Mechanisms for successful implementation of AI-based medical devices in Dutch hospitals, an exploratory study

By

Iris Overgaauw

s1794329

i.j.overgaauw@student.utwente.nl

Submitted in partial fulfillment of the requirements for the degree of Master of Science, Public Administration, University of Twente

2021-2022

Supervisors:

Prof. dr. R. Torenvlied, University of Twente

dr. M.A. Lindenberg, University of Twente

UNIVERSITY OF TWENTE.

CONTENTS

1 INTRODUCTION

Artificial Intelligence(AI) is on the top of the research agenda within different institutions and companies. The technology has even caught the attention of the European Union(EU). In response to growing advancements in AI-technology, the EU has set up a committee, the European AI Alliance, to promote research and development of AI-based devices [Commission, 2018]. A specific aim of the European AI Alliance is to guide European health institutions and companies during research and development of AI-based medical devices [Commission, 2018]. The AI Alliance strives to change EU legislation and collect funds to facilitate proper research in order to turn Europe into "the global hub for trustworthy AI" [Commission, 2018]. In response to the European AI Alliance, the Dutch Government created the Strategic Action Plan of Artificial Intelligence [Rijksoverheid, 2019]. In this policy document, the Dutch Government expresses the aim to reach the same goals as the European AI Alliance, but on a national level via an entity called the Dutch AI Coalition. This Coalition advocates within the Ministry of Finance for the reservation of funds and subsidies for AI research in the health sector [Rijksoverheid, 2019].

One of the reasons why AI-technology reached such a prominent place on the agenda of both the EU and the Netherlands, is due to its potential to change, for example, healthcare as we know it. With the use of AI-technology, tasks are standardized and automated [Shaw et al., 2019, Tartar et al., 2021]. Task automation means the task is not executed by humans anymore, which causes a decrease in the influence of human errors [van den Berge and Mamede, 2013, Finn et al., 2014]. To illustrate, humans are influenced by their surroundings and perform less consistent than an AI-based medical device. Therefore, task automation, by AI-based medical devices, increases the accuracy of a diagnostic intervention, for example.

AI-technology allows the analysis of enormous amounts of data. The results of this analysis are immediately deployed by the algorithm of the AI-based medical device. AI contains a collection of techniques, from which machine-learning and deep-learning are frequently used in healthcare applications [Yin et al., 2021, Gerke et al., 2020, Shaw et al., 2019]. Machine learning(ML) is a subset of AI and is used most frequently for healthcare applications [Yin et al., 2021, Gerke et al., 2020]. ML "allows computational systems learn from past data and improve their performance without being explicitly programmed" [Gerke et al., 2020]. Deep-learning is a subset of ML. This technique "employs artificial neural networks with multiple layers to identify patterns in very large datasets" [Gerke et al., 2020]. To specify, in this research an AI-based medical device is defined as a medical device which contains an algorithm, based on either machine-learning or deep-learning techniques, and assists during the treatment or diagnosis of patients in a Dutch hospital.

An example of an AI-based medical device is the aeoNose, developed by the e-Nose company. The aeoNose is able to detect multiple types of cancer while only examining the breath of patients. Types of cancer which are detected with the aeoNose are colorectal cancer and lungcancer [Krauss et al., 2018, Krauss et al., 2020]. In 2018, 2.1 million patients were diagnosed with lungcancer and 1.1 million patients were diagnosed with colorectal cancer worldwide [Bray et al., 2018]. Using current diagnostic interventions, these cancer types have already developed to a late stadium before diagnosed, which worsens the prognosis of these patients tremendously [Steenhuis et al., 2020]. The aeoNose is developed to improve diagnosis for these patients and its performance shows great promises. It is highly sensitive and

specific, is minimally invasive and less expensive compared to current interventions [Krauss et al., 2018, Krauss et al., 2020, Waltman et al., 2020, van Keulen et al., 2020, Steenhuis et al., 2020].

However, the aeoNose has never been successfully implemented in a Dutch hospital, which is common for developed, promising, AI-based medical devices. Throughout academic literature concerns are expressed about the increasing gap between the number of publications about the development of AI-based medical devices and the minimal number of publications about its implementation [Yin et al., 2021, Felmingham et al., 2020, Manlhiot et al., 2021]. Besides, multiple articles state hospitals and medical departments, as organizations, are not ready to implement an AI-based medical device and can not foresee what is needed to do so [He et al., 2019, Kelly et al., 2019, Manlhiot et al., 2021]. Even though, many articles put emphasis on the numerous opportunities AI-technology brings to improve healthcare [Thomasian et al., 2021, Yin et al., 2021, Felmingham et al., 2020, Bélisle-Pipon et al., 2021, van Baalen et al., 2021].

In order to enjoy the full potential of AI-based medical devices, implementation of these devices needs stimulation. Throughout academic literature, emphasis is put on the influence of the unique device characteristics, caused by AI-technology, which influences its users and other medical professionals within a medical department [Asan and Choudhury, 2021, Sujan et al., 2019, Felmingham et al., 2020, Gerke et al., 2020]. Examples of these device characteristics are changeability, controllability and explainability. Furthermore, these characteristics also have an influence on the medical department, defined as organization, in which the device is implemented. Organizational routines and workflows are changed during the implementation and these changes must be accepted [Greenhalgh and Abimbola, 2019, He et al., 2019, Chua et al., 2021]. Because of the unique characteristics of AI-based medical devices, these changes are needed in administrative, technical and medical routines [Arora, 2020]. Therefore, during implementation, the AI-based medical device, the user and other medical professional within the medical department interact with each other. Such an interaction is defined as a mechanism. The mechanisms which occur during the implementation of AI-based medical devices is unique for such devices. Consequently, current literature on the implementation of medical devices in general is less applicable to AI-based medical devices [Bélisle-Pipon et al., 2021, Shaw et al., 2019, He et al., 2019].

To increase the number of successfully implemented AI-based medical devices in hospitals, to ultimately improve quality of care, mechanisms which are present during these implementations need research. Therefore, the following questions are researched:

1.1 General research question

How do mechanisms, present on an organizational level, influence the successful implementation, by users, of AI-based medical devices in medical departments of Dutch hospitals?

1.2 Sub-questions

1. How can we define and measure the successful implementation, by users, of AI-based medical devices? *(Conceptual)*
2. How are device-, user- and organizational characteristics expected to interact? *(Theoretical)*
3. How do mechanisms on an organizational level result from the interaction between device-, user- and organizational characteristics? *(Exploratory)*

To answer these research questions, a qualitative research design is most suitable to explore processes and mechanisms as empirical phenomena. A case-study is the most suitable to investigate what mechanisms are present on an organizational level during a real-life implementation of an AI-based medical device [Yin, 2011]. Therefore, a case-study is done in which two cases are included. First a theoretical framework is constructed, from this theoretical framework several general expectations are deduced. An expectation contains a mechanism, present on the level of the organization, which is expected to stimulate successful implementation of an AI-based medical device. Consequently, the expectations are compared to empirical data, which is gathered during the case-study via interviews and collected documents.

1.3 Relevance

The answers to the previously stated research questions are relevant to both society and science. First, the societal relevance of the research results are related to the clinical potential of AI-based medical devices and the agenda of both the EU and the Netherlands. To start, the clinical potential of AI-based medical devices is emphasized and research by multiple studies [Shaw et al., 2019, Tartar et al., 2021, Thomasian et al., 2021, Yin et al., 2021, Felmingham et al., 2020, Bélisle-Pipon et al., 2021, van Baalen et al., 2021]. This research contributes to offering practical information about successful implementation of AI-based medical devices. Ultimately, when AI-based medical devices improve accuracy and consistency of therapeutic and diagnostic interventions, these interventions become more cost-effective. This is especially relevant for the Dutch healthcare system, in which costs are regulated by the market. To continue, AI-technology has a prominent place on the agenda within the EU as well as the Netherlands [Commission, 2018, Rijksoverheid, 2019]. This research promotes successful implementation of AI-based medical devices in hospitals, which contributes to these internationally set goals.

Scientifically, this research contributes by filling a science-gap and offering new theoretical insights through combining existing scientific literature. To start, literature concerning AI-based medical devices is often solely focused on development and performance of AI-based medical devices [Sujan et al., 2019, Chua et al., 2021, Yin et al., 2021, Manlhiot et al., 2021, Kelly et al., 2019]. Research concerning the successful implementation of these devices in healthcare is scarce [Yin et al., 2021, Felmingham et al., 2020, Manlhiot et al., 2021]. Therefore, this research contributes to filling this science-gap. Moreover, this research contributes scientifically by using a unique combination of theoretical concepts and models. To specify, the Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability (NASSS) framework is combined with research about human factors and the unique device characteristics caused by AI-technology. By comparing the combination of theoretical concepts to empirical data from the case-study, new theoretical insights are attained about mechanisms which are present during successful implementation of AI-based medical devices in Dutch hospitals. Furthermore, after this research, a statement is made about the applicability of the NASSS framework in this context.

2 THEORETICAL BACKGROUND

The goal of this section is to identify important concepts discussed in academic literature which are then used to analyze influential mechanisms. As a product, a theoretical framework is developed. Next to this framework, two expectations are formulated. These expectations are based on the framework and contain mechanisms which are expected to occur during successful implementations of AI-based medical devices in Dutch hospitals. The expectations are used to compare data, gathered during the case-study, to theoretical concepts included in the theoretical framework, based on academic literature.

To explore mechanisms on the organizational level, two other levels are identified from which theoretical concepts interact with the organizational level. These are the level of the AI-based medical device and the user level. To describe the configuration of theoretical concepts within each level, the theoretical concepts are defined and expressed by characteristics. A level is described by a variable with characteristics. To illustrate, the user level is described using the variable 'user characteristics'. Together the three variables, device-, user- and organizational characteristics, form the conceptual model of this research, presented by figure 2.1. The corners of the pyramid represent the variables. The straight arrows, or edges of the pyramid, represent mechanisms, which are interactions between variables. Lastly, an interaction between variables is dependent on time, therefore mechanisms are dynamic. This is in line with Greenhalgh et al. (2019), who state that implementation is a dynamic process. Therefore, the curved arrows and multiple triangles on top of each other show the dynamic character of the model and how the configuration of characteristics within the model can change through time.

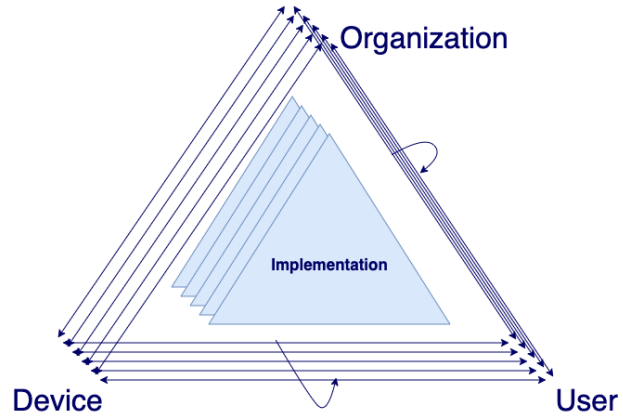

Figure 2.1: *Conceptual framework of implementation process of AI-based medical devices and the influence of device-, user- and organization characteristics.*

Translating the conceptual model of figure 2.1 into the context of this research, the user is defined as a medical professional, the device as an AI-based medical device and the organization as a medical department within a Dutch hospital. The degree of successful implementation is placed in the centre of the pyramid as this is the dependent variable. The dependent variable is used to investigate what influence a mechanism has on the implementation of an AI-based medical device, which is expressed in the degree of success. No arrows point from the corners of the pyramid to the middle, as the influence of device-, user- and organizational characteristics on the degree of successful implementation remains unknown.

2.1 Successful implementation

In academic literature there is not one definition for a successful implementation of a medical device. This complicates the evaluation and measurement of successful implementation. According to Rogers(1995), an implementation is successful when the device is used. Van Beest et al. (2020) add that an implementation is successful when the device is integrated in daily routines of users, the device is brought on the market and the use of the device is monitored by the organization in which it is implemented. Greenhalgh et al. (2017) add a different perspective to successful implementation. They state, an implementation is successful when there is a low degree of complexity within the seven domains of the NASSS framework. If there is a high degree of complexity within one domain, then it is more difficult to successfully implement the medical device.

These three definitions illustrate how measuring successful implementation is complicated without a uniform definition. To solve this, within this research multiple studies are combined to identify a definition. The degree of successful implementation is expressed by using two dimensions, organizational goals and use. A graphical overview of these two dimensions, and their concepts, is shown in table 2.1. To clarify, according to multiple studies, an implementation is more successful when goals, set by members of the organization, are reached during or after the implementation [Lundmark et al., 2021, Greenhalgh et al., 2017]. The inter-relatedness of these goals influences the degree of successful implementation. If goals are contradictory, then it is more difficult to reach all goals. Additionally, an implementation is more successful when workflow changes are accepted and the AI-based medical device is used daily by multiple medical specialists within one medical department [Rogers, 1995, Greenhalgh et al., 2017]. By defining these two dimensions, the user level as well as the organizational level is included.

Successful implementation

Dimension	Concepts		
Organizational goals[1,2]	Inter-relatedness		
Use of the devices[2,3]	Number of users	Frequency of use	Acceptance workflow changes

1. Lundmark et al. (2021), 2. Greenhalgh et al. (2019), 3. Rogers (1995)

Table 2.1: *Graphical overview of the conceptualization of the dependent variable 'Successful implementation'.*

2.2 The NASSS framework

The defined variables are further conceptualized using a combination of multiple theoretical constructs. One of which is the widely-used NASSS framework [Greenhalgh et al., 2017]. To introduce, the NASSS framework is a determinant framework which is used to explore determinants within and outside of the organization in which a medical device is implemented. Determinants are factors which influence the degree of successful implementation [Greenhalgh et al., 2017]. Determinant frameworks are the most suitable to analyze hindering and stimulating factors during the implementation of medical devices [Nilsen, 2015]. When applying the NASSS framework, the investigated concepts are evaluated in terms of complexity [Greenhalgh and Abimbola, 2019].

To adjust the NASSS framework to the scope of this research, five out of the seven NASSS domains are selected: condition, technology, adopter system, organization and embedding and adaptation over time [Greenhalgh et al., 2017]. The excluded domains go beyond the scope of this research as one focuses on financial or economic determinants, which is the value proposition domain. And the other contains determinants which exceed the level of the organization, which is the wider system domain. The remaining domains fit in the conceptual model of this research. The domains 'technology' and 'condition' are combined into the device characteristics. The domain 'adopter system' is utilized in user characteristics and the domain 'organization' in organizational characteristics. Although the NASSS framework is very applicable to e-Health innovations, it needs adjustment to fit AI-based medical devices. AI-based medical devices consist of unique characteristics, caused by AI technology. Therefore, the NASSS framework is complemented with research that investigated implementation of AI-based medical devices specifically. This research is introduced in the following sections.

2.3 Expectations

In this section the two expectations of this research are deduced. These expectations are deduced using device-, user- and organizational characteristics and mechanisms which re-occur in academic literature. The expectations contain mechanisms which are expected to influence the degree of successful implementation of AI-based medical devices. Tables 2.2, 2.3 and 2.4 show a graphical overview of these characteristics per variable. The following sections describe the two expectations, using the concepts from tables 2.2, 2.3 and 2.4, in more detail.

Device characteristics

Dimension	Concepts			
Technical features[1]	Changeability[3-4, 7]	Controllability[3-4, 7]	Explainability[3-6]	Validity and reliability[3-6]
Condition[1, 2]	Stability	Knowledge	Co-morbidities	

1. Greenhalgh et al. (2019), 2. Chua et al. (2021), 3. Arora et al. (2020), 4. Gerke et al. (2020), 5. Shaw et al. (2019), 6. Asan & Choudbury (2020), 7. COCIR (2020)

Table 2.2: *Graphical overview of the conceptualization of the independent variable 'Device characteristics'.*

User characteristics

Dimension	Concepts			
Intended users[1]	Profession	Knowledge background	Role development and/or implementation process[3-5, 8]	
Human factors[2, 6-8]	Workload[2]	Perception of AI-based medical device[2, 8]	Trust	Understanding

1. Greenhalgh et al. (2019), 2. Sujan et al. (2019), 3. Strohm et al. (2020), 4. Damschroder et al. (2009), 5. Smuck et al. (2021), 6. Phansalkar et al. (2010), 7. Felmingham et al. (2020), 8. Asan & Choudbury (2021)

Table 2.3: *Graphical overview of the conceptualization of the independent variable 'User characteristics'.*

Organizational characteristics

Dimension	Concepts	
Innovation capacity[1]	Resource availability	Knowledge translation[5,6]
Organizational readiness[1]	Leadership[1-3, 4]	Change needed to organizational routines

1. Greenhalgh et al. (2017), 2. Shaw et al.(2019), 3: Strohm et al.(2020), 4. Laukka et al. (2020), 5. Dadich and Doloswala (2018), 6. Kislov et al. (2014)

Table 2.4: *Graphical overview of the conceptualization of the independent variable 'Organizational characteristics'.*

2.4 Expectation 1: Training

AI-based medical devices distinguish themselves from medical devices in general due to the presence of the so-called 'black-box'. This 'black-box' causes certain device characteristics, like changeability and controllability, which are further explored in the following section. Users and other professionals within the organization are influenced by these device characteristics, which in term has an effect on the degree of successful implementation. Therefore, users and other professionals within the medical department need training to promote trust and understanding about device. Who needs training specifically is also determined in the following section. To determine which professional is best suitable to take this responsibility, multiple roles which are present within a medical department are explored. Ultimately, an expected organizational mechanism is deduced into an expectation to express why training is necessary, who needs training and who needs to facilitate it in order to successfully implement an AI-based medical device.

2.4.1 Device characteristics

Throughout the literature, five device characteristics are defined as important when implementing AI-based medical devices: changeability, controllability, explainability, validity and reliability. These characteristics are also shown in table 2.2. To start, changeability refers to the AI algorithm its ability to learn in an online fashion based on new input data through time [van Baalen et al., 2021]. The part of the algorithm which changes is the so-called 'black-box', which contains computed patterns and relationships between input variables [van Baalen et al., 2021]. Multiple studies define changeability as an important device characteristic when implementing an AI-based medical device [Gerke et al., 2020, Arora, 2020, Chua et al., 2021]. Here, three val-

ues are identified, which are in line with the definition of the European Coordination Committee of the Radiological, Electromedical and Healthcare IT Industry(COCIR) [COCIR, 2021]. These are (1) locked, (2) continuous and (3) discrete change. Locked change indicates algorithms where only the manufacturer can change the algorithm by releasing an update. The algorithm does not change during use. Continuous change indicates algorithms which learn and change during use without explicit manufacturer or user interaction. Discrete change indicates algorithms which learn during use with explicit manufacturer or user interaction.

The controllability of the AI-based medical device influences how the device is integrated into the workflow of the user. This integration has an effect on the user. To clarify, implementing an AI-based medical device causes tasks automation [Shaw et al., 2019, Tartar et al., 2021]. Controllability refers to the extent to which medical specialists can interfere in the functioning of the AI-based medical device [COCIR, 2021]. To clarify, whether or not it is possible for medical specialists to stop or edit the outcome generated by the algorithm of the AI-based medical device. According to COCIR, three values of the concept controllability are identified: (1) direct control, (2) supervisory and (3) no control. When an algorithm is directly controlled, the algorithm performs a task, waits for the user to approve before continuing to the next task. Supervisory control means the algorithm can sense, decide and act on its own. The user supervises its operation and can intervene when required. Lastly, no control means the user cannot intervene.

To continue, the explainability of the AI-based medical device is negatively influenced by the so-called 'black-box', which is present in the AI algorithm [Gerke et al., 2020, Shaw et al., 2019, Asan and Choudhury, 2021, Arora, 2020, Felmingham et al., 2020]. Consequently, outcomes generated by these devices are more difficult to explain by users. The 'black-box' refers to the operating part of AI algorithm, within the device, which is concealed. Knowing exactly what the algorithm is basing an outcome on, is impossible. However, explaining outcomes from medical interventions in Dutch hospitals is very important. Due to laws and regulations in the Netherlands, lacking explainability of outcomes can cause a problem in the context of Dutch hospitals. Medical specialists are obligated to inform the patient about study results and risks of a certain treatment to promote co-decision-making between specialist and patient [Staten Generaal, 1994]. Additionally, when a patient is well informed it stimulates trust between medical specialist and patient [Gerke et al., 2020, Shaw et al., 2019].

Finally, validity and reliability are device characteristics which are important as they influence trust among users and other professionals within the organization. AI-technology is still in the early-adopter phase as a medical technology, because more articles are published concerning the development of these devices compared to articles concerning successful implementation [Yin et al., 2021, Manlhiot et al., 2021]. Therefore, few medical specialist have experience with AI-based medical devices. Multiple articles state that the validity and reliability of a device are important to promote trust among medical specialists [Chua et al., 2021, Helman et al., 2022, Strohm et al., 2020]. Additionally, professionals within the organization trust the device more, if outcomes generated by the device are perceived as valid and reliable [Gerke et al., 2020, Sujan et al., 2019]. Without trust, successful implementation is less likely. Consequently, training about the changeability, controllability, explainability, validity and reliability of an AI-based medical device is needed to successfully implement it.

2.4.2 Who needs training?

Professionals from various professions within a medical department need training to successfully implement an AI-based medical device. To determine who specifically needs training, it is important to identify intended users, investigate what role other professionals within the organization have, investigate their background knowledge and what knowledge they need to fulfill

their responsibilities during the implementation. To start, to identify intended users, their profession and knowledge background are determined. The user's profession is important to identify as skills and knowledge vary between medical professions [Greenhalgh and Abimbola, 2019]. Additionally, within a profession group, background knowledge also varies. Often, medical professionals have limited technical background knowledge [Gerke et al., 2020]. Understanding the functioning of this AI-based medical device is more complicated for medical professionals compared to professionals with technical professions [Strohm et al., 2020]. However, this understanding is important. As Chua et al.(2021) states, users must understand the outcome generated by the AI-based medical device, trust the evidence on which the device is based and find it relevant. Furthermore, this article puts emphasis on the importance of adapting the threshold for understanding the AI-based medical device to the use case and the utilized AI-technique.

Additionally, by engaging intended users during the development process of the AI-based medical device, the usability of the device improves which leads to a more successful implementation of the device [Asan and Choudhury, 2021, Phansalkar et al., 2010]. Additionally, users gain a more thorough understanding about the AI-based medical device when they engage during this process [He et al., 2019].

Users not only need training about the device, but also training which promotes trust in their own skills needed to use the AI-based medical device [Gerke et al., 2020, Asan and Choudhury, 2021]. If users perceive their abilities as sufficient, their trust also increases.

Not only users, but other professionals within the organization need to have sufficient knowledge about the AI-based medical device, therefore they need training. Who needs what kind of training depends on the background knowledge and the role this professional has during the development and/or implementation process. An implementation often causes re-organization of the medical department [Arora, 2020, Greenhalgh and Abimbola, 2019]. To add, due to the implementation of AI-based medical devices tasks are often redistributed due to task automation [Strohm et al., 2020]. Consequently, changes are necessary for organizational routines and workflows. If professionals are responsible for managing or executing these changes, they need training in order to make a proper estimation during the implementation of such a device [He et al., 2019, Chua et al., 2021].

As AI-based medical devices contain unique characteristics, changes needed to organizational routines are different. For example, when a medical department develops an AI-based medical device internally, professionals within the department have to manage the development and implementation of the device in line with regulations as stated by the Medical Device Regulation(MDR) [European Parliament, 2017]. This is a very labour-intensive responsibility about which professionals within the medical department need training [Beckers et al., 2021, Shaw et al., 2019].

Next to professionals who manage and execute necessary changes, medical specialists specifically also need training when implementing an AI-based medical device. Even when they are not intended users of the device. According to Strohm et al.(2020), medical specialists are apprehensive towards the implementation of AI-based medical devices due to low trust in its clinical relevance and performance of the device. When looking at the organizational structure of Dutch hospitals, medical specialists have a high hierarchical position and have multiple liabilities [Staten Generaal, 1993, Staten Generaal, 1993, Lange et al., 2011]. This means that medical specialists have a powerful position within the hospital and medical department. Their liabilities make them feel responsible and cautious when implementing a new device when used during diagnostic or therapeutic interventions.

2.4.3 The organizational level

On the organizational level, multiple actors are operating to facilitate and organize the implementation of an AI-based medical device. Within the organization the necessary resources and infrastructure is needed to facilitate such an implementation. To analyze these concepts within a medical department, two dimensions: innovation capacity and organizational readiness are identified. Within innovation capacity, resource availability and knowledge translation are discussed. To describe organizational readiness, leadership in terms of role division is discussed.

Innovation capacity

Before the start of an implementation, professionals within the organization need to estimate what resources are necessary to facilitate training and change organizational routines and workflows. Therefore, the available resources within the organization are identified [Greenhalgh and Abimbola, 2019]. A part of these resources are the available infrastructure within the organization that is necessary to spread knowledge and organize training activities.

To estimate what type of training is needed to support professionals within a department during the implementation of an AI-based medical device, knowledge translation within that department is critical. Although there has been critique on the exact term, knowledge translation describes how research-based evidence is translated and implemented into practice within an organization [Greenhalgh and Wieringa, 2011]. Training is an activity which promotes knowledge translation within an organization.

Several studies investigated how knowledge translation occurs within a medical department or hospital. To start, a study unveiled that efforts to change a medical department which are accommodated to the potential interrelated influence of the professional and the medical department are more successful compared to single initiatives to promote evidence-based care [Dadich and Doloswala, 2018]. Additionally, another study advocates for more strategic thinking about how new capabilities or knowledge are utilized, maintained and updated as a part of the organization [Kislov et al., 2014]. Next to this, they emphasise the importance of integrating 'learning from colleagues' and 'learning by doing' into a training program. This enables to adapt, absorb and modify knowledge and capabilities through repeated practice.

Organizational readiness

The capability of an organization to facilitate knowledge translation is bound to organizational readiness. To clarify, the degree to which an organization is 'ready' to implement a device [Greenhalgh et al., 2017, Damschroder et al., 2009]. During an implementation, involved professionals have different responsibilities. For example, to lead the implementation or to manage changes to organizational routines. These different responsibilities are referred to as roles. Together, these roles shape the organization which carries the responsibility to successfully implement a device. Important roles to distinguish are policy entrepreneurs, decision-makers, leaders, managers, key-users, change agents and facilitators.

At the beginning of the implementation, a policy entrepreneur is responsible to get more professionals actively involved in order to initiate the implementation. During this phase, where, for example, the implementation of that innovation has to be put on the agenda, the policy entrepreneur plays a prominent role [Cairney, 2018]. After the efforts of the policy entrepreneur, a professional has to make the decision to initiate the implementation. This professional has often a leading position within the medical department [Laukka et al., 2020]]. This professional is described as the decision-maker.

Once the implementation is initiated, an implementation leader is appointed. While leading and giving direction to the department during the implementation, an implementation leader can

also operate via the other described roles [Laukka et al., 2020]. Consequently, a manager is assigned. This professional is actively participating in the planning of the implementation, creating implementation teams and scheduling the implementation. Additionally the manager is responsible for offering clear communication towards all involved professionals about changes and identifying where resistance exists among professionals towards the implementation. A key-user is an intended user who is more involved compared to other intended users. The key-user contributes actively to the implementation as a whole, by for example articulating needs of the user group.

During an implementation, not all professionals within a medical department may be enthusiastic or supportive of the implementation. This can cause resistance within the department towards the implementation. In these situations a change agent is important [Laukka et al., 2020, Strohm et al., 2020, Felmingham et al., 2020]. A change agent carries the responsibility to promote acceptance of the implementation, and all its necessary changes, among colleagues. With the ultimate goal to convince colleagues. In academic literature change agent and local champion are two terms used interchangeably [Strohm et al., 2020, Felmingham et al., 2020, Greenhalgh and Abimbola, 2019]. According to Laukka et al.(2020) this is also in line with the role they identify as advocate and supporter. In this research, change agent is used and defined as "an enthusiastic visioneer willing to initiate and drive services forward" [Kujala et al., 2019]. Important attributes which determine the success of a change agent are a large circle of influence, ownership of the implementation, physical presence at the point of change, grit, persuasiveness, and a participative leadership style [Bonawitz et al., 2020].

Lastly, a facilitator tries to close promote cooperation between professions and professionals [Laukka et al., 2020]. The implementation of AI-based medical devices demands changes in many different area's [Arora, 2020]. Therefore, a lot of professions, who are not used to cooperating, now have to cooperate. Between these different profession groups there are differences in interests and priorities. Therefore, the role of facilitator is extra important during the implementation of AI-based medical devices.

Due to these interactions between device-, user- and organizational characteristics, the following mechanism is present on the organizational level during the successful implementation of an AI-based medical device, which is defined as expectation 1:
If implementation leaders facilitate training about characteristics of the AI-based medical device for professionals of multiple disciplines within the medical department, trust in the device is promoted and all necessary changes are made and accepted. Therefore, the degree of successful implementation increases.

2.5 Expectation 2: Multi-disciplinary involvement

Multi-disciplinary involvement is needed to facilitate adequate training, to detect what changes are needed to organizational routines and workflows and to promote task redistribution, which are described in the first subsection. However, actively involving professionals not always easy. Multiple strategies are discussed in the second subsection which promote active involvement of medical professionals during an implementation.

2.5.1 Who to involve?

First of all, to give training about AI-based medical devices, software know-how is needed [He et al., 2019, Helman et al., 2022, Shaw et al., 2019]. This know-how is often not present among medical specialists [Gerke et al., 2020]. Therefore other disciplines need to be involved to facilitate adequate training.

Second, according to Greenhalgh et al. (2019), changes to organizational routines and workflows are necessary to implement a medical device successfully. One article even states to implement an AI-based medical device necessary changes are needed in more areas of a medical department, for example administrative area's as well as technical areas [Arora, 2020]. Multiple studies advocate for multi-disciplinary involvement to estimate is what changes are necessary to integrate the AI-based medical device into the workflow of multiple professionals [He et al., 2019, Chua et al., 2021]. Many articles emphasize the importance of seamless and intuitive integration of AI-based medical devices into the workflow of medical professionals [Chua et al., 2021, Sujan et al., 2019, He et al., 2019, Greenes et al., 2018]. AI-based medical devices often automate tasks, therefore sufficient integration into the workflow is extra important [Sujan et al., 2019, van Baalen et al., 2021, Shaw et al., 2019, Arora, 2020]. A task is first standardized, before it is automated. This standardization process requires a medical perspective and a technical software perspective to estimate how the task is standardized so that the AI-based medical device functions optimally in line with desires of professionals within the medical department.

Third, as AI-based medical devices enable the task automation, task redistribution is promoted [Shaw et al., 2019, Tartar et al., 2021]. Tasks are redistributed from professionals with a higher task complexity, like a medical specialist, to professionals with a lower task complexity, for example a physician assistant. Therefore task redistribution is a multi-disciplinary act in itself. In the next paragraphs, strategies to involve professionals from important disciplines are described. Additionally, obstacles which complicate this multi-disciplinary involvement are addressed.

2.5.2 Strategies to involve multiple disciplines

Strategies to involve professionals from multiple disciplines during the implementation of AI-based medical devices is needed to involve professionals with a high workload and to find a balance between costs and benefits gained from involving more professionals. To begin, medical specialists are essential during the implementation of an AI-based medical device due to their hierarchical position, liabilities and medical expertise [Lange et al., 2011]. Some studies state that medical specialists feel apprehensive towards task redistribution due to their liabilities and fear of losing professional autonomy [Gerke et al., 2020, Tartar et al., 2021, Strohm et al., 2020]. Because of this it is important to involve medical specialists, but these professionals are often burdened with a high workload, which causes a high cognitive load [Asan and Choudhury, 2021, Tartar et al., 2021]. Consequently, medical specialists may not prioritize active involvement during the implementation of a device. To promote involvement among medical specialists, multiple studies recommend to appoint a change agent within the medical department [Gerke et al., 2020, Strohm et al., 2020, Smuck et al., 2021, Shaw et al., 2019]. More specifically, among the group of medical specialists [Felmingham et al., 2020].

To continue, involving professionals from multiple disciplines in order to successfully implement an AI-based medical device is time- and resource-consuming. Involving more professionals costs many resources and decision-making processes are often more complex [Feiock, 2007]. However, in some situations involving professionals promotes resource collection. So, a balance is needed between the number of involved professionals, resource costs and optimal decision-making circumstances. One aspect which is specifically important for AI-based medical devices is the distinction between externally developed devices and internally developed devices. Due to the Medical Device Regulation (MDR), professionals within the medical department need to develop an AI-based medical device in line with extensive rules and regulations [European Parliament, 2017]. To document and organize the development of the AI-based medical devices in accordance with the MDR is a very resource costly procedure [Shaw et al., 2019]. Therefore, internally developed AI-based medical devices have a different resource-

consumption compared to externally developed AI-based medical devices. In this situation, involvement of professionals from multiple disciplines must be promoted to collect resources.

According to the literature, other strategies can be used by implementation leaders to involve professionals actively. For example, involving users during the development process of the AI-based medical device [Asan and Choudhury, 2021, Phansalkar et al., 2010]. Furthermore, communication plays an important role. Through communicating the right information, professionals become motivated and involved during the implementation process. The organizational structure and organization of a medical department can facilitate the means which are necessary to promote open and transparent communication among professionals. This is important during implementation processes as collaboration is required between professionals of varying disciplines who possibly have diverting interests [Lundmark et al., 2021, Greenhalgh and Abimbola, 2019]. An implementation leader can also appoint a facilitator to promote cooperation between disciplines or appoint a change agent to promote involvement among professionals [Laukka et al., 2020].

Due to these interactions between device-, user- and organizational characteristics, the following mechanism is present on the organizational level during the successful implementation of an AI-based medical device, which is defined as expectation 2:
If implementation leaders, facilitators and/or change agents involve stakeholders and professionals from all disciplines whose workflow changes or who manage or execute these changes, during the development or implementation of an AI-based medical device, the degree of successful implementation increases.

3 RESEARCH DESIGN

To check the two expectations, regarding organizational mechanisms which stimulate successful implementation of AI-based medical devices, a case study was performed. Two cases, containing the implementation of two different AI-based medical devices, were included in this case-study. A case is defined as the event during which an AI-based medical device is implemented on a medical department in a Dutch hospital. This section elaborates on how the case study, data-collection and data-analysis are organized.

Two types of data were gathered to gain information about the case. First, several semi-structured interviews were held with multiple medical professionals who are member of the medical department of that case. These professionals either used the AI-based medical device or contributed to the implementation process. Second, internal documents, which have been created or altered during the implementation of the AI-based medical device, were collected and analyzed.

Consequently, a qualitative data-analysis was executed to evaluate the interaction between device-, user- and organizational characteristics and how this interaction influences the degree of successful implementation. Referring back to figure 2.1, a bottom-up approach was used to analyze the data. To clarify, to gain insight about mechanisms on the organizational level, first mechanisms present on lower levels were explored as these mechanisms interact with mechanisms on the organizational level. This led to identifying the influence of mechanisms on the organizational level on the degree of successful implementation of AI-based medical devices, which is the goal of this research.

In the following sections several components of the research design are elaborated upon: in- and exclusion criteria, case description of both cases, data-collection, bias and data-analysis.

3.1 Inclusion- and exclusion criteria of case selection

Cases were selected through the educational network of the researcher. Based on several criteria, cases were either in- or excluded. To sum up, these criteria are: country, availability of professionals, finalized implementation process, type of user, software purpose and variety.

The scope of this research is limited to the Dutch healthcare system, constrained and tied to Dutch and European laws and regulations. Therefore, only medical departments of Dutch hospitals were selected. No distinction was made between an outbound hospital and an academic hospital. Moreover, to guarantee collection of sufficient and reliable data, only cases were selected if members of the medical department were available and willing to cooperate with the case study. To add, the implemented AI-based medical device needed to be used by users on a daily basis and by multiple users. This also contributes to sufficient and more reliable data-collection. Reliable is the sense that more professionals can be interviewed about their experiences.

Cases were selected if the medical department had succeeded the implementation process of an AI-based medical device. This AI-based medical device was ought to be used during diagnosis and/or treatment of diseases. Furthermore, the implemented AI-based medical device had to consist of a Conformité Européenne (CE) mark or, if the device was developed by

members of the department, had be in line with requirements as stated by the Medical Device Regulation (MDR) from the European Union (EU) [European Parliament, 2017]. This means AI-based medical device were past the end-development stage, tested and validated.

Cases were only included if the implemented AI-based medical devices assisted medical professionals during interventions, who treated or diagnosed patients. Furthermore, cases were included if users of the device are medical professionals. By only selecting implemented AI-based medical devices which are used by medical professionals for medical interventions, these devices are bound to special laws and regulations like the MDR [European Parliament, 2017]. By narrowing down to one type of user and one type of software-purpose, the variety of device- and user characteristics is limited to some extent. To illustrate, when patients use the AI-based medical device, there is a larger variety of background knowledge between patients. Within a population group of medical professionals there is a lower variety of background knowledge. Limiting variety in device- and user characteristics promotes the gathering of more in-depth data about organizational mechanisms, which are a consequence of the interaction between device- and user characteristics.

Nevertheless, enough variation is needed between and within cases to make a comparison when confirming or rejecting the expectations. Between cases there is variety between the implemented AI-based medical device and type of medical department. To guarantee this, each case concerns the implementation of a different AI-based medical device, a different medical department and a different Dutch hospital. Within the case there is variety between users. To illustrate, in both cases multiple professionals use the AI-based medical device, either within the same profession or multiple professions.

AI-based medical devices which assisted during interventions for the diagnosis or treatment of COVID-19 patients were excluded. Due to the COVID-19 pandemic, a crisis situation is created which is incomparable to usual health care routines and processes. If these devices were included in this research, its conclusions would be less applicable to 'usual' health care situations.

3.2 Cases

Two cases are included in this research, which are referred to as case 1 and case 2. Hereafter the cases will be described in terms of the type of medical department, the implemented AI-based medical device, why this case is selected, how professionals were selected for the interviews and what professions these professionals have.

3.2.1 Case 1

For case 1, a radiotherapy department was selected which is situated in an outbound hospital located in the Netherlands. The radiotherapy department is a relatively small department, which means treating patients is prioritized above research. The department consists of a department manager, a group of CT-technologists, a group of medical physicists, a group of radiotherapist and a technical medicine specialist. The radiotherapy department bought the AI-based medical device from a company which distributes this device on an international scale. The AI-based medical device was already approved by the Federal Drug Administration (FDA) in the USA. Therefore, the device was CE-marked and tested and validated according to FDA standards.

When contacting this case, members of the department were enthusiastic to cooperate with this research. At that moment, the AI-based medical device was already implemented and used on a daily basis by CT-technologists. These professionals work with the CT-scanners that execute radiotherapy treatments. CT-technologists are responsible for giving the treatment to the patient, which is in line with the radiotherapist's plan. Patients diagnosed with cancer, who are able to receive radiotherapy, are referred to these professionals. Radiotherapists are medical

specialists who create treatment plans, for patients who suffer from cancer, according to scan data from these patients. Patients also consult radiotherapists when suffering from side-effects caused by radiotherapy treatments. The AI-based medical device assists CT-technologists during a treatment intervention where a patient receives a radiotherapy treatment.

This case description fits the inclusion criteria of this research, the case was therefore selected for the case-study. When planning the interviews with members of the radiotherapy department, the goal was to include as many key-actors of the implementation process as possible. To identify key-actors in advance, a meeting was planned with the contact of that case. For this case, this was the technical medicine specialist. By informing this specialist about the goal of the research and the goal of the interviews, the specialist estimated which professionals were key-actors during the implementation process. After discussing, the following professions were selected: technical medicine specialist, radiotherapist, senior-CT-technologist and the department manager.

For the interviews, professionals from these professions were selected. Some profession groups consisted of one professional, namely technical medicine specialist and department manager. These professionals were automatically selected for an interview. Within the group of CT-technologists and radiotherapists, a selection was made based on several criteria. The professional was selected if the professional was involved during the organization of the implementation process of that AI-based medical device or if the professional used the device.

With regard to CT-technologists, two senior-CT-technologists were involved during the organization of the implementation process and were the first users of the AI-based medical device. Therefore, both senior-CT-technologists were selected for an interview. Considering the group of radiotherapists, a radiotherapist was selected who was involved during the organization and execution of the implementation of the AI-based medical device for prostate cancer patients.

3.2.2 Case 2

For case 2, a nuclear medicine department was selected. This department is part of an outbound hospital located in the Netherlands. Similar to case 1, this department does not prioritize research over patient care. The department consists of nuclear medicine specialists, technical medicine specialists, medical physicians, physician assistants and scan-technologists. The AI-based medical device assists nuclear medicine specialists and physician assistants during a diagnostic intervention. When contacting this case, members of the department were enthusiastic to cooperate with this research. The AI-based medical device was already used on a weekly basis in the medical department.

The implemented AI-based medical device was developed by students during an internship at the medical department. The project was initiated and guided by members of the department. The medical physicist is responsible for documentation needed to satisfy rules and regulations of the Medical Device Regulation(MDR) [European Parliament, 2017]. Additionally, the medical physicist was responsible for setting up all internal routines and procedures so that the device could operate properly. The executive board of the hospital gave permission to the medical department, which is necessary according to the MDR before the AI-based medical device can be implemented. Therefore this case fits the stated inclusion criteria.

When selecting professionals to interview for this research, the same criteria were applied as compared to case 1. This resulted in the selection of a medical physicist, a technical medicine specialist and a nuclear medicine specialist. These three professionals initiated the development of the AI-based medical device. Also a physician assistant was selected for an interview because this professional is a user. Lastly, another nuclear medicine specialist was selected as this professional was an intended user who was informed during the development process.

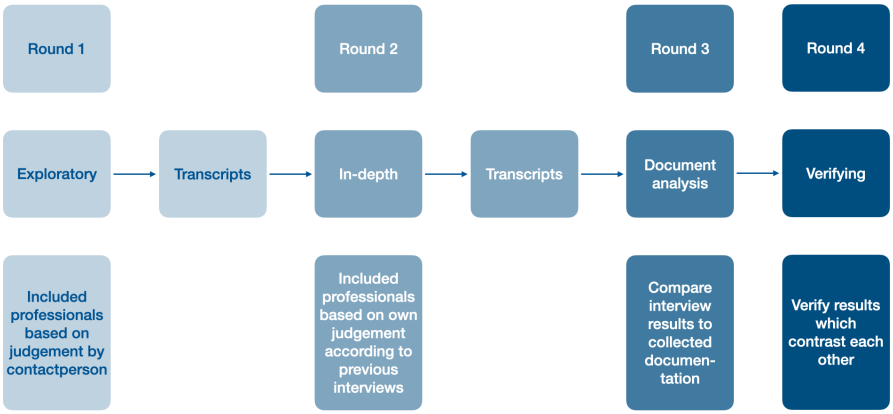

Figure 3.1: *Graphical overview of organization and timeline of the case interviews for both cases.*

3.3 Data-collection

To get a complete image of how device-, user- and organizational characteristics interact during the implementation process, data was collected via semi-structured interviews and documentation. The interviews, of approximately one hour, were held individually with professionals who contributed to the implementation process or are users of the AI-based medical device. Second, internal documents, which have been created or altered during the implementation of the AI-based medical device, were collected and analyzed. These documents consisted, for example, of protocols. In figure 3.1 a graphical overview is given of the timeline and organization of these two data-collection methods.

3.3.1 Interviews

The interviews were held in December 2021. Due to government measures taken to mitigate the COVID-19 pandemic, it was not always possible to visit the hospital. So, the interviews for one case were held online. The interviews for the other case were held in person at the medical department of that case. For case 1, 7 professionals were interviewed. For case 2, 5 professionals were interviewed, which resulted in a total of 12 interviews. Each interview took approximately one hour. Before the start of the interview, the interviewees signed an informed consent in which they consented to (1) their audio being recorded, (2) their audio being transcribed into a transcript and (3) their data being anonymously used in this research. No personal information is mentioned except for their profession. People were only included in the research if they offered a new perspective on the implementation process of the AI-based medical device of that case. Interviewees either played an important role on a user or organizational level during the implementation of the AI-based medical device. By interviewing both types of roles, information was gathered from two perspectives: the organizational and user perspective.

To prepare for the interviews, a list of subjects was created instead of a list of questions. This preparation suited best for semi-structured interviews. At the start of each interview, the aim of the interview was clearly introduced. During this introduction, the subjects of interests were mentioned. The goal of the interviews was to get a complete image of the implementation process of a case in terms of who was involved and how the process was organized. But also, what changes and resources were necessary and what obstacles needed to be faced to implement the AI-based medical device successfully. After this introduction, the interviewee started by describing his or her role during the implementation process and the interviewer steered the

conversation so that all subjects were covered in detail. The interviewer also tried to gather in-depth information about why certain decisions were made and why certain organizational structures are in place.

As the goal of the interviews was to get a complete image of the implementation process of a case, it is important to include as many relevant perspectives as possible. By organizing the interviews in several rounds, as graphically shown in figure 3.1, more perspectives could be included based on results gathered during the first round of interviews. Next to this, figure 3.1 shows how the analysis of important documents was incorporated in the timeline of the interview rounds. The interviews for both cases were organized according to the same timeline. During the first round of interviews, the exploratory interviews, only professionals were interviewed from professions who were selected by the contact person. For both cases, the contact person was the technical medicine specialist who was part of the medical department. The contact person was informed about the research and the goal of the interview. According to this information, the contact person selected professionals.

After this first round, the results of the interviews were worked out into a transcript. During this process, attention was paid to people or professions who were involved in the implementation process but who were not included in the interviews. In both cases, more people were included to get a more complete image of the implementation process of that case. These people were contacted and interviewed during the second round of interviews, the in-depth interviews. As during the first round of interviews more general information about the case was gathered, during the second round of interviews important aspects could be discussed in more detail. Hence the name, 'in-depth interviews'. Following these in-depth interviews, the results were worked out into a transcript. The third round consisted of a document analysis from collected documents. These documents were created or altered during the implementation process of the AI-based medical device. These results combined were used during the fourth round of data-collection, the verifying interviews. These interviews were shorter and no new people were included. This last round of interviews was used to verify any contradictory results or to find missing information.

3.3.2 Document analysis

To gather more empirical data about each case, documents were gathered which were created or altered during the implementation process. Only documents were included which the medical department wanted to disclose and which were necessary for organizing and monitoring the implementation process and use of the AI-based medical device in routine care. Examples of collected types of documents are protocols, teaching schedules, teaching sessions and design and develop documentation. To use this information as empirical data, a document analysis was executed. The results from the document analysis are used for triangulation with the results from the interviews. This increases the reliability and validity of the results of this research.

3.4 Bias

When collecting data, it is important to be aware of bias and how this bias can influence the collected data. In this qualitative research which uses interviews and documents as main data sources, it is important to be aware of the following types of biases in order to adjust research methods in such a way that the influence of these biases is minimized. By minimizing the influence of bias, the validity and reliability of the research and research results is increased.

In this qualitative research, two types of biases can be encountered: (1) participant bias and (2) researcher bias [Yin, 2011]. Participant bias concerns how participants respond to the ques-

tions based on their own thoughts and interpretations. But also on what is a socially acceptable or acceptable within the hospital type of answer. Asking open-ended question is needed to minimize the influence of these biases. By asking open-ended questions and avoiding leading questions the participant is more tended to give an answer which is not influenced by socially accepted assumptions. Furthermore, by guaranteeing anonymity, the participant has more freedom in talking about concepts if they are not in line with hospital interests.

Another important and more specific type of participant bias is retrospective bias, which is especially important for this qualitative research as most of the events which are talked about are in the past. Retrospective bias refers participants to have the tendency to remind past events more positively than they actually were [Yin, 2011]. Furthermore, people are more likely to attribute success to themselves and blame others. In this sense, people can lie and twist the truth. To minimize the influence of this type of bias, two methods are used: (1) common sense, (2) triangulation, (3) asking questions strategically to multiple participants and (4) data-collection timeline. To start, when participants are lying or giving strange answers, it is important as a researcher to remain critical and use common sense to assess if this participants is telling the full truth.

Second, triangulation refers to using two types of methods to collect data. For this research, this corresponds to the collection of data via semi-structured interviews and via document collection. This enables the researcher to use documents to verify data gathered via the interviews and vice versa. This promotes the reliability and validity of the final results. To illustrate, if certain concepts are found in documents and in transcripts, the results are more reliable. If a result is only found in transcripts and the opposite is found in documents, then a result is less reliable. In these situations, the concept is questioned during the verifying interviews.

Third, asking questions strategically refers to not asking certain questions to certain participants and asking the certain questions to multiple participants. To clarify, if a participant has a very strong positive opinion about a concept or has a very strong interest in a certain outcome, this participant is tended to give a biased answer.

Fourth, the interviews are held in three rounds. By using this structure via which data is collected, retrospective bias is deduced. By asking multiple participants, at different moments in time about a specific concept, this concept could have developed or evolved through time. In some cases more information may be known to the participants which alters their perception of a concept. As a researcher it is important to track these processes to see if retrospective bias was present, how the perception of the interviewee changes through time and what led to this change. To continue, a researcher or interviewer is also prone to bias, like confirmation bias. Confirmation bias refers to looking for data which suits the formulated expectations of the research. The influence of this bias is reduced by analyzing the data with a clear mind, while re-evaluating impressions and responses. This is done by using multiple types of coding techniques, like in-text coding, thematic coding and framework coding. In this way, the researcher is not only analyzing the data based on the pre-existing expectations based on the theoretical framework, but is also able to maintain a clear mind and look for other re-occurring themes or concepts in the data.

3.5 Data-analysis

All interviews were transcribed and the relevant documents were summarized. The following steps illustrate the data-analysis per case. After transcribing, a coding scheme was created. An overview of the used coding scheme is presented in Appendix A. Two methods of coding have been used to create the coding scheme. First, codes were created in advance and based on the two formulated expectations, which are in line with the theoretical framework. Second, during the coding, thematic coding was used to extract multiple theme's which re-occurred within the transcripts. For both cases, the same code-book was used. In order to code the transcripts systematically, Atlas.ti was used. Results from the coding of the transcripts were compared to documentation summaries to look for discrepancies. These discrepancies were used to prepare for the verifying round of the interviews.

4 RESULTS

The results are discussed by elaborating on the two cases separately. Each case is discussed in light of the configuration of device-, user- and organization characteristics during the implementation. This configuration is structured in a chronological order and elaborates on the implementation as a process. A brief overview of the variation between the two cases is shown in figure 4.1. After the configuration, the expectations are compared to the configuration of that case. The two expectations are either confirmed or rejected. The confirmation or rejection of each expectation is supplied with arguments based on the configuration of that case.

Variation in characteristics between cases

Concept	Case 1	Case 2
Development device	External	Internal
Purpose device	Treatment	Diagnosis
Outcome device	Image	Number
Condition	Prostate cancer	Parkinson's disease
Number of professions who use device	1	2
Number of professionals who use device	8	2
How often used	Daily	Weekly
User involvement	Implementation	Medical specialist: development phase, PA: none
Organization implementation	Team-based	One key-actor, who operated multiple roles
Change needed to organizational routines	Workflow changes for four professions, set-up training program, technical medicine specialist as new profession within the department	Workflow changes for two professions, documentation MDR, student guidance, collecting resources
Task redistribution	Medical specialist —> CT-technologist Via task standardization	Medical specialist —> PA, supervised by medical specialist
Change agent	Yes	No
Collected resources via	Company and hospital	Hospital network and hospital
Important resources	Emulator, technical medicine specialist, training	Students, hardware, research funds

Figure 4.1: *Graphical overview of variation in concepts between case 1 and case 2.*

4.1 Case 1

To gather information about the implementation of the AI-based medical device, seven interviews were organized with professionals from the following professions: technical medicine specialist, senior-CT-technologist, radiotherapist, department manager and medical physicist. Within each profession, one professional was interviewed, excluding radiotherapist and senior-CT-technologist. From these professions, two professionals were interviewed. During the interviews, another radiotherapist was found who functioned as a change agent during the implementation process of the device. Therefore, this radiotherapist was also included.

Information was gathered on how the implementation process was organized, how the implementation process happened, what events occurred along the way and what effect the implementation had on the department. Furthermore, relations between the multiple involved professions were explored next to the relation between the company and the department. Finally, the organization of the department and the hospital were discussed.

Besides interviews, also documentation was collected from the department. The collected documents are: training program, training presentation and a published article about the implementation. From these documents the organization of the training program, which was an important part of the implementation, was investigated. From the article new insights were collected about the organization of the implementation of the same AI-based medical device on different radiotherapy departments across the globe and what the effect of the implementation has been on the department.

In the upcoming subsections the following aspects are discussed: the situation before the implementation, the AI-based medical device and the chronological process description. This process description is described by elaborating on several segments of this process. Namely, the purchasing process, the organization of the implementation, the implementation process, training CT-technologists, training radiotherapists and the situation after the implementation. Finally, the degree of successful implementation of this case is discussed and the two expectations are either confirmed or rejected according to the case results.

4.1.1 Before the implementation

The following paragraphs elaborate on the situation before the implementation within the department and the intervention. Normally, to implement a new device, the radiotherapy department would copy other, larger, radiotherapy departments. To illustrate, if another department bought a new machine and implemented it, the radiotherapy department would contact the other department to learn about their experiences and how they organized the implementation.

Patients who visit the radiotherapy department are treated for cancer. During a radiotherapy treatment, the patient lies still in a CT-scanner for about 15 minutes. Most radiotherapy treatment plans consist of 20 separate radiation treatments, referred to as fractions. During each fraction the tumor is radiated. The radiation kills tumor cells by delivering an adequate dose on the tumor location estimation. In order to kill as many tumor cells and as few healthy cells as possible, a plan is made beforehand to locate the tumor and the surrounding, healthy, organs. This plan is uploaded on the CT-scanner and the CT-scanner automatically alters settings so that it fits the plan. Consequently, the patient receives the planned amount of radiation on the right location.

Before the AI-based medical device was implemented, one plan was made for one patient for the total amount of fractions which the treatment plan contained. However, this is not the most effective and efficient way to treat patients using radiotherapy. A patient's anatomy changes from day to day and from moment to moment. This depends on factors like: bladder volume, the amount of air in the intestines and body position. Although a patient's anatomy varies, the plan could not be adapted to these variations. This can lead to more and heavier side-effects for the patient and sub-optimal patient outcomes. This was the reason members of the department started looking for an improvement of the radiotherapy treatments.

4.1.2 The AI-based medical device

The radiotherapy department bought the AI-based medical device, which is a software package, from a company which distributes this software on an international scale. The department already bought a specific CT-scanner from the company on which the AI-based medical device can be installed. The AI-based medical device exists entirely out of software, one part which is based on AI-technology and another part which is not. The AI-part of the software is based on machine-learning methods and changes through time discretely. Which means the software is only changed when the company releases an update and the update is installed by the users.

Using the AI-based medical device, radiotherapy changes to adaptive radiotherapy. The AI-based medical device exists out of software which is installed on a specific type of CT-scanner. With this software, CT-technologists are able to adapt the treatment plan from a patient to the patient's anatomy in that moment. The AI-part of the software is able to recognise and locate certain anatomical structures, so-called influencers, on a CT-scan. Normally, this task executed manually by radiotherapists and would take a few hours. Due to the implementation, the task is executed in a few minutes by CT-technologists while assisted by the AI-based medical device.

Therefore, due to the implementation of the AI-based medical device tasks are redistributed from radiotherapists to CT-technologists. Radiotherapists are medical specialists. Patients who are diagnosed with cancer are referred to them and these specialists create a suitable treatment plan. The CT-technologists work with the CT-scanners which execute these treatments, they are responsible for giving the treatment to the patient as planned by the radiotherapist. After influencers are located, an adapted radiotherapy plan is generated by the AI-based medical device and the treatment continues as planned. This is the non-AI part of the software. CT-technologists are able to change the location of the influencers as generated by the AI-based medical device, however they can not change the generated plan. This plan is always compared to and confirmed by a reference plan.

4.1.3 Chronological process description

Purchasing the AI-based medical device

Long before the AI-based medical device was implemented, the department had identified the struggle of the accuracy of their radiotherapy treatments. Consequently, the management-team started looking for an innovative device which could improve patient outcomes for patients receiving radiotherapy. The management-team consists of the department manager, the managing radiotherapist and a medical physicist. The management-team started looking for an AI-based medical device, as AI-technology lends itself perfectly for automatic adaption of a treatment plan to the anatomy of the patient in that moment. When debating about which AI-based medical device was the best fit to implement adaptive radiotherapy on the department, the management-team took a few aspects into consideration. To start, the duration of the radiotherapy treatment. The duration had to be minimized as this is more convenient for the patient and improves patient outcomes. It also enables the department to treat more patients in a shorter period of time. Furthermore, the quality of imaging needs to be sufficient to treat a wide range of soft tissue cancers. Lastly, the decision-makers wanted to buy the device from a company with whom they had a good relation.

The AI-based medical device which was eventually selected was applicable for CT-scans via which radiotherapy treatments were given. The advantage of CT-scans is the shorter duration of the scan compared to an MRI, which is the main competitor in the field of adaptive radiotherapy. However, the imaging quality is lower compared to the MRI. Despite this difference, the imaging quality of the CT-scan was still sufficient to treat a wide range of cancer types. The management-team also chose to buy the AI-based medical device from a company with

whom they cooperated with for several years already. They had bought multiple machines from the company and liked working together. In this sense, the management-team functioned as decision-maker.

When the management-team made the decision to buy the AI-based medical device, the device still needed approval of the Federation for Drug Administration(FDA). Because of this, the management-team had to sign a contract of confidentiality. As a consequence, the management-team was not able to share any details about the AI-based medical device with other co-workers on the department. Their colleagues had to trust the management-team in their judgement of the AI-based medical device. Members of the department described this as an exciting but anxious period. They trusted their management-team, but were also eager to get more information about the new device.

Consequently, the department manager and the company had to discuss about the terms of the selling contract. This was a clear responsibility of the department manager as this professional was also in charge of collecting all necessary funds within the hospital and putting new innovations on the agenda of the hospital. In that sense, the department manager functioned as a policy entrepreneur. One important aspect which was incorporated in this selling contract were the resources supplied by the company in order to guarantee the radiotherapy department a successful implementation of the AI-based medical device. The company agreed to subsidize a year salary from a technical medicine specialist. The technical medicine specialist operated as the implementation manager. This deal was of great importance for the department, without this resource the implementation would have failed. One important aspect to note here is that it was only possible to hire the technical medicine specialist, because a medical physicist within the department retired.

Organizing the implementation

When the AI-based medical device eventually arrived, the implementation of the device was organized. The department is relatively small and has limited resources. Normally, this department used to copy implementation methods from other departments. For this implementation, copying was impossible. Only one other radiotherapy department in the world had just started the implementation of this device. With support from the company, the department organized the implementation from scratch. Examples of guidance supplied by the company are presentations to inform members of the department, a simulator with which CT-technologists could train themselves and hiring a subsidizing of a year salary from a technical medicine specialist.

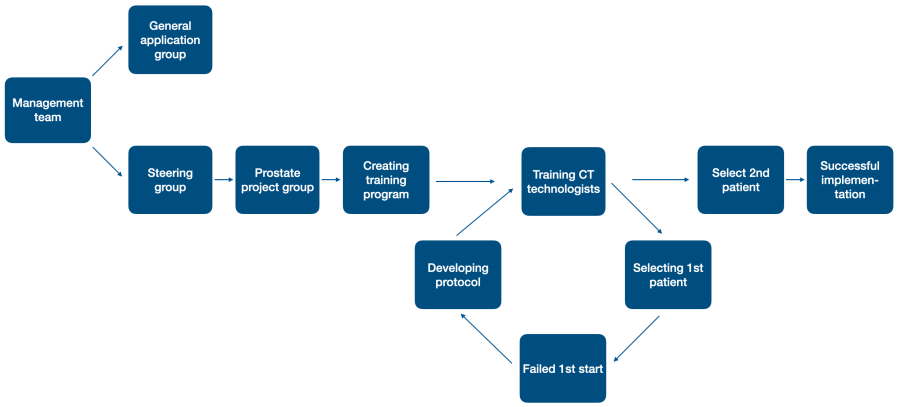

Figure 4.2: *Graphical overview of the organization and course of the implementation process of case 1.*

The implementation process was organized by setting up different project groups, which each had their own responsibility. A graphical overview of the organization of the implementation process of case 1 is illustrated by figure 4.2. These groups met frequently and a clear deadline was communicated by the management-team to the project-groups. By forming multiple groups containing professionals from multiple professions, tasks were distributed and a representative of each profession was involved during the implementation process. This representative had the responsibility to inform and update colleagues within the same profession in order for all professionals to be up to date. Not all professionals were involved as this slowed down decision-making processes and cost too many resources.

The implementation

As soon as the AI-based medical device arrived at the department, the management-team formed the steering group and the general application group, which is a manager task. The role of the steering group was to organize the implementation process and decide to which type of cancer the AI-based medical device is first introduced. The steering group consisted of a radiotherapist, a technical medicine specialist, two senior-CT-technologists and a medical physicist. Within the steering group, the technical medicine specialist was appointed as leader and manager of the group. This specialist organized the meetings, set-up the agenda's and kept track of progress. Also, the two senior-CT-technologists who were involved functioned as key-users. They were intended users of the device and contributed to the organization of the implementation. The role of the general application group was to install the AI-based medical device and do the commissioning. This means to get everything up and running within legally safe limits. This group consisted of a technical medicine specialist, a medical physicist and two senior-CT-technologists. After the AI-based medical device was fully installed and commissioned, this group stopped existing as their job was done.

The steering group decided the AI-based medical device was first introduced to patients suffering from prostate cancer. This enabled the formation of the prostate project-group. The prostate project-group executed the implementation of the AI-based medical device for prostate cancer patients. People from the following professions were involved: a technical medicine specialist, a radiotherapist, two senior-CT-technologists and a medical physicist. In this group, the technical medicine specialist operated in the same way compared to the steering group. Additionally, the technical medicine specialist functioned as a facilitator as cooperation between professionals from the multiple involved disciplines had to be promoted and ensured. This is discussed in more detail in upcoming sections. The main focus of the prostate project group

was to arrange the training program and to select which CT-technologist were trained, using this program.

To clarify, not all CT-technologists were trained to use the AI-based medical device. This meant too many CT-technologists were trained compared to how many actually use these skills in their daily work. This is applicable for many specific tasks and skills within the realm of CT-technologists. Therefore, within the group of CT-technologists there is a division of tasks. Every CT-technologist is specialized in one type of task, referred to as a specialization. This can either be planning, managing or a specific type of radiotherapy treatment. Adaptive radiotherapy using the AI-based medical device is also a type of specialization.

To select CT-technologists, several criteria were used to base this decision on. To start, it depended on the type of contract the CT-technologist had, part-time or full-time. Priority was given to full-timers, as this made it easier to make a schedule. Within this schedule always one CT-technologist must be present who is specialized in adaptive radiotherapy. Also, priority was given to CT-technologists who had experience with the planning specialization. Last, only CT-technologists could be selected if they did not have specialization yet or if they wanted, and if possible, to switch to another specialization.

Training CT-technologists

After a group of four CT-technologists was chosen, the training program started. The goal of this program was to train CT-technologists on the following areas: human anatomy, human anatomy on a CT scan, how to use the AI-based medical device and how to interpret outcomes generated by the AI-based medical device and adjust them when necessary. Therefore, the program consisted of four steps: (1) application training hosted by the company, (2) anatomy lessons hosted by radiotherapists, (3) emulator training which was done individually and (4) treatment sessions executed by CT-technologists and supervised by radiotherapists. CT-technologists were qualified to use the AI-based medical device individually if all four parts of the program had been completed.

When all CT-technologists had completed the first three steps of the training program, a patient was selected to continue with step four: supervised treatment of real patients. The first patient was selected and the treatment plan was prepared according to the new plan protocol using the AI-based medical device. Before the patient arrived for the first treatment, the radiotherapists checked the treatment plan generated by the AI-based medical device. However, the radiotherapists were not satisfied with the treatment plan. Especially the radiotherapist who was responsible for the patient saw some strange choices being made when locating the influencers. After discussing the treatment plan, the radiotherapists realized they all applied different rules when locating influencers. However, they had to reach consensus to start adaptive radiotherapy for this patient.

Retrospectively, the radiotherapists felt a lot of pressure to perform. They felt the whole world was looking at them because they were the second department worldwide to implement this device. Furthermore, it appeared the radiotherapists were not as actively involved as thought. The rules by which the influencers are located using the device were known and discussed beforehand. Due to high workload and an agenda filled with patient consults, it was hard for radiotherapists to prioritize the implementation. After a long night of discussing, the radiotherapists could not reach consensus and the treatment of that patient was canceled.

Training radiotherapists

Consequently, radiotherapists had the responsibility to reach consensus and make a decision which rules were going to be used when locating influencers on a CT-scan. Apparently, there have always been little nuances and differences between the radiotherapists. Normally, radiotherapists would frequently look at the plans their colleagues made, but the discussion about how influencers were located never started as they trusted their colleagues in their skills and responsibility. A reason for this change is that during the implementation, the plan was created by a CT-technologist and not a fellow radiotherapist. Radiotherapists did not fully trust the skills of the CT-technologist yet. Additionally, the radiotherapists, as a group, felt pressure and responsibility to perform. This forced them to start the discussion about how influencers must be located. Consequently, the group of radiotherapists started acting more pro-active and were more actively involved during the implementation.

Now that the task, locating influencers, needed to be redistributed to CT-technologists, the group radiotherapists had two responsibilities: the standardization of the task and the creation of a protocol with uniform rules. During this process, radiotherapists had to gain trust in the skills of the CT-technologists. To stimulate this process, a comparative research was organized and executed within the department. The skills of the radiotherapists were compared to the skills of the CT-technologists when locating the influencers. The results from this study show that CT-technologists are equally skilled or scored even better compared to the radiotherapists. Ultimately, this led to the convincing of radiotherapists that CT-technologists were competent to execute this task.

Next to promoting trust in the skills of CT-technologists among radiotherapists, some radiotherapists were also afraid more tasks would be taken away from them. Therefore, these radiotherapists had to be convinced that the task which was redistributed was a standardized task which could in fact be redistributed. Other tasks can not be standardized and will therefore not be redistributed. A radiotherapist with a lot of experience and authority within the group of radiotherapists took the lead during this process. The radiotherapist showed a lot of commitment to convince fellow radiotherapists about the potential of the device. Therefore, this radiotherapist is considered as a change agent.

As a result of the persuasiveness and commitment of the change agent, and several protocol meetings supported by the technical medicine specialist and the CT-technologists, the group of radiotherapists reached consensus after a few weeks. Together with anatomy lessons, the radiotherapists were guided by the technical medicine specialist who informed them how their decisions in locating influencers could affect the final treatment plan generated by the AI-based medical device. The protocol was finally formulated. In this situation the technical medicine specialist again operated as a facilitator as cooperation between different disciplines was ensured.

After the development of the protocol, anatomy training was reshaped using this protocol and the CT-technologists had to be retrained. This retraining was also led and executed by the change agent. After this training session was done, a second patient was selected and successfully introduced to adaptive radiotherapy. This type of radiotherapy, assisted by the AI-based medical device, has become routine for the CT-technologists. The department is now introducing more cancer types to adaptive radiotherapy and uses the same organizational structure with every implementation.

After the implementation

The radiotherapy department evolved due to the implementation process. According to the interviewees, the implementation caused members of the department to be more involved with professionals from other professions, besides their own work. Furthermore, before the imple-

mentation, the initiation of an innovation always started with a medical specialist. As discussed, this profession has a high workload, and as a consequence is not able to prioritize innovation and research. Therefore, innovations used to develop slowly on the department. Due to the implementation of the AI-based medical device, professionals from other professions became more assertive and started taking initiative more frequently. Professionals do not wait to receive a task from a medical specialist. But work out their ideas and bring this to the medical specialists in small operable pieces. This enables working on projects simultaneously per profession. Information streams and projects are managed parallel to each other per profession. Professionals from involved professions meet and update each other frequently. Consequently, professionals can work on the same project at the same time in stead of one after another. Altogether, interviewees described frequently how the implementation process has affected the whole department positively.

Degree of successful implementation

To evaluate the degree of successful implementation, three factors are discussed: goals, number of users and frequency of use. To start, the initial goal of the implementation was to improve patient outcomes who are treated with radiotherapy and decrease toxic side-effects of this treatment. These goals are reached.

To continue, around eight CT-technologists use the AI-based medical device and the device is used on a daily basis. Furthermore, patients suffering from other types of cancer are treated by adaptive radiotherapy, so radiotherapy assisted by the AI-based medical device. This shows the use of the AI-based medical device is adopted by more users. Consequently, this implementation is considered as successful.

4.1.4 Expectations

Expectation 1: If implementation leaders facilitate training about characteristics of the AI-based medical device for professionals of multiple disciplines within the medical department, trust in the device is promoted and all necessary changes are made and accepted. Therefore, the degree of successful implementation increases. Training was facilitated by the company and the technical medicine specialist who operated as a leader, manager and facilitator. The change agent and key-users also assisted during the facilitation of the training program. Training was necessary to improve skills of CT-technologists and to promote trust in the device among radiotherapists. Extra training was not necessary for CT-technologists to promote trust as their trust was stimulated through practical experiences with the device. Radiotherapists needed training to trust the skills of the CT-technologists and trust the device in terms of validity and reliability. Therefore, this expectation is confirmed.

Expectation 2: If implementation leaders, facilitators and/or change agents involve stakeholders and professionals from all disciplines whose workflow changes or who manage or execute these changes, during the development or implementation of an AI-based medical device, the degree of successful implementation increases. One of the obstacles faced during the implementation process occurred when a workflow change was forgotten. Consequently, professionals, for whom this workflow change was needed, were not involved. Therefore these professionals did not understand and accept the workflow change. As soon as these professionals were involved, the workflow change was accepted.

Efforts were made by the technical medicine specialist, who operated as leader, manager and facilitator, to involve radiotherapists. However, this profession group was not actively involved before a change agent was appointed within this profession group. Radiotherapists did not prioritize the implementation due to a high workload. However, their involvement was

needed to execute workflow changes and task redistribution. The change agent was able to actively involve the radiotherapists due to persuasiveness, grit and a sense of authority. Together, the facilitator, the change agent and the key-users were able to convince the radiotherapists of the validity and reliability of the device and the skills of the CT-technologists. Once all necessary professionals were involved, the device was successfully implemented. Therefore, the expectation is confirmed.

4.2 Case 2

To gather information about the implementation of the AI-based medical device, five interviews were organized with professionals from the following professions: technical medicine specialist, two nuclear medicine specialists, medical physicist and physician assistant.

The implemented AI-based medical device was developed by students and members of the medical department. Therefore, information was gathered about the development and implementation process. During the interviews, information was gathered about how these processes were organized, how these processes were executed, what events occurred and what effect the processes had on the department. Furthermore, relations between the multiple involved professionals were explored, next to relation between involved professionals and the rest of the hospital. The organization of the department and the hospital were addressed as well.

Collected documents were changed or developed due to or to facilitate the development and/or the implementation of the AI-based medical device. The collected documents for this case are: a scan acquisition protocol, quality documentation in line with Medical Device Regulation (MDR) criteria, two published articles about the development and implementation of the AI-based medical device and a master thesis from a technical medicine student. From these documents the organization of the development process was investigated which also clarify certain aspects of the implementation process.

In the following subsections the following aspects are discussed: the situation before the implementation, the AI-based medical device and the chronological process description. This process description is described by elaborating on several segments of this process. Namely, the initiation of the development process, the development process, workflow integration and the situation after the implementation. Finally, the degree of successful implementation of this case is discussed and the two expectations are either confirmed or rejected according to the case results.

4.2.1 Before the implementation

The nuclear medicine department receives requests from other departments within and outside of the hospital in which the nuclear medicine department is situated. Via these requests the nuclear medicine department is asked to execute a diagnostic intervention on a patient who is treated or to be treated by the department from which the request originates. Once the diagnostic intervention has been executed, the nuclear medicine specialist writes a report based on results from the intervention. This report, and the results from the intervention, are sent to the requesting department. Consequently, the referring medical specialist makes a decision concerning a diagnosis or treatment for the patient, which is based on the report from the nuclear medicine specialist.

The hospital, in which the nuclear medicine department is situated, is part of a hospital network which consists of multiple hospitals which are located in the same region. Nuclear medicine specialists, who work for one of these institutions, are tied by the same specialist section. This means they do not have one hospital in which they work, but they travel between multiple hospital within the hospital network. This only applies for the nuclear medicine specialists.

One of the diagnostic interventions, in which the nuclear medicine department is specialized, is the Dopamine Transporter-scan (DAT-scan). A DAT-scan is requested by a neurologist and aims to give more diagnostic certainty if a patient is suffering from Parkinson's or not. A neurologist only requests this scan if it is unclear whether a patient is suffering from Parkinson's disease or not. These situations occur when the patient is suffering from symptoms which are in line with Parkinson's disease, but do not respond to prescribed medication as expected.

Before, the implementation of the AI-based medical device, only highly experienced medical specialists could review the scans. This is caused by multiple factors. Reporting the DAT-scan consists of a visual part, looking at the scan image, and a quantitative part, looking at ratio's calculated by a scan algorithm. However, nuclear medicine specialists only used the visual part, whereas the quantitative part was neglected. By only using the visual part, only a highly experienced medical specialist was able to report the DAT-scan. Furthermore, variability between medical specialists was present when reporting the DAT-scan. This variability is present due to differences between risk-adversity, previous experiences and education.

4.2.2 The AI-based medical device

The department wanted to explore how AI-technology could help them improve diagnostic quality of the DAT-scan. To decrease the variability between medical specialists and increase the diagnostic certainty of the DAT-scan, the quantitative part of the scan was integrated into the workflow and automated by an AI-based medical device. To do this, the AI-based medical device is based on machine learning methods. These methods use the quantitative part of the DAT-scan to calculate the probability of Parkinson's disease for that patient. To use the AI-based medical device, users have to copy the ratio's from the DAT-scan into the AI-based medical device. The AI-based medical device generates an outcome. Consequently, users have to copy the outcome into the Electronic Health Record (EHR), which is sent to a neurologist who takes it into account during diagnosis. As an effect of the implementation, medical specialists, who used to report "scan could possibly fit description of Parkinson", now report: "increased risk…" or "very high suspicion of…". This report is accompanied with the outcome of the AI-based medical device. Together, this information gives more diagnostic certainty to neurologists which they can base their final conclusion on.

4.2.3 Chronological process description

Initiating the development process

To improve diagnostic quality and explore the possibilities of AI-technology for the nuclear medicine department, three members of the department decided to start a project to develop an AI-based medical device. This development process was initiated by the following three professionals; a technical medicine specialist, a nuclear medicine specialist and a medical physicist. Together this group functioned as decision-makers. Each professional had a different role. The technical medicine specialist managed the project and investigated the needs of the neurology department and the status quo of the current DAT-scan. Therefore, this professional operated as a manager. The nuclear medicine specialist was responsible for putting AI-technology and the development of this device on the agenda within the hospital, locating necessary resources within the hospital and the implementation process. Therefore the nuclear medicine specialist operated as policy entrepreneur and implementation leader.

The AI-based medical device was developed inside of the nuclear medicine department. Consequently, the department had to change, develop and implement multiple routines in accordance with Medical Device Regulation (MDR) criteria. Examples of such routines are, monitoring the use and functionality of the device and a manual for neurologists, nuclear medicine

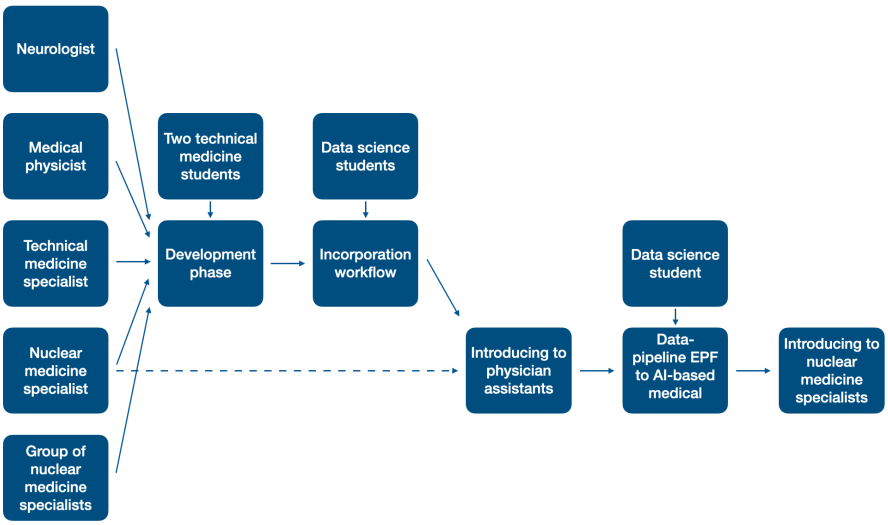

Figure 4.3: *Graphical overview of the organization and course of the implementation process of case 2, the dashed line relates to how the nuclear medicine specialist was responsible for the implementation phase starting from the introduction of the AI-based medical device to the PA's.*

specialists and PA's . The medical physicist was responsible for keeping track of and documenting to be in line with the rules and regulations of the MDR. Next to that, this professional investigated what internal routines were necessary to monitor the functioning of the device, for example. In that sense, the medical physicist also operated as a manager, but on a different aspect compared to the technical medicine specialist.

The project-group, consisting of the technical medicine specialist, the nuclear medicine specialist and the medical physicist, needed more professionals actively involved before starting the development process. According to the project-group, the nuclear medical specialists group was the most important group to get involved. The group of nuclear medicine specialists only engaged if they recognized the development of the device as clinically relevant. Furthermore, nuclear medicine specialists were intended users. In the eyes of the project-group, involving users during the development process was important, because this would make the development and implementation more successful. Involving the nuclear medicine specialists marked the start of the development and implementation process, represented by figure 4.3.

Once the nuclear medicine specialist group was convinced about the clinical relevance of the device, the three project-groups involved one more actor in the development process. This actor is a neurologist from another hospital within the same hospital network. Neurologists from the neurology department of the same hospital as the project-group were not interested in cooperating in this project as their field of research and interest did not match this project. The neurologist, and the belonging department, who were involved were specialized in Alzheimer and other neurological movement disorders, like Parkinson's disease. Therefore the field of research and interests matched this project and involvement would benefit the neurologist as well. The role of the neurologist was mainly to articulate the needs and wants of the neurology department. So what outcome of the project would fit these needs and wants. Together, the neurologist and the nuclear medicine specialist who also operated as a policy entrepreneur were the key-users of the project. These two professionals would ultimately use the device or the outcomes of the device and contributed to the development and implementation process as a whole.

No professional within the nuclear medicine department contained the know-how or the resources to develop an AI-based medical device, the department attracted technical medicine students for an internship at the department while developing the device. Two technical medicine students did an internship, sequentially, and developed the device. A data science student eventually incorporated the device into the workflow of the DAT-scan.

The development process

The development process started when the first student started an internship. To guide this student, all members of the project-group had their own responsibility. To start, the technical medicine specialist took the lead in guiding the student. The nuclear medicine specialist assisted and supervised the student during clinical activities. The medical physicist guided the student during questions concerning the MDR. The neurologist did not have a specific role during the internship of the student as this neurologist had to leave the project prematurely. The group of nuclear medicine specialists were only updated every three to four months about the progress. Especially the changeability, controllability, validity, reliability and interpretability of the AI-based medical device were important factors about which the medical specialists wanted information.

The student and the project-group ran into some issues which they could not solve themselves. Therefore, they reached out to an academic hospital. During several meetings with professionals from the academic hospital, the project-group and the student got advice. This hospital had more experience and people with more AI expertise. Ultimately, the development of the device reached the validation and test phase. The device was validated and tested internally, with data from the department, as well as externally, with data from another hospital.

Once the device was validated and tested, the internship of the technical medicine student came to an end and the student left the department. Before the implementation could start, the device needed some final adjustments. To do this, another technical medicine student started an internship. This student worked on two projects, to translate the device into a web-application which was accessible for all users and to develop another AI-based medical device for the nuclear medicine department. When this internship ended, this student left as well. Now that two students left, a lot of knowledge and understanding about the device was lost. Moreover, because no student needed guidance anymore, the project-group lost part of their purpose and involvement in the project.

Workflow integration

Once the device was translated to a web-application, the implementation process could start. Before nuclear medicine specialists started using the AI-based medical device, the usability of the device needed improvement. In the eyes of the medical specialists, too many manual actions were needed to use the device. To be more specific, ratio's had to be manually copied into the device. These actions decreased the usability of the device and was prone to human errors. Usability was very important because medical specialists have a very high workload. Therefore, they only accept workflow changes if it decreases workload. To adjust this, a data science student was attracted for an internship. When guiding this student, the nuclear medicine specialist realized that guiding a data science student asked for a different type of guidance compared to a technical medicine student. This was caused by the lacking medical knowledge and way of reasoning of this student. Therefore, this process was more time-consuming than expected. During this process, the nuclear medicine specialist, who also operated as policy entrepreneur, key-user and decision-maker, now also operated as implementation leader. This professional was responsible for satisfying the demands of the profession group who were intended users

of the device, namely the medical specialists. With the ultimate responsibility to successfully implement the device.

The implementation process

In the meanwhile, the implementation leader introduced a physician assistant(PA) to the AI-based medical device. At first, this profession, PA, was not indicated as an intended user. But during the process, they were interested and an opportunity occurred as one of the PA's was personally educated by the implementation leader. PA's have a lower workload compared to nuclear medicine specialists. Furthermore, this profession required less training about validity and reliability of the device to promote trust. The PA's do not carry any liabilities, which nuclear medicine specialists do [Staten Generaal, 1993]. PA's only had to trust nuclear medicine specialists in their judgement and responsibility. Before specializing to become a PA, these professionals often have several years of experience as a scan-technologist. This results in a technical and pragmatic knowledge background. According to a PA, there was no issue with the usability of the device. Also, the nuclear medicine specialist checks the reports written by the PA's as a second check.

To train PA's, the nuclear medicine specialist explained how a high quality scan is created, how to estimate if the quality of the scan is sufficient, how to adjust scan parameters to increase scan quality and how to report the DAT-scan using the device. Also PA's were taught how to interpret the outcomes generated by the AI-based medical device. No manual was shared, although this is obligated by the MDR. Although PA's are able to work with the AI-based medical device, there is a desire to get more information about AI-technology.

During training more emphasis was put on evaluating scan quality and scan acquisition compared to interpreting the outcomes generated by the device. The quality of the outcomes generated by the device are highly dependent on scan quality and scan acquisition. The device generates an outcome based on numbers calculated from the result of the scan. To calculate the numbers correctly, certain settings need to be applied consistently by the scan-technologists. However, this does not always happen. Therefore, the person who reports the scan always needs to check the settings which are applied by the scan-technologists. If these are the wrong settings, they need to be changed afterwards and the numbers need recalculation, based on the corrected settings. This is possible without requiring the patient receive another scan. However, PA's and nuclear medicine specialists do need to be aware of this. During the internship of the first technical medicine student a protocol was developed which should standardize the application of settings of the DAT-scan by the scan-technologists. This has improved the scan acquisition. However, sometimes scan-technologists forget to use the protocol. Hence, people who report the scan still need to be aware of the quality of the scan acquisition.

Now that PA's use the AI-based medical device, the usability still needs to be improved before nuclear medicine specialists are introduced to the device. To manage this, the initiating medicine specialist wants to attract another data science student who is able to create a data pipeline between the EHR and the AI-based medical device. With this data pipeline, the numbers needed for the AI-based medical device are automatically transferred from the EHR to the device. Therefore, no manual steps need to be taken by users. The initiating medical specialist estimates this will be accomplished in March 2022.

Effect of the implementation

One of the effects of the development and implementation process of the AI-based medical device has been the formation of an AI project group in the imaging department and the promotion of AI technology among different departments inside the hospital. The development process of the AI-based medical device has constructed relations between members of different professions. This resulted in the formation of the AI project-group on the imaging department. This project-group consists of a researcher, two technical medicine specialists, two radiologists, a nuclear medicine specialist and a medical physicist. They have monthly meetings in which they discuss current on-going projects and new developments within the two imaging departments, nuclear medicine and radiology. Furthermore, they discuss how all necessary resources can be collected in order to facilitate these projects. There are two ways in which this is done. First, by involving the right professionals. This is done via the formation of the AI project-group, but they also invest in a close relationship to the Chief Medical Information Officer (CMIO). This professional is responsible for the organization and management of all ICT related projects inside of the hospital. Examples are software applications, the EHR and AI-related research projects. By investing in a close relationship between the AI project-group and the CMIO, the group is able to collect financial resources, facilities like ICT access and hardware.

The second approach used to collect resources is via promoting the AI technology inside the hospital via the internal website and in-congress meetings. By promoting the possibilities and benefits of the use of AI-based medical devices the project-group puts AI technology on the agenda. To add, the AI project-group shows the contribution of the device to the nuclear medicine department. This should stimulate future investments in AI-related research. Moreover, this promotes the AI project-group as well so that other AI research groups are familiar with this AI project-group, this promotes collaboration.

Degree of successful implementation

To evaluate the degree of successful implementation, three factors are discussed: implementation goals, amount of use and number of users. To start, the initial goal of the project-group was to learn more about AI technology and how the nuclear medicine department could benefit from the use of AI-based medical devices. This goal has definitely been reached successfully. Additionally, the goal of the device is to increase diagnostic quality of the DAT-scan. This goal as also been reached according to the implementation leader.

To continue, the number of users is more than 1. This number is expected to increase according to the implementation leader, but will not increase according to the intended users. According to the implementation leader, intended users do not use the device due to usability issues. However, intended users are not convinced about the clinical relevance of the device. The professionals which use the AI-based medical device, use it on a weekly basis. To conclude, the implementation of the AI-based medical device succeeded as the device is used regularly and implementation goals have been reached. However, the intended users do and will not use the AI-based medical device. Therefore, the implementation is not considered as 'successful' at this moment.

4.2.4 Expectations

Expectation 1: If implementation leaders facilitate training about characteristics of the AI-based medical device for professionals of multiple disciplines within the medical department, trust in the device is promoted and all necessary changes are made and accepted. Therefore, the degree of successful implementation increases. In this case, a distinction is made between intended users and users. Intended users were trained during their involvement in the development process of the device. During this training, trust was promoted and intended users were convinced about the validity and reliability of the device. However, intended users were not convinced about the clinical relevance of the device and its usability. Other professionals, who were involved during the development process, were convinced about the device.

To continue, users were trained how to use the device. Users did not receive any training about other device characteristics, but these users trusted the judgement and responsibility of their superior. Therefore, it was not needed to promote trust through training about the device characteristics of the AI-based medical device. As the intended users were not convinced to use the device, despite their training, and the users had a different training demand, this expectation is rejected.

Expectation 2: If implementation leaders, facilitators and/or change agents involve stakeholders and professionals from all disciplines whose workflow changes or who manage or execute these changes, during the development or implementation of an AI-based medical device, the degree of successful implementation increases. Stakeholders were to some extent actively involved during the development process of the device as they cooperated during resource collection. However, their involvement was limited to the development phase. Professionals whose workflow would change during the implementation of the device were involved during the development process. However, due to lacking clinical relevance and usability these professionals did not accept workflow changes. Professionals who manage or execute workflow changes were involved during the development and implementation process. As a consequence of unaccepted workflow changes from one profession group, the implementation leader involved another profession group. This group did accept workflow changes, which made implementation successful. These professionals eventually used the device and were not involved during the development and implementation process. Therefore, involvement was not necessary in this case to implement the device among one profession group and involvement did not stimulate successful implementation for another profession group. Therefore, the expectation is rejected.

5 DISCUSSION

This section elaborates on the validity of the research by evaluating the used academic literature and methods to collect and analyze empirical data. By evaluating the validity and reliability of the research, an estimation about the generalizability of research conclusions is made. Additionally, research results are further interpreted in light of academic literature. After interpreting the results, limitations of the research are discussed. Consequently, implications for practical situations and the academic sphere are elaborated on. Ultimately, suggestions for future research are made.

5.1 Validity and reliability

The validity and reliability of this research is evaluated by critically reviewing the NASSS framework and other academic literature used to form the theoretical framework. Consequently, the presence of bias, and its influence, during data-collection and analysis is evaluated. Lastly, the data-collection and analysis are evaluated as a whole.

To start, the backbone of the theoretical framework is based on five domains of the NASSS framework as defined by Greenhalgh et al. (2017). This framework lends itself perfectly when investigating implementations of medical devices in general as a complex system. Although some NASSS domains were excluded from this research, the framework was applicable and valid. Only concepts within included domains were used if they re-occurred in other academic literature, concerning the implementation of AI-based medical devices, and the empirical data. Concepts were excluded if they were not measurable with the collected data. By focusing on a limited number of domains and concepts, the usability of the NASSS framework improved.

To improve validity and reliability of the theoretical framework, the NASSS framework was complemented by other academic literature. This literature mainly consisted of implementation studies which specifically evaluated implementation successes and failures of AI-based medical devices, implemented in hospitals. Additionally, studies were included which explored AI-based medical device characteristics which influenced users and influenced how these devices were integrated in a medical department with all its organizational routines and workflows. AI-technology is in the early-adaptor stage, therefore most cited studies were recently published. This makes the theoretical framework contemporary.

When doing qualitative research, which uses interviews and internal documents as a data source, bias is present within the data-set. The used research methods aimed to minimize the influence of bias. However, during the data-collection some events occurred which had an influence on the data-collection. In light of the COVID-19 pandemic, government restrictions made it difficult to organize interviews physically. For case 2, all interviews were online. Consequently, a number of issues arose which influenced data-collection. During the online interviews, less information was gathered about the medical department. It was harder to estimate the atmosphere within the department, the organization of the department and to reach out to other professionals. Additionally, questions were more often misinterpreted. As a consequence, the online interviews often took longer as more time was needed to gather all necessary information

and to get an understanding about the organization of the department and which professionals were involved. Therefore, the results from case 2 might be less reliable compared to case 1.

During the interviews for case 1, which were in-person, some issues occurred which influenced data-collection. Namely, some of these interviews were interrupted. The influence of these interruptions was minimized by noting down the specific topic when the interview ended and by re-playing audio recordings. If a professional had to leave the interview, another interview was planned to resume. Before the follow-up interview started, the interviewer gave a short summary of discussed topics to start the conversation where it had ended.

The influence of retrospective bias was noticed during the data-collection process. During the online interviews, gathered data from participants consisted of discrepancies. To illustrate, most participants were very enthusiastic and positive about the implementation process. However, when going into detail about certain events during this implementation process, they were less positive. The verifying interviews and document analysis were useful to solve discrepancies and misconceptions. During these verifying interviews, the interviewer prioritized asking about facts instead of experiences to minimize the influence of retrospective bias. Also the discrepancies where discussed with participants to investigate their perception of the cause of these discrepancies. An important note to make here is that the cause of these discrepancies and misconceptions is either retrospective bias or the fact that these interviews were online instead of in-person.

Moreover, participants were influenced by their surrounding. Some participants were interviewed twice. To illustrate, when a participant was stressed questions were more often misapprehended and shorter answers were given. By interviewing participants more often, the influence of their surroundings was limited. Therefore, reliability and validity of data-collection, data-analysis and results improved.

Overall, the interviews were successful. The collected data was useful to analyze. After the analysis, the results contained enough information to reject or confirm expectations. Participants made an effort to provide the interviewer with as much information as possible. Within both cases participants were very helpful to reach out to colleagues, when thought necessary for the research. Nevertheless, in some situations it took long to get a response from these colleagues, which prolonged the data-collection phase of the research.

Although used theories and methods prove to be valid to some extent, when generalizing the conclusions three aspects need consideration. First, a qualitative case-study does not lend itself to base causal explanations on [Yin, 2011]. Second, as only two cases were included in this research the variation between and within the cases with respect to device-, user- and organizational characteristics is limited. Therefore, conclusions can not be generalized to all implementations of AI-based medical devices in medical departments of Dutch hospitals. Third, in accordance with Strohm et al. (2020), the organization of health care and medical departments is very much bound to the social, economic, political and cultural context of a country. Therefore, the results of this research can not be generalized to other countries or academic hospitals.

5.2 Interpreting results

In this section, both expectations are elaborated on in light of the case results, the differences between the two cases and academic literature. In figure 5.1, an overview is given about the confirmation and rejection of expectations per case. Both expectations consist of combined elements, which are referred to as layers. These layers are evaluated separately. With this strategy, the validity of the expectations is evaluated as well.

Results expectations per case

Expectation	Case 1	Case 2
1. Training		
2. Multi-disciplinary involvement		

Figure 5.1: *Overview of rejection and confirmation of expectations per case.*

5.2.1 Expectation 1: Training

This expectation consisted of the following layers: the implementation leader as actor who facilitates training, the training concerns device characteristics, professionals from multiple disciplines within the department are trained and due to this training trust in the device is promoted and all necessary adaptations are made and accepted. To start, both cases consisted of an implementation leader who also took on different roles, for example in case 1 the leader also operated as facilitator. These implementation leaders were responsible for giving direction to the medical department and think of all necessary steps which were needed to take during the implementation, which is in line with the definition stated by Laukka et al.(2020). An example of such a step is training all members of the department about the implementation. In that sense, this element is satisfied by results from both cases.

Although in different form, in both cases training possibilities were offered to professionals from multiple disciplines. In case 1, professionals from all disciplines within the medical department received training from either the company or their colleagues. This training mainly considered how the implemented device would change existing workflows, which was a consequence of device characteristics in some situations, which is in line with several studies [Arora, 2020, He et al., 2019, Chua et al., 2021]. To give an example, in case 1 the workflow of the medical physicist changed due to the changeability of the algorithm within the AI-based medical device. Due to this changeability, the medical physicist had less freedom when altering radiation settings. In case 2, professionals from most disciplines within the medical department were trained. This training occurred during the development phase of the device. In that sense, the internally developing an AI-based medical device is a training in itself from which professionals within the department gained knowledge about AI-technology, which is also reported by He et al. (2019). During the implementation, users were trained to some extent about the usability of the device. Users were not trained about other device characteristics, which was no issue due to different liabilities and their trust in their superior.

As a consequence of facilitated training, trust in the device was promoted in both cases. As professionals from multiple disciplines and intended users engaged during the test and validation phase of the device from case 2, they were convinced of its validity and reliability. However, the intended users were not convinced about the clinical relevance and usability of the device. This contradicts results from several studies who state that user involvement during the development phase of a device improves usability and convinces users of clinical relevance [Asan and Choudhury, 2021, Phansalkar et al., 2010, Felmingham et al., 2020]. As a consequence, workflow changes, which were needed to implement the device, were not accepted by intended users. In case 1, trust in the device was promoted through training. In line with several studies, in this case the implementation of the device causes task redistribution due to automation [Shaw

et al., 2019, Tartar et al., 2021]. Between professionals from the two professions among which tasks were redistributed, trust needed to grow during the implementation. Additionally, professionals mainly needed training to assist during adapting workflows and standardizing tasks in order to redistribute them successfully. This is supported by the literature [He et al., 2019, Chua et al., 2021]. When trust was eventually promoted, workflow changes were accepted by medical specialists. Lastly, the apprehensive attitude of medical specialists towards the redistribution of tasks is in accordance with several studies [Tartar et al., 2021, Strohm et al., 2020]. However, these studies focus on losing autonomy and liability issues as a cause for apprehensiveness whereas in this research trust was mentioned as the main cause.

In both cases, most elements of this expectation were present and influenced the degree of successful implementation. This increases the validity of the expectation. However, within case 2, training about device characteristics was not sufficient to promote acceptance of workflow changes as intended users where not convinced about the usability and clinical relevance of the device. Therefore, the expectation was rejected for this case.

However, these two aspects are device characteristics, so three considerations are important. First, the facilitated training was not enough to convince intended users about the usability and clinical relevance of the device. Second, the usability and clinical relevance of the device indeed were not enough to improve diagnostic quality of the DAT-scan. However, physician assistants do use the device and are convinced of its added value. Between these two profession groups there is a difference in working standards, which may explain this difference. Physician assistants are less busy compared to medical specialists, which could cause a more apprehensive attitude towards workflow changes among the group of medical specialists. Third, there is another factor which moved intended users towards rejecting workflow changes. One possible factor is fear of losing autonomy and when AI-technology is introduced, which is reflected by academic literature [Gerke et al., 2020, Tartar et al., 2021, Strohm et al., 2020]. If the last consideration is true, then training must focus more on promoting trust among medical specialist. In academic literature, trust is a concept influenced by perception and promoted through training and spread of information [Asan and Choudhury, 2021, Chew and Achananuparp, 2022]. Where Asan and Choudbury (2021) state that trust is promoted if it is more clear where the AI based its outcome on and ifs performance is sufficient. Chew and Achananuparp (2022) state that trust engendered if the maturity of AI-technology increases and if patient safety is guaranteed. These concepts are described by these articles in light of how professionals perceive these concepts. However, according to the interviews trust is a concept promoted through experiences.

5.2.2 Expectation 2: Multi-disciplinary involvement

This expectation consisted of the following elements: the implementation leader, facilitator and change-agent as actors who actively involve other professionals, stakeholders and professionals become actively involved, professionals whose workflow changes or who manage and execute these changes are actively involved and professionals are involved during the development and/or implementation phase of the device. To start, in both cases the implementation leader showed commitment to actively involving other professionals from multiple disciplines. In case 1, the implementation leader tried to actively involve other professionals via meetings, updates and project-groups. However, before medical specialists were actively involved, appointing a change-agent among medical specialists was necessary. In case 2, both the implementation leader and the facilitator contributed to involving other professionals. Especially the technical medicine specialist who involved neurologists. How professionals operated via their roles is in accordance with definitions stated by Laukka et al. (2020).

Despite efforts made by involving actors, in both cases actively involving other professionals was difficult. In accordance with multiple studies, case 1 illustrated the importance of informing and updating all professionals whose workflow changes [He et al., 2019, Chua et al., 2021, Arora, 2020]. When workflow changes and professionals were forgotten, the implementation process stagnated. Moreover, appointing a change-agent was needed to involve medical specialists actively. This is also reported by Strohm et al. (2020) who state involving medical specialists is difficult due to high workload and Felmingham et al. (2020) who advocate for appointing a change-agent to involve colleagues. Apart from professionals within the department, also stakeholders were involved during the implementation in case 1. Members of the company operated as stakeholders as they invested resources to stimulate successful implementation of their device. Apart from professionals within the department and the company, no other important stakeholders were involved.

Within case 2, multiple stakeholders became involved during the development process of the device. This involvement was needed to collect all necessary resources. Developing an AI-based medical device which fit MDR criteria costed case 2 a lot of resources, slowed down the development process and caused uncertainty. Therefore, the MDR is seen as a factor which complicated and slowed down implementation, which, according to the NASSS framework, is a hindering factor [Greenhalgh et al., 2017]. However, this is not line with findings from Strohm et al. (2020), who state that the criteria of the MDR were no hindering factor for successful implementation for their seven included cases. Nevertheless, two studies state that developing an AI-based medical device internally requires many resources [Shaw et al., 2019, Beckers et al., 2021]. One of which advocates for medical physicists as the best suitable profession group to take the responsibility of managing all changes needed to develop a device in accordance with the MDR [Beckers et al., 2021]. This is in line with the task distribution from case 2.

Apart from stakeholder involvement, involving all professionals whose workflow changes was less successful in case 2. Medical specialists were not convinced about the clinical relevance and usability, therefore they rejected workflow changes. Furthermore, communication and collaboration between the neurology and nuclear medicine department lacked according to users. In contrast, Strohm et al. (2020) state that collaboration between the diagnostic department and the department from the referring medical specialists is necessary to implement an AI-based medical device successfully.

One other important aspect is deducted from the results of case 2. Although literature states involvement during the development process is needed to increase usability and to promote clinical relevance of a device, the results from case 2 illustrate differently [Asan and Choudhury, 2021, Phansalkar et al., 2010, Felmingham et al., 2020]. The results of this case presented that involvement is not enough to promote acceptance of workflow changes. Involvement in this case was more useful for the collection of resources. Therefore, expectation 2 for case 2 is rejected.

In both cases, most elements of this expectation were present. Also, all elements showed to influence the degree of successful implementation to some extent. However, the specifically defined role of the actor, who actively involves other professionals, may not be as valid. The responsibility to involve other professionals during an implementation fits the definition of multiple roles as stated by Laukka et al. (2021). For example, the policy entrepreneur, the facilitator or the change-agent [Laukka et al., 2020]. Also, the two expectations show interrelatedness. Involvement is not enough to promote workflow changes, training is necessary in addition. Therefore expectation 2 is less valid than expectation 1.

5.2.3 Successful implementation

In this research, the degree of successful implementation is used to express how successful an implementation of an AI-based medical device is. To measure the degree of successful implementation, three concepts are used: organizational goals, number of users and frequency of use. During the interviews, the number of users and frequency of use were well discussed concepts. Organizational goals was less of a concern for professionals who were included in this research. Only professionals who had a management function within the department were concerned about the goal attainment.

Regarding the scope of this study, which aimed to explore mechanisms present during the implementation of AI-based medical devices, by users, it would have been more fitting to limit the definition of successful implementation to concepts on a user level. Additionally, the concept 'organizational goals' was less informative as these goals were often influenced by the national agenda or the agenda within the hospital. Within this scope, the number of users and the frequency of use were valid and reliable concepts to measure the degree of successful implementation.

5.3 Other relevant observations

Within the cases, three other interesting mechanisms were present, which concerned the relation between understandability and explainability, applicability of human factor research and the monitoring of the AI-based medical device after implementation. To start, within both cases, a question was raised to what extent the user and the medical specialist, if not the same, need to understand the AI-based medical devices in order to interpret and explain the outcomes generated by it. According to Chua et al. (2021), the degree to which understanding is necessary about the outcomes generated by an AI-based medical device needs adaptation to the intervention for which it is used. They state for most applications of AI-based medical devices, full understanding is not necessary to explain outcomes generated by the device. However, explainability is an important factor for medical specialists due to their liabilities [Staten Generaal, 1993, Amann et al., 2020, Gerke et al., 2020]. Especially in case 1, the medical specialists had a very strong urge to understand the AI-based medical device. Throughout the implementation process, this was a hindering factor about which other professionals expressed their doubts. Especially CT-technologists wondered whether full comprehension about the device is necessary in order to interpret and explain its outcomes. These results are in line with results from multiple studies [Amann et al., 2020, Gerke et al., 2020].

Additionally, the literature mainly focuses on human factors which influence users during the implementation of AI-based medical devices [Phansalkar et al., 2010, Asan and Choudhury, 2021]. However, this research implies human factors apply to all involved professionals within the medical department during the implementation of the AI-based medical device. To illustrate, organizing an implementation process of an AI-based medical device is a labour-intensive task. Especially as workflow changes are needed for professionals from multiple professions. Therefore, it is emphasized to consider the influence of human factors for users and other professionals within that department when implementing an AI-based medical devices.

Lastly, in both cases the monitoring and evaluating of the use of the AI-based medical device was an important factor after the implementation process, which is in accordance with two studies [Manlhiot et al., 2021, van Beest et al., 2020]. During this phase, the implementation is improved if necessary. Regarding case 2, these monitoring and evaluating tasks have not been distributed among the department, which concerns some professionals. Furthermore, this makes the use and the performance of the AI-based medical device untraceable, which might have caused the different perceptions of the implementation of this device.

For case 2, the collection of necessary resources was important to successfully develop and implement the AI-based medical device. Resources were collected on the level of the hospital, which exceeds the level of the medical department. Although beyond the scope of this research, these results show the role of the hospital during the development and implementation of an AI-based medical device.

5.4 Limitations

Apart from exploring interesting and relevant mechanisms which are present during implementation of AI-based medical devices, this research also contains limitations. This research is based on a case-study, in which two cases were included. Therefore, the amount of variety within the data is limited. Per case about six interviews were held and documents were collected which provide a complete and detailed overview of the implementation of the AI-based medical device in that department.

Moreover, no academic hospitals were included in the research. These hospitals often have more resources to facilitate research and have different implementation strategies. Which could create an environment in which change is more welcomed by professionals. Including academic hospitals in future research is recommended as it offers a larger variety of organizational structures, available resources and implementation strategies. However, as this research did not include academic hospitals, research conclusions are not generalizable for academic hospitals.

5.5 Implications

The results of this study imply that the NASSS framework is applicable in the context of this research, ergo the successful implementation of AI-based medical devices in Dutch hospitals. It illustrates that domains within the NASSS framework can be selected to fit the scope of the research. Besides the NASSS framework, also human factor research proves to be applicable in this context. Moreover, the two complement each other beneficially in light of characteristics unique for AI-based medical devices, such as trust and explainability.

Additionally, the literature mainly focuses on human factors which influence users during the implementation of AI-based medical devices [Phansalkar et al., 2010, Asan and Choudhury, 2021]. However, this research implies human factors apply to all involved professionals within the medical department during the implementation of the device.

Lastly, this research offers medical departments tools to organize the implementation of AI-based medical devices in order to increase the degree of successful implementation. An important note to make here is that the device-, user- and organizational characteristics of that medical department need to match the characteristics as described in the case-study of this research. Also, results of this research can be used as an example for medical departments who want to develop an AI-based medical device internally or implement such a device.

5.6 Suggestions future research

On the basis of this study, three suggestions are made for future research. To start, during the formation of the theoretical framework and empirical analysis, other organizational mechanisms which influence the degree of successful implementation of AI-based medical devices were found. However, these mechanisms were not explored as the collected data contained not enough variety to base conclusions on. Examples of such mechanisms are the influence of the relationship between the patient and the medical specialist. To illustrate, in both cases the patient was not informed about the use of the AI-based medical device during the intervention as the intervention itself was not changed and professionals of the department decided it was not necessary to inform patients. In the literature, questions are raised in the context of AI-based medical devices when a patient should be informed about its use [Gerke et al., 2020, Amann et al., 2020].

Lastly, future research should use other research methods to evaluate the influence of the organizational mechanisms, as described by the two expectations, and the degree of successful implementation. By using other research methods, like quantitative research, conclusions can be formulated about the causality between the concepts described in this research. Additionally, these expectations should be tested on a larger variety of data to improve generalizability of the expectations. Improving generalizability promotes the applicability of the organizational mechanisms to a larger variety of practical situations. This offers professionals within hospitals tools when implementing AI-based medical devices.

6 CONCLUSION

The aim of this study was to explore how mechanisms, on an organizational level, influence the successful implementation of AI-based medical devices in medical departments of Dutch hospitals. To explore these mechanisms, first, a conceptual definition of successful implementation, by users, of AI-based medical devices was formulated. Within the scope of this study, successful implementation is defined when the device is used by more than 1 professional and used on a daily weekly basis, minimally.

With the second and third sub-question, this study explored how device-, user- and organizational characteristics are, theoretically, expected to interact. And how mechanisms on an organizational level result from the interaction between device-, user- and organizational characteristics. As a result of answering these questions, two expectations were deduced from the theoretical framework created during this study. The two expectations, based on a theoretical analysis, show how two organizational mechanisms result from the expected interaction between device-, user- and organizational characteristics. To repeat, expectation 1 states: *"If implementation leaders facilitate training about characteristics of the AI-based medical device for professionals of multiple disciplines within the medical department, trust in the device is promoted and all necessary changes are made and accepted. Therefore, the degree of successful implementation increases."*. Expectation 2 states: *"If implementation leaders, facilitators and/or change agents involve stakeholders and professionals from all disciplines whose workflow changes or who manage or execute these changes, during the development or implementation of an AI-based medical device, the degree of successful implementation increases."*.

To answer the main research question, a case-study was executed. Within the two included cases, the presence of the two expected organizational mechanisms was investigated. First, training is necessary to promote trust among professionals within the medical department, but it is not always needed for users. To add, trust among medical specialists was critical in both cases. If this profession group does not trust the device, workflow changes are not accepted. Trust is either promoted through the spread of information and experiences when using the device. Internally developed AI-based medical devices demand a different type of training compared to externally developed devices. Also, the type of profession is related to the type of training needed to implement the device successfully. Second, multi-disciplinary involvement is needed to estimate what workflow changes are necessary to implement an AI-based medical device. However, active involvement does not always lead to acceptance of workflow changes. Training is more effective to stimulate acceptance of workflow changes. Moreover, multi-disciplinary involvement exceeds the level of the medical department. This involvement mainly concerns the collection of resources.

These expectations were applicable to the two cases included in this study. However, data of one case was less reliable due to circumstances created by the COVID-19 pandemic. To further analyze the applicability of these organizational mechanisms to a larger variety of AI-based medical devices, users and medical departments, quantitative research on a larger variety of data is needed. Additionally, only two organizational mechanisms were analyzed during this study. Other organizational mechanisms should be analyzed in future research, for example how the relationship between the patient and the medical specialist influences the degree of successful implementation of an AI-based medical device.

REFERENCES

Amann, J., Blasimme, A., Vayena, E., Frey, D., and Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20(1):1–9.

Arora, A. (2020). Conceptualising artificial intelligence as a digital healthcare innovation: an introductory review. *Medical Devices (Auckland, NZ)*, 13:223.

Asan, O. and Choudhury, A. (2021). Research trends in artificial intelligence applications in human factors health care: Mapping review. *Journal of Medical Internet Research, Human Factors*, 8(4):e28236.

Beckers, R., Kwade, Z., and Zanca, F. (2021). The eu medical device regulation: Implications for artificial intelligence-based medical device software in medical physics. *Physica Medica*, 83:1–8.

Bélisle-Pipon, J.-C., Couture, V., Roy, M.-C., Ganache, I., Goetghebeur, M., and Cohen, I. G. (2021). What makes artificial intelligence exceptional in health technology assessment? *Frontiers in Artificial Intelligence*, 4.

Bonawitz, K., Wetmore, M., Heisler, M., Dalton, V. K., Damschroder, L. J., Forman, J., Allan, K. R., and Moniz, M. H. (2020). Champions in context: which attributes matter for change efforts in healthcare? *Implementation science*, 15(1):1–10.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6):394–424.

Cairney, P. (2018). Three habits of successful policy entrepreneurs. *Policy & Politics*, 46(2):199–215.

Chew, H. S. J. and Achananuparp, P. (2022). Perceptions and needs of artificial intelligence in health care to increase adoption: Scoping review. *Journal of medical Internet research*, 24(1):e32939.

Chua, I. S., Gaziel-Yablowitz, M., Korach, Z. T., Kehl, K. L., Levitan, N. A., Arriaga, Y. E., Jackson, G. P., Bates, D. W., and Hassett, M. (2021). Artificial intelligence in oncology: Path to implementation. *Cancer Medicine*.

COCIR (2021). Artificial intelligence in eu medical device legislation. https://www.cocir.org/media-centre/publications/article/cocir-analysis-on-ai-in-medical-device-legislation-may-2021.html.

Commission, E. (2018). A european approach to artificial intelligence.

Dadich, A. and Doloswala, N. (2018). What can organisational theory offer knowledge translation in healthcare? a thematic and lexical analysis. *BMC health services research*, 18(1):1–20.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., and Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation science*, 4(1):1–15.

European Parliament, E. U. (2017). Regulation on medical devices. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745.

Feiock, R. C. (2007). Rational choice and regional governance. *Journal of urban affairs*, 29(1):47–63.

Felmingham, C. M., Adler, N. R., Ge, Z., Morton, R. L., Janda, M., and Mar, V. J. (2020). The importance of incorporating human factors in the design and implementation of artificial intelligence for skin cancer diagnosis in the real world. *American journal of clinical dermatology*, pages 1–10.

Finn, Y., Cantillon, P., and Flaherty, G. (2014). Exploration of a possible relationship between examiner stringency and personality factors in clinical assessments: a pilot study. *BMC medical education*, 14(1):1–6.

Gerke, S., Minssen, T., and Cohen, G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. In *Artificial intelligence in healthcare*, pages 295–336. Elsevier.

Greenes, R. A., Bates, D. W., Kawamoto, K., Middleton, B., Osheroff, J., and Shahar, Y. (2018). Clinical decision support models and frameworks: seeking to address research issues underlying implementation successes and failures. *Journal of biomedical informatics*, 78:134–143.

Greenhalgh, T. and Abimbola, S. (2019). The nasss framework-a synthesis of multiple theories of technology implementation. *Stud Health Technol Inform*, 263:193–204.

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., Hinder, S., Fahy, N., Procter, R., Shaw, S., et al. (2017). Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of medical Internet research*, 19(11):e367.

Greenhalgh, T. and Wieringa, S. (2011). Is it time to drop the 'knowledge translation'metaphor? a critical literature review. *Journal of the Royal Society of Medicine*, 104(12):501–509.

He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., and Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature medicine*, 25(1):30–36.

Helman, S., Terry, M. A., Pellathy, T., Williams, A., Dubrawski, A., Clermont, G., Pinsky, M. R., Al-Zaiti, S., and Hravnak, M. (2022). Engaging clinicians early during the development of a graphical user display of an intelligent alerting system at the bedside. *International journal of medical informatics*, 159:104643.

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., and King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC medicine*, 17(1):1–9.

Kislov, R., Waterman, H., Harvey, G., and Boaden, R. (2014). Rethinking capacity building for knowledge mobilisation: developing multilevel capabilities in healthcare organisations. *Implementation Science*, 9(1):1–12.

Krauss, E., Haberer, J., Barreto, G., Degen, M., Seeger, W., and Guenther, A. (2020). Recognition of breathprints of lung cancer and chronic obstructive pulmonary disease using the aeonose® electronic nose. *Journal of breath research*, 14(4):046004.

Krauss, E., Zoelitz, J., Wagner, J., Barretto, G., Degen, M., Seeger, W., and Guenther, A. (2018). The use of electronic nose technology for the detection of lung cancer (lc): analysis of exhaled volatile compounds by aeonose®.

Kujala, S., Heponiemi, T., and Hilama, P. (2019). Clinical leaders' self-perceived ehealth competencies in the implementation of new ehealth services. In *MEDINFO 2019: Health and Wellbeing e-Networks for All*, pages 1253–1257. IOS Press.

Lange, J., Dekker-van Doorn, C., Haerkens, M. H., and Klein, J. (2011). A safety culture in hospitals. *Nederlands Tijdschrift Voor Geneeskunde*, 155.

Laukka, E., Huhtakangas, M., Heponiemi, T., and Kanste, O. (2020). Identifying the roles of healthcare leaders in hit implementation: a scoping review of the quantitative and qualitative evidence. *International journal of environmental research and public health*, 17(8):2865.

Lundmark, R., Hasson, H., Richter, A., Khachatryan, E., Åkesson, A., and Eriksson, L. (2021). Alignment in implementation of evidence-based interventions: a scoping review. *Implementation Science*, 16(1):1–14.

Manlhiot, C., Van den Eynde, J., Kutty, S., and Ross, H. J. (2021). A primer on the present state and future prospects for machine learning and artificial intelligence applications in cardiology. *Canadian Journal of Cardiology*.

Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science*, 10(53).

Phansalkar, S., Edworthy, J., Hellier, E., Seger, D. L., Schedlbauer, A., Avery, A. J., and Bates, D. W. (2010). A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. *Journal of the American Medical Informatics Association*, 17(5):493–501.

Rijksoverheid (2019). Strategisch actieplan voor artificiële intelligentie.

Rogers, E. (1995). Diffusion of innovations.

Shaw, J., Rudzicz, F., Jamieson, T., and Goldfarb, A. (2019). Artificial intelligence and the implementation challenge. *Journal of medical Internet research*, 21(7):e13659.

Smuck, M., Odonkor, C. A., Wilt, J. K., Schmidt, N., and Swiernik, M. A. (2021). The emerging clinical role of wearables: factors for successful implementation in healthcare. *NPJ Digital Medicine*, 4(1):1–8.

Staten Generaal, N. L. (1993). Wet op de beroepen in de individuele gezondheidszorg. https://wetten.overheid.nl/BWBR0006251/2021-07-01.

Staten Generaal, N. L. (1994). Wet geneeskundige behandelingsovereenkomst. https://wetten.overheid.nl/jci1.3:c:BWBR0007021&z=2006-02-01&g=2006-02-01.

Steenhuis, E., Schoenaker, I., de Groot, J., Fiebrich, H., de Graaf, J., Brohet, R., van Dijk, J., van Westreenen, H., Siersema, P., et al. (2020). Feasibility of volatile organic compound in breath analysis in the follow-up of colorectal cancer: A pilot study. *European Journal of Surgical Oncology*, 46(11):2068–2073.

Strohm, L., Hehakaya, C., Ranschaert, E. R., Boon, W. P., and Moors, E. H. (2020). Implementation of artificial intelligence (ai) applications in radiology: hindering and facilitating factors. *European radiology*, 30:5525–5532.

Sujan, M., Furniss, D., Grundy, K., Grundy, H., Nelson, D., Elliott, M., White, S., Habli, I., and Reynolds, N. (2019). Human factors challenges for the safe use of artificial intelligence in patient care. *BMJ health & care informatics*, 26(1).

Tartar, M., Le, L., Watanabe, A. T., and Enomoto, A. J. (2021). Artificial intelligence support for mammography: in-practice clinical experience. *Journal of the American College of Radiology*, 18(11):1510–1513.

Thomasian, N. M., Kamel, I. R., and Bai, H. X. (2021). Machine intelligence in non-invasive endocrine cancer diagnostics. *Nature Reviews Endocrinology*, pages 1–15.

van Baalen, S., Boon, M., and Verhoef, P. (2021). From clinical decision support to clinical reasoning support systems. *Journal of evaluation in clinical practice*, 27(3):520–528.

van Beest, W., Boon, W. P., Andriessen, D., Moors, E. H., van der Veen, G., and Pol, H. (2020). Successful implementation of self-management health innovations. *Journal of Public Health*, pages 1–15.

van den Berge, K. and Mamede, S. (2013). Cognitive diagnostic error in internal medicine. *European journal of internal medicine*, 24(6):525–529.

van Keulen, K. E., Jansen, M. E., Schrauwen, R. W., Kolkman, J. J., and Siersema, P. D. (2020). Volatile organic compounds in breath can serve as a non-invasive diagnostic biomarker for the detection of advanced adenomas and colorectal cancer. *Alimentary pharmacology & therapeutics*, 51(3):334–346.

Waltman, C. G., Marcelissen, T. A., and van Roermund, J. G. (2020). Exhaled-breath testing for prostate cancer based on volatile organic compound profiling using an electronic nose device (aeonose™): a preliminary report. *European urology focus*, 6(6):1220–1225.

Yin, J., Ngiam, K. Y., and Teo, H. H. (2021). Role of artificial intelligence applications in real-life clinical practice: Systematic review. *Journal of medical Internet research*, 23(4):e25759.

Yin, R. K. (2011). *Applications of case study research*. sage.

APPENDIX A

Coding scheme

Before the implementation process	Organizing the implementation process	During the implementation process	Factors
Communication	Organization of the development process	Training (validity/ reliability)	Trust
Department-hospital	Organization of the implementation process	Training (improve skills)	Variability
Hospital network	User role during development	Training (interpret/ explain outcomes)	Usability
Company-hospital	Knowledge background medical specialists	Financial resources	Understandability
Manager-medical specialist	Knowledge background users	Other resources	Responsbility
Manager-user	Role of the AI-based medical device	Discussions	Workload
Medical specialist-user	Complexity of tasks	Response to implementation	Publicity
Organization of the department	Redistribution of tasks		
Relation technical medicine	Local champion		
Medical specialists group	Leadership		
Condition	Technical medicine		
Status quo			
Goal of the AI-based medical device			

Figure 1: *Table showing the coding scheme used to analyze data gathered via the interviews, the right column via thematic coding, the three left columns via frame-work analysis.*

Figure A.1 shows the coding scheme used to code the transcripts of the interviews, which was part of the data-analysis. The codes shown in the left three columns are based on the theoretical framework. The codes shown in the rightmost column is based on thematic coding.