that machine learning methods are one of the most rapidly expanding fields of computer science, having widespread applications. Where machine learning methods are methods used to identify meaningful patterns in data. This study elaborates on the classification and comparison between different machine learning methods. The present machine learning approaches described in this paper will be elaborated upon first.

- **Linear Regression:** A supervised learning algorithm that models the linear relationship between a dependent variable and one or more independent variables.

- **Logistic Regression:** Supervised learning algorithm used for classification problems where the goal is to estimate the probability of a binary or categorical outcome. Logistic regression is a simple and interpretable method that is easy to implement since the model equation is a linear combination of the predictor variables, making it easy to understand and interpret the role of each variable in the model.

- **Decision Trees:** A supervised learning algorithm that creates a tree-like model of decisions and their possible consequences. The goal is to create the most homogeneous subsets possible until all subsets are as homogeneous as possible or until a stopping criterion is met. Once a decision tree is trained, it might be used to make predictions on new data by following the splits in the tree until reaching a leaf node and then returning the prediction for that leaf node.

- **Random Forest:** An ensemble learning method that combines multiple decision trees to reduce overfitting and improve prediction accuracy. Random forest uses bagging: creating several decision trees using the random selection of data and attributes. A random forest estimates each tree, and a majority vote determines the ultimate estimation. Compared to logistic regression, random forest is more complex but more accurate since the random forest can model complex non-linear relationships between the predictor variables and the dependent variable, which makes it well-suited for large datasets with complex features.

- **Naive Bayes:** A supervised learning algorithm that uses Bayes' theorem to calculate the probability of a particular class based on the input features. Naive Bayes is called "naive" because a strong assumption is made regarding the conditionally independent features given the class. In other words, it assumes that the presence or absence of one characteristic in a class is separate from the presence or absence of any other characteristic. This is often not true in practice, but the algorithm can still work well even when this assumption is violated.

- **K-Nearest Neighbors (KNN):** A supervised learning algorithm used for classification and regression. It is based on the assumption that labels for related data points tend to be similar. The method compares a new data point to the k nearest training set data points. The most often occurring label of the k closest data points is then used to decide the label of the new data point. The value of 'k' is a user-defined parameter often determined by testing.

- **Neural Networks:** A family of models that can be used for supervised and unsupervised learning tasks inspired by biological neurons' structure and function. Once a neural network is trained, it can predict new data by transmitting the input using the network and returning the output from the output layer. Overall, neural networks are powerful, flexible machine learning algorithms that model complex non-linear relationships between inputs and outputs. They have achieved state-of-the-art performance in many areas, including image and speech recognition and natural language processing.

Osisanwo et al. (2017) compares various classification algorithms with large data sets based on different parameters: the time required to execute the algorithm, percentage of correctly classified observations, test mode, error terms, and precision statistics. Also, the various classification algorithms are compared using small data sets. The study concludes that precision and correct classification are leading parameters when using machine learning classification methods. The study also shows that using a machine learning method for a particular dataset does not guarantee the outcomes of another dataset. More specifically, the comparison outcomes for large datasets differed from those for small datasets. Using machine learning methods requires an extensive finetuning of parameters.

2.3.2 Incorperating Machine Learning into decision-making

Machine learning algorithms allow humans to do things beyond their capabilities (Marsland, 2011). For example, analyzing large datasets to study behavioral patterns would take humans multiple years, and computers using machine learning algorithms provide the output in seconds. As Marsland (2011) already stated in 2011, "If Data Had Mass, the Earth Would Be a Black Hole", which illustrates the enormous scope of the applicability of data. Machine learning algorithms, therefore, can also be used in a decision-making process.

In general, as explained in the introduction, the machine learning process consists of data preprocessing, the definition of training data, algorithms selection, training, evaluation with test data, parameter tuning, and classification (Osisanwo et al., 2017). In all steps of the machine learning process, humans are, to a greater or lesser extent, involved. Osisanwo et al. (2017) explain that the process starts after humans identify what data is required to solve a pre-defined problem. Defining a problem and interpreting the classification output are steps humans are required (Maadi et al., 2021). Maadi et al. (2021) specified why humans should be involved in the machine learning process: adding knowledge to the learning process, evaluating machine learning outputs, and deciding to refine the method or accept the outcomes. Maadi et al. (2021) also indicates that human satisfaction indicates the importance of including humans in the process, also called the acceptance of machine learning approaches. The study proposes that domain experts evaluate the performance of machine-learning approaches based on the model outputs and the extent to which those outcomes are trusted, and thus accepted.

Alahmari et al. (2018) also suggests that humans need to be involved in the process. This study indicates that domain expertise is required for the health application field, which ordinary people and crowd workers cannot perform as this might result in a lack of quality of the data samples and labels. In the machine learning process, the definition of training data requires humans to specify data samples, subsets of the full population representative of the phenomenon studied, and label the data based on the targeted behavior of the phenomenon studied. For example, this study required humans to specify the dataset used to study appointment behavior in the context of the Dutch youth health care (JGZ) appointment, where no-show behavior should be labeled using available data fields: the information collected by health professionals and included in the data of the GGD Brabant Zuid-Oost.

Defining the knowledge and experience of domain experts is important considering the decision-making process. As Alahmari et al. (2018) explains, ordinary people and crowd workers cannot use machine learning methods to assist decision-making. A so-called trade-off between 'black box' approaches and easily interpretable methods plays a role (Topuz et al., 2018). The literature discusses a trade-off between 'black box' approaches, i.e., sophisticated analytical methods that would be incomprehensible to most healthcare managers, and more easily interpretable but less accurate approaches such as regression models that can be easily interpreted. Including 'black box' approaches requires domain experts with a background in data analysis and machine learning approaches to be included in the decision-making process. Using easily interpretable approaches might be interpretable by healthcare managers and healthcare professionals themselves, but provides the results' lower quality, i.e., less accurate classification methods. Topuz et al. (2018) suggest that, if possible, 'black box' approaches are used, which acquire domain experts with a background in data analysis and machine learning approaches. In the end, this adds more value to the decision-making process. Shin (2021) suggest that domain experts are more likely to trust the machine learning models' accuracy, validity, and relevance to accept and incorporate them into their decision-making process. Indicating that 'black box' approaches can be used if these methods are trusted by domain experts, which depends on the study context and the familiarity and experience of domain experts with machine learning methods. The interaction between domain expertise and machine learning involves a balance between using easily understandable and trustworthy methods versus methods that may be more accurate but difficult to comprehend for most healthcare professionals. The former approach is likely to be accepted and relied upon, while the latter is seen as a 'black box' and requires the involvement of domain experts to aid in the decision-making process.

Chapter 3

Method

3.1 Approach

This study aims to improve the effectiveness of healthcare delivery in the context of the GGD by mapping no-show behavior, supporting the decision-making process of improving appointment attendance, and minimizing the operational impact of no-show behavior. This includes a secondary data prediction study to predict non-attendance at Dutch youth healthcare (JGZ) appointments.

The study's approach contains three parts, including four sub-questions contributing to answering the main research question: **How can no-show behavior at Dutch youth healthcare (JGZ) appointments be predicted?** The three parts consist of:

1. **Explanatory model** in which the behavioral and contextual factors that influence an individual's probability of missing appointments are studied.
   SQ1: What are the behavioral and contextual factors influencing the likelihood of individuals missing their appointments?

2. **Most accurate classification**, machine learning method, to classify individuals most likely to exhibit no-show behavior.
   SQ2: Which method accurately predicts no-show behavior in Dutch youth healthcare (JGZ) appointments?

3. **Machine learning and decision-making** in which the prediction model is evaluated for its ability to accurately predict independently of data subgroups.
   SQ3a: Does the prediction model perform accurately independently of the data subgroups?
   Next, the study intends to demonstrate whether domain experts are likely to accept the assistance of machine learning methods in the decision-making process.
   SQ3b: Is there a deviation between the beliefs of domain experts and the prediction model's factors?

3.2 Measures

This study predicts no-show behavior at Dutch youth healthcare (JGZ) appointments. First, the predictor variables of the dataset will be explained, followed by the measures used within this study.

Predictor variables
Within this study, the GGD Brabant Zuid-Oost collects information about newborns in the 21 municipalities of the GGD Brabant Zuid-Oost. The Dutch youth healthcare (JGZ) has a deviation between 0 to 4-year-olds' appointments and 5 to 18-year-olds'. Appointment frequency must be sufficient to conclude appointment behavior. Since most visits are made between ages 0 and 4, this age range is selected for this study. This study assumes that the appointment behavior of the children in the data is fully influenced by the parent who accompanies the child toward the appointment. The section on the literature highlights the determinants of no-show behavior and the psychological constructs that influence those predictors. Section 2.1.1 demonstrated that several demographic factors, such as gender, ethnicity, travel distance, etcetera, affect no-show behavior. The demographics given in this section serve as rough indicators of no-show behavior. Section 2.1.2 demonstrated that the behavioral characteristics associated with no-show behavior have consistent psychological constructs causing the observed outcomes across all studies, namely

self-efficacy, motivation, perceived importance, and perceived control.

Table 3.1 illustrates all variables included in the dataset. Potential predictors of no-show behavior for Dutch youth health care (JGZ) appointments within the GGD Brabant Zuid-Oost are categorized into demographic and behavioral characteristics. Table B.1 of Appendix B at page 50 gives a more elaborate view of the dataset. This table includes each variable, the variable type, the type of observations it holds, and the variable's frequencies or range of potential values.

Demographics	**Behavioral characteristics**
Gender of the child *Gender*	Appointment status *Appointment status*
Age of the child *Age*	Amount of past (scheduled) appointments *Appointment number*
Municipality living in *Municipality*	Appointment Date *Date*
Having Dutch nationality *Dutch_yes_no*	Appointment Type *BDS_Naam*
Socioeconomic status* *SES_WOA*	Child had an intervention before *Intervention_yes_no*
	Past appointment behavior *Not_attending*
	*= neighbourhood level

Table 3.1: Variables of the dataset

First, all accessible demographics are included in the data as predictor variables to test whether the demographic markers of no-show behavior a generalizable to the context of Dutch youth healthcare (JGZ) appointments. This means including age, gender, municipality living in, having Dutch nationality, and socioeconomic status. For socioeconomic status subscales include data on private capital, employment over the last four years, level of education, and standardized income are also available and included in the dataset.

The dataset aims to encompass all psychological constructs relevant to behavioral characteristics, as shown in figure 2.1. To achieve this, potential factors are used to identify the presence of these constructs. Subsequently, a machine learning model will analyze whether these factors affect Dutch youth healthcare (JGZ) appointments. This section includes all relevant information about potential factors in the data. Initially, appointment status is integrated into the data, as it serves as an indicator for defining a no-show, which is the study's dependent variable. For self-efficacy, the **amount of rescheduling an appointment** should be included in the data (Goffman et al., 2017). However, in the context of the GGD, this variable includes too many (70%) missing variables since the variable was not available among all municipalities. Therefore, the amount of rescheduling an appointment is not included in the data. Considering the psychological construct motivation, the **time between two consecutive appointments** indicates whether the coping mechanisms used are directly influenced by motivation (McComb et al., 2017). The variable appointment date is included in the dataset to calculate the time between two consecutive appointments. Next, the psychological construct of perceived control is linked to having **multiple appointments scheduled at one day** (Hu et al., 2020). Since the GGD data only includes the appointments with the Dutch youth health care (JGZ), and the frequency of these appointments is not daily, information to create this variable cannot be included in the dataset used for this study. Examining the psychological construct of perceived importance is connected to **type of appointment** (Chariatte et al., 2008). Within the Dutch youth health care (JGZ), 14 appointment types can be found in the codebook in table B.1 on page 50. Since extra interventions are only planned when health professionals redirect patients, it is assumed that the perceived importance of such an appointment is high. As C. Kelly et al. (2016) explained, the sense of responsibility created by realizing the importance of an appointment causes patients to act responsibly towards future appointments. Therefore, a variable indicating whether a patient **had an intervention before** is included in the dataset. Several studies also indicated **past appointment behavior** as the main predictor of no-show behavior, where the **number of past (scheduled) appointments** explained the strength of this behavior an individual had before (Chariatte et al., 2008; Ajzen, 1991; McComb et al., 2017). Both the amount of past (scheduled) appointments and past appointment behavior are included in the dataset.

Appointment status
The main measure within this study includes appointment status since this variable tells something about whether a person attended the appointment. The appointment statuses that appear in this study can be summarized as follows:

- **Attending**: individuals attending the appointment
- **Not attending, no notice**: individuals not attending the appointment and failed to provide any notice of cancellation
- **Not attending, late notice**: individuals not attending the appointment, who have chosen to cancel their appointments at late notice
- **Not attending, notice in time**: individuals not attending the appointment, who have chosen to cancel their appointments timely

Within the context of the GGD Brabant Zuid-Oost, timely notice must be given at least 48 hours in advance; thus, prolematic no-show behavior represents the dependent variable in this study and is formed by appointments with the appointment statuses: 'not attending, no notice', and 'not attending, late notice'.

The dataset
The data in the Data Warehouse of the GGD Brabant Zuid-Oost is collected by health professionals administering a system. Since March 1, 2021, all municipalities have adopted a centralized administration system. This study's data collection period is one year, from June 1, 2021, to May 31, 2022. The dataset contains every appointment that occurred during this timeframe. The dataset includes 110174 observations of 30319 unique patients, including information on 20 variables.

3.3 Data Analysis: Machine Learning Process

This analysis is divided into an explanatory model and machine learning methods. The implementation of this analysis will be explained using the steps of the machine learning process. Osisanwo et al. (2017) visualized this process in their study. Specifically for this study, the process is visualized in figure 3.1. Since the study is divided into an explanatory model and machine learning methods, the link with the explanatory model is included in this visualization. The data pre-processing is the same for the explanatory model and the machine learning methods. So, after the data pre-processing, the implementation of the explanatory model will be explained.

This analysis used in this study is retrieved using the Data Warehouse of the GGD Brabant Zuid-Oost: a query-language-based (SQL) database in which all information about the GGD Brabant Zuid-Oost is stored. Secondly, Statline, a database of the Centraal Bureau Statistiek, is used to retrieve information at the neighborhood level of individuals (*Centraal Bureau voor Statistiek: Statline*, n.d.). Query language links all data sources and generates a raw dataset. The query used

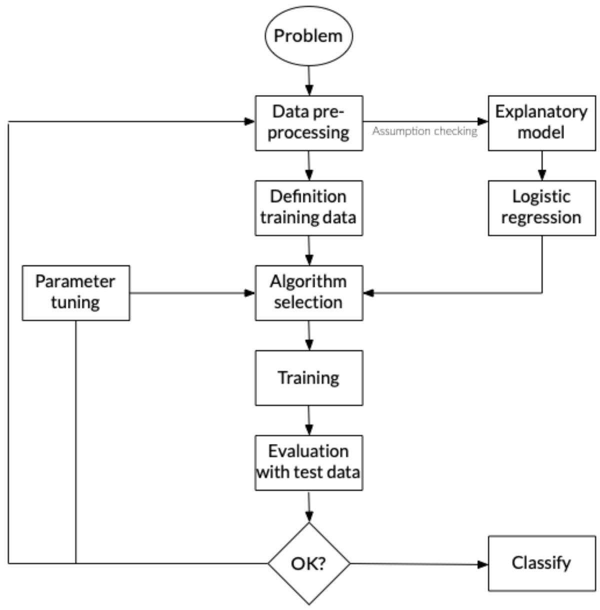


Figure 3.1: Machine Learning Process for this study

can be found in appendix A on page 48. Next, R Studio, an open-source data science platform, is used to analyze data and creates machine learning algorithms. Before conducting data analysis, data processing occurred.

Data Pre-Processing

The data contains 110174 observations, of which 9241 were eliminated during both data pre-processing stages. Data processing is executed at two stages: retrieving the data from the Data Warehouse of the GGD Brabant Zuid-Oost and inspecting and cleaning the data in R Studio. During the collection of data from the Data Warehouse of the GGD Brabant Zuid-Oost, the following modifications were made: excluding observations with appointment statuses equal to 'Unknown,' as these were used for test records; excluding observations with municipalities outside the GGD Brabant Zuid-Oost region, excluding variables with inconsistent or more than 80% missing values, and excluding appointment types other than the 14 appointment types within the 0-4 years old scope of this study. During this stage, it was also ensured that a large enough sample per municipality would be used. During the second data pre-processing stage, observations were eliminated for the following reasons. Initially, duplicate observations were eliminated. Next, all variables are inspected for probable outliers: observations containing impossible information, i.e., appointment numbers less than one, number of prior no-shows less than zero, data with information about future events, and ages outside the scope of the research, i.e., between zero and four years old. Last, observations with missing client numbers were removed since these also imply test records. After both stages of data pre-processing, the dataset consists of 110933 observations, 8.4% of the dataset was eliminated.

Explanatory model

Some assumptions were tested during data pre-processing: ensuring a large enough sample and accounting for outliers. For the explanatory model, logistic regression is used. Utilizing five models, each of which is conducted in a logistic regression model, generates a final logistic regression model. This model tests for all assumptions and identifies which behavioral and contextual factors influence the probability of persons missing appointments. A common approach is to develop a general model that includes all potential factors. This initial model provides a broad view of the relationships between the dependent and independent variables. However, issues such as multi-collinearity and data clustering may arise, leading to inaccurate or unreliable results. To address the issue of multicollinearity, a second model is used. This model examines the mutual correlation between variables using the Pearson method and checks the variance inflation factors to identify variables that may be causing multicollinearity. By eliminating these variables, the accuracy and reliability of the results can be improved. In addition, exploring interaction effects can provide insight into how the relationship between variables changes depending on the value of a third variable. Therefore, the third model examines several interaction effects to determine if they are significant and should be included in the model. When the data includes repeated measures, such as when clients are measured multiple times, it is necessary to correct for group-level differences. In this case, a fourth model, the Random Intercept Model, is used. This model provides information about the percentage of variance a patient can explain in the data. The Random Slope Model is used as a fifth model for data with high variability or when the relationships between variables differ between different groups. This model extends the Random Intercept Model by allowing the model to have varying slopes for different groups in the data.

The definition of training data

To use machine learning methods, the dataset is split into training and validation or test data. The training data (80%) of the data is used to build the machine learning model, while the validation data (20%) is used to evaluate the model's performance.

Algorithms selection

The GGD Brabant Zuid-Oost pre-selected the following machine learning methods: logistic regression, random forest, and K-Nearest Neighbour due to their experience with these methods.

Training and parameter tuning

The selected algorithms are applied to the training data set during this step to create machine learning models. This step also involves adjusting the algorithm's hyperparameters and other configuration settings to optimize performance. The goal of the most accurate classification method is that the GGD Brabant Zuid-Oost can accurately categorize the majority of the population. Those with a chance of non-attendance of more than 90 percent are categorized as a no-show, which indicates the most important threshold. This threshold is the standard threshold the GGD Brabant Zuid-Oost operates with. The threshold might be changed if the model has too many false negatives (i.e., misclassifying someone who does not show up as someone who will show up) since these have a high cost. The threshold may be lowered to reduce the rate of false negatives. For the random forest method, the number of trees, the maximum tree depth, the minimum samples per split, and the maximum features are the parameters that can be adjusted to improve the model during fine-tuning. The range of randomly selected features to be considered for each split in a decision tree within the forest, also called MTRY, is based on the cross-validated performance of the model. The number of trees is, by default, set to 500. The performance of the random forest model generally improves up to a certain point, after which it starts to level off. The visualization of the performance might indicate a lower number of trees, after which the model is constant. R Studio's 'caret' package gives standard values for maximum tree depth and minimum sample per split. These values might be adjusted after the performance of the random forest model is evaluated. For K-Nearest Neighbour, the number of neighbors (K), the distance metric used to calculate the distance between the test instance and its neighbors, and the weight function can be adjusted to improve the model. Again, R Studio's 'caret' package gives a standard for the distance metric used to calculate the distance between the test instance and its neighbors. These values might be adjusted after the performance of the N-Nearest Neighbour model is evaluated.

Evaluation with test data

In the end, the validation dataset is used to evaluate each model's prediction ability once it has been created using the training dataset. To comment on the performance of the model, accuracy metrics are examined. Where performance is the models' ability to correctly predict non-attendance at new data, i.e., the validation dataset. Cross-validation is utilized to compare differences between the confusion matrices of the training and validation datasets. As a consequence of cross-validation, the f1-, sensitivity-, precision-, recall-scores, and confusion matrices reported in this thesis are thus mean values of all matrices established for every cross-validation fold. The differences between the training and validation datasets will be tested using a t-test on the summary statistics of the confusion matrices established during cross-validation. In addition, the performance of each model is then shown among various thresholds using a receiver operating characteristic (ROC) curve as a graphical representation. The area represents the classifiers' performance under the curve (AUC). The most accurate classification method can be established by combining accuracy metrics and the receiver operating characteristic (ROC) curve.

Also, to ensure a reasonable model, the data is also tested for equal performance independently of data subgroups such as socioeconomic status and urban versus rural areas. There are three different **fairness metrics** that are used to compare subgroups:

- The equal opportunity difference (also called odds difference) is a measure of fairness that indicates the difference in the true positive rate (TPR) between two groups.
- The statistical parity difference (also called demographic parity difference) is a measure of fairness that indicates the difference in the predicted positive rate (PPR) between two groups.
- The average odds difference (also called equality of odds difference) is a measure of fairness that indicates the difference in the false positive rate (FPR) and the false negative rate (FNR) between two groups.

For all three metrics, having a value closest to 0 means no difference between the two groups, and the model used can be assumed equal among data groups. If the disparities are less than the 5% threshold, the model does not need adjustment. Otherwise, the model is valid when comparing the two subgroups.

Parameter tuning
After evaluation, the model's performance can be improved by fine-tuning its parameters, which might be a repeating step till the requirements of the performance and optimizing the methods used are met. After the model has been tuned and tested it can be used for classification.

3.4 Acceptance of machine learning

The final sub-question of this study aims to compare the prediction model to the beliefs of domain experts to identify any potential mismatches in decision-making processes. In healthcare, health professionals guide decision-making to prevent no-shows. To ensure these domain experts accept machine learning methods, their beliefs must align with the significant factors predicted by the model. A questionnaire was distributed to over 400 health professionals and domain experts to assess this, presenting them with a list of probable determinants based on current literature. This list includes the following factors: travel distance, marital status, nationality, gender of the parents, socioeconomic status, past appointment behavior, appointment type, time of day, and consequences of no-show. The domain experts were then asked to identify relevant indicators by answering simple 'yes-no' questions.

Fifty-one employees completed the survey, and the factors indicated by these individuals were compared to the significant variables of the final model. Suppose the factors indicated by the domain experts align with those indicated by the prediction model. In that case, it is assumed that there is no deviation between their beliefs and the model's significant factors. However, if there is a deviation, the study strives to find an explanation to understand this misinterpretation.

If no interpretation can be found, domain experts might be hampered by their acceptance of machine-learning methods in decision-making. If an explanation is found, their acceptance might not be hampered by their beliefs.

Chapter 4

Results

This chapter presents an all-encompassing summary of the analysis's results. Appendix B supplements this chapter and provides the complete results for all different analyses.

4.1 Descriptive statistics

Before the main results concerning the research questions are discussed, some interesting characteristic elements of the study will be explained to give a visual overview of the GGD data.

General findings

In total, the data included 101993 observations. Among the observations, there are 30319 unique clients (50.9% males, 49.1% females). These clients reside in 21 municipalities and have attended, on average, eight distinct appointments, divided over 14 distinct types of appointments, between the age category of 0 and 4. All variables of the data may be found in Appendix B's codebook at page 50, which contains their descriptions, variable types, the type of observation it holds, and the variable's frequencies or range of potential values.

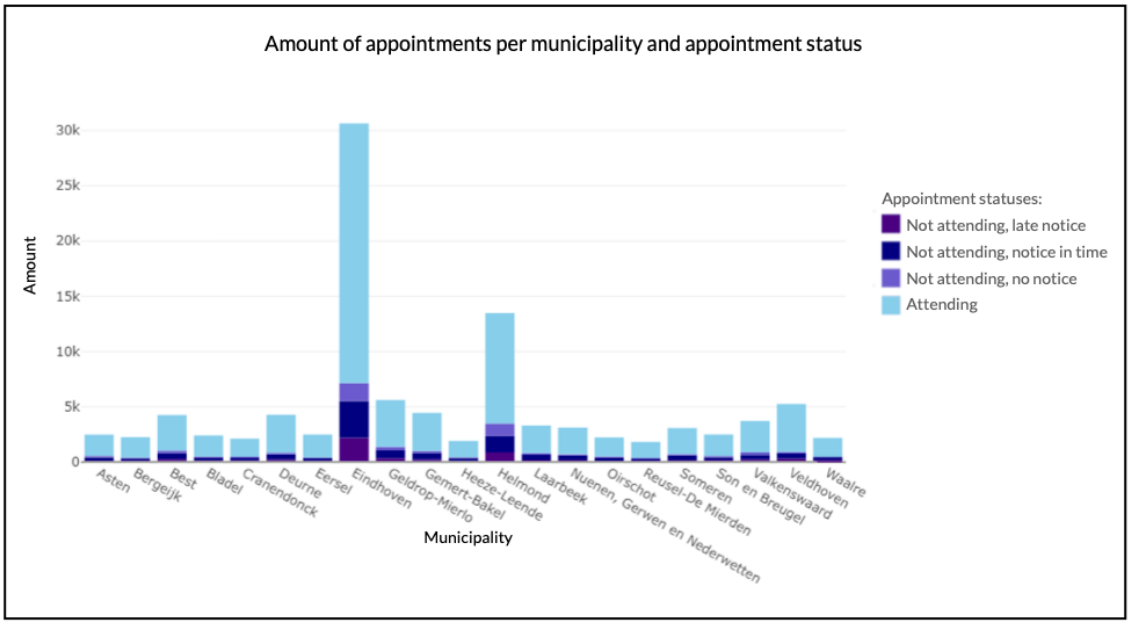


Figure 4.1: Number of appointments per appointment status per appointment type

Non-attendance rate

Seventy-seven percent of all appointments are attended, 22.32% (*SD* = 2.50%, range = 16.90% - 24.67%) of all appointments are not attended, where 11.80% (*SD* = 1.61%, range = 8.79% - 14.00%) canceled their appointment in time (more than 48 hours in advance), 6.39% (*SD* = 0.98%, range = 4.43% - 8.33%) canceled their appointment too late (within 48 hours), and 4.13% (*SD* = 1.58%, range = 1.19% - 8.33%) did not cancel their appointment and did not show up. This study focuses on the 10.52% (*SD* = 1.84%, range = 7.20% - 14.85%) cases that either did not up or canceled their appointment late.

Two visualizations take a deeper look at the data, namely at the municipality level and at the type of appointment level. In figure 4.1, the data is broken down at the municipality level, divided among the 21 municipalities in the GGD Brabant Zuid-Oost region. The two urban areas, the cities Eindhoven (30545 appointments) and Helmond (13279 appointments), have the highest rates of clients visiting Dutch youth healthcare (JGZ) appointments.
Also, this graph represents the division among municipalities. The municipality with the highest attendance rate is Eersel (83.10%), whereas the municipality with the lowest attendance rate is Helmond (74.12%). Also, the municipality with the highest problem behavior is Helmond (14.85%).

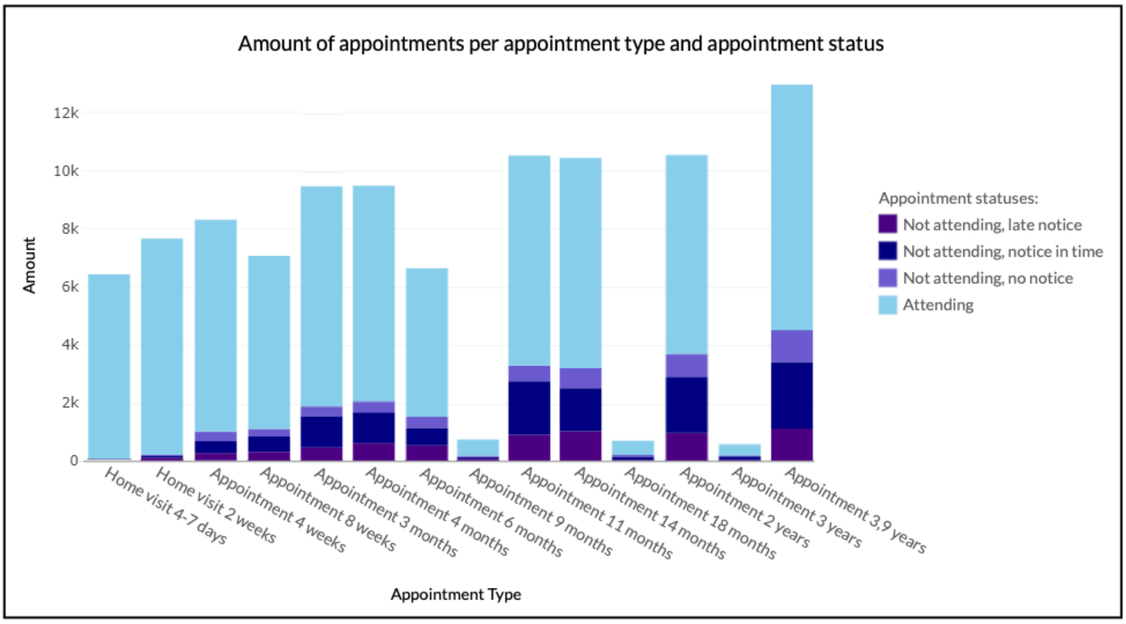


Figure 4.2: Number of appointments per appointment status at municipality level

In figure 4.2, the data is broken down at the appointment type level, divided among the 14 appointment types. This graph depicts the number of appointments per appointment status, with appointment type represented on the x-axis. Remarkable are the three appointment types (18 months, three years, and nine months) are barely used compared to the other appointment types. Also, appointment type 3,9 years is most represented in the data (12994 appointments). Besides that, two appointment types (Home visit 4-7 days and Home visit two weeks) have the highest attendance rate, respectively 99.7% and weeks 97.3%. This may be because these appointments occur in a person's house immediately after the birth of their child. Also, this graph represents the division among appointment types, where the appointment status attended is highest for every type of appointment (M = 77.02%, SD = 11.79%, range = 63.64% - 99.67%). The percentage of individuals who have chosen to cancel their appointments timely represents the second large group of appointments (M = 11.30%, SD = 6.43%, range 0.00% - 20.03%). The smallest group represents the percentage of individuals who have failed to provide any notice of cancellation (M = 4.99%, SD = 2.54%, ranges = 0.02% - 8.48%). Lastly, the percentage of individuals who have chosen to cancel their appointments at late notice completes that data (M = 6.69%, SD = 3.28%, range = 0.31% - 10.16%). The appointment types with the most problem behavior are 14 months (16.77%) and 3.9 years (17.20%).

Lastly, the data is divided by Dutch nationality into two groups: those with Dutch nationality and those without Dutch nationality. Interestingly, there is a slight variation in the percentage of observations per appointment status. Appointment attendance varies between 77.08% for Dutch nationals and 77.09% for non-Dutch nationals, with 11.42% to 11.81% for timely cancellations, 6.54% to 6.51% for late cancellations, and 4.96% to 4.69% for failure to provide any notice of cancellation. These results are rather too small to concern.

4.2 Results of the explanatory model

After a first look at the data, the data set is examined to uncover patterns and trends that can provide insight into the underlying phenomena being studied. Five models have been used to create a final model that gives an understanding of the data. The five models' formulas and individual results are in Appendix B on page 52. Five different logistic regression models were used to test what behavioral factors influence the likelihood of individuals not missing appointments.

First, a **general** model including all potential factors was used to make a first distinction. The output can be found in table B.3 at page 55. The variables gender ($p = 0.834$), time of day ($p = 0.220$), the spread of socioeconomic status (0.487), percentage of people employed during the last four years ($p = 0.652$), had an intervention before ($p = 0.397$) and having the Dutch nationality ($p = 0.141$) were found not significant, thus not influencing the likelihood of individuals missing appointments. The significant variables that negatively affect no-show behavior are age, time of day, past appointment behavior, the spread of socioeconomic status, the standardized income of a neighborhood, and the percentage of people with various degrees (master's, bachelor's, practical). Each variable has a negative effect on no-show behavior, meaning that a one-unit increase in these variables leads to increased odds of missing appointments. This percentage increase per variable is summarized as follows: age (71%), time of day (3%), past appointment behavior (84%), the spread of socioeconomic status (17%), the standardized income of a neighborhood (5%), percentage of people with a master's degree (85%), percentage of people with a bachelor's degree (87%), and percentage of people with a practical degree (87%). The significant variables affecting no-show behavior are appointment number, socioeconomic status, and private capital. Each of these variables has a positive effect on no-show behavior, meaning that a one-unit increase in these variables leads to a decrease in the odds of missing appointments. This percentage decrease per variable is summarized as follows: appointment number (23%), socioeconomic status (83%), and private capital (2%). Time of day, neighborhood, and private capital has a moderate impact on the probability of missing an appointment, but the other factors have a major impact.

A second model **corrects for multi-collinearity** by checking the variables in a correlation matrix using the Pearson method with listwise deletion. In figure B.1 at page 52 the correlation matrix can be found. The socioeconomic status subscales (spread, private capital, employed last 4 years, standardized income, education low, education high) were discovered to be multi-collinear, with correlations greater than 0.8. The only sub-scale of socioeconomic status, not multi-collinear, was the percentage of people with bachelor's degrees ($r = -0.296$). The variance inflation factor is the second statistic employed for testing multicollinearity. The variance inflation factors for each variable are presented in table B.6. This table demonstrates, in addition to the correlation matrix, that age and appointment type have a large variance inflation factor. This suggests that age and appointment type are highly correlated. According to the GGD, appointment type is more explanatory for the model. Hence it is agreed to eliminate age from the model. The model is modified to eliminate all multi-collinear subscales and variables. This is done by keeping the appointment type in the data but removing age from the data. Also, all subscales of socioeconomic status, except for the percentage of people having a bachelor's degree, were removed from the data.

The third model looks for possible **interactions** that could be added to the model. The output can be found in table B.7 at page 57. The interaction effects tested in the model were: appointment number and past no-show behavior, appointment type and previous no-show behavior, socioeconomic status and time of day, time of day, and percentage employed in the last four years. The

GGD commissioned these interaction effects and therefore tested them. First, as single effects, socioeconomic status was found significant ($p < 0.001$), and time of day was found not significant for both categories (afternoon compared to morning: $p = 0.220$, other compared to morning: $p = 0.142$). The interaction effect between socioeconomic status and time of day ($p = 0.279$) and time of day and percentage of people employed in the last four years ($p = 0.235$) were insignificant. This means that the effect of socioeconomic status on the likelihood of individuals missing appointments is not different depending on the time of day, and the effect of time of day on the probability of individuals missing appointments is not different depending on the percentage of people employed in the last four years.
On the other hand, the effect of the appointment number on the likelihood of individuals missing appointments differs depending on past no-show behavior, which indicates a negative effect on no-show behavior. Meaning a one-unit increase in the interaction between appointment number and past appointment behavior leads to an increase of 2% in the odds of missing appointments. Also, the type of appointment affects how likely an individual is to show up, which changes based on their past appointment behavior and the number of missed appointments. A one-unit increase in the interaction between a certain appointment type and past appointment behavior decreases the odds of missing appointments ranging from 39% for an appointment at eight weeks to 76% at 3.9 years. The interaction effects between the number of appointments and past no-show behavior and the type of appointment past no-show behavior were included in the model.

The fourth model, **random intercept model**, aims to account for the intercept variation across groups in the data. The cause of aiming to execute such a model is the firm belief that the data is clustered on a client level, meaning 30319 different groups in the data. A random intercept model may answer the question of to what extent variation in the data can be explained at an individual level, which means at a client level. For example, individual-level predictors might include patient characteristics such as age and gender, whereas group-level predictors might include variables such as appointment type. The random slope intercept would allow for the estimation of both fixed effects (i.e., the average effect of a predictor variable across all observations) and random effects (i.e., the deviation of the effect of a predictor variable for each group (client) in the data. Within the model, the group-level predictor client number is used, and the formula of the random intercept model can be found in Appendix B in section B.2.4.

The fifth model, **random slope model**, in contrast to a random intercept model, permits each group to have a distinct slope, which implies that the predictor variable might have a varied influence on each group. It allows the connection between the explanatory variable and the response to vary among groups. Within the model, the group-level intercept past appointment behavior is used; the formula of the random slope model can be found in Appendix B in section B.2.5.

Ultimately, both models can be combined as an addition to the third model. The random slope and random intercept model allow more flexibility in predicting attendance by accounting for variation in both the slope and the intercept across groups. In the end, this increases the accuracy and precision of the predictions made by the model. Due to the complexity of the model, which had 110933 observations with 30319 groups of client numbers, and 19 groups of past appointment behavior, both models could not give a result. The model failed to converge after 14 hours of execution, despite the use of various optimizers. Therefore, no random intercept or random intercept is included in the final model.

4.2.1 Final Model

The final model is the model combining the output of all five models. The goal of the final model is to understand the data, which is explained by the output of the final model. The final model uses the following formula, including variables and interaction effects:

$$\mathit{Not\_attending\_problematic} = \mathit{Gender} + \mathit{Municipality} + \mathit{Appointment\ type} + \mathit{SES\_WOA} + \mathit{Appointment\_number} + \mathit{not\_attending} + \mathit{Dutch\_yes\_no} + \mathit{Intervention\_yes\_no} + \mathit{Education\_middle} + \mathit{Appointment\_number}{*}\mathit{not\_attending} + \mathit{Appointment\ type}{*}\mathit{not\_attending}$$

The final logistic regression model results are presented in tables 4.1 and 4.2. First, the **demographics** are considered, i.e., the features that cannot be (easily) altered by an individual. The outcome indicates that for the variable gender, being female or male does not affect the likelihood of people missing their appointments. This indicates that the gender of the child has no bearing on their behavior. The output also shows that within the variable municipality, some municipalities (Bergeijk; p = 0.235, Best; p = 0.394, Bladel; p = 0.516, Cranendonck; p = 0.360, Geldrop-Mierlo; p = 0.126, Heeze-Leende; p = 0.896, Nuenen, Gerwen en Nederwetten; p = 0.357, Reusel-De Mierden; p = 0.456, Veldhoven; p = 0.076, Waalre; p = 0.282) do not affect the likelihood of individuals missing appointments compared to the municipality Eindhoven. The other municipalities show a small effect on no-show behavior. Compared to the municipality of Eindhoven, the odd ratios of the other municipalities range from 0.69 for the municipality of Someren to 1.15 for Valkenswaard. Living in Someren reduces the odds of missing appointments by 31% compared to living in Eindhoven, whilst living in Valkenswaard increases the odds of missing appointments by 15% compared to living in Eindhoven. This implies that, in the interest of this study, appointment behavior cannot be considered the same throughout all GGD Brabant Zuid-Oost municipalities and either positively or negatively affects no-show behavior.

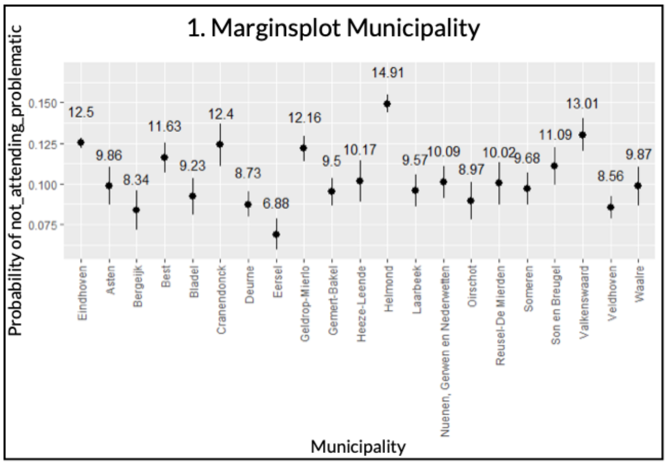


Figure 4.3: Marginsplots: Municipality

However, to get a better overview of the differences between all municipalities, instead of only comparing it to the municipality of Eindhoven, a marginsplot is used. The marginsplot in figure 4.3 depicts the likelihood of appointment non-attendance for each municipality, maintaining all other predictor variables at their respective mean values. Suppose a municipality, such as Bladel, is insignificant. The predicted probabilities for the municipality Bladel and the reference category Eindhoven will be comparable, and they will not differ significantly in terms of the likelihood of missing appointments. Figure 4.3 demonstrates the probability of appointment non-attendance (M = 10.34%, SD = 1.87%, range = 6.88% - 14.91%). Eindhoven and Helmond are more than one standard deviation off the mean, meaning they are farther away from the average value than would be expected by chance; this is remarkable. The individuals living in Eindhoven and Helmond account for 43 percent of the data, which might be problematic because it may not accurately represent the diverse population of all GGD Brabant Zuid-Oost municipalities. This offers an additional argument for model validation, which will be explained in section 4.4. The municipality of Eersel has the lowest likelihood of appointment non-attendance, at 6.88 percent, and might serve as an example to other municipalities regarding practical consequences.

	not_attending_problematic		
Predictors	*OR*	*CI*	*p*
(Intercept)	0.04	0.04-0.06	<**0.001**
Gender [Female]	1.00	0.96-1.05	0.882
Municipality [Asten]	0.73	0.65-0.95	**0.001**
Municipality [Bergeijk]	0.87	0.69-1.09	0.235
Municipality [Best]	1.06	0.95-1.25	0.394
Municipality [Bladel]	0.94	0.77-1.14	0.516
Municipality [Cranendonck]	1.09	0.91-1.31	0.360
Municipality [Deurne]	0.78	0.67-0.91	**0.002**
Municipality [Eersel]	0.77	0.63-0.94	**0.010**
Municipality [Geldrop-Mierlo]	0.91	0.81-1.03	0.126
Municipality [Gemert-Bakel]	0.75	0.65-0.88	<**0.001**
Municipality [Heeze-Leende]	1.01	0.83-1.21	0.896
Municipality [Helmond]	1.10	1.01-1.10	**0.036**
Municipality [Laarbeek]	0.72	0.61-0.85	<**0.001**
Municipality [Nuenen, Gerwen en Nederwetten]	0.93	0.80-1.08	0.357
Municipality [Oirshot]	0.81	0.67-0.98	**0.035**
Municipality [Reusel-De Mierden]	0.92	0.74-1.14	0.456
Municipality [Someren]	0.69	0.58-0.82	<**0.001**
Municipality [Son en Breugel]	1.23	1.04-1.45	**0.014**
Municipality [Valkenswaard]	1.15	1.00-1.32	**0.049**
Municipality [Veldhoven]	0.89	0.78-1.01	0.076
Municipality [Waalre]	0.91	0.76-1.08	0.282
Appointment type [Home visit 4-7 days]	0.09	0.04-0.12	<**0.001**
Appointment type [Home visit 2 weeks]	0.44	0.33-0.59	<**0.001**
Appointment type [Appointment 8 weeks]	1.19	0.80-1.58	0.216
Appointment type [Appointment 3 months]	1.33	1.01-1.74	0.126
Appointment type [Appointment 4 months]	1.61	1.21-2.15	**0.001**
Appointment type [Appointment 6 months]	1.97	1.41-2.74	<**0.001**
Appointment type [Appointment 9 months]	5.38	2.78-9.69	<**0.001**
Appointment type [Appointment 11 months]	2.85	2.15-3.78	<**0.001**
Appointment type [Appointment 14 months]	3.51	2.63-4.69	<**0.001**
Appointment type [Appointment 18 months]	5.40	2.75-9.98	<**0.001**
Appointment type [Appointment 2 year]	3.87	2.78-5.39	<**0.001**
Appointment type [Appointment 3 year]	3.28	1.65-6.15	<**0.001**
Appointment type [Appointment 3.9 year]	3.04	2.00-4.58	<**0.001**
Appointment number	0.72	0.71-0.73	<**0.001**
Not attending	4.84	0.71-0.73	<**0.001**
SES WOA	0.41	0.37-0.45	<**0.001**
education middle	1.00	1.00-1.01	0.141
Dutch yes no [1]	1.07	0.99-1.15	0.096
Intervention yes no [1]	0.99	0.94-1.05	0.796
appointment_number x not attending	1.02	1.02-1.03	<**0.001**
Appointment type [Home visit 4-7 days] x not attending	0.65	0.39-1.02	0.068
Appointment type [Home visit 2 weeks] x not attending	1.07	0.81-1.41	0.634

Table 4.1: Final Model: Data output (1)

	not_attending_problematic		
Predictors	*OR*	*CI*	*p*
Appointment type [Appointment 8 weeks] x not attending	0.60	0.53-0.69	**<0.001**
Appointment type [Appointment 3 months] x not attending	0.49	0.43-0.55	**<0.001**
Appointment type [Appointment 4 months] x not attending	0.42	0.37-0.47	**<0.001**
Appointment type [Appointment 6 months] x not attending	0.40	0.35-0.45	**<0.001**
Appointment type [Appointment 9 months] x not attending	0.31	0.27-0.36	**<0.001**
Appointment type [Appointment 11 months] x not attending	0.32	0.29-0.36	**<0.001**
Appointment type [Appointment 14 months] x not attending	0.30	0.23-0.33	**<0.001**
Appointment type [Appointment 18 months] x not attending	0.27	0.24-0.31	**<0.001**
Appointment type [Appointment 2 year] x not attending	0.26	0.23-0.29	**<0.001**
Appointment type [Appointment 3 year] x not attending	0.25	0.22-0.28	**<0.001**
Appointment type [Appointment 3.9 year] x not attending	0.24	0.22-0.27	**<0.001**
Observations	101993		
R^2 Tjur	0.289		

Table 4.2: Final Model: Data output (2)

Another demographic is a person's socioeconomic status, which also has a small effect on their likelihood of missing appointments. Table 4.1 demonstrates that a one-unit increase in socioeconomic status ($p < 0.001$) decreases the odds of missing appointments by 59 percent. The marginsplot 'SES_WOA' in figure B.2 on page 54 of Appendix B depicts the probability of appointment non-attendance for each socioeconomic status, ranging from 48% to 13% as socioeconomic status increases. In other words, the greater a person's socioeconomic status, the lesser their likelihood of missing appointments. Another demographic in the model is Dutch nationality. The output in table 4.1 indicates no significant difference in attendance rates between Dutch and non-Dutch nationals ($p = 0.096$). The last demographic is the proportion of individuals with a bachelor's degree. Table 4.1 demonstrates that having a bachelor's degree ($p = 0.141$) does not affect the probability that individuals would miss appointments.

Second, the **behavioral characteristics** are considered, i.e., the features that are impacted by an individual's behavior. The output at table 4.1 indicates that past appointment behavior ($p < 0.001$) significantly influences individuals' probability of keeping future appointments. The number of missed appointments in the past ranges from 0 to 19. This variable has the biggest effect on no-show behavior, as a one-unit increase in prior appointment behavior enhances the odds of missing an appointment by 4.84 times. The marginsplot 'not_attending' in figure B.2 on page 54 of Appendix B demonstrates that the model identifies six prior no-shows as a one hundred percent non-attendee. Thus, it may be concluded that past appointment significantly impacts appointment non-attendance. The data in table 4.1 further reveal that the number of appointments has a small significant positive effect ($p < 0.001$) on the likelihood of appointment non-attendance. The appointment number is the number of scheduled appointments an individual has, which may exceed the 14 appointment types because a canceled appointment also has an appointment number. The result suggests that a one-appointment increase decreases the odds of missed appointments by 28%. The marginsplot 'number of appointments' in figure B.2 on page 54 of Appendix B demonstrates that the probability of missing an appointment reduces as the number of scheduled appointments increases. This means that the more appointments an individual has, the less likely these appointments result in a no-show. The last behavioral characteristic examined is the interaction effect of past appointment behavior and appointment number. Both past appointment behavior and appointment number individually have a significant effect on appointment attendance. Together past appointment behavior just slightly negatively influences appointment non-attendance. A one-unit increase in the interaction between appointment number and past appointment behavior leads to an increase of 2% in the odds of missing appointments. The marginsplot 'not_attending x number of appointments' in figure B.2 on page 54 of Appendix B does not reveal a clear effect, as the lines per appointment do not fluctuate. Next, the effect of the type of appointment ranges from 0.07 to 5.40, meaning that for having a home visit at 4-7 days ($p < 0.001$), the probability of missing appointments decreases by 93% compared to having a 4-week appointment ($p < 0.001$), whereas having an appointment at 18 months increases the odds of missing an appointment with 5.4 times compared to having an appointment at four weeks. The magnitude and direction of the effect vary among different appointment types.

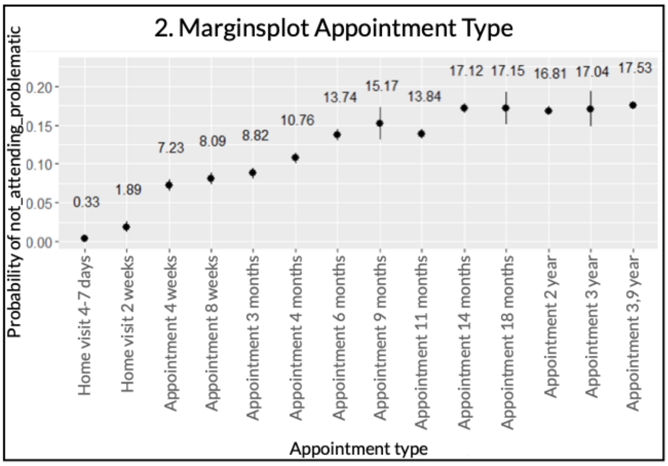


Figure 4.4: Marginsplots: Appointment Type

However, to get a better overview of the differences between all appointment types, instead of only comparing it to the appointment type 4 weeks, a marginsplot is used. The marginsplot in figure 4.4 displays the probability of missing appointments for each type, with all predictor variables maintained at their respective mean values. The probability of missing an appointment is depicted in 4.4 ($M = 11.82\%$, $SD = 5.79\%$, range = 0.33% - 17.53%). The types of appointments are listed in chronological order. Figure 4.4 demonstrates that the likelihood of missing appointments rises with time, except for 11-month, 2-year, and 3-year appointments, which means that appointments with a later chronological time are more likely to have no-shows, particularly the two years, three years, and 3.9-year appointments.
Other than appointment numbers, appointment types represent a specific appointment type, and appointment numbers specifically zoom in on the frequency of scheduled appointments among these types. Another behavioral characteristic would be whether an individual had an intervention before; this does not significantly affect missing appointments ($p = 0.796$). The final behavioral characteristic evaluated is the interaction effect of past appointment behavior and appointment type. Table 4.2 reveals that the interaction has a statistically significant small effect on the majority of appointment types ($p < 0.001$), excluding home visits for 4 to 7 days ($p = 0.068$)

and home visits for two weeks (p = 0.634). This effect reduces individuals' odds of keeping their scheduled appointments by 40% to 66%. The marginsplot 'not_attending x type of appointment' in Appendix B picture B.2 on page 54 demonstrates the relationship between the interaction effect of appointment type and past appointment behavior. The likelihood of a no-show is influenced separately by appointment type and past behavior. The two variables demonstrate that the influence of past no-show behavior on an individual's likelihood of missing an appointment varies depending on the specific types of appointments by diminishing the effect with time. The marginsplot illustrates this effect, with the appointment type home visit 4-7 days displaying a steeper line than appointment type 3.9 years. This demonstrates that the influence diminishes with time since the probability of a 100 percent no-show increases as the number of missed appointments increases.
Overall, past appointment behavior, the number of appointments, appointment type, and the interaction between appointment type and past appointment behavior are the behavioral characteristics that influence individuals' tendency to miss appointments.

Performance final model
For the final model, the accuracy (0.83), precision (0.93), recall (0.90), f1-score (0.91), and confusion matrix (74128 true positives, 16017 false positives, 10787 true negatives, and 861 false negatives) indicate the model's performance. The threshold for classifying a no-show is set at 90% probability of missing an appointment due to the GGD Brabant Zuid-Oost guidelines. The model can classify individuals with a 90% probability of missing an appointment. The model classifies 26.33% of the total data, of which 40.24% is correctly classified as non-attender. These non-attenders represent 92.60% of the individuals missing appointments according to the model. All in all, this model performs well since it uses a small part of the data to detect 92.60% of non-attenders.

4.3 Results most accurate classification method

This section discusses three classification methods to determine the most accurate classification method for use in practice. The three classification methods used are the classification methods the GGD Brabant Zuid-Oost uses a standard: logistic regression, K-Nearest Neighbour, and Random Forest. Since these algorithms are widely used and well-established in the field of machine learning, there is a wealth of resources available for understanding and implementing them. Which, for practical implication, is very useful. This section aims to find the classification method that can correctly classify most individuals. The data is split into a training dataset (80% of the data) and a validation dataset (20% of the data). The model's performance is measured by comparing the training dataset with the validation dataset using accuracy metrics.

Logistic Regression
The output of the logistic regression can be found in table B.10 and B.11 one page 59. The significant predictors, including the ones with the highest effect, thus the odd ratios, are the municipality, type of appointment, past appointment behavior, socioeconomic status, and appointment number. These are also the variables that showed a significant effect in the final model of the explanatory section.

Random Forest
The random forest output can be found in figure B.3 on page 60. In order of highest predictive value, the significant predictors are past appointment behavior, appointment number, appointment type, the time between appointments, and socioeconomic status.

K-Nearest Neighbour
The output of the K-Nearest Neighbour can be found in figure B.4 on page 61. This illustrates that

the model needs five neighbors to correctly classify 88.2% of the data, seven neighbors to classify 88.3%, and nine neighbors to classify 88.4%. This does not say anything about the significant predictors the model uses.

4.3.1 Most accurate classification method

A few metrics are used to compare the machine learning methods: accuracy, sensitivity, f1-score, and confusion matrices. These metrics are created for the training as well as for the validation data. Table 4.3 illustrate these metrics for the training data, whereas table 4.4 illustrates these metrics for the validation data.

Model	*Accuracy*	*Precision*	*Recall*	*F1-score*	*Confusion matrix*		
Logit	0.83	0.82	0.99	0.90		0	1
					0	59303	689
					1	12814	8629
KNN	0.68	0.68	0.90	0.88		0	1
					0	45110	97
					1	27007	9221
RF	0.88	0.83	0.99	0.91		0	1
					0	61051	149
					1	10856	9369

Table 4.3: Summary Output Models: Training dataset

Model	*Accuracy*	*Precision*	*Recall*	*F1-score*	*Confusion matrix*		
Logit	0.82	0.81	0.96	0.89		0	1
					0	14678	179
					1	3382	2149
KNN	0.63	0.61	0.88	0.75		0	1
					0	11008	423
					1	6979	1905
RF	0.83	0.83	0.97	0.89		0	1
					0	14905	148
					1	3109	2180

Table 4.4: Summary Output Models: Validation dataset

All models identify non-attendees as individuals with at least a 90% probability of missing an appointment. When looking at the metrics of the training data (see figure 4.3), logistic regression classifies 26.33% of the total data, of which 40.24% is correctly classified as non-attender. These non-attenders represent 92.60% of the non-attenders according to the model. K-Nearest neighbor classifies 48.49% of the total data, of which 25.45% is correctly classified as non-attender. These non-attenders represent 94.70% of the individuals missing appointments according to the model. Random forest classifies 24.83% of the total data, of which 46.28% is correctly classified as non-attender. These non-attenders represent 98.43% of individuals missing appointments according to the model.

In general, looking at the accuracy scores, high accuracy scores indicate that the models are able to correctly predict the outcome of the majority of the training data. Meaning that logistic regression and random forest have good accuracy, whereas K-Nearest Neighbours' accuracy score can be considered bad. The precision of the models shows the rate of false positives, individuals that have been indicated as non-attenders, but actually are not. Meaning that logistic regression and random forest have a good precision score, whereas K-Nearest Neighbours' precision score can be considered bad. This is also indicated by the fact that in the group of individuals indicated by K-Nearest Neighbour as non-attenders, only 25.45% is correctly classified as non-attenders. Next, it is remarkable that all methods have excellent recall values, which indicates that the model has a low false negative rate. In other words, a high recall means that the model can correctly identify a high proportion of people missing appointments. Last, the f1-scores indicate the balance between the precision and recall in the harmonic mean between them, indicating the model has a low false positive rate and a low false negative rate, identifying positive instances while avoiding false predictions. Logistic regression and random forest have excellent f1-scores, where K-Nearest Neighbours' f1-score can be indicated well. Looking at these results, K-Nearest Neighbour performs worse than logistic regression and random forest considering the training data, where random forest performs slightly better than logistic regression. Next, these values are compared with the validation data to see how well the models can predict non-attendance among new data.

When looking at the metrics of the validation data (see figure 4.4), logistic regression classifies 27.38% of the total data, of which 39.19% is correctly classified as non-attender. These non-attenders represent 92.31% of the non-attenders according to the model. K-Nearest neighbor classifies 43.73% of the total data, of which 21.44% is correctly classified as non-attender. These non-attenders represent 81.83% of the non-attenders according to the model. Remarkable is the drop in accuracy of the K-Nearest Neighbour from 68% to 63%. A drop in accuracy between the training and validation data generally indicates that the model overfitted the training data. This means the model has learned to fit the training data too closely and has not generalized well to new, unseen data. Random forest classifies 26.03% of the total data, of which 40.11% is correctly classified as non-attender. These non-attenders represent 93.64% of the non-attenders according to the model. The results show that K-Nearest Neighbour performs worse than logistic regression and random forest considering the training data. K-Nearest Neighbours compared to logistic regression, and the random forest was most overfitted at the training data, resulting in a drop in all metrics. Comparing the metrics from the training in validation data, it is found that the confusion matrices of the logistic regression method ($p = 0.254$) do not show statistically significant differences. The confusion matrices of the K-Nearest Neighbour ($p = 0.837$) also do not show statistically significant differences. The confusion matrices of the random forest method ($p = 0.463$) also do not show statistically significant differences. The validation data is a subset of the original data set aside for testing. This illustrates that the models used can be generalized to new data. Random forest, compared to logistic regression, slightly outperforms logistic regression. This is indicated by indicating a smaller part of the total data as non-attenders (26.03% compared to 27.38%), finding a larger amount of actual non-attenders (93.64% compared to 92.31%).

Cross-validation Area Under the Curves (AUCs) for the logistic regression model, K-Nearest Neighbor, and Random Forest are 0.89, 0.79, and 0.92, respectively. Plots of the performance of a classification model at all classification thresholds are displayed in figure 4.5. This model visualizes that the AUC of K-Nearest Neighbour is lower than that of logistic regression and random forest, and random forest slightly outperforms logistic regression.

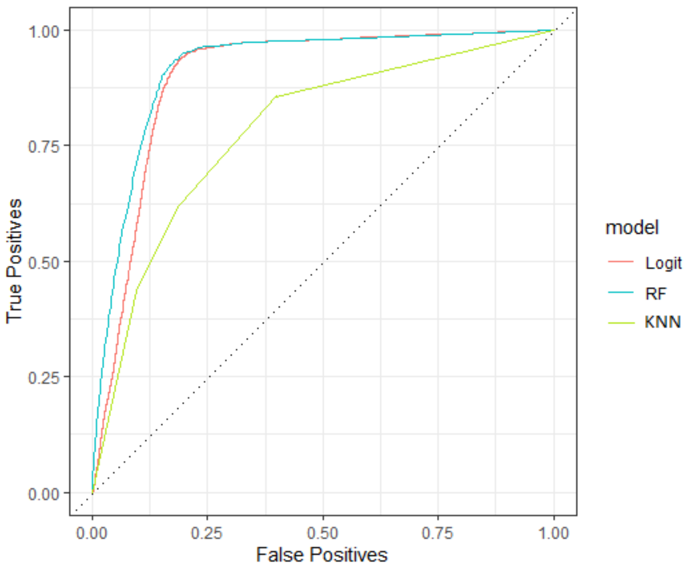


Figure 4.5: ROC curve

4.4 Results of machine learning and decision-making

This section tests for equality among data groups to guarantee that the model works equally well for subgroups. This tests how well the predictive model meets its intended purpose and specifications across demographics. This is done by comparing four groups: urban versus rural areas and low-income versus high-income neighborhoods. Classification accuracy and fairness metrics are used to compare the groups.

4.4.1 Urban versus rural areas

The data is split into urban areas (43824 appointments) and rural areas (57969 appointments). In table B.12, the **classification accuracy** metric can be found. Comparing urban and rural areas, the accuracy score is higher. Comparing urban and rural areas, the accuracy score is greater in rural areas, the sensitivity score is greater in urban areas, and the f1-score is greater in rural areas. Considering confusion matrices, urban areas require more data to classify 98% of non-attendees accurately, which both of the subgroups do. A t-test compares the two confusion matrices using data from all cross-validation folds. There is no significant difference between urban and rural locations ($p = 0.681$). Indicating that these differences are not statistically significant and that the observed disparities across subgroups are attributable to random chance.

The outcomes of comparing urban and rural locations for **fairness metrics** are depicted in figure B.13. The equal opportunity difference is -0.0049, which indicates that the model labels 0.49% more urban residents as non-attenders. The statistical parity difference is -0.0050, which indicates that the model identifies 0.50% more urban residents as non-attenders. Last, the average odds difference is -0.0026, which indicates that the model identifies 0.26% more urban residents as non-attenders. Overall, urban areas have an average 0.42 percent difference greater probability of non-attendance. The disparities are far less than the 5% threshold normally used to determine random chance, so they do not need adjustment.

4.4.2 Low- versus high-income areas

The data is split into low-income areas (37618 appointments) and high-income areas (64175 appointments). In table B.12, the **classification accuracy** metric can be found. Comparing low and high-income areas, the accuracy score is higher for high-income areas, the sensitivity scores are equal for both groups, and the f1-score is higher for low-income areas. Comparing low-income and high-income areas reveals that the accuracy score is greater in high-income areas, the sensitivity score is the same for both groups, and the f1-score is higher in low-income areas. Considering the confusion matrices, low-income areas require a greater portion of the data to classify 98% of non-attenders accurately, which both subgroups do. A t-test compares the two confusion matrices using data from all cross-validation folds. There is no significant difference ($p = 0.481$) between low-income and high-income locations. These differences are not statistically significant, and the observed disparities across subgroups are related to chance.

The findings of comparing low- and high-income locations for **fairness metrics** are depicted in figure B.13. The equal opportunity difference is -0.0095, which indicates that the model classifies 0.95% more low-income individuals as non-attenders. The statistical parity difference is -0.031, which indicates that the model identifies 3.1% more low-income individuals as non-attenders. Last, the average odds difference is -0.0018, indicating that the algorithm identifies 0.18% more low-income individuals as non-attenders. Low-income areas had an average 1.41 percentage point greater likelihood of being labeled non-attendance. The disparities are far less than the 5% threshold normally used to determine random chance, so they do not need adjustment. Despite these modest differences, the model is valid when comparing low- and high-income subgroups.

4.4.3 Acceptance of machine learning

This part includes how well the predictive model meets its intended purpose and specifications across domain expertise. This section aims to identify potential mismatches in decision-making by testing deviation between domain experts' beliefs and the prediction model's factors.
The prediction model's significant predictor variables are:

- Municipality
- Appointment type
- Appointment number
- Past appointment behavior
- Socioeconomic status

The health professionals and domain experts indicated if potential predictor variables predict non-attendance. The outcomes are depicted in figure 4.6. This figure displays the combined number of health professionals and domain experts, with a possible predictor indicated on the x-axis. In figure 4.6, the positive number on the y-axis represents the number of people who believe a factor impacts predicting no-show behavior. In contrast, the negative amount on the y-axis represents those who believe a factor has no impact on predicting no-show behavior. All the predictor variables in figure 4.6 derive from the literature framework. Before the completion of the dataset, these variables were presented to health professionals and domain experts. Consequently, distance, marital status, the gender of the parents, and the consequences of a no-show are not accounted for in this data analysis.

An equal division between people who believe a predictor variable has an effect and people who believe the predictor does not indicate there is no proof that health professionals and domain experts have an unequivocal opinion about the predictor. This is the case for distance, marital status, and time of day. The predictor variable 'consequences of a no-show' are commonly mentioned as an

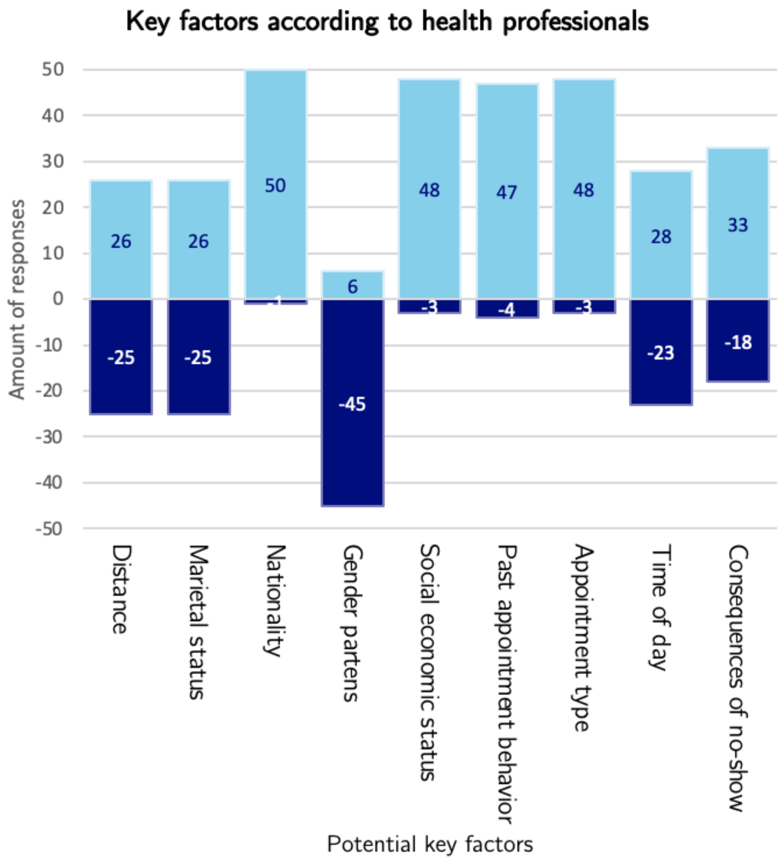


Figure 4.6: Potential predictor variables according to domain expertise

important element, although only infrequently. Nationality, gender of the parents, socioeconomic status, past appointment behavior, and appointment type are the characteristics with definite indications. Most health professionals and domain experts indicate the gender of the parents as a predictor variable that does not influence no-show behavior. The results show that most health professionals and domain experts indicate nationality, socioeconomic status, past appointment behavior, and appointment type impact the ability to predict appointment non-attendance.

The deviation between what the prediction model predicts and what domain experts believe is that nationality has no impact on appointment attendance. This indicates that the beliefs of domain experts and the prediction model variables mismatch. From the perspective of health professionals and domain experts, nationality is a predictor for no-show behavior, whereas, in the prediction model, this effect is not found. The domain experts' understanding of the problem is not fully aligned with how the model interprets the data.

Chapter 5

Conclusion

This thesis project studies no-show behavior in the context of Dutch youth healthcare (JGZ) appointments, part of GGD Brabant Zuid-Oost. This study aims to improve the effectiveness of healthcare delivery in the context of the GGD Brabant Zuid-Oost by mapping no-show behavior, supporting the decision-making process of improving appointment attendance, and minimizing the operational impact of no-show behavior. This is achieved by addressing the main research question: **How can attendance at Dutch youth healthcare (JGZ) appointments be predicted?** To address the primary research question, the study is divided into three different parts: the exploratory model examines the behavioral and environmental factors that determine a person's probability of missing appointments, the most accurate classification method that classifies those who are most likely to exhibit no-show behavior, and machine learning methods and decision-making in which the model is evaluated for its capacity to predict independently of data subgroups effectively and to indicate a potential deviation between the beliefs of domain experts and the prediction model's significant features.

Summary and interpretation of the results
This study contributes to the scientific literature by identifying predictors for no-show behavior unique to the Dutch youth healthcare (JGZ) setting. However, by comparing this study to another study that examines no-show behavior in other healthcare settings, it is possible to determine whether the predictor variables of this study are unique to the Dutch youth healthcare (JGZ) setting or whether these results are generalizable. First, the main findings are summarized and interpreted using the explanatory model, the most accurate classification, machine learning, and decision-making.

The first part, the **explanatory model** found that the municipality where individual lives and their socioeconomic status are the demographics that significantly affect the likelihood of not attending appointments. Many municipalities, relative to Eindhoven, had a small positive or negative effect on no-show behavior, depending on the municipality, indicating that the place where you live, something that cannot be easily changed, already affects no-show behavior positively or negatively. This could be explained by the difference in the availability of public transport, and the effort patients must undertake for visiting appointments (Morera-Guitart et al., 2002). For example, the difference between Eindhoven and Helmond might be explained as, on average, patients from Eindhoven need to travel shorter distances to visit appointments than patients from Helmond due to the number of health clinics in each municipality. The required effort for patients from Helmond might be higher than for patients from Eindhoven. This increases the likelihood of employing maladaptive coping techniques, leading to missed appointments. This is consistent with the findings of the studies that explained that an increased travel distance increased the likelihood of exhibiting no-show behavior (Morera-Guitart et al., 2002; Daggy et al., 2010; Dantas et al., 2019).

Another demographic that plays a role in no-show behavior is socioeconomic status, which had a small effect on no-show behavior. This study found that a lower socioeconomic status increases the likelihood of individuals missing appointments, illustrating a negative effect of socioeconomic status on no-show behavior. The magnitude and direction of the effect are consistent with the finding of previous studies (C. Kelly et al., 2016; Meléndez-Ackerman et al., 2014). Different psychological constructs can explain the effect depending on the research context. C. Kelly et al. (2016) contribute the results of socioeconomic status as a predictor for no-show behavior to language and cultural barriers, and Meléndez-Ackerman et al. (2014) contribute the results to the willingness to seek help. These explanations, language, cultural barriers, and the willingness to seek help might explain why the effect of socioeconomic status might be found. However, the more supported and logical explanation is based on the paper of Adler and Ostrove (1999). They described that socioeconomic status includes an individual's access to resources. Worse availability of health clinics leads to an increased demand for effort, thereby increasing the probability of individuals missing their appointments (Morera-Guitart et al., 2002), as explained before. Therefore, the effort required to attend an appointment might be an overarching construct explaining the role of demographics in no-show behavior.

The study found that various behavioral characteristics significantly influence a person's likelihood of missing appointments. Previous research indicates that self-efficacy, motivation, perceived control, and perceived importance influence the behavioral characteristics that cause no-show behavior. Moreover, the results of this study show that past appointment behavior is the biggest predictor of no-show behavior, showing a large effect negative effect on no-show behavior. Chariatte et al. (2008) discovered that past appointment behavior significantly impacted no-show behavior, where at least four appointments are required to discuss a behavioral pattern. As an explanation, they propose that individuals repeat past behavior in comparable contexts, commonly known as behavioral consistency. Past appointment behavior in this study might also be driven by behavioral consistency.

The number of appointments also had a small effect on non-appearance. Where an increase in appointments reduces the likelihood of missing them. Patients have eight appointments on average, and referrals are required to arrange additional appointments. People with more than 14 appointments attend health interventions since there are only 14 appointment types within this study's Dutch youth health care (JGZ) scope. C. Kelly et al. (2016) explain that when patients feel the health situation is more serious, people are more eager to come. A serious health situation increases the perceived importance of the appointment, which might explain why the appointment number affects no-show behavior.

Last, this study found that no-show behavior varies across appointment types. The different appointment types, relative to the four-week appointment, had a large positive or negative effect on no-show behavior. The effect found is consistent with the finding of previous studies (Goffman et al., 2017; C. Kelly et al., 2016). Both studies suggest that perceived importance explains why appointment type is a significant predictor. Since the different appointments have different purposes, it is assumed that perceived importance might also, in this context, explain why appointment type predicts no-show behavior. The interaction effect between past appointment behavior and appointment time shows that if a patient has missed appointments in the past, their likelihood of missing their next appointment is reduced if the appointment is of a certain type. This suggests that different types of appointments may have varying levels of importance to patients, which can affect their behavior in attending them. Since the appointments are chronologically ordered, the effect of past appointment behavior diminishes over time.

All in all, this answers the first sub-question as the demographics municipality living in and socioeconomic status, and the behavioral characteristics of past appointment behavior, appointment number, appointment type, and the interaction between past appointment behavior and appointment type predict no-show behavior. The results illustrate that the predictors of no-show behavior found in the Dutch youth healthcare (JGZ) context are not unique in this context due to literature in other healthcare contexts supporting these results. The predictors of no-show behavior found within this study are generalizable. In contrast, the psychological constructs causing the effects are understudied and suggest potential reasons why the effect in this study might be found.

The **most accurate classification** aimed to identify the most accurate classification method for predicting non-attendance at new data. Three classification methods were compared: logistic regression, K-Nearest Neighbor, and random forest. The models were evaluated based on their accuracy metrics and the number of individuals that could be correctly classified. The study found that random forest was the most accurate classification approach compared to logistics regression. Random forest also required a smaller percentage of data to forecast non-attendees compared to logistic regression. Ultimately, the differences between random forest and logistic regression are rather small. The second sub-question, which asks which method predicts no-show behavior most accurately, could be answered by recommending the random forest method. However, considering incorporating machine learning methods into the decision-making process Topuz et al. (2018) suggests the preferable use of easily interpretable but less accurate approaches such as regression models. Within this study, using logistic regression compared to the more sophisticated random forest model hardly drops accuracy, i.e., both models work excellently. Therefore it is recommended to use logistic regression instead of random forest since logistic regression would be comprehensible to most healthcare managers themselves, which might increase the probability they will accept such methods assisting the decision-making process.

The last section, **machine learning and decision-making**, consisted of two parts. The first part investigated whether the most accurate prediction method can effectively categorize data subgroups, such as urban and rural areas and low- and high-income areas. The study found that there were no statistically significant differences between these subgroups. This indicates that the model is reasonable and answers the third sub-questions by confirming the prediction model performs equally independently of data subgroups.

The second part examined the discrepancy between the beliefs of domain experts and the factors used in the prediction model. The study found that the domain experts' understanding of the problem did not fully align with how the model interprets the data, specifically regarding the impact of nationality on appointment attendance. This can be explained by a confounding variable, namely socioeconomic status. Literature shows that neighborhoods with a low socioeconomic status more often have people with a migration background (Adler & Ostrove, 1999). Domain experts might see people with different migration backgrounds thinking that missing appointments is caused by their nationality, where the low socioeconomic status of the neighborhood might actually causes this effect. In the end, the fourth sub-question can be answered by stating that there is a deviation between the beliefs of the domain experts and the prediction model's factors. However, domain experts might accept machine learning algorithms' assistance within this study's context. This can be argued by the fact that a confounding variable can explain the deviation in the literature, and recommending logistic regression enables health experts to use a comprehensible insightful method.

5.1 Discussion

Limitations and future work

This study has a few limitations. First, all the factors and their proposed explanation suggest how the psychological constructs involved are connected to no-show behavior. However, this study did not measure self-efficacy, motivation, perceived importance, and perceived control levels. For future work, participants' self-efficacy, motivation, perceived importance, and control levels can be measured to understand their connection to no-show behavior better. This could be done by conducting a study in which a questionnaire is developed to assess the participants' self-efficacy, motivation, perceived importance, and perceived control levels. This survey could be administered by email after the appointment has taken place to measure how these psychological constructs affect the patient for attending or missing the appointment. The outcomes of this questionnaire can be used to more extensively understand the psychological constructs that drive no-show behavior within the context of Dutch youth health care (JGZ) appointment of the GGD Brabant Zuid-Oost.

Another limitation of the study is that the data quality could be improved since it is collected by health professionals administering a system. More specifically, due to missing values and inconsistencies, the data could not include different variables: family members and detailed information about the parents. Past appointment behavior is the biggest predictor of no-show behavior within this study. As Chariatte et al. (2008) explained, data about at least four appointments are required to describe a behavioral pattern and thus say something about past appointment behavior. For the Dutch youth healthcare (JGZ) context, it is desirable to predict whether parents with a newborn child will attend an appointment; it may not be able to do so accurately if the child is just born and there is no data available on past appointment behavior. Future work could address this limitation by connecting patients in the data via their relationships with parents and siblings. Due to behavioral consistency, this information might include historical behavioral patterns that might likely be repeated. Also, personal information about the parents, such as a home address, might include information about travel distance: one of the unavailable demographics. Aditionally, controlling past appointment behavior in the statistical analysis may be necessary to assess other variables' impact on no-show behavior accurately and estimate the model's functionality on data of parents with a newborn child new to the healthcare system.

The last issue is the extent to which this study can say about embracing the assistance of machine learning methods. Domain experts will likely embrace such procedures, but this has not been thoroughly researched. Furthermore, domain experts and health professionals identified plausible causes of no-show behavior. In practice, health professionals and domain experts have completely different roles. The various functions of healthcare managers, health professionals, domain experts, and other concerned positions should be evaluated for practical consequences. This can also be accomplished by sending out a questionnaire to identify how different roles are involved in the decision-making process, mapping those people's values and beliefs towards machine learning approaches, and directly asking them their opinion on using machine learning approaches in the decision-making process. The questionnaire results reveal the conditions under which individuals are most inclined to adopt machine learning algorithms and the extent to which those various roles should be involved in decision-making.

Practical implications

As explained before, Ferro et al. (2020) and Millhiser and Veral (2019) show two approaches for dealing with no-show behavior in healthcare. For both recommendations, it is required to carefully identify patients groups in the data, as Zebina et al. (2019) explains that this is the key element in identifying the patients that should be targeted with each strategy. This study can correctly identify patient groups, as explained in the results section. For example, identifying patients with the highest probability of no-show behavior and inhabitants of municipalities where no-show behavior is highest. The first approach improves appointment attendance. It can be recommended that the GGD Brabant Zuid-Oost should send out reminders to specific patient groups seven days before an appointment occurs. Previous studies show that this decreases no-show behavior (Wu et al., 2019; Zebina et al., 2019; Lin et al., 2012). Also, the GGD Brabant Zuid-Oost should educate people about the added value and importance of attending appointments, as supported by current literature (Wu et al., 2019; Weaver et al., 2019).

The second approach aims to minimize the operational impact on the organization. Currently, no-shows take away the ability to use resources effectively. In practice, several appointment types in different municipalities got canceled due to staff shortages. Part of this problem is that people are at least one time automatically scheduled for each appointment type. This means that when people are unaware that an appointment has been automatically scheduled, this will result in a no-show. Current literature suggests scheduling appointments more effectively reduces the burden for the organization (Brailsford et al., 2012). This study showed that people with six no-shows are classified as non-attendees, where after missing three appointments, the probability of no-show behavior is already 60%. It is proposed to personally contact all patients who miss three appointments to announce that they are no longer automatically scheduled. Additionally, it would be considering the above be recommended to use 'open access' scheduling (Rose et al., 2011). All patients will be automatically scheduled for the first two appointments, the home visits, and afterward, patients can schedule the appointments themselves. Via this way, people that do not want to attend appointments can no longer (un)consciously block a time slot withholding other patients the ability to attend an appointment. Also, patients with work difficulties and other obliged occupations can more easily book a timeslot that suits their schedule.

5.2 Conclusion

This study answers the main research question as follows. Within the context of Dutch youth healthcare (JGZ) appointments, it can be concluded that several demographics of municipality living and socioeconomic status predict no-show behavior, which the psychological construct of required effort might cause. Also, it can be concluded that several behavioral characteristics such as past appointment behavior predict no-show behavior, which might be caused by behavioral consistency. Last, the behavioral characteristics, appointment number, and appointment type predict no-show behavior, which might be caused by perceived importance. In conclusion, the predictors of no-show behavior found within the context of Dutch youth healthcare (JGZ) appointments are generalizable. In contrast, the psychological constructs causing the effects are understudied and suggest potential reasons why the effect in this study might be found.

Next, logistic regression can be used to most accurately predict no-show behavior in the context of the GGD Brabant Zuid-Oost. In the end, domain experts might accept this method as supporting decision-making processes.

Regarding practical implications, this study recommends sending out reminders to patients and educating people about the importance of attending appointments to improve appointment attendance and use ‘open access‘ scheduling to minimize the operational impact of no-shows. Using this study’s results enables classifying different patient groups that should be targeted with different strategies, i.e. sending reminders and educating people. All practical implications strive to improve healthcare effectiveness in the context of Dutch youth health care (JGZ) appointments.

References

Adler, N. E., & Ostrove, J. M. (1999). Socioeconomic status and health: what we know and what we don't. *Annals of the New York academy of Sciences*, *896*(1), 3–15. 10, 40, 41

Ahmadi-Javid, A., Jalali, Z., & Klassen, K. J. (2017). Outpatient appointment systems in healthcare: A review of optimization studies. *European Journal of Operational Research*, *258*(1), 3–34. 5

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179–211. 12, 20

Alahmari, S., Goldgof, D., Hall, L., Dave, P., Phoulady, H. A., & Mouton, P. (2018). Iterative deep learning based unbiased stereology with human-in-the-loop. In *2018 17th ieee international conference on machine learning and applications (icmla)* (pp. 665–670). 7, 17

Begg, S. (2007). The burden of disease and injury in australia 2003. *http://www. aihw. gov. au/publication-detail/? id= 6442467990&libID= 6442467988*. 10

Brailsford, S. C., Harper, P. R., & Sykes, J. (2012). Incorporating human behaviour in simulation models of screening for breast cancer. *European Journal of Operational Research*, *219*(3), 491–507. 15, 42

Centraal Bureau voor Statistiek: Statline. (n.d.). Retrieved from `https://opendata.cbs.nl/statline/` 20

Chariatte, V., Berchtold, A., Akré, C., Michaud, P. A., & Suris, J. C. (2008, 7). Missed appointments in an outpatient clinic for adolescents, an approach to predict the risk of missing. *Journal of Adolescent Health*, *43*, 38-45. doi: 10.1016/J.JADOHEALTH.2007.12.017 5, 12, 19, 20, 40, 42

Daggy, J., Lawley, M., Willis, D., Thayer, D., Suelzer, C., DeLaurentis, P.-C., . . . Sands, L. (2010). Using no-show modeling to improve clinic performance. *Health informatics journal*, *16*(4), 246–259. 8, 15, 39

Dantas, L. F., Hamacher, S., Cyrino Oliveira, F. L., Barbosa, S. D., & Viegas, F. (2019). Predicting patient no-show behavior: a study in a bariatric clinic. *Obesity surgery*, *29*(1), 40–47. 8, 39

Dunnink, G., & Lijs-Spek, W. (2008). Activiteiten basistakenpakket jeugdgezondheidszorg 0-19 jaar per contactmoment. *RIVM rapport 295001001*. 6

Faisal-Cury, A., Quayle, J., Marques, T., Menezes, P. R., & Matijasevich, A. (2015, 8). The relationship between socioeconomic indicators during pregnancy and gynecological appointment at any time after childbirth. *International Journal for Equity in Health*, *14*. doi: 10.1186/S12939-015-0191-X 5, 10

Fan, G., Deng, Z., Ye, Q., & Wang, B. (2021). Machine learning-based prediction models for patients no-show in online outpatient appointments. *Data Science and Management*, *2*, 45–52. 8

Ferro, D. B., Brailsford, S., Bravo, C., & Smith, H. (2020, 11). Improving healthcare access management by predicting patient no-show behaviour. *Decision Support Systems*, *138*, 113398. doi: 10.1016/J.DSS.2020.113398 5, 14, 42

Fiscella, K., Franks, P., Doescher, M. P., & Saver, B. G. (2002). Disparities in health care by race, ethnicity, and language among the insured: findings from a national sample. *Medical care*, 52–59. 9, 10

Goffman, R. M., Harris, S. L., May, J. H., Milicevic, A. S., Monte, R. J., Myaskovsky, L., . . . Vargas, D. L. (2017). Modeling patient no-show history and predicting future outpatient appointment behavior in the veterans health administration. *Military medicine*, *182*(5-6), e1708–e1714. 5, 11, 12, 19, 40

Goldsmith, J. C. (1980). The health care market: can hospitals survive? *Harvard Business Review*, *58*(5), 100–112. 9

Hu, M., Xu, X., Li, X., & Che, T. (2020). Managing patients' no-show behaviour to improve the sustainability of hospital appointment systems: Exploring the conscious and unconscious determinants of no-show behaviour. *Journal of Cleaner Production*, *269*, 122318. 12, 19

Jennings, N., Boyle, M., Mahawar, K., Balupuri, S., & Small, P. (2013). The relationship of distance from the surgical centre on attendance and weight loss after laparoscopic gastric bypass surgery in the united kingdom. *Clinical Obesity*, *3*(6), 180–184. 8

Kaur, D., Uslu, S., Rittichier, K. J., & Durresi, A. (2022). Trustworthy artificial intelligence: a review. *ACM Computing Surveys (CSUR)*, *55*(2), 1–38. 6

Kelly, C., Hulme, C., Farragher, T., & Clarke, G. (2016). Are differences in travel time or distance to healthcare for adults in global north countries associated with an impact on health outcomes? a systematic review. *BMJ open*, *6*(11), e013059. 10, 12, 20, 39, 40

Kelly, S. R., Loiselle, A. R., Pandey, R., Combes, A., Murphy, C., Kavanagh, H., ... others (2021). Factors associated with non-attendance in the irish national diabetic retinopathy screening programme (indear study report no. 2). *Acta diabetologica*, *58*(5), 643–650. 10

Lin, H., Chen, W., Luo, L., Congdon, N., Zhang, X., Zhong, X., ... others (2012). Effectiveness of a short message reminder in increasing compliance with pediatric cataract treatment: a randomized trial. *Ophthalmology*, *119*(12), 2463–2470. 14, 42

Maadi, M., Akbarzadeh Khorshidi, H., & Aickelin, U. (2021). A review on human–ai interaction in machine learning and insights for medical applications. *International journal of environmental research and public health*, *18*(4), 2121. 7, 17

Margolis, H., & McCabe, P. P. (2006). Improving self-efficacy and motivation: What to do, what to say. *Intervention in school and clinic*, *41*(4), 218–227. 11

Marsland, S. (2011). *Machine learning: an algorithmic perspective.* Chapman and Hall/CRC. 16

McComb, S., Tian, Z., Sands, L., Turkcan, A., Zhang, L., Frazier, S., & Lawley, M. (2017). Cancelled primary care appointments: a prospective cohort study of diabetic patients. *Journal of medical systems*, *41*, 1–8. 11, 12, 19, 20

McQueenie, R., Ellis, D. A., McConnachie, A., Wilson, P., & Williamson, A. E. (2019). Morbidity, mortality and missed appointments in healthcare: a national retrospective data linkage study. *BMC medicine*, *17*(1), 1–9. 5

Meléndez-Ackerman, E. J., Santiago-Bartolomei, R., Vila-Ruiz, C. P., Santiago, L. E., García-Montiel, D., Verdejo-Ortiz, J. C., ... Hernández-Calo, E. (2014). Socioeconomic drivers of yard sustainable practices in a tropical city. *Ecology and Society*, *19*(3). 10, 39

Mikhaeil, J. S., Celo, E., Shanahan, J., Harvey, B., Sipos, B., & Law, M. P. (2019). Attend: A two-pronged trial to eliminate no shows in diagnostic imaging at a community-based hospital. *Journal of Medical Imaging and Radiation Sciences*, *50*(1), 36–42. 5

Millhiser, W. P., & Veral, E. A. (2019). A decision support system for real-time scheduling of multiple patient classes in outpatient services. *Health care management science*, *22*, 180–195. 5, 14, 15, 42

Mitchell, A. J., & Selmes, T. (2007). Why don't patients attend their appointments? maintaining engagement with psychiatric services. *Advances in psychiatric treatment*, *13*(6), 423–434. 10

Mitchell, R. B. (2014). Compliance theory: An overview 1. *Improving compliance with international environmental law*, 3–28. 13

Morera-Guitart, J., Mas-Server, M., & Mas-Sese, G. (2002). Analysis of the patients who missed their appointments at the neurology clinic of the marina alta. *Revista de neurologia*, *34*(8), 701–705. 8, 14, 39, 40

Neal, R. D., Lawlor, D. A., Allgar, V., Colledge, M., Ali, S., Hassey, A., ... Wilson, A. (2001). Missed appointments in general practice: retrospective data analysis from four practices. *British journal of general practice*, *51*(471), 830–832. 5

Ofei-Dodoo, S., Kellerman, R., Hartpence, C., Mills, K., & Manlove, E. (2019). Why patients miss scheduled outpatient appointments at urban academic residency clinics: a qualitative evaluation. *Kansas journal of medicine*, *12*(3), 57. 10

Osisanwo, F., Akinsola, J., Awodele, O., Hinmikaiye, J., Olakanmi, O., Akinjobi, J., et al. (2017). Supervised machine learning algorithms: classification and comparison. *International Journal of Computer Trends and Technology (IJCTT)*, *48*(3), 128–138. 6, 15, 16, 17, 20

Pandey, K. R., Yang, F., Cagney, K. A., Smieliauskas, F., Meltzer, D. O., & Ruhnke, G. W. (2019). The impact of marital status on health care utilization among medicare beneficiaries. *Medicine*, *98*(12). 9

Pell, C., Menaca, A., Were, F., Afrah, N. A., Chatio, S., Manda-Taylor, L., ... others (2013). Factors affecting antenatal care attendance: results from qualitative studies in ghana, kenya and malawi. *PloS one*, *8*(1), e53747. 9

Percac-Lima, S., Cronin, P. R., Ryan, D. P., Chabner, B. A., Daly, E. A., & Kimball, A. B. (2015). Patient navigation based on predictive modeling decreases no-show rates in cancer care. *Cancer*, *121*(10), 1662–1670. 14

Pillai, R., Bhangu, N., Narayanan, M., & Yoong, W. (2012). A demographic study to profile non-attenders at a gynaecology outpatient clinic. *Journal of Obstetrics and Gynaecology*, *32*(2), 156–158. 9, 10

Rabideau, S. T. (2005). *Effects of achievement motivation on behavior.* 11

Rose, K. D., Ross, J. S., & Horwitz, L. I. (2011). Advanced access scheduling outcomes: a systematic review. *Archives of internal medicine*, *171*(13), 1150–1159. 15, 42

Schwebel, F. J., & Larimer, M. E. (2018). Using text message reminders in health care services: a narrative literature review. *Internet interventions*, *13*, 82–104. 5

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable ai. *International Journal of Human-Computer Studies*, *146*, 102551. 17

Shrestha, M. P., Hu, C., & Taleban, S. (2017). Appointment wait time, primary care provider status, and patient demographics are associated with nonattendance at outpatient gastroenterology clinic. *Journal of Clinical Gastroenterology*, *51*, 433-438. doi: 10.1097/MCG.0000000000000706 9, 11

Skinner, E. A., Edge, K., Altman, J., & Sherwood, H. (2003). Searching for the structure of coping: a review and critique of category systems for classifying ways of coping. *Psychological bulletin*, *129*(2), 216. 11

Skinner, E. A., Zimmer-Gembeck, M. J., Connell, J. P., Eccles, J. S., & Wellborn, J. G. (1998). Individual differences and the development of perceived control. *Monographs of the society for Research in Child Development*, i–231. 12

Staddon, J. E., & Cerutti, D. T. (2003). Operant conditioning. *Annual review of psychology*, *54*(1), 115–144. 13

Stormacq, C., Van den Broucke, S., & Wosinski, J. (2019). Does health literacy mediate the relationship between socioeconomic status and health disparities? integrative review. *Health promotion international*, *34*(5), e1–e17. 10

Topuz, K., Uner, H., Oztekin, A., & Yildirim, M. B. (2018). Predicting pediatric clinic no-shows: a decision analytic framework using elastic net and bayesian belief network. *Annals of Operations Research*, *263*, 479–499. 6, 17, 41

Vallée, J., Cadot, E., Grillo, F., Parizot, I., & Chauvin, P. (2010). The combined effects of activity space and neighbourhood of residence on participation in preventive health-care activities: The case of cervical screening in the paris metropolitan area (france). *Health & place*, *16*(5), 838–852. 10

Wallace, D. J., Ray, K. N., Degan, A., Kurland, K., Angus, D. C., & Malinow, A. (2018). Transportation characteristics associated with non-arrivals to paediatric clinic appointments: a retrospective analysis of 51 580 scheduled visits. *BMJ quality & safety*, *27*(6), 437–444. 5

Wammes, J., Jeurissen, P., Westert, G., & Tanke, M. (2020). The dutch health care system. *International profiles of health care systems*, 137. 6

Weaver, K. R., Talley, M., Mullins, M., & Selleck, C. (2019). Evaluating patient navigation to improve first appointment no-show rates in uninsured patients with diabetes. *Journal of Community Health Nursing*, *36*(1), 11–18. 5, 14, 42

Wu, Y., Liang, Y., Zhou, Q., Liu, H., Lin, G., Cai, W., ... Gu, J. (2019). Effectiveness of a short message service intervention to motivate people with positive results in preliminary colorectal cancer screening to undergo colonoscopy: a randomized controlled trial. *Cancer*, *125*(13), 2252–2261. 14, 42

Zebina, M., Melot, B., Binachon, B., Ouissa, R., Lamaury, I., & Hoen, B. (2019). Impact of an sms reminder service on outpatient clinic attendance rates by patients with hiv followed-up at pointe-à-pitre university hospital. *Patient preference and adherence*, 215–221. 5, 14, 15, 42

Appendix A

Appendix A: Query

Query used to retrieve data from the Data Warehouse (DWH) of the GGD.

```
/****** Script for SelectTopNRows command from SSMS  ******/
WITH afspraken AS (
SELECT *
                    ,DATEPART(HOUR, Afspraakdatum) as uur
                    ,case when afspraakstatus = 'Afgezegd op verzoek client, laat bericht' or afspraakstatus = 'Afgezegd op
verzoek client, tijdig bericht' or afspraakstatus = 'Niet verschenen' then 'niet verschenen' else 'nvt' end as
niet_verschenen
                    ,case when afspraakstatus = 'verschenen' then 'ja' else 'nee' end as verschenen
                    ,case when afspraakstatus = 'Afgezegd op verzoek client, laat bericht'  or Afspraakstatus = 'Onbekend'
then '1' else '0' end as niet_verschenen_zonder_bericht
  FROM [dsc].DM_Afspraken
  where datum >= '2021-06-01' and datum <= '2022-05-31'
                    and rijnummer =1
                    and afspraakstatus <> 'Verzet door JGZ'
                    and afspraakstatus <> 'gepland'
                    and afspraakstatus <> 'onbekend'
                    and BZOJaNEE <> '0'
                    and (BDS_Naam = 'Huisbezoek 4-7 dagen'
                      or BDS_Naam = 'Huisbezoek 2 weken'
                      or BDS_Naam = 'Contactmoment 4 weken'
                      or BDS_Naam = 'Contactmoment 8 weken'
                      or BDS_Naam = 'Contactmoment 3 maanden'
                      or BDS_Naam = 'Contactmoment 4 maanden'
                      or BDS_Naam = 'Contactmoment 6 maanden'
                      or BDS_Naam = 'Contactmoment 9 maanden'
                      or BDS_Naam = 'Contactmoment 11 maanden'
                      or BDS_Naam = 'Contactmoment 14 maanden'
                      or BDS_Naam = 'Contactmoment 18 maanden'
                      or BDS_Naam = 'Contactmoment 2 jaar'
                      or BDS_Naam = 'Contactmoment 3 jaar'
                      or BDS_Naam = 'Contactmoment 3,9 jaar'
                      )
                    and (WijzeVerzetten <> 'nvt' or WijzeVerzetten is NULL)
  )

 , aantalafsprakengehad as (
  select Clientnummer
                    ,datum as datumhuidigeafspraak
                    ,afspraakstatus
                    ,count(*) as aantal
                    ,ROW_NUMBER()over(Partition by Clientnummer order by datum) as afspraaknummer
                    ,dense_rank()over(Partition by Clientnummer,sum(case when Afspraakstatus = 'Niet verschenen' or
Afspraakstatus = 'Afgezegd op verzoek client, laat bericht' or Afspraakstatus = 'Afgezegd op verzoek client, tijdig
bericht' then 0 else 1 end) order by datum, afspraakstatus) as verschenen
  from [dsc].DM_Afspraken
  where  afspraakstatus <> 'Verzet door JGZ'
                    and afspraakstatus <> 'gepland'
                    and afspraakstatus <> 'onbekend'
                    and BZOJaNEE <> '0'
                    and (BDS_Naam = 'Huisbezoek 4-7 dagen'
                      or BDS_Naam = 'Huisbezoek 2 weken'
                      or BDS_Naam = 'Contactmoment 4 weken'
                      or BDS_Naam = 'Contactmoment 8 weken'
                      or BDS_Naam = 'Contactmoment 3 maanden'
                      or BDS_Naam = 'Contactmoment 4 maanden'
                      or BDS_Naam = 'Contactmoment 6 maanden'
                      or BDS_Naam = 'Contactmoment 9 maanden'
                      or BDS_Naam = 'Contactmoment 11 maanden'
                      or BDS_Naam = 'Contactmoment 14 maanden'
                      or BDS_Naam = 'Contactmoment 18 maanden'
                      or BDS_Naam = 'Contactmoment 2 jaar'
                      or BDS_Naam = 'Contactmoment 3 jaar'
                      or BDS_Naam = 'Contactmoment 3,9 jaar'
                      )
                    and (WijzeVerzetten <> 'nvt' or WijzeVerzetten is NULL)
         group by clientnummer
                    ,datum
                    ,afspraakstatus
  )

  , tijdtussen as (
  select Clientnummer
                    ,datum as datumvorigeafspraak
                    ,ROW_NUMBER()over(Partition by Clientnummer order by datum) as afspraaknummer1
  from [dsc].DM_Afspraken
  group by clientnummer
                    ,datum
  )
```

Figure A.1: Query: Part I

```
  , woa as (
  select SES_WOA
            ,[WK_NAAM]
       ,[vermogen]
       ,[werkzaam_afl_4jaar]
       ,[GestandaardiseerdInkomen]
       ,[opleiding_hoog]
       ,[opleiding_middel]
       ,[opleiding_laag]
  FROM [dsc].[Wijk_WOA_2019]
  )

  , interventie as (
  SELECT [Clientnummer]
       ,Datum
            ,sum(case when Indicatie is not null or AdviesEnVerwijzingNaar is not null or Interventie is not null  then 1
else 0 end) as interventie_ja_nee
  FROM [dsc].[Interventies]

  group by Datum
                  ,Clientnummer
)

  select afspraken.[Clientnummer]
          ,afspraken.[Geslacht]
          ,afspraken.[Leeftijd]
          ,afspraken.[Nationaliteit]
          ,afspraken.[Gemeente]
          ,afspraken.[Buurt]
          ,afspraken.[Datum]
          ,afspraken.[uur]
          ,afspraken.[BDS_Naam]
          ,afspraken.[Afspraakstatus]
          ,afspraken.[WijzeVerzetten]
          ,afspraken.[AantDgnVoorafAfgezegd]
          ,afspraken.[rijnummer]
          ,aantalafsprakengehad.afspraaknummer
          ,aantalafsprakengehad.verschenen
          ,aantalafsprakengehad.afspaaknummer - aantalafsprakengehad.verschenen as niet_verschenen
          ,afspraken.niet_verschenen_zonder_bericht
          ,datediff(day, tijdtussen.datumvorigeafspraak, aantalafsprakengehad.datumhuidigeafspraak) as tussentijd
                  ,woa.SES_WOA
                  ,woa.spreiding
                  ,woa.opleiding_laag
                  ,interventie.interventie_ja_nee
  from afspraken
                  left join aantalafsprakengehad on afspraken.Clientnummer = aantalafsprakengehad.Clientnummer and
afspraken.Datum =aantalafsprakengehad.datumhuidigeafspraak
                  left join tijdtussen on aantalafsprakengehad.afspraaknummer = tijdtussen.afspraaknummer1+1 and
tijdtussen.Clientnummer = aantalafsprakengehad.clientnummer
                  left join dsc.buurt_WOA_2019 woa on cast(case when Buurtcode = 'Onbekend' then 0 else Buurtcode end as
int) = woa.RegiocodeGemeenteWijkBuurt_2
                  left join interventie on interventie.Clientnummer = afspraken.Clientnummer and interventie.Datum <
aantalafsprakengehad.datumhuidigeafspraak
  order by afspraken.clientnummer, afspraken.datum, afspraken.bds_naam
```

Figure A.2: Query: Part II

Appendix B

Appendix B: Data Output

B.1 Data exploration

This section of the appendix B contains the codebook for all variables. This codebook describes each variable, the variable type, the type of observations it holds, and the variable's potential values.

Variable	*Type*	*Description or dummy*	*Frequency or range*
Client number	Number	30319 unique numbers	
Gender	Character	Male or Female	M:51823, F:49970
Age	Number	Age between 0-4	0: 63743 1: 14110 2: 10875 3: 11421 4: 1635
Municipality	Factor - 21 levels	Municipalities part of GGD Brabant Zuid-Oost	
		Eindhoven	30545
		Asten	2376
		Bergeijk	1741
		Best	4244
		Bladel	2263
		Cranendonck	2032
		Deurne	4262
		Eersel	2412
		Gelderop-Mierlo	5594
		Gemenrt-Bakel	4357
		Heeze-Leende	1852
		Helmond	13279
		Laarbeek	3198
		Nuenen, Gerwen en Nederwetten	3114
		Oirschot	2199
		Reusel-De Mierden	1786
		Someren	3048
		Son en Breugel	2460
		Valkenswaard	3630
		Veldhoven	5208
		Waalre	2193
Appointment type	Factor - 14 levels	Appointment type	
		Appointment 4 weeks	8324
		Appointment 11 months	10571
		Appointment 14 months	10490
		Appointment 18 months	732

Table B.1: Codebook: Dataset Variable (1)

Variable	*Type*	*Description or dummy*	*Frequency or range*
Appointment type		Appointment 2 year	10562
		Appointment 3 year	585
		Appointment 3 months	9500
		Appointment 3.9 year	12994
		Appointment 4 months	9516
		Appointment 6 months	6649
		Appointment 8 weeks	7072
		Appointment 9 months	743
		Home visit 2 weeks	7664
		Home visit 4-7 days	6391
Appointment status	Factor - 4 levels	Appointment status	
		Not attending, late notice	6626
		Not attending, notice in time	11617
		Not attending	5021
		Attending	78529
Appointment number	Number	Clients appointment n	1-28
Not_attending	Number	Clients amount of no-show	0-19
Not_attending_problematic	Factor - 2 levels	Client amount of problematic no-show	0: 90146, 1:11647
Time between	Number	Time between consecutive appointments in days	0-996
SES_WOA	Number	socioeconomic status*	Between -1 and 1
Spread	Number	How well the socioeconomic status is spread*	Between -1 and 1
Private capital	Number	Average personal capital*	Between -1 and 1
Employed_last_4_years	Number	% of people employed*	0-100
Standardized Income	Number	Amount x1000 people earned*	21 - 87.2
Education_high	Number	% of people with masters degree*	0-100
Education_middle	Number	% of people with bachelor degree*	0-100
Education_low	Number	% of people with practical degree*	0-100
intervention_yes_no	Number	If the client had interventions before before	0-1
Dutch_yes_no	Factor - 2 levels	Having the Dutch nationality yes or no	1: 89946, 0: 10847
Intervention_yes_no	Factor - 2 levels	Have had an intervention before yes or no	1: 23230, 0: 78563
Time_morning_afternoon_other	Factor - 3 levels	Time of day	
		morning	72478
		afternoon	27989
		other	1326
Municipality_Ehv_Hlmd_yes_no	Factor - 2 levels	Living in urban or rural area	rural: 57969, urban: 43824
* = at neighbourhood level			

Table B.2: Codebook: Dataset Variable (2)

B.2 Understanding

B.2.1 Model 1: General Model

The first model includes all variables and is made using the following command:

Model1 <- glm(not_attending_problematic = Gender + Age + Municipality + time_morning_afternoon_other + Appointment type + Appointment_number + not_attending + time between + SES_WOA + Spread + private capital + employed_last_4_years + Standardized Income + Education_high + Education_middle + Education_low + Dutch_yes_no + intervention_yes_no, data=data10, family = "binomial"(link="logit"))

In table B.3 the output of model 1 can be found. For the first the model the accuracy, precision, recall, f1-score and confusion matrix can be found in table B.5.

B.2.2 Model 2: Correlation / multi-collinearity

The second model is equal to the first model, but checks for correlation and multi-collinearity. The second model corrects for multi-collinearity by checking the variables in a correlation matrix using peasron-method with listwise-deletion. The output of the correlation matrix can be found in figure B.1.

	Leeftijd	afspraaknummer	SES_WOA	spreiding	vermogen	werkzaam_afl_4jaar	GestandaardiseerdInkomen	opleiding_hoog	opleiding_middel	opleiding_laag	interventie_hoeveelheid
Leeftijd		0.690***	0.002	-0.008*	0.011***	0.002	0.006	-0.014***	0.012***	0.012***	0.323***
afspraaknummer	0.690***		-0.002	-0.003	0.008*	-0.001	0.003	-0.014***	0.010**	0.015***	0.346***
SES_WOA	0.002	-0.002		-0.912***	0.883***	0.807***	0.985***	0.699***	-0.296***	-0.854***	0.011***
spreiding	-0.008*	-0.003	-0.912***		-0.864***	-0.733***	-0.903***	-0.480***	0.090***	0.676***	-0.027***
vermogen	0.011***	0.008*	0.883***	-0.864***		0.571***	0.890***	0.425***	-0.066***	-0.609***	0.021***
werkzaam_afl_4jaar	0.002	-0.001	0.807***	-0.733***	0.571***		0.757***	0.514***	-0.102***	-0.719***	0.013***
GestandaardiseerdInkomen	0.006	0.003	0.985***	-0.903***	0.890***	0.757***		0.698***	-0.340***	-0.818***	0.015***
opleiding_hoog	-0.014***	-0.014***	0.699***	-0.480***	0.425***	0.514***	0.698***		-0.833***	-0.901***	-0.033***
opleiding_middel	0.012***	0.010**	-0.296***	0.090***	-0.066***	-0.102***	-0.340***	-0.833***		0.511***	0.043***
opleiding_laag	0.012***	0.015***	-0.854***	0.676***	-0.609***	-0.719***	-0.818***	-0.901***	0.511***		0.018***
interventie_hoeveelheid	0.323***	0.346***	0.011***	-0.027***	0.021***	0.013***	0.015***	-0.033***	0.043***	0.018***	

Computed correlation used pearson-method with listwise-deletion.

Figure B.1: Correlation Matrix

Variables with a correlation larger than 0.8 and marked significant are removed from the model. The second model is made using the following command:

The second model is also tested for the variance inflation factor (VIF). The values are summarized in table B.6

B.2.3 Model 3: Including interaction effects

The first model includes all variables and is made using the following command:
model3a <- glm(not_attending_problematic = Gender + Municipality + Appointment type + Appointment_number + SES_WOA + Education_middle + Appointment_number*not_attending + Appointment type*not_attending + SES_WOA*time_morning_afternoon_other + employed_last_4_years*time_morning_afternoon_other, data=data10, family = "binomial")

In table B.7 the output of model 1 can be found. For the first the model the accuracy, precision, recall, f1-score and confusion matrix can be found in table B.9.

B.2.4 Model 4: Random Intercept Model

The forth model includes all variables and is made using the following command:
model4 <- glmer(not_attending_problematic = Municipality + Appointment type + Appointment_number + not_attending + SES_WOA + (1 | Client number), data=data10, family = "binomial")

Due to the complexity of the model, which had 30319 levels of client numbers, it was unable to execute.

B.2.5 Model 5: Random Slope Model

The fifth model includes all variables and is made using the following command:
model5 <- glmer(not_attending_problematic = Municipality + Appointment type + Appointment_number + not_attending + SES_WOA + (1 + not_attending | Client number), data=data10, family = "binomial", control = lmerControl(optimizer ="Nelder_Mead"))

Due to the complexity of the model, which had 30319 levels of client numbers and 19 levels of previous no-show behavior, it was unable to execute.

B.2.6 Final Model

The final model includes all variables and is made using the following command:
logit <- glm(not_attending_problematic = Gender + Municipality + Appointment type + Appointment_number + not_attending + Dutch_yes_no + Intervention_yes_no + SES_WOA + Education_middle + Appointment_number*not_attending + Appointment type*not_attending,data = train.df,family = 'binomial')
In table 4.1 the output of model 1 can be found.
The marginsplots for all significant predictors and interaction effects can be found in figure B.2.

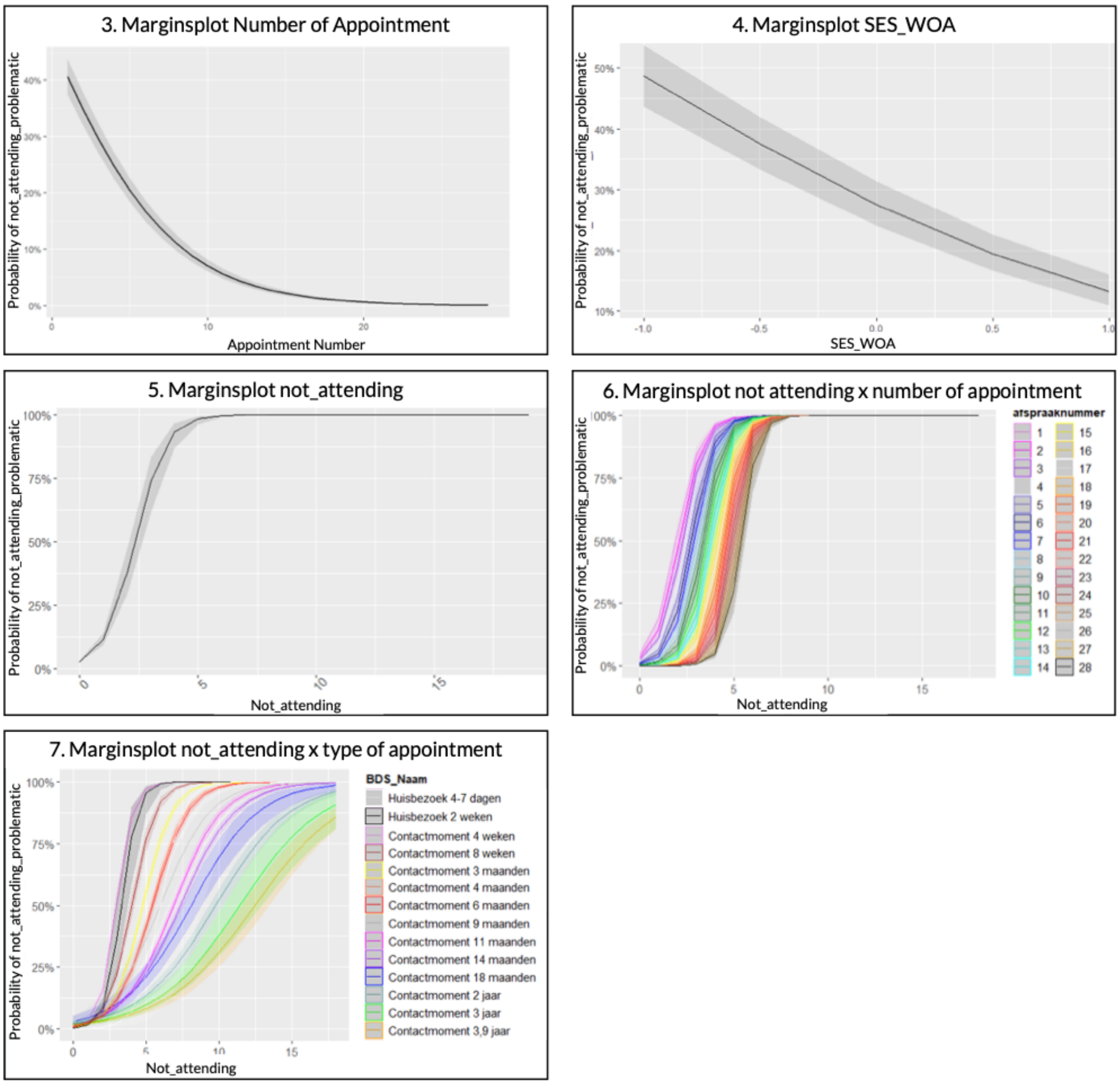


Figure B.2: Marginsplots Significant Predictors and Interaction Effects

	not_attending_problematic		
Predictors	*OR*	*CI*	*p*
(Intercept)	0.00	0.00-0.00	**0.007**
Gender [Female]	1.00	0.95-1.04	0.834
Age	1.71	1.54-1.90	<**0.001**
Municipality [Asten]	0.78	0.64-0.95	**0.013**
Municipality [Bergeijk]	0.83	0.66-1.05	0.124
Municipality [Best]	0.93	0.81-1.06	0.263
Municipality [Bladel]	0.93	0.76-1.14	0.513
Municipality [Cranendonck]	0.98	0.81-1.18	0.818
Municipality [Deurne]	0.64	0.54-0.75	<**0.001**
Municipality [Eersel]	0.70	0.57-0.86	**0.001**
Municipality [Geldrop-Mierlo]	0.81	0.71-0.92	**0.001**
Municipality [Gemert-Bakel]	0.70	0.59-0.82	<**0.001**
Municipality [Heeze-Leende]	0.89	0.72-1.08	0.242
Municipality [Helmond]	0.92	0.82-1.02	0.105
Municipality [Laarbeek]	0.62	0.52-0.74	<**0.001**
Municipality [Nuenen, Gerwen en Nederwetten]	0.79	0.68-0.93	**0.005**
Municipality [Oirshot]	0.80	0.65-0.99	**0.037**
Municipality [Reusel-De Mierden]	0.92	0.73-1.15	0.449
Municipality [Someren]	0.70	0.58-0.84	<**0.001**
Municipality [Son en Breugel]	1.05	0.89-1.25	0.561
Municipality [Valkenswaard]	1.08	0.94-1.25	0.283
Municipality [Veldhoven]	0.83	0.72-0.95	**0.008**
Municipality [Waalre]	0.77	0.63-0.93	**0.006**
Time morning afternoon other [afternoon]	1.03	0.98-1.09	0.220
Time morning afternoon other [other]	0.75	0.51-1.09	0.142
Appointment type [Appointment 11 months]	0.82	0.72-0.93	**0.002**
Appointment type [Appointment 14 months]	0.68	0.58-0.80	<**0.001**
Appointment type [Appointment 18 months]	0.57	0.42-0.76	<**0.001**
Appointment type [Appointment 2 year]	0.26	0.21-0.34	<**0.001**
Appointment type [Appointment 3 year]	0.20	0.14-0.29	<**0.001**
Appointment type [Appointment 3 months]	1.12	0.99-1.26	0.067
Appointment type [Appointment 3.9 year]	0.11	0.08-0.16	<**0.001**
Appointment type [Appointment 4 months]	1.21	1.07-1.36	**0.002**
Appointment type [Appointment 6 months]	1.58	1.39-1.80	<**0.001**
Appointment type [Appointment 8 months]	1.15	1.02-1.31	**0.025**
Appointment type [Appointment 9 months]	1.44	1.10-1.87	**0.007**
Appointment type [Home visit 2 weeks]	0.21	0.17-0.25	<**0.001**
Appointment type [Home visit 4-7 days]	0.04	0.02-0.06	<**0.001**
Appointment number	0.77	0.76-0.78	<**0.001**
Not attending	1.84	1.81-1.86	<**0.001**
Time between	1.00	1.00-1.00	<**0.001**
SES WOA	0.17	0.07-0.42	<**0.001**
Spread	1.17	0.74-1.84	0.487
Private capital	0.98	0.98-0.99	<**0.001**
Employed last 4 years	1.00	0.99-1.00	0.652
Standardized Income	1.05	1.03-1.08	<**0.001**
Education high	1.85	1.14-3.00	**0.013**
Education middle	1.87	1.16-3.04	**0.011**

Table B.3: Model 1: Data output (1)

(continuation)	**not_attending_problematic**		
Predictors	*OR*	*CI*	*p*
Education low	1.87	1.15-3.02	**0.011**
Dutch yes no [1]	1.06	0.98-1.14	0.141
intervention yes no [1]	0.98	0.92-1.03	0.397
Observation	101793		
R^2 Tjur	0.268		

Table B.4: Model 1: Data output (2)

Model	*Accuracy*	*Precision*	*Recall*	*F1-score*	*Confusion matrix*		
1	0.88	0.96	0.91	0.93		0	1
					0	88222	9019
					1	3628	2799

Table B.5: Accuracy, precision, recall, f1-score and confusion matrix for model 1

Variable	*(G)VIF*	*Df*
Gender	1.00	1
Age	27.36	1
Municipality	1.49	20
Time morning afternoon other	1.08	2
Appointment type	45.96	13
Appointment number	4.12	1
Not attending	3.34	1
Time between two appointments	1.21	1
SES WOA	131.09	1
Spread	9.08	1
Privat capital	20.89	1
Employed last 4 years	7.29	1
Education high	66.53e3	1
Education middle	15.69e3	1
Education low	29.76e3	1
Dutch yes no	1.00	1
Intervention yes no	1.19	1

Table B.6: Variance Inflation Facotrs (VIF) per variable

	not_attending_problematic		
Predictors	*OR*	*CI*	*p*
(Intercept)	0.04	0.03-0.06	<**0.001**
Gender [Female]	1.00	0.96-1.05	0.862
Municipality [Asten]	0.74	0.61-0.89	**0.001**
Municipality [Bergeijk]	0.89	0.71-1.12	0.332
Municipality [Best]	1.06	0.93-1.21	0.387
Municipality [Bladel]	0.95	0.78-1.16	0.630
Municipality [Cranendonck]	1.11	0.92-1.33	0.269
Municipality [Deurne]	0.79	0.67-0.92	**0.003**
Municipality [Eersel]	0.78	0.64-0.96	**0.018**
Municipality [Geldrop-Mierlo]	0.92	0.92-1.04	0.189
Municipality [Gemert-Bakel]	0.76	0.65-0.88	<**0.001**
Municipality [Heeze-Leende]	1.04	0.85-1.26	0.725
Municipality [Helmond]	1.09	1.00-1.19	**0.050**
Municipality [Laarbeek]	0.73	0.62-0.86	<**0.001**
Municipality [Nuenen, Gerwen en Nederwetten]	0.95	0.81-1.11	0.518
Municipality [Oirschot]	0.82	0.67-1.00	0.051
Municipality [Reusel-De Mieren]	0.93	0.75-1.15	0.500
Municipality [Someren]	0.70	0.59-0.83	<**0.001**
Municipality [Son en Breugel]	1.24	1.05-1.46	**0.010**
Municipality [Valkenswaard]	1.17	1.01-1.35	**0.032**
Municipality [Veldhoven]	0.90	0.78-1.02	0.111
Municipality [Waalre]	0.93	0.77-1.11	0.400
Appointment type [Home Visit 4-7 days]	0.09	0.05-0.14	<**0.001**
Appointment type [Home Visit 2 weeks]	0.48	0.37-0.62	<**0.001**
Appointment type [Appointment 8 weeks]	1.13	0.88-1.46	0.326
Appointment type [Appointment 3 months]	1.31	1.03-1.67	**0.028**
Appointment type [Appointment 4 months]	1.62	1.25-2.09	<**0.001**
Appointment type [Appointment 6 months]	1.99	1.48-2.66	<**0.001**
Appointment type [Appointment 9 months]	4.85	2.66-8.30	<**0.001**
Appointment type [Appointment 11 months]	3.51	2.75-4.49	<**0.001**
Appointment type [Appointment 14 months]	4.98	3.91-4.64	<**0.001**
Appointment type [Appointment 18 months]	7.39	4.05-12.77	<**0.001**
Appointment type [Appointment 2 year]	7.80	6.22-9.78	<**0.001**
Appointment type [Appointment 3 year]	8.84	5.19-9.50	<**0.001**
Appointment type [Appointment 3.9 year]	10.04	8.19-12.34	<**0.001**
Appointment number	0.73	0.71-0.74	<**0.001**
SES WOA	0.39	0.31-0.49	<**0.001**
Education middle	1.00	1.00-1.01	0.296
Dutch yes no [1]	1.07	0.99-1.15	0.098
Intervention yes no [1]	0.99	0.94-1.05	0.736
Not attending	4.82	4.40-5.30	<**0.001**
Time morning afternoon other [afternoon]	0.75	0.47-1.19	0.227
Time morning afternoon other [other]	0.19	0.01-4.97	0.338
Employed last 4 years	1.00	1.00-1.01	0.732
Appointment number x not attending	1.02	1.02-1.02	<**0.001**
Appointment type [Home Visit 4-7 days] x not attending	0.69	0.45-1.03	0.070
Appointment type [Home Visit 2 weeks] x not attending	1.07	0.85-1.36	0.567

Table B.7: Model 3: Data output (1)

	not_attending_problematic		
Predictors	*OR*	*CI*	*p*
Appointment type [Appointment 8 weeks] x not attending	0.61	0.55-0.69	<**0.001**
Appointment type [Appointment 3 months] x not attending	0.48	0.43-0.53	<**0.001**
Appointment type [Appointment 4 months] x not attending	0.42	0.37-0.46	<**0.001**
Appointment type [Appointment 6 months] x not attending	0.40	0.36-0.45	<**0.001**
Appointment type [Appointment 9 months] x not attending	0.32	0.28-0.37	<**0.001**
Appointment type [Appointment 11 months] x not attending	0.31	0.28-0.35	<**0.001**
Appointment type [Appointment 14 months] x not attending	0.30	0.27-0.33	<**0.001**
Appointment type [Appointment 18 months] x not attending	0.27	0.24-0.31	<**0.001**
Appointment type [Appointment 2 year] x not attending	0.26	0.24-0.29	<**0.001**
Appointment type [Appointment 3 year] x not attending	0.25	0.22-0.27	<**0.001**
Appointment type [Appointment 3.9 year] x not attending	0.24	0.22-0.27	<**0.001**
SES WOA x time morning afternoon other [afternoon]	0.92	0.66-1.28	0.626
SES WOA x time morning afternoon other [other]	1.31	0.12-13.73	0.825
Time morning afternoon other [afternoon] x employed last 4 years	1.01	1.00-1.01	0.185
Time morning afternoon other [other] x employed last 4 years	1.02	0.96-1.09	0.460
Observation	101579		
R^2 Tjur	0.288		

Table B.8: Model 3: Data output (2)

Model	*Accuracy*	*Precision*	*Recall*	*F1-score*	*Confusion matrix*		
3	0.87	0.96	0.90	0.93		0	1
					0	86524	9060
					1	3664	2545

Table B.9: Accuracy, precision, recall, f1-score and confusion matrix for model 3

B.3 Most accurate classification

B.3.1 Logistic Regression

The outcome of the logistic regression model is summarized in tables B.10 and B.11.

	not_attending_problematic		
Predictors	*OR*	*CI*	*p*
(Intercept)	0.05	0.04-0.06	<**0.001**
Municipality [Asten]	0.83	0.69-0.99	**0.046**
Municipality [Bergeijk]	0.96	0.76-1.22	0.807
Municipality [Best]	1.15	0.95-1.32	**0.044**
Municipality [Bladel]	1.06	0.87-1.28	0.960
Municipality [Cranendonck]	1.21	1.00-1.46	**0.046**
Municipality [Deurne]	0.85	0.73-0.99	**0.032**
Municipality [Eersel]	0.84	0.68-1.03	0.091
Municipality [Geldrop-Mierlo]	0.94	0.83-1.06	0.307
Municipality [Gemert-Bakel]	0.83	0.72-0.95	**0.006**
Municipality [Heeze-Leende]	1.12	1.09-1.38	0.299
Municipality [Helmond]	1.14	1.05-1.23	**0.002**
Municipality [Laarbeek]	0.73	0.62-0.86	<**0.001**
Municipality [Nuenen, Gerwen en Nederwetten]	0.99	0.84-1.17	0.917
Municipality [Oirshot]	0.88	0.72-1.07	0.219
Municipality [Reusel-De Mierden]	0.97	0.78-1.21	0.817
Municipality [Someren]	0.75	0.63-0.88	<**0.001**
Municipality [Son en Breugel]	1.29	0.97-1.38	0.110
Municipality [Valkenswaard]	1.21	0.99-1.32	0.060
Municipality [Veldhoven]	0.94	0.78-1.04	0.404
Municipality [Waalre]	0.89	0.80-1.17	0.264
Appointment type [Appointment 11 months]	2.90	2.20-3.82	<**0.001**
Appointment type [Appointment 14 months]	3.04	2.28-4.05	<**0.001**
Appointment type [Appointment 18 months]	4.03	2.00-7.55	<**0.001**
Appointment type [Appointment 2 year]	2.78	1.99-3.88	<**0.001**
Appointment type [Appointment 3 year]	2.66	1.39-4.88	**0.002**
Appointment type [Appointment 3 months]	1.23	0.94-1.61	0.126
Appointment type [Appointment 3.9 year]	2.03	1.33-3.10	<**0.001**
Appointment type [Appointment 4 months]	1.58	1.19-2.10	**0.001**
Appointment type [Appointment 6 months]	1.96	1.41-2.71	<**0.001**
Appointment type [Appointment 8 weeks]	1.06	0.80-1.40	0.216
Appointment type [Appointment 9 months]	5.20	2.75-9.20	<**0.001**
Appointment type [Home visit 2 weeks]	0.43	0.32-0.57	<**0.001**
Appointment type [Home visit 4-7 days]	0.08	0.04-0.13	<**0.001**
Appointment number	0.72	0.71-0.73	<**0.001**
Not attending	4.84	4.30-5.27	<**0.001**
SES WOA	0.39	0.35-0.44	<**0.001**
Appointment number x not attending	1.02	1.02-1.03	<**0.001**
Appointment type [Appointment 11 months] x not attending	0.32	0.29-0.36	<**0.001**

Table B.10: Logistic Model: Data output (1)

	not_attending_problematic		
Predictors	*OR*	*CI*	*p*
Appointment type [Appointment 14 months] x not attending	0.30	0.23-0.33	<**0.001**
Appointment type [Appointment 18 months] x not attending	0.28	0.24-0.31	<**0.001**
Appointment type [Appointment 2 year] x not attending	0.27	0.23-0.29	<**0.001**
Appointment type [Appointment 3 year] x not attending	0.25	0.22-0.28	<**0.001**
Appointment type [Appointment 3 months] x not attending	0.49	0.44-0.55	<**0.001**
Appointment type [Appointment 3.9 year] x not attending	0.25	0.22-0.27	<**0.001**
Appointment type [Appointment 4 months] x not attending	0.42	0.36-0.46	<**0.001**
Appointment type [Appointment 6 months] x not attending	0.41	0.36-0.46	<**0.001**
Appointment type [Appointment 8 months] x not attending	0.61	0.53-0.70	<**0.001**
Appointment type [Appointment 9 months] x not attending	0.32	0.28-0.37	<**0.001**
Appointment type [Home visit 2 weeks] x not attending	1.02	0.79-1.32	0.634
Appointment type [Home visit 4-7 days] x not attending	0.62	0.37-0.95	0.068
Observation	101793		
R^2 Tjur	0.289		

Table B.11: Logistic Model: Data output (2)

B.3.2 Random Forest

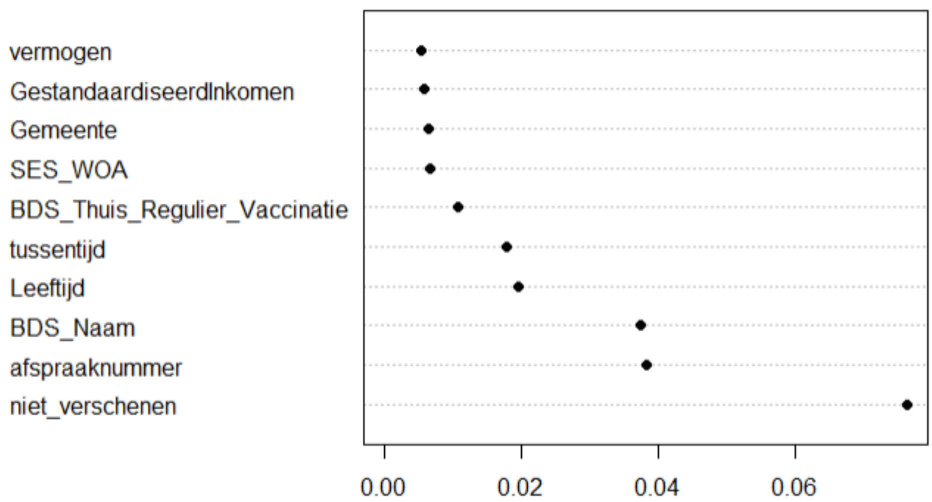


Figure B.3: Random Forest Output: *Significant predictors*

B.3.3 K-nearest Neighbour (KNN)

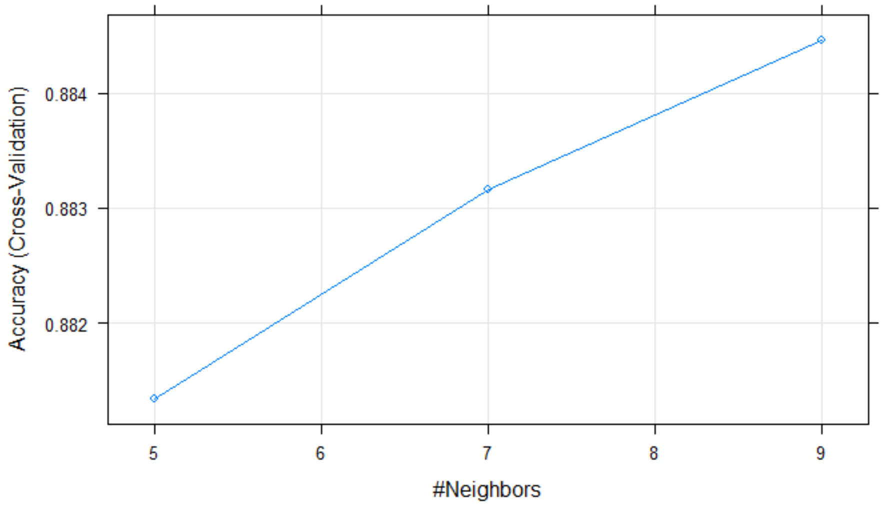


Figure B.4: K-Nearest Neighbour: *Output*

B.4 Human-technology interaction

B.4.1 Classification Accuracy

	Urban		*Rural*	
Accuracy	0.81		0.84	
Sensitivity	0.99		0.98	
F1-score	0.63		0.65	
Confusion Matrix	0	1	0	1
0	23013	49	32223	94
1	6437	5561	7294	6765

	Low income		*High income*	
Accuracy	0.81		0.84	
Sensitivity	0.99		0.99	
F1-score	0.64		0.63	
Confusion Matrix	0	1	0	1
0	19546	33	36050	67
1	5601	4915	8107	7116

Table B.12: Classification Accuracy Metrics

B.4.2 Fairness metrics

	Urban	*vs*	*Rural*	
Equal opportunity difference	(TP2-TP1)/TP1		**-0.0049**	
Statistical parity difference				
Prob Confusion Matrix	0	1	0	1
0	65.64%	0.14%	70.21%	0.14%
1	18.36%	15.85%	15.79%	13.86%
	gives	TN2-TN1	**-0.0050**	
Average odds difference	(FP1-FP2)+(FN1-FN2)		**-0.0026**	

	Low income	*vs*	*High income*	
Equal opportunity difference	(TP2-TP1)/TP1		**-0.0095**	
Statistical parity difference				
Prob Confusion Matrix	0	1	0	1
0	64.95%	0.11%	66.20%	0.73%
1	18.61%	16.33%	19.81%	13.25%
	gives	TN2-TN1	**-0.031**	
Average odds difference	(FP1-FP2)+(FN1-FN2)		**-0.018**	

Table B.13: Fairness Metrics

PO Box 513
5600 MB Eindhoven
The Netherlands
tue.nl

DEPARTMENT OF INDUSTRIAL ENGINEERING & INNOVATION SCIENCES

EINDHOVEN UNIVERSITY OF TECHNOLOGY