MASTER'S THESIS – MASTER INNOVATION SCIENCES

THE AUGMENTED RADIOLOGIST

CHALLENGES AND OPPORTUNITIES FOR WIDESCALE IMPLEMENTATION OF AI-BASED APPLICATIONS IN DUTCH RADIOLOGY DEPARTMENTS

Lea Strohm

6278760

l.g.d.strohm@students.uu.nl

+31628537619

Supervisor: Prof. Dr. Ellen H.M. Moors

Copernicus Institute of Sustainable Development

Utrecht University

12.08.2019

SUMMARY

Clinical radiology in the Netherlands is being flooded with digital data, mainly in the form of medical images. Software applications that perform computerized automated image analysis, so-called artificial intelligence (AI) applications, are becoming increasingly accurate, showing better performance than trained radiologists for certain tasks. Yet, very few AI applications are currently implemented in routine clinical use in radiology departments of Dutch hospitals. The practical implementation of new technologies in the medical field, especially in hospital settings, depends on a range of different factors, such as the large variety of stakeholders involved, the rigid routines and strong professional identities, as well as the strict legal and regulatory standards to be abided. These factors hinder or facilitate the implementation process and often interact in dynamic ways, as demonstrated by the recently published nonadoption, abandonment, scale-up, spread and sustainability (NASSS) framework, which focuses explicitly on determinants of unsuccessful adoption. This research aimed at identifying facilitating and hindering factors to the successful implementation of AI applications in Dutch radiology departments and how the hindering factors could be overcome. Due to the early stage of adoption of AI applications in radiology, an exploratory, qualitative research design was followed, based on an embedded multiple case study. In a first deductive step, guiding propositions were derived from the existing NASSS framework. In a second inductive step, the framework was refined for the case of AI applications in radiology. The results showed a wide array of facilitating and hindering factors to successful implementation of AI applications in Dutch radiology departments. Among the most important facilitating factors is the presence of a ‘local champion’, an individual with a strong personal interest in AI applications, which most often initiated and actively pushed forward the implementation of AI applications in their respective organization. Among the most prominent hindering factors are the uncertain added-value for clinical practice of AI applications, which causes low acceptance of AI applications among adopters and complicates the mobilization of funds to acquire AI applications. Furthermore, the failure to include all relevant stakeholders in the planning and execution phase of the implementation of AI applications was found a major hindering factor. To increase low acceptance among adopters, more evidence of the added-benefit of their AI applications in the clinical setting is needed. Also, all affected stakeholders (most notably radiologists and referring clinicians) should be included in the decisions and the design of implementation processes of AI applications.

"There is nothing more difficult to take in hand, more perilous to conduct, or more uncertain in its success, than to take the lead in the introduction of a new order of things."

Niccolo Machiavelli – The Prince

ACKNOWLEDGEMENTS

I would like to thank those, who have made this research possible. First of all, thanks go to my supervisor Prof. Dr. Ellen Moors for her valuable and frequent feedback and her continuous support during the research process, especially when external circumstances suddenly forced me to finish the research in a shorter time than planned. Furthermore, I would like to thank Charisma Hehakaya and Dr. Mohammad Rezazade for repeatedly providing me with helpful inputs to my research. Special thanks also go to my interviewees for the interesting conversations and their willingness to share their opinions and their personal experiences with the topic of my research. Finally, I would like to thank Christopher Schulz for his feedback and assistance during the entire research process.

CONTENTS

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AZ	Algemeen Ziekenhuis *(General Hospital)*
CAD	Computer Aided Diagnosis
CDSS	Computerized Decision Support System
CT	Computer Tomography
DBC	Diagnose Behandelcombinatie *(Diagnosis-Therapy Combination)*
GDPR	General Data Protection Regulation
IT	Information Technology
MDR	Medical Device Regulation
NASSS	Nonadoption, Abandonment, Scale-up, Spread, and Sustainability
NVvR	Nederlandse Vereniging voor Radiologie *(Radiological Society of the Netherlands)*
PACS	Picture Archiving and Communication System
TKZ	Topklinisch Ziekenhuis *(Hospital Center of Clinical Excellence)*
UMC	Universitair Medisch Centrum *(University Medical Center)*

1. INTRODUCTION

Like many other industries, the health care sector is being flooded with data. In clinical practice this translates into an increase of data generated on individual patients, for example in the form of biomarkers and medical images (Obermeyer & Emanuel, 2016; Ottes, 2016). In theory, this data should allow for more accurate and efficient diagnosing and treatment. In practice, however, doctors and other medical professionals often do not have the time or competences to take the whole set of available data into account when diagnosing and treating patients. A potential solution is provided by mechanisms of automated data analysis in the form of advanced computer algorithms.

One of the medical fields particularly affected by this development is radiology, the medical specialty in which "trained physicians visually [assess] medical images for the detection, characterization and monitoring of diseases" (Hosny, Parmar, Quackenbush, Schwartz, & Aerts, 2018, p. 500). Across the health care field, radiologists have the most digitized work environment (Nawrocki, Maldjian, Slasky, & Contractor, 2018). The potential of computers to assist radiologists in repetitive tasks was recognized over fifty years ago, when the term Computer-Aided Diagnosis (CAD) was first coined (Lodwick, 1966). CAD can be broadly defined as "the use of computer algorithms to aid the image interpretation process" (van Ginneken, Schaefer-Prokop, & Prokop, 2011, p. 720). In clinical terms, a CAD system produces a diagnostic output, which is used as a 'second opinion' for the radiologist in the diagnostic process. In the 1980s, first CAD systems based on early artificial intelligence (AI) algorithms were being developed in certain radiology subfields, such as chest imaging and mammography. Yet, by the early 2010s CAD systems were still not used routinely across the radiology profession. Most literature from the medical or data science field associated the failure of early CAD systems to reach widespread adoption to their suboptimal technical performance (Kohli & Jha, 2018; van Ginneken et al., 2011) or the way CAD was put into practice (Nishikawa & Bae, 2018). Other potential barriers to adoption, such as organizational or social aspects, were largely ignored.

Conveniently, since the late 2000s, research in the field of AI, which forms the technical basis of CAD, has made particularly large advances in the sub-field of automated image analysis, using techniques known as machine learning and deep learning.[1] These new techniques allow for large improvements in the accuracy of object detection in images, i.e., the detection of patterns and anomalies, as is done by most clinical CAD applications. They also render possible automated image recognition, i.e., classifying and quantifying objects in medical images (Nawrocki et al., 2018; Schmidhuber, 2015). Accordingly, first software applications using more advanced AI for radiology have been developed and are already being deployed in clinical environments[2] (Bluemke, 2018; van Ginneken, 2017).

---

[1] Machine learning is a subfield of the much broader term artificial intelligence and includes a technique called *deep neural network learning* (also known as deep learning), where algorithms learn to identify relevant features without explicit selection by a human programmer. The years 2011 to 2012 are seen as a crucial breakthrough moment for the field of automated image analysis, bringing forward a new type of neural network learning, *deep convolutional neural networks,* which enabled the large advances in computer vision (Nawrocki et al., 2018; Schmidhuber, 2015).

[2] E.g. *BoneXpert*, by Danish firm Visiana. The program runs automatic bone-age detection from X-rays and has been on the market with since 2009 (H. Lee et al., 2017). A more recent example is the program *ContaCT*, made by US firm Viz.ai. It analyzes CT scans for signs of strokes and received regulatory approval for the US market in February 2018 (Bluemke, 2018)

From a clinical perspective, advanced AI techniques can potentially be used at all points of the radiologists' workflow, as shown in figure 1 below.

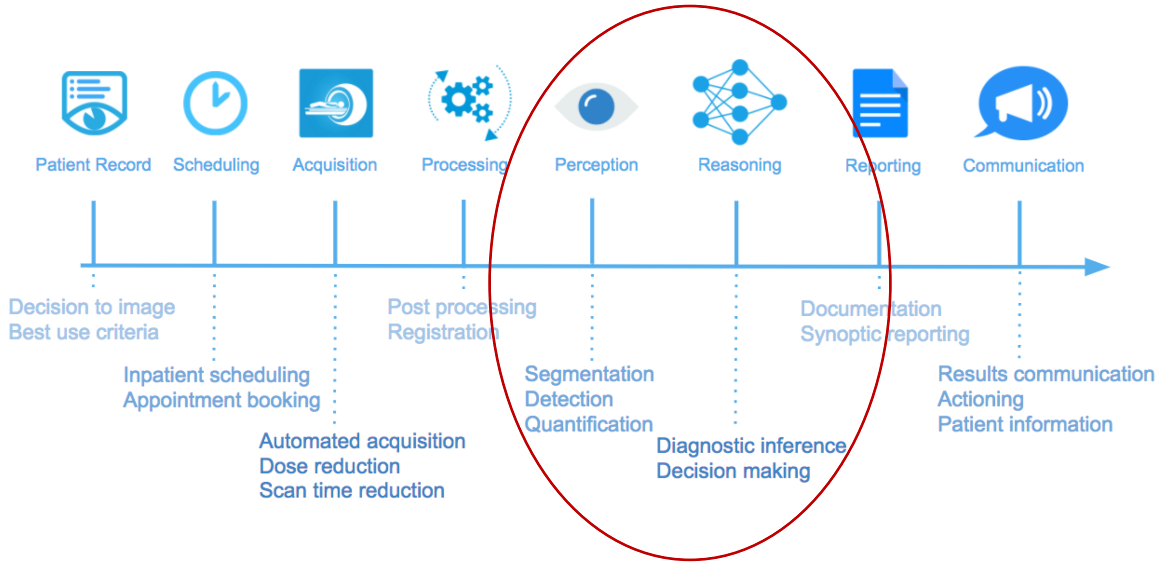

FIGURE 1: SIMPLIFIED SCHEMATIC OF THE DIAGNOSTIC RADIOLOGY WORKFLOW (HARVEY, 2018), RED CIRCLE ADDED MANUALLY BY AUTHOR TO HIGHLIGHT FOCUS AREA OF THIS RESEARCH.

The two use-cases most commonly referred to when talking about AI in radiology are image interpretation (segmentation/detection/quantification) and diagnostic inference, as indicated by the red circle in figure 1 (European Society of Radiology, 2019). Segmentation, detection and quantification are advanced forms of computer vision, where algorithms perform some form of image interpretation. This can be partitioning an image in certain areas, e.g. organs (segmentation), measuring and comparing over time the size and density of certain elements in an image, such as length of bones (quantification), or identifying anomalies from an image, e.g. tumors (detection). Detection algorithms help the radiologist in the process of analyzing an image, but the diagnostic is made by the radiologists. Diagnostic inference is the most advanced functionality, as it uses one or several of the former functionalities in image interpretation in order to produce some form of automated diagnosis. This can be an algorithm, which detects a stroke in a Computer Tomography (CT) of the head, and based on the analysis of the surrounding area can produce a score that indicates if a surgical intervention is adequate for the patient.

Within the field of radiology, the term CAD is mostly used to name early lesion detection systems, which were developed mainly for breast, lung, colon and prostate cancer (Doi, 2007). Because the type of functionalities and the technical performance of algorithms increased with the surge of more advanced machine learning techniques, the broader term 'artificial intelligence (AI) applications' was coined (Bluemke, 2018; European Society of Radiology, 2019; Pesapane, Volonté, Codari, & Sardanelli, 2018). Within the scope of this research, AI applications are defined as computer programs for clinical radiological practice, which use statistical learning techniques to perform automated image interpretation and/or support the diagnostic decision making.

At present, these AI applications only cover a small subset of radiologists' tasks, which means that the impact on the day-to-day work of radiologists is still rather limited. Nonetheless, the large technological improvements have created high expectations on the potential of AI applications and triggered strong responses from radiologists. Since 2017, major conferences and scientific publications in the field of radiology have been flooded with discussions and opinion papers on the potential impact of AI for the radiology profession (Choy, Samir, & Brink, 2018; European

Society of Radiology, 2019; Obermeyer & Emanuel, 2016). Some voices envision the radiologist of the future less as a physical practitioner and more as a data wrangler (Bluemke, 2018; Nawrocki et al., 2018). The fear of 'getting replaced by AI' is thereby a recurrent theme, also due to controversial statements from known data scientists. One example is Stanford's Andrew Ng, who stated that "a highly trained and specialised radiologist may now be in greater danger of being replaced by a machine than his own executive assistant" (Morgenstern, 2016, para. 3). While this statement has repeatedly been contested, there seems to be little doubt on the major impact AI will have on the radiological profession. No consensus, however, exists on the timeframe for AI applications to start being widely adopted in the field of radiology. While some experts speak of years, others expect decades to pass by (European Society of Radiology, 2019; Harvey, 2018; Nawrocki et al., 2018).

When trying to situate AI applications for radiology within a broader perspective of innovation in health care, the innovation can be argued to fall at the intersection of digitization of health care, automation and medical technology. Naturally, most radiologists lack understanding of the technical functioning of advanced computer algorithms and see these algorithms as a "black box" (Choy et al., 2018; Nishikawa & Bae, 2018). While data scientists are conducting R&D efforts to improve algorithms for an increasing number of radiologists' tasks, the products that have already received regulatory approval must currently prove their added value in clinical practice. Their promised efficiency and quality gains are particularly interesting in the context of increasing workloads of radiologists in bigger medical institutions, such as hospitals (Bluemke, 2018; van Ginneken et al., 2011). In these complex environments the implementation of a new technology is not only influenced by and impacts its direct user, that is the radiologist, but also a range of other professionals and organizational structures. From the point of view of the concerned organization, the hospital, it is critical to have in-depth knowledge of these processes in order to prepare for and execute the change processes induced by AI in radiology.

The process of intentionally and purposively putting an innovation into action is known as implementation. It refers to "the transition period during which targeted organizational members ideally become increasingly skillful, consistent, and committed in their use of an innovation" (Klein & Speer Sorra, 1996, p. 1057). Identifying and understanding the factors that influence this process in the health care field is at the core of *implementation research* literature (Proctor et al., 2011). Based largely on Rogers' diffusion of innovation theory (Rogers, 1962, 2003), a number of implementation frameworks have been developed. (Damschroder et al., 2009; Fleuren, Wiefferink, & Paulussen, 2004; Greenhalgh, Robert, Macfarlane, Bate, & Kyriakidou, 2004; Greenhalgh et al., 2017; Liberati et al., 2017). These frameworks comprise varying sets of *micro-level* variables linked to the (technical) nature of the innovation and the user of the innovation, as well as *meso-level* variables, linked to the implementation process within an organizational context, and finally *macro-level* variables, linked to the larger external context of the innovation, e.g. the readiness of the health system. Arguably, the most prominent of these frameworks is the framework on diffusion of innovations in service organizations by Greenhalgh et al. published in 2004. Based on vast empirical evidence, this framework was recently developed further to cover cases, where technological innovations experience non-adoption or abandonment by individuals or fail to spread up, a phenomenon largely ignored by implementation researchers. The findings were integrated in a new framework: the *Nonadoption, Abandonment, Scale-up, Spread, and Sustainability* (NASSS) framework (Greenhalgh et al., 2017).

Because AI applications in radiology are at an early stage of the adoption and implementation process, it is too early to assess the success or failure of their entire implementation process. However, considering the unsuccessful widespread adoption of earlier CAD systems, it can be assumed that AI applications will encounter substantial barriers to adoption. Due to its focus on nonadoption and failed sustainable scale-up of innovations, the NASSS framework by Greenhalgh et al. (2017) appears to be suitable to understand the determinants and dynamics of a successful implementation process of AI applications in clinical radiology, leading to the following research question:

What are facilitating and hindering factors for the successful implementation of AI-based applications in radiology departments in Dutch hospitals and how can they be overcome?

From a practical perspective, the results of this research will contribute to give hospitals guidance in preparing and managing the change processes linked to the emergence of AI in the field of radiology. In case of success, these change processes promise better and more efficient health care delivery, relevant not only for the service-providing organization but also for society as a whole. Additionally, new insights into the broader theme of digitization and big data in health care will provide answers to existing societal concerns in the Netherlands. Among these concerns are privacy, related to the use of sensitive health data by third parties, or risks related to unreliable or inaccurate AI algorithms used in medical diagnostics (Lupton, 2014; Ottes, 2016; WRR, 2016).

From a theoretical perspective, the research will add to the existing empirical evidence on the implementation challenges of technological innovation in health, more specifically of big data in medicine. Existing evidence on the barriers to adoption of other forms of big data applications in the medical field, such as Computerized Decision Support Systems (CDSS), affirms the intricate nature of these implementation processes (Craig et al., 2008; Liberati et al., 2017). AI-based applications in radiology have the potential to not only support, but potentially *automate* certain medical decision processes, thereby calling into question the jobs of highly educated individuals. This element of job displacement due to automatization adds to the complexity of adoption and implementation process in the field of health digitization, not yet researched to present. Often, innovation studies are biased towards researching successful innovations and fail to appropriately discuss and research cases of failed widespread adoption, as was the case for early versions of CAD systems. Hence, this research will add to the empirical evidence on failed scaling up of innovation or non-sustainable adoption and will refine the NASSS framework by taking into account insights from the adoption and implementation process of a complex innovation, which converges the disciplines of medical technology and big data in the health sector.

This report is structured as follows: Chapter 2 addresses the theoretical framework of the research. Chapter 3 provides the methodological approach used for the empirical part of the research. Chapter 4 presents the results, complemented by an in-depth analysis of the results in chapter 5. In chapter 6, conclusions are drawn regarding the results and analysis. Finally, chapter 7 discusses the theoretical and practical contribution of this research, gives an overview on its limitations and presents suggestions for further research.

2. THEORY

Due to its technological, legal and organizational complexity, the topic of health care and medical innovation has been the object of study for various strains of innovation researchers. Rogers' diffusion of innovation theory (Rogers, 1962), for example, is based partially on empirical work on the spread of innovation in the public health and medical technology field.

Rogers (1962) separated the innovation process in an organization in different stages: Leading up to the adoption decision is the initiation phase, which includes agenda setting, i.e., determining where and which innovation is needed, and matching, i.e., finding the right innovation for the identified need. The adoption decision is followed by the implementation phase, which includes the processes of redefining, clarifying and routinizing. Redefining refers to the process, in which, on the one hand, the innovation is molded to fit the organization, and on the other hand, the organization alters certain structures to accommodate the innovation. This stage is followed by clarifying, when the place of the new innovation within the organization is delineated more clearly and finally the stage of routinizing, when the innovation becomes an integral part of the organization. In other words: "Implementation is the critical gateway between the decision to adopt the innovation and the routine use of the innovation within an organization" (Klein & Speer Sorra, 1996, p. 1057).

Based on the fundamental work of Rogers, Greenhalgh et al. (2004) developed a highly cited, unifying model on the implementation of innovations in health service organizations, such as hospitals. Although vastly used, the model does not account for cases of unsuccessful adoption. Consequently, a new framework was created: the Nonadoption, Abandonment, Scale-up, Spread, and Sustainability (NASSS) framework (Greenhalgh et al., 2017), which includes ideas from different theoretical approaches, such as actor-network theory (May et al., 2007; Robert, Greenhalgh, MacFarlane, & Peacock, 2010), organizational routinization theory (Chaudoir, Dugan, & Barr, 2013; Cresswell & Sheikh, 2013) and complexity theory (Abbott, Foster, Marin, & Dykes, 2014; Plsek & Greenhalgh, 2001). It aims at informing the design and planning of complex innovations and programs, as well as explaining and preventing nonadoption, abandonment or the failure of achieving long-term sustainability of innovations. While the NASSS framework highlights several possible *unsuccessful* implementation outcomes, it does not explicitly state criteria for successful implementation. This reflects the lacking consensus in implementation literature on what qualifies as a successful implementation outcome, an issue discussed elaborately by Proctor et al. (2011). Among the different criteria for successful implementation proposed by Proctor et al. (2011) are perceived acceptability by adopters, fidelity to how an innovation was prescribed to work and sustainability. The concept of sustainability also appears within the NASSS framework, where it is understood as the continuous use of the innovation after initial implementation efforts are concluded (Greenhalgh et al., 2017). Accordingly, an innovation can be considered successfully implemented, if it achieves its intended aim, i.e., the problem(s) it is supposed to solve, in a long-term sustainable way.

The NASSS framework is built around seven domains (see figure 2), which each include several sub-domains. It is important not to regard these components individually, but to consider how they interact with each other over time.

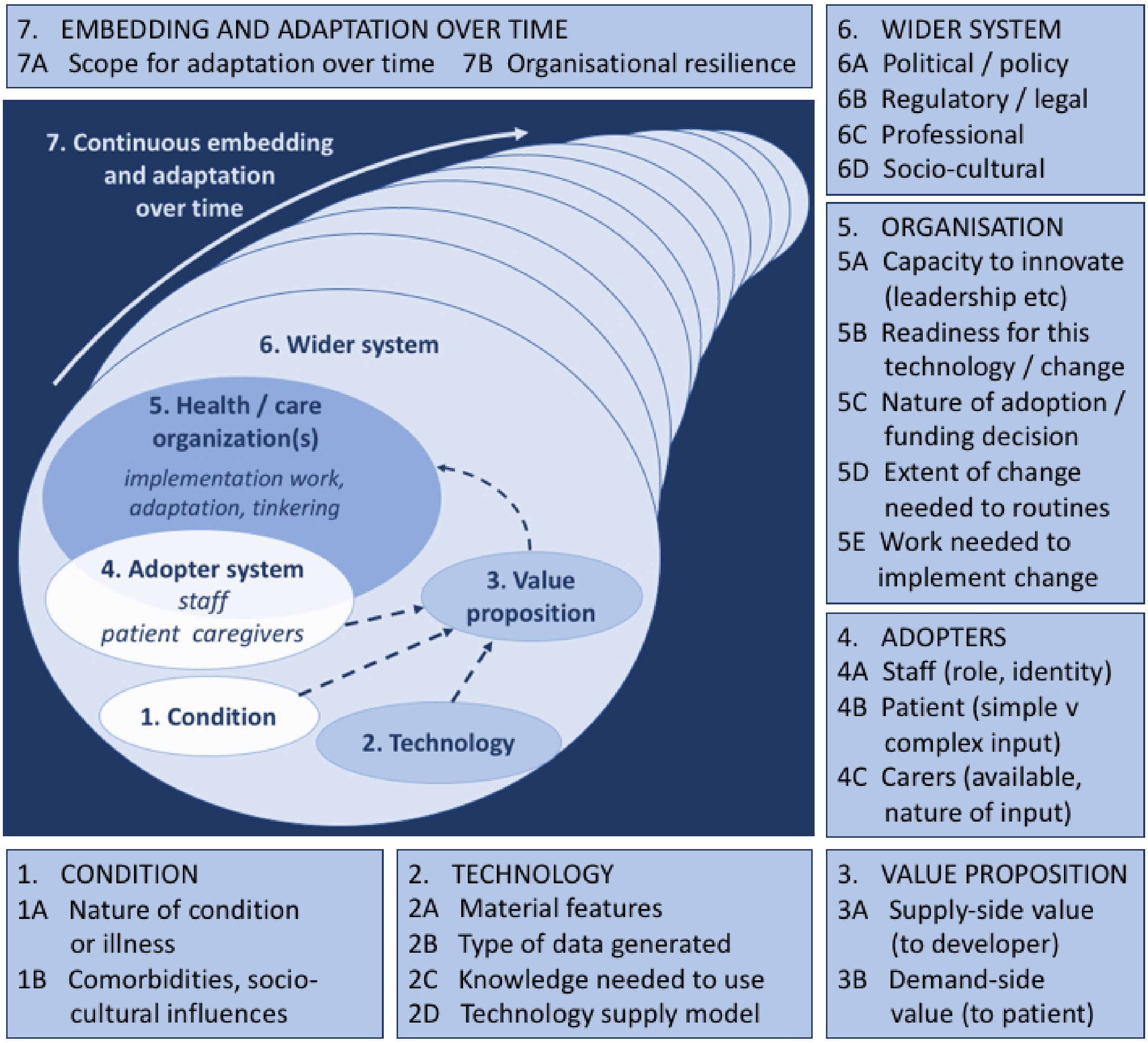

FIGURE 2: THE NASSS FRAMEWORK BY GREENHALGH ET AL. (2017)

The first domain, *the condition*, refers to the clinical situation, in which the innovation is going to be used. Within the framework, condition is used to describe the nature of the illness and comorbidities of the patient, which is targeted by the innovation. These elements change for each condition, and were found to be insufficiently addressed in previous frameworks. Since AI applications in radiology can be used for innumerous clinical situations, the condition will also change for each particular application. Radiologists are generally specialized in one or several body-parts, called the radiological subspecialties.[3] This means depending on the clinical situation addressed by a specific AI application, a specific sub-group of radiologists will be involved. Since different radiology subspecialties use different imaging techniques and have different diagnostic routines, it can be assumed that these subspecialties will approach the implementation of AI applications differently. This leads to the following proposition: *For the implementation of AI*

[3] The recognized subspecialties in the Netherlands are: abdominal radiology, emergency radiology, cardiovascular radiology, forensic and postmortem radiology, head-neck radiology, interventional radiology, pediatric radiology, breast radiology, musculoskeletal radiology, neuroradiology and thoracic radiology (Nederlandse Vereniging voor Radiologie, 2019).

applications to be successful, the heterogeneity of the diagnostic process across different subspecialties in radiology needs to be taken into account (proposition 1).

The second domain, *the technology*, covers the technical properties of the technology: on the one hand these include material properties, such as physical aspects, functionality and compatibility with existing technologies. On the other hand, questions relating to relevant knowledge are invoked. Both the knowledge base of the technology itself as well as the relevant knowledge to use and support the technology need to be taken into account. For the case of AI applications in radiology, the knowledge base of the technology (data science) is distant from the knowledge base of its users (radiologists). The most obvious material properties of AI applications are the algorithm's performance and the user-friendliness of the application. Algorithmic performance is most commonly determined by its sensitivity and specificity[4] (Doi, 2007). User-friendliness relates to the smooth integration in existing IT systems and the easy to interpret user-interface (Liberati et al., 2017; Nishikawa & Bae, 2018). Thus, *it can be assumed that in order to successfully implement AI applications for radiology, the applications need to display high performance and user-friendliness (proposition 2)*.

The third domain, *the value proposition*, draws attention to the question if the innovation creates or adds value to patients, medical professionals, health organizations (the 'demand-side') and the technology provider/developer (the 'supply-side'). Within the NASSS framework, the concept of value goes beyond a simply financial measure, but also comprises non-financial benefits, such as quality or efficiency gains. Evidently, value will be defined differently by the involved stakeholders. While the technology providers need to generate revenue in order to cover their development costs, the hospital needs to find internal or external sources of financing to cover the purchasing costs. The need to clearly demonstrate the added value is particularly important in the context of renegotiating reimbursement with external sources of income, such as health insurers (C. I. Lee et al., 2018). Finally, in order to engage with the technology, the medical professionals need to be convinced of its added value for clinical practice. This leads to the following guiding proposition: *in order to achieve successful implementation, AI applications need to provide clear added value as perceived by the relevant stakeholders, mainly medical professionals, the health organizations and the technology provider (proposition 3)*.

Unlike the value proposition, *the adopters*, the fourth domain of NASSS, are at the core of most implementation frameworks. Considering the evidence reviewed by the authors, "acceptance by professional staff may be the single most important determinant of whether a new technology-supported service succeeds or fails at a local level" (Greenhalgh et al., 2017, p. e367). This acceptance is not only based on usability, related to the material properties (see above), but also on the impact on professional identities and scope of action, related to the symbolic properties. As was found by Liberati et al. (2017), "health professionals sometimes discard the use of evidence out of fear of compromising their critical reasoning, medical judgments, and professional autonomy" (Liberati et al., 2017, p. 113). Within an organization, acceptance can be influenced by local "champions", meaning influential individuals who openly support an innovation within an organization (Rogers, 2003). These individuals are central in convincing their peers of the usefulness and safety of the technology and demonstrate that the use of the latter is appropriate (C. I. Lee et al., 2018; Marcial et al., 2019; Wade & Eliott, 2012). As mentioned, radiologists have shown strong responses towards AI applications, related to the fear of

[4] High sensitivity means a high true positive rate, i.e. the algorithm identifies actual anomalies at a high rate (e.g. 90% of actual anomalies are recognizes and 10% are missed by the algorithm), while a high specificity means a low false-positive rate (there are few cases in which the algorithm recognizes an anomaly, which is actually none).

reduction of radiologists' scope of practice (Obermeyer & Emanuel, 2016). In addition to the radiologists, critical professional staff includes other medical doctors (the so-called 'referring clinicians', who request the medical images), technical staff and members of the (innovation) management of the adopting organizations. Presumably, patients and caregivers are less relevant for AI in radiology, because they do not have direct contact or decision power over the technology. Hence, it can be expected that *in order to reach successful implementation of AI applications for radiology, acceptance of relevant medical professionals, technical staff and hospital management needs to be achieved, a process that can be facilitated by local champions in organizational contexts (proposition 4)*.

The fifth domain considered in the NASSS framework is *the organization.* Meant are health organizations, such as hospitals, general practitioner's or medical specialist practices. The organizational setting was at the core of the Greenhalgh et al. (2004) framework and builds in large parts on organizational change theories not specific to the health field. Several organizational characteristics are known to positively influence the capacity to innovate at organizational level: strong managerial leadership that supports and encourages innovation and risk-taking (Anderson & West, 1998; Gustafson et al., 2003; Van de Ven, Polley, Garud, & Venkataraman, 1999). Management needs to liberate spare resources to support innovation projects and make sure that the scope of and reasoning behind the innovation project is clearly and extensively communicated throughout the organization (Greenhalgh et al., 2004). For the case of AI applications for radiology, the organizations in question are primarily hospitals with sufficiently large radiology departments. Radiology departments employ different types of medical professionals: radiologists, nuclear physicians, technicians, lab technicians and clinical physicists. They work for (or in collaboration with) other departments in the hospital, for example pediatrics for suspected growth disorders in children. This means that AI applications indirectly also impact the work of other departments than just the radiology department within the organization. Thus, management has an important function in connecting the involved sub-units and make sure they are aligned on a strategic level (Liberati et al., 2017). It is expected that *for successful implementation of AI applications for radiology, management needs to create an organizational culture open to innovation, for example through innovation strategies that serve as a shared vision for all involved stakeholders (proposition 5a)*.

Although hospitals are characterized by strong hierarchies, an organizational structure that allows departments some autonomy in their decision making can help increase acceptance among adopters. This also means that frontline staff impacted by the innovation participates in the decision processes, the implementation process and the remodeling of work routines (Liberati et al., 2017). Thus, it can be assumed that *in order to successfully implement AI applications for radiology, adoption decisions and implementation strategies need to be taken into account and developed in a participatory way, by and for the relevant stakeholders in the organization (proposition 5b)*.

Furthermore, hospitals are highly regulated organizations with clear-cut professional boundaries and rigid organizational routines (Liberati et al., 2017). The redesigning of existing work-routines and creation of new routines is inherent to the introduction of new technologies (Edmondson, Bohmer & Pisano, 2001). As pointed out in normalization process theory[5], this routinization process involves a lot of work from individuals and organizations (May & Finch, 2009).

[5]Normalization process theory aims to "understand the conditions in which new technologies, techniques, working practices, and organizational interventions – complex interventions – *can become embedded as routine elements* of clinical and organizational work in health care" (May, 2006, p. 68).

Furthermore, Liberati et al. (2017) argue that involving relevant clinicians and information technology (IT) staff in the process of remodeling work routines, can create legitimacy and enhance the usability of the new technology, leading to the following guiding proposition: *In order to successfully implement AI applications for radiology, all relevant stakeholders need to be included in (re)designing of routines, on the individual and the organizational level (proposition 5c)*.

The sixth domain, *the wider system*, includes aspects related to the broader societal context in which the organizations are embedded. In a highly regulated field, such as health care, safety, quality and efficiency/efficacy aspects are dealt with by (supra-)national regulators. The same goes for legal topics around liability and intellectual property. Less straight-forward are questions related to financial aspects, e.g. who covers the costs for the innovations. These policy issues are highly political and involve different stakeholders, such as health insurance companies, professional associations, regulatory agencies on national and international levels, political parties etc., who have differing if not competing interests. Interesting insights into how some of these actors actively try to speed up or slow down the adoption of certain innovations, can be gained from the field of institutional entrepreneurship (Kukk, Moors, & Hekkert, 2016). These policy and regulatory issues regarding digital health technologies have already shown to create controversy, for example privacy concerns related to lacking transparency in the gathering and use and the security of big data (Lupton, 2014, 2016). Because AI applications are located at the intersection of medicine and information technology, new legal and regulatory questions have arisen concerning the regulatory regime applicable to algorithms or the doubt on who is legally responsible for damages caused through mistakes committed by algorithms. While the technology is developing rapidly in the field of AI in medicine, regulatory and legal questions are known to respond with a significant time-lag, creating ambiguity for adopters and developers alike (French-Mowat & Burnett, 2012; Pesapane et al., 2018). Thus, *to achieve sustainable adoption of AI applications in radiology, legal and regulatory certainty on how the technology will be integrated in the existing health system needs to be created and expected institutional opposition needs to be overcome (proposition 6)*.

The seventh and last domain is *the embedding and adaptation over time*. At different stages of the adoption process, other challenges related to the previous six domains will appear. Because of the non-linear nature of the implementation process, it is essential to include moments of reflection within an organizational context in order to assure smooth progress and recognize potential problems of the implementation process (May & Finch, 2009). Based hereon, the process can be adapted, both on the technology level, but also on the organizational level. Currently, AI applications are at an early stage of development and there is substantial uncertainty on the time-frame and scope of technological advances even in the near and medium-term future. However, attention to the adaptation over time needs to be given already during the planning phase of implementation processes, indicating that *in order to achieve sustainable adoption of AI applications, implementation processes need to include moments of reflection and allow for continuous adaption (proposition 7)*.

Because the NASSS framework is very recent (Greenhalgh et al., 2017), little empirical evidence has been collected yet. One exception is a study published by the authors of the framework, which looked at the adoption process of six technology-based innovations, four patient-centered technological devices[6] and two organizational-level IT innovations[7] (Greenhalgh et al., 2018). The authors' analysis of large amounts of primary data suggest that those programs (i.e., implementation of innovations) which are characterized by complexity in various of the NASSS domains, were not successfully adopted over time. Consequently, this research will adapt and refine the NASSS framework for the case of a complex innovation, which converges the fields of medical technology and automated (big) data analysis in the health sector.

[6] Video outpatient consultations, GPS tracking for mentally impaired patients, pendant alarms, telehealth for heart failure.
[7] Care organizing software for relatives of patients, shared data infrastructure project between health and social care services.

3. METHODOLOGY

The following section elaborates on the research design (3.1) and sampling strategy (3.2), the method of data collection (3.3), data analysis (3.4) and the quality indicators of the research (3.5).

3.1 RESEARCH DESIGN

This research had two aims: Firstly, identifying which factors facilitate and hinder the successful implementation of AI applications in Dutch radiology departments. Secondly, specifying how the hindering factors could be overcome. Due to the early stage of the technological development, the implementation processes of AI applications in Dutch radiology departments are not advanced enough to perform deductive, exploitative research, i.e., by testing to what extent the NASSS model and accompanying factors hold. Therefore, at this stage, an explorative research design is more adequate, in order to gain an in-depth understanding of the underlying processes which influence the implementation of AI applications.

Because the institutional set-up and regulation of health systems differs strongly between countries, empirical studies of health innovation processes are often limited to one country, in this case the Netherlands. The Dutch health care system stands out among its European neighbors through the high health care costs. With 10.5% of the gross domestic product spent on health care, it is among the top five in European Union in 2015 (Eurostat, 2018). The national expenses on health care have strongly increased in the last 30 years. Accordingly, cost containment has been one of the key priorities of the Dutch health care policies (Kroneman et al., 2016; Ministerie van Volksgezondheid Welzijn en Sport, 2018). The rate of growth of health care expenses has been lowered substantially since the beginning of the 21st century. This is partially thanks to agreements between hospitals, insurers and medical specialists, among other actors, on limiting the growth of costs in so-called medical specialist health care. Medical specialist care accounts for approximately 30% of total health care costs and is mainly provided in hospitals (Volksgezondheidenzorg.info, 2019). In the field of medical imaging, the Netherlands provides a particularly interesting case. Comparably low availability of certain medical imaging devices (in comparison to other EU countries) has caused large investments in these technologies over the last decade and more than doubled the number of PET scans, MRI and CT scans between 2004 and 2016 (Kroneman et al., 2016; OECD, 2018). This makes efficiency gains in medical image analysis in the Netherlands particularly relevant.

This research followed a two-step deductive-inductive approach in order to answer the research question, namely which are facilitating and hindering factors for the successful implementation of AI-based software applications for radiology in Dutch hospitals and how can they be overcome. In a first deductive step, the existing NASSS framework was taken as a starting point and adapted to the case of AI applications in clinical radiology. Following the domains of the NASSS framework, a number of guiding propositions were devised. This was achieved by first consulting technical literature from the radiology and data science field, as well as elaborately discussing the framework with two experts, one data scientist and one radiologist, who have familiarity with the topic. In a second inductive step, the adapted framework and the guiding propositions were used to direct an embedded multiple case study approach, involving seven Dutch hospitals as cases, in order to explore and identify hindering and facilitating factors to the implementation of AI applications in the selected Dutch hospitals. The findings were then consolidated in a refined framework, including important interrelations between the individual factors. The final framework was again validated by two experts (one developer of AI applications for radiology and one radiologist).

In general, adopting a case study approach permits to gain in-depth insights into the determinants and processes that impact the implementation of AI based applications in radiology (Bryman, 2012; Liberati et al., 2017; Moja et al., 2014). Embedded case studies allow for more than one unit of analysis. For this research, the main units of analysis were the organizations as a whole, the hospitals, while the smaller units of analysis were the departments (the radiology departments), and groups of individuals (the radiologists) (Scholz & Tietje, 2002). By investigating multiple cases, the implementation of AI applications for radiology could be analyzed under consideration of different contextual conditions. Gathering multiple observations of the same phenomenon allows for comparison and therefore a critical assessment of the data (Scholz & Tietje, 2002; Yin, 2003).

Data triangulation based on different empirical data collection approaches was used: (1) semi-structured interviews with experts for the initial and final validation, (2) qualitative document analysis and semi-structured interviews with members of the participating radiology departments for the case studies, and (3) qualitative document analysis and semi-structured key informant interviews to cover the external environment of the cases, i.e., the wider health system.

3.2 SAMPLING STRATEGY

As common in inductive, qualitative research, a general purposive sampling strategy was followed (Bryman, 2012). Sampling happened on two levels: on context-level, meaning the sampling of radiology departments in Dutch hospitals, and on participant level, meaning the individuals that are considered to be related to the different radiology departments. Overall, the purposive sampling strategy aimed at achieving data saturation, both on context- as well as participant level.

On context level, criterion sampling was followed to identified the Dutch hospitals, which are currently using the AI based diagnostic application BoneXpert in their radiology departments (Greenhalgh et al., 2018; Moja et al., 2014; Powell et al., 2013). BoneXpert, a software-only medical device commercially distributed since 2009, runs automated bone maturity assessments based on X-rays of the hand. It is developed and used for pediatric patients. BoneXpert is currently used in over 70 European hospitals, of which eight are located in the Netherlands (Visiana, 2018). It was chosen as the sampling criterion because it is one of the first commercial applications of an AI-based application in radiology (H. Lee et al., 2017) and because it is the only AI application in clinical use across several hospitals in the Netherlands. From the eight hospitals, seven were included in the sample, due to non-response of the eighth hospital. Of these seven sampled hospitals, four are academic (Universitair Medisch Centra, UMC), two are so-called ‘hospital centers of clinical excellence’ (Topklinische Ziekenhuisen, TKZ) and one is a general hospital (Algemeen Ziekenhuis, AZ). An overview of the seven hospitals can be found in appendix 2.

On the participant level, both documents as well as interviewees were selected in line with the research question and the guiding propositions, following a purposive sampling strategy (Moja et al., 2014). Using a maximum variability logic, individuals occupying different organizational positions were sampled for. Within the cases, individuals were contacted based on their experience with BoneXpert in particular or the implementation process of AI applications for radiology more generally, as indicated by publicly available information or internal referral. The number of participants varied across the seven cases, from 1 to 4 participants per case (see table 1), depending on the availability of interviewees.

Four key informant interviews were used to gather data on the external context of the cases (i.e., the societal & regulatory environment). Therefore, purposive sampling identified one individual per organization of the following professional organizations: the Radiological Society of the Netherlands (NVvR), the Federatie Medisch Specialist (FMS) and the European Society of Medical Imaging Informatics (EuSoMII). Additionally, a representative of a large medical imaging technology provider was included in the sample. These individuals were able to provide more generalized information on the current discussions surrounding AI applications in the radiology profession on the national level (NVvR) and the international level (EuSoMII), as well as on legal-regulatory aspects (FMS, NVvR). Also, interviewing a member of a large medical imaging provider allowed to cover the perspective of developers of AI applications for radiology.

TABLE 1: OVERVIEW OF PARTICIPANTS AND CASES

Cases (7 hospitals studied)	*Number of Interviews*	*Roles of Interviewees*
TKZ1	4	Senior radiologist, Legal consultant, Clinical physicist, Operational department manager
TKZ2	4	Senior radiologist (2), Junior technical physician, Innovation advisor
UMC1	4	Senior radiologists (3), Innovation and Valorization Officer
UMC2	3	Junior radiologist (2), Senior data scientist,
UMC3	3	Senior radiologist (2), Senior data scientist
UMC4	1	Senior radiologist
AZ1	1	Senior radiologist
External Organizations		
NVvR	1	Managing director
FMS	1	Implementation advisor
EuSoMII	1	Chairman
Imaging Technology Provider	1	Innovation lead
Total Number of Interviews	*24*	

3.3 DATA COLLECTION

Data collection was based on document analysis and semi-structured interviews. For the document analysis, the following type of documents were considered: on the one hand, publicly available documents, such as academic literature, grey literature, press articles and policy/regulatory documents. These documents were used to understand the external context, such as regulatory aspects and the societal stance towards the technology. Additionally, official documents and strategy documents from professional organizations and other active organizations in the health care sector provided insights into official stances of concerned organizations. On the other hand, when available, internal documents of participating radiology departments were included in the document analysis. Internal documents, policy and regulatory documents, scientific literature and grey literature and other official documents were purposively sampled for. For internal documents, sampling identified documents related to innovation strategies and implementation processes of AI applications in the participating organizations. For grey literature, policy and regulatory documents, the sampling criterion was the topic of AI technology in health care in The Netherlands. For scientific literature, articles that covered implementation or adoption of AI applications in radiology were sampled for. Due to the low number of documents in these categories, it was attempted to include all identified documents published until the May 2019 that provided relevant information. The sampling approach for press articles was structured. In order to gain an overview of the current societal environment for AI applications in the Netherlands, Dutch press articles, published between the 01.01.2018 and the 24.05.2019, were searched on NexisUni, using the following keywords: kunstmatige intelligentie, geneeskunde, zorg, gezondheidszorg, nederland.[8] From the initial 348 documents a manual scan of title and abstract identified 71 articles specifically related to the research topic.

TABLE 2: OVERVIEW OF SAMPLED DOCUMENTS

Type of Document	# of documents
Internal documents	4
Governmental regulation/Policy documents	6
Official documentation & strategy documents from organizations active in the Dutch health care sector	8
Grey literature	8
Press articles	71
Scientific literature	5

From February to June 2019 a total of 24 interviews were conducted. As is common in inductive research, interviews were held in a semi-structured way, in order to allow for flexibility with regard to topics and emphasis, following the interviewees' inputs. This applies both for interviews within the cases, as well as for the key-informant interviews. Based on the theoretical considerations, an interview guide, including a list of questions on topics to be covered, was used and adapted for each interviewee (see appendices 4 & 5 for generic interview guides). In addition to discussing pre-identified topics, the semi-structured nature of the interviews allowed for new themes to emerge during the conversations. While the application BoneXpert was used as a sampling criterion for the cases, the semi-structured interviews covered not only the interviewees' experience with BoneXpert, but with AI applications in general. Interviews were conducted until the point of thematical saturation was reached, meaning when no new themes appeared during additional interviews. Depending on availability, interviews were conducted personally (21 IVs) or alternatively by telephone (3 IVs). With the exception of one interview, interviews were held in English and lasted between 20 minutes and 80 minutes. Oral permission for recording was granted by all interviewees. The interviews were subsequently transcribed and coded. A complete overview of interviewees can be found in appendix 1 (excluding names to guarantee anonymity).

[8] The exact bolean search term was: ("kunstmatige intelligentie") and (geneeskunde or zorg or gezondheidszorg) and nederland*

3.4 DATA ANALYSIS

Based on the initial theoretical framework and propositions, a preliminary analytical framework was developed. Following the main domains of the NASSS framework (Greenhalgh et al., 2017), sub-domains were adapted to the case of AI applications for radiology. Because the domains and sub-domains of the framework are very broad, specific concepts for AI applications in radiology were derived from the theory and the initial expert interviews. The preliminary analytical framework can be seen in table 3.

The transcribed interviews and documents were subjected to several rounds of coding using NVivo. Due to the deductive-inductive approach followed in this research, coding was done as an iterative process aiming at thematical saturation. The coding rounds started open and gradually turned more focused (Corbin & Strauss, 1990). In the first round of coding, open coding was used to identify initial concepts from the interviews. In a second round of coding, additional concepts were identified from the interviews and the documents. Axial coding was employed to combine initial concepts in subcategories and categories (following the NASSS framework terminology, these are called domains and sub-domains) and compare them to the preliminary analytical framework. After a third round of axial coding, which did not lead to new concepts, but allowed to redefine categories and subcategories and explore relationships among categories, the point of theoretical saturation was reached. The final refined analytical framework, including coding rules, can be found in appendix 5. The output of the data analysis process is presented in the chapters 4, 'Results', and 5, 'Analysis'.

TABLE 3: INITIAL ANALYTICAL FRAMEWORK, BASED ON GREENHALGH (2017)

Domain	Sub-domain	Concepts
1. Condition	1.1 Nature of condition	Specificities of radiology subspecialties which implement AI applications (e.g. routines/collaborations with other departments)
2. Technology	2.1 Material Properties	Technical performance of algorithm (inability of program to interpret images/ other unforeseen errors)
		Integration of innovation in existing workflows/ PACS
	2.2 Knowledge/data generated	Easy Interpretation of application's output (radiologists and non-radiologists?)
		Basic understanding of the type of analysis and mechanism behind AI application by user of technology
	2.3 Knowledge to use/ necessary support	User-friendly design of application.
		Understanding of limitations of algorithm and associated risks (i.e., false negatives/ false positives)
		Availability of IT support (internal and/or external)
	2.4 Supply Model	Level of customization necessary to implement application (e.g. "plug & play"/ pay by use applications)
3. Value Proposition	3.1 Supply-side value (developer)	Clear business case for developers
	3.2.1 Demand-side value (to organization/management)	Clear desirability: responds to a concrete organizational problem
		Clear efficacy gains
		Quality gains
		Assured safety
		Cost-effectiveness

	3.2.2. Demand-side value (to radiologists)	Clear efficacy gains
		Desirability of solution for concrete problem
4. Adopters	4.1.1 Staff: Radiologists	Amount of new knowledge/skills needed
		Effort of adaption of workflows and practices
		Changes in role of radiologist with regard to other staff members
		Changes in professional identity of radiologist
		Perceived risks associated to technology
	4.1.2 Staff: Technicians	Amount of new knowledge/skills needed
		Effort of adaption of workflows and practices
		Changes in role of radiologist with regard to other staff members
		Changes in professional identity of technician
	4.1.3 Staff: other	Additional staff members affected by technology
	4.2 Local Champion	Presence of local champion that pushes implementation agenda forwards
5. Organization	5.1 Capacity to Innovate	Availability of slack resources for implementation efforts
		Strength of leadership/relations between management and concerned staff
		Innovative organizational climate: favoring of risk-taking
	5.2 Readiness for this technology	Tension for change: perceived organizational problem which needs solution (e.g. too high workloads in radiology)
		Innovation System-fit: technology aligns within broader innovation strategy
		Level of support across organizational sub-units.
	5.3 Nature of adoption/funding decision	Sufficient resources on organizational level for implementation and maintenance/continuous support of technology
		Clear internal financing (how are costs of technology covered within organization)
		Clear external financing (how are cost for technology covered by insurers)
		Additional/New staff needs (IT support/data scientists)
	5.4 Extent of change needed to organizational routines	Establishment of new team routines necessary within radiology departments
		Establishment of new organizational routines/ best-practices across departments (e.g. radiology-pediatrics)
	5.5 Work needed to implement and evaluate the change	Building of shared vision easily done
		Enacting of new practices easy to monitor (task clearly assigned to individual/group)
		Monitoring of impact uncomplicated (task clearly assigned to individual/group)
6. Wider System	6.1 Political/Policy context	Political context around technology uncontested.
	6.2 Regulatory/legal issues	Regulation of specific technology existing and clear (on national and/or supranational level)
		Legal certainty with regard to responsibilities (individual/organizational)
	6.3 Professional bodies	Positioning and guidance from professional body (Nederlandse Vereniging voor Radiologie)
	6.4 Socio-cultural context	Societal opinion towards AI technology

7. Embedding & Adaptation over time	7.1 Scope for adaptation over time	Expectations of technological development are positive
		Intention for adapting and increasing use of technology over time
	7.2 Organizational resilience	Encouragement of collective reflection & continuous adaptation to context by the organization

3.5 RESEARCH QUALITY INDICATORS

Ensuring measurement validity and internal reliability was pursued in two ways: On the one hand the data collection approach was based on data triangulation of document analysis, qualitative interviews and expert interviews. One the other hand, during the data collection and the data analysis process findings were repeatedly discussed with a secondary researcher, amounting to investigator triangulation.

Internal validity was pursued by applying a rigorous data analysis approach, including constant comparison between the empirical data and the emerging theoretical concepts and by applying strict coding rules (Bryman, 2012). In order to avoid inconsistent inferences, the description of the results and analysis stayed as close to the original data as possible, and identified concepts were supported by quotes from the interviews. By closely documenting the empirical and analytical approach followed during all stages of the research process, it was attempted to make the research as replicable as possible (Bryman, 2012). The interview guides, as well as the final coding framework can be found in the appendix (appendix 3-5). Transcripts of the interviews can be obtained on request from the researcher.

As is common for qualitative case-study based research, external validity is limited by the non-probability sampling strategy. Due to the contextual particularity of each case, conclusions from this research may not necessarily hold for the larger population of hospitals in the Netherlands. However, triangulation across the multiple cases in the sample allowed for the distinction between findings that were common to a majority of the cases and findings that were unique to one particular case. Also, external validity was pursued by including data, which was not especially created for this research (i.e., the document analysis).

4. RESULTS

The seven Dutch hospitals studied vary strongly in the number and type of internally available applications, which fall under the broad scope of AI-based applications in radiology and which were therefore investigated. These range from the detection and quantification of lung nodules in CT scans, mammography CAD systems, to stroke detection and to automated bone-age assessment (BoneXpert). Because the hospitals were sampled for based on their use of BoneXpert, this application was investigated more in-depth. However, the implementation processes of other AI applications are equally covered.

The results are presented following the lines of the seven domains of the NASSS framework by Greenhalgh et al., (2017). As a result of the data analysis process, a refined list of concepts for each domain and subdomain emerged, highlighted in brackets. The concepts show the specific aspects that influence the implementation of AI applications for radiology for each subdomain. The results are first and foremost based on the interviews with members of the hospitals, who provided insights on different aspects of the implementation process based on their respective expertise. Only when explicitly mentioned, the findings resulted from the document analysis. Section 4.1 deals with clinical condition targeted by AI applications, section 4.2 explains the technological aspects surrounding AI applications in radiology, followed by the value proposition of AI applications in section 4.3. In section 4.4, the adopter system of AI applications in radiology is discussed, followed by the organizational context surrounding the implementation of AI applications in radiology in section 4.5. Section 4.6 displays findings concerning the wider system around AI applications in radiology. Finally, section 4.7 looks at the implementation over time.

4.1 CLINICAL CONDITION TARGETED BY AI APPLICATIONS FOR RADIOLOGY

This part will discuss both the technical requirements of AI applications depending on the targeted clinical condition in section 4.1.1, as well as the different adopters that are involved depending on the condition in section 4.1.2.

4.1.1 Technical requirements for AI applications for radiology

The medical conditions targeted by different AI applications are situated in different clinical scenarios, which vary with regard to their degree of clinical emergency and the potential treatment decisions made based on the diagnosis. Scanning for and quantifying tumors, for example, is a rather lengthy diagnostic process, in a non-emergency setting, which requires extreme attention to detail and precision (IV 3). In contrast, detecting a stroke case happens in an emergency-setting and requires a very quick diagnosis.

> *"Time is one thing. Because, when a stroke patient comes, time means saving brain for the patient. When a patient comes here with a stroke, you have to determine if the patient is eligible for intervention or not. And the software can do that for you. So, it's more accurate and it's faster than when I do it." (IV 8)*

It occurs that different AI applications can and are expected to provide different forms of added-value (e.g. time-gain, increased accuracy) to clinical radiological practice (IV 3, 8, 22). These depend on the medical condition targeted by the AI application. *(Desired clinical benefit for condition)*

Furthermore, the clinical situation and especially the degree of emergency of the targeted condition also appear to require differing standards of technical performance (IV 1, 14, 22). While the risk of an algorithm missing a hand fracture can lead to a prolonged period of pain for the patient, an algorithm missing a stroke can be a question of life and death. Thus, an AI application used in an emergency setting requires higher technical quality standards than an AI application in a non-emergency setting (IV 13). *(Quality standards required for condition)*

> *"We will rather accept missing 2 cases of cancer per thousand, than increasing the number of false positives [in the breast cancer screening program]. So, it's a very delicate balance. […] But also the anxiety: How much do women believe in the quality of the screening program if everyone will be recalled with a false positive result." (IV 22)*

Overall, the medical condition targeted by a specific AI application was found to determine the required quality standards for that AI application and its desired clinical benefit.

4.1.2 Adopters involved in the implementation of AI applications for radiology

Depending on the clinical condition targeted by an AI application, different individuals will be involved in the diagnostic process. On the one hand, the sub-set of radiologists, who will directly use the AI application, varies depending on the radiological subspecialty in question by the application. Also, depending on the condition, the request for the medical image and diagnosis comes from a different group of referring clinicians (e.g. neurologists for brain-scans, oncologists for detection of tumors, etc.). Thus, the clinical condition targeted by a specific AI application determines both the group of direct users (radiologists) and the group of indirect users (referring clinicians). For the case of BoneXpert, it became apparent that the diagnostic process of conducting a bone age assessment (which BoneXpert aims to automate), does not exclusively happen at the radiological department. In fact, pediatric endocrinologists account for most requests for bone age assessments received by the radiology department. It appeared from the interviews that although radiologists perform the scan and the bone age assessment and communicate the results in their report, endocrinologists often redo the bone age reading themselves, meaning the diagnostic work is done twice (IV 5, 10, 21). This has historic reasons, as bone age assessment has been part of the training of pediatric endocrinologists (IV 21). In two cases (UMC1, UMC4) it was mentioned that this practice did not change with the introduction of BoneXpert. *(Direct and indirect users)*

Thus, it appears that the group of direct and indirect users of an AI application is determined by the medical condition it targets. Mostly, the delineation of the user groups follows the lines of medical subspecialties.

4.2 AI APPLICATIONS FOR CLINICAL RADIOLOGY

This section concerns the technological aspects of AI applications for radiology and starts with discussing the technical features of AI applications in section 4.2.1, followed by the comprehension of the output of AI applications in section 4.2.2, the use in clinical practice in section 4.2.3 and, finally, the supply models of current AI applications in section 4.2.4.

4.2.1 Technical features of AI applications

The term 'artificial intelligence' has been around in medical imaging for several decades and, until present, does not have a uniform definition. A common technical distinction is made between deep-learning and non-deep learning algorithms. This distinction, however, does not easily translate to clinical practice, where distinctions are rather made on the type of functionality and/or the position in the diagnostic workflow.[9] From the technical side, algorithms are usually assessed by their performance, i.e., the sensitivity (number of false positives) and specificity (number of false negatives). In clinical terms, having too many false positives means that the radiologist is forced by the application to analyze a large number of lesions that are in-fact nothing. This is very time-consuming and can lead to low confidence in the programs' accuracy, a problem mentioned for the mammography CAD systems available in the analyzed hospitals (IV 2, 8, 10). Although available in several hospitals in the sample, these CAD systems were not routinely used by the radiologists (IV 2, 22). Having a large number of false negatives is even more dangerous, because it means that a potential lesion is missed (IV 6). Even though avoided at any cost, missed or erroneous diagnoses are common in clinical radiology, and can be circumvented by having two radiologists read every scan ('double-reading'). Interestingly, it appeared that radiologists expect a computer program to have a much lower incidence of false negatives than human readers do (IV 11). (*Technical performance)*

While performance-metrics are usually provided by the developers of the software, information on the nature and origin of the data used to train the algorithm can serve as another important quality-indicator (IV 3, 20). Yet, only a validation in the local context can trustworthily assess the technical performance of the algorithm for that specific hospital (IV 6, 10, 14, 16). For BoneXpert, it was found that in-house validation is not consistently done across hospitals. While it was given high importance in some cases (e.g. TKZ1), other cases relied on the metrics provided by the developers or simply the presence of CE mark (e.g. TKZ2). The CE mark is a certification mark required for medical devices in the European Economic Area. It indicates that these devices are conform with relevant regulatory directives. *(Local validation)*

Performing local validation requires the application to be available for use on the hospital's computers and can therefore not be done up front. Thus, radiologists and other actors involved in the adoption process rely on published scientific evidence on the technical performance of the algorithm to assess its quality before engaging in negotiations with developers. For example, there are several peer-reviewed publications, which show the technical performance of BoneXpert (Geldermann, Grouls, Kuhl, Deserno, & Spreckelsen, 2013; Martin et al., 2009; Van Rijn, Lequin, & Thodberg, 2009). Interviewees, however, noted that most of the published validation studies were not based on evidence from testing the algorithms in a clinical setting (IV 3, 5). Also, a lack of evidence on the potential positive benefits for clinical practice was observed, for example attempts to scientifically measure time gains which result from an AI application (IV 22). *(Empirical evidence)*

The smooth integration of AI applications in existing IT infrastructure showed to be one of the crucial elements for its successful implementation, mentioned by all interviewed radiologists and technical staff. Radiologists use a series of different IT applications in their clinical workflow, of which the most important are the electronic health record, the radiological information system

[9] Van Ginneken (2018) proposed the following classification: AI r – replacing a task of a radiologist: BoneXpert; AI a – assisting with a task for the radiologist (make it more accurate, faster, less tedious); AI x – extending (doing something a radiologist/clinician does not do todays and will/can never do).

and the Picture Archiving and Communication System' (PACS). While the radiological information system manages the workflow in the imaging department (scheduling, reporting, resource management) and is complementary to the hospital information system, the PACS is the key interface for image visualization used while reporting. Therefore, a smooth integration in the PACS is imperative for AI applications, and was mentioned as a key concern by all interviewed radiologists. The integration of BoneXpert in the PACS was perceived as very smooth by all interviewed users, a central reason to its perceived user-friendliness (IV 1, 2, 5, 21). In practical terms, this means that the output of the AI software needs to be displayed within the PACS interface with the least extra clicks as possible. When direct PACS integration is not possible, it needs to be clear and easy to find the AI applications in the IT system (IV 5, 14). *(Integration in existing IT systems)*

> *"But one of the things are that there are so many different programs in the package. And because we do all body parts for the children, we would have to know all that software, but that's so complicated and so much that we basically don't use it now. It's not very user-friendly so far." (IV 5)*

The responsibility of aspects of implementing the applications within the existing IT infrastructure lies with the IT department and the medical physicists, who, in consultation with the radiologists, define how to ideally integrate the application in the existing workflow and IT systems and provide support if needed. *(IT support)*

Often, however, it is the readiness of the existing IT infrastructure of hospitals which proves as a bottleneck for the implementation of AI applications (IV 3, 5, 7, 8, 12). Because this type of software package can use a lot of computing space and/or internet connection for cloud-based services, it can happen that the results of the algorithm take a long time to be visible in the PACS, or in the worst case that the workstations crash when opening a heavy software package. In one case (UMC3), the results of BoneXpert sometimes take hours to show in the PACS, which delays the reporting of these cases significantly. Although this problem is not new and is usually well known by the hospital management, it becomes increasingly problematic as more separate software packages are added. In the worst case, it can lead to the non-use of the applications. *(Readiness of IT infrastructure)*

In summary, it can be stated that the technical performance of AI applications and their integration in existing IT infrastructure are particularly important elements for the implementation and were mentioned by all interviewed radiologists and technical staff.

4.2.2 Comprehension of output from AI applications

As became evident from the interviews, computer science and programming knowledge needed to develop AI algorithms does not fall under the typical competences of radiologists. However, interviewees consider it important and useful to have some understanding of the technical aspects, for example in order to assess its quality and safety and therefore create trust (IV 1, 11). *(Understanding of the technology)*

A proactive minority of radiologists searches to understand the technical aspects by reading scientific literature. It was mentioned that it is necessary to teach all radiologists some basic knowledge of computer science (IV 3, 13). There are ongoing efforts to include AI in the national curriculum for future radiology residents, whereas more senior radiologists can receive trainings or participate in courses (IV 2, 23). An attempt of such a course was done in the TKZ1, but was not positively received by the group of radiologists, who displayed little interest in learning about the technical aspects of AI.

> *"So, I asked [a professor for data science] to come over to the hospital to talk with my colleagues. He did a marvelous presentation. He explained, in a simple way, how it [AI/Machine Learning] works. And then, you can react positively: 'oh it's marvelous, we have to do something with it', or you can react quite negatively. And it's mostly the latter that happened. My colleagues asked him a couple of not so nice questions and that was it." (IV 1)*

While it is yet to be decided how technical these courses ought to be, they are seen as a potential form of building trust in AI technology among radiologists. The lack of understanding of the data science behind the algorithms makes it difficult for radiologists to reconstruct how the algorithm comes to its output and therefore limits radiologists' trust in the outputs (IV 8, 11, 13, 14). *(Technical training)*

This issue is particularly pronounced in cases where radiologists disagree with the output of the algorithm. Because they cannot reconstruct, radiologists cannot explain or resolve these contradictions (IV 15). On the one hand, radiologists fear that referring clinicians might begin to trust the algorithm more than the expertise of the radiologists (IV 5, 11). On the other hand, it is feared that radiologists might become hesitant to challenge the algorithm in the first place (IV 3, 11). *(Dealing with human-machine contradictions)*

It follows that radiologists do not have sufficient understanding of the technical aspects of AI applications to assess their quality and to fully trust the applications.

4.2.3 Use of AI applications in clinical practice

Besides lacking knowledge on the technical mechanisms behind AI applications, interviewees mentioned another important knowledge-gap, namely lack of know-how on the implementation of the application in the form of guidelines or best practices (IV 3, 5, 9, 12, 16). Because the overall adoption of AI applications in clinical practice is at an early stage, there is no scientific evidence available on these aspects. *(Guidance on Implementation)*

The lack of guidance on how to implement AI can result in unstructured or unguided implementation of AI in the clinical workflow. Just as the technical integration mentioned in 4.2.1, the integration of AI applications in the clinical workflow is one of the crucial elements for successful implementation, because it directly affects perceived user-friendliness of the AI application by radiologists. There are two aspects to be taken into account: firstly, the degree to which the integration in the workflow is standardized within the department, meaning whether all relevant radiologists use the application in the same fashion in their diagnostic process (IV 2, 14, 16). It appeared that even within teams of four radiologists (e.g. four pediatric radiologists), the frequency of use is not the same: while one radiologist uses the AI application every time he performs a certain diagnosis, another radiologist only uses it, if he has a doubt and another one does not use the AI application at all for the same diagnosis.

> *"Still one of my colleagues doesn't use [BoneXpert]. And one was sort of hesitant, she's now convinced that she can use it. And my other colleague uses it. Says it's easy to use." (IV 21)*

Secondly, the role the application performs in the diagnostic workflow also differs within and across departments. Two different roles were identified: supporting the diagnostic process of the radiologists (either as a first or a second reader of the scan) or (fully) automating the diagnosis. It was found that in the case of BoneXpert, the same application was used in a different role in different hospitals (IV 21). *(Implementation in workflow)*

Overall, it can be stated that currently the use of AI applications in clinical practice is not well standardized and therefore varies strongly between members of the same radiology department, and even more so across radiology departments in different hospitals.

4.2.4 Supply models of AI applications for radiology

AI applications are currently being sold to hospitals in two types of business models: on the one hand 'reading as a service', meaning the developer receives a license fee for the scans read. This fee is either charger per analysis ('pay per view') or as a fixed amount per year ('package deal'). On the other hand, AI applications can be integrated in existing platforms or software packages (such as a PACS). The 'reading as a service' model requires large distribution and sales activities, since it necessitates negotiations with each individual client, for example the hospital or clinic (IV 20). Many smaller developers, such as BoneXpert developer Visiana started off with this model. The larger medical IT companies generally include AI applications within existing products, either as a new functionality in their products or as a (costly) add-on to these products (IV 1, 20). Simultaneously, these larger IT companies are developing platforms, where developers can upload their algorithm. These platforms are intended to function as a sort of 'app-store 'for AI algorithms (IV 3, 6, 13, 20). *(Business models)*

> *"During the two years we negotiated with the startup, our existing PACS system developed more in the way of what the startup offered. So, there was no real use anymore to go through with buying their solution. And we hope that the next version of the PACS system gets close to what they [the startup] can do." (IV 8)*

It can thus be said that although not one single business model for AI applications for radiology has emerged as dominant, there is a tendency towards integrating these algorithms in existing products or larger platforms. This responds to the user demands of smoothly integrated solutions and can potentially overcome barriers to purchasing individual applications (see domain 4.3.2).

4.3 VALUE PROPOSITION FOR DEVELOPERS AND ADOPTERS OF AI APPLICATIONS FOR RADIOLOGY

The third domain, the value proposition, highlights the findings concerning the added value of AI applications for clinical radiology for the supply-side, namely the developers, and the demand-side, namely hospitals and clinicians. The discussion of the business case for AI applications in 4.3.1 is followed by considerations regarding the added value for clinical practice of AI applications for radiology in 4.3.2.

4.3.1 Business case for AI applications in radiology

Different types of actors were found to be active in the development of AI applications for medical image analysis. These range from a large number of small start-up companies (often spin-offs from universities or research institutes), to existing radiology and medical IT companies (such as PACS providers), to medical technology companies, to large non-medical technology companies (such as Google, IBM, Amazon and Facebook). These developers have very different financial and non-financial resources (IV 20). *(Different providers)*

Like all software, AI applications display economies of scale, meaning they have very high initial development costs and very low costs per extra unit. The high development costs are mainly due to the timely task of annotating training data for the algorithms and designing the user interface. Therefore, one of the key concerns, especially for the smaller developers on the market, is how to retrieve the initial investments (IV 1, 3, 20). It was found that investments in the development of AI applications for medical imaging have been strongly increasing (Harris, 2019; He et al., 2019). *(Development)*

Currently, there are two main types of customers for developers of AI applications for radiology: health service providers (e.g. hospitals), which are the focus of this research, and pharmaceutical companies. As mentioned, the hospitals in this research do not have many fully commercial programs implemented yet. Most had one or several test versions running, but these test versions mostly do not generate revenue for the developer. Thus, at this stage, an important source of revenue for developers of the algorithms is the pharmaceutical industry, which uses AI algorithms to gain more precise and much cheaper analyses for its clinical trials (IV 20, 21). Depending on the application, another potential source of revenue for certain developers are screening organizations, which have a very high volume of scans and therefore a potentially large revenue stream (IV 3). Existing screening programs in the Netherlands are among the most efficient and performing in the world. This means AI applications would have to demonstrate very tangible efficiency or quality gains, to be adopted by screening organizations in the Netherlands (IV 22). *(Revenue streams)*

Besides having hospitals as their target client-group, developers of AI algorithms also cooperate with hospitals and clinics during the development and mainly the validation phase of their products. They need data and clinical testing from hospitals and in return often make test versions available for free. This type of collaboration happens in all four academic centers, which are part of this research, as well as two of the non-academic hospitals. Small developers are also collaborating with the bigger incumbents, such as *Philips Healthcare* or *Siemens Healthineers*, to get their algorithms or applications integrated in their platforms and/or get access to the larger companies' networks of clients and partner hospitals. *(Cooperation)*

Overall, it appeared that the business case for AI applications is not yet consolidated, since AI application providers lack consistent revenue streams to cover their high initial development costs.

4.3.2 Value of AI applications for clinical practice in radiology

Because very few AI applications are being used standardly in clinical practice and because there has been very little research on the impact of the applications in clinical practice, the actual added-value for the demand side is still hard to assess. This means that many of the elements in this section are *potential* benefits, which AI applications may have, as promoted by developers or expected by clinicians. Even for BoneXpert, which has been in use for several years in certain cases, a systematic assessment of its added-value has not been done in any of the cases researched for this study.

The list of *potential* benefits of AI-based applications for radiological clinical practice is long. These benefits can be broadly divided in two categories: improved diagnostic practice and operational benefits. In the category improved diagnostic practice the following benefits of AI applications were mentioned (from most to least frequently mentioned): more precise diagnosis (mentioned by 13 interviewees), avoid mistakes (mentioned by 11 interviewees), automate cumbersome tasks (mentioned by 9 interviewees), more objective diagnosis and gaining additional information. In the category operational benefits, the following benefits were mentioned: time-saving (mentioned by 15 interviewees), more consistent reporting across radiologists (mentioned by 12 interviewees), remedy against increasing workload (mentioned by 11 interviewees), always available, i.e., the software does not get sick or need sleep like humans. All of these benefits were perceived as desirable and respond to existing clinical needs (see domain 4.5.2).

> *"But BoneXpert is a good example of where this can help you because it's much faster and more reproducible and more precise. And if that's the case, you should immediately leave it to such a program." (IV 13)*

AI applications can have more than one potential benefit and often mix operational and diagnostic benefits. It appeared that clear improvements in efficiency might outweigh improvements in quality of diagnoses for two reasons: first, their return on investment is more easily demonstrated (IV 11, 12). Second, the usefulness of applications which provide (marginal) improvements in diagnostic quality was questioned by some radiologists (IV 2, 5, 6, 8). *(Clinical benefit of AI applications)*

Next to efficiency and quality, gaining experience with AI was mentioned as an added-value by the hospital management in the TKZ1 and AZ1 (IV 1, 9, 16). From a management perspective, the potential clinical benefits need to be contrasted to the costs of the applications. On the one hand, these applications promise cost savings through more efficient operations or through lowering personnel costs by being able to hire less radiologists (IV 4, 13, 16, 22). Just like the clinical benefits for radiologists, these cost saving benefits need yet to be proven and are currently of a more hypothetical nature. On the other hand, these applications have a purchasing price. This additional expense needs to be incorporated in the budget (see section 5.3) (IV 1, 2, 22). Interestingly, opinions differ across interviewees about what constitutes high or low costs for AI applications.

> *"It's expensive. So that's at least what I know that the cost of BoneXpert are approximately similar to what we are getting for the image." (IV 6)*
>
> *"So then [BoneXpert] was implemented, so and these costs are not very high. […] You get a package deal of 2600 Euros per year. And this is just a small amount of money for the department." (IV 2)*

Although on the first glance, the main added-value on the demand side lies within the radiology department, potential benefits for referring clinicians, patients and society were also mentioned by interviewees. As the final recipient of the diagnoses, the referring clinicians also benefit from more precise and objective diagnoses through the use of AI applications (IV 1, 2, 5, 6). For example, more precise diagnoses of mammograms can potentially avoid expensive and painful biopsies, leading to more precise treatment and better overall care to the patient (IV 3, 5, 8). Finally, AI applications can potentially also have positive outcomes on societal level, such as a reduction in costs and a better access to high-quality specialist care. *(Indirect demand-side benefits)*

While the list of potential benefits for several stakeholders is long, real evidence, especially in quantitative form, is scarce. One reason is that measuring these benefits on the micro-level is difficult, for example how to measure increases in the quality of diagnosis (IV 4, 6, 12). To present, there is no standard methodology on how these benefits can and ought to be measured (IV 3). This is also related to the fact that no quantitative information needs to be submitted in order to receive regulatory approval in the EU. *(Measuring demand-side value)*

Potentially, AI applications for radiology have substantial direct benefits for clinical practice (e.g. saving time) and indirect operational benefits (e.g. saving costs). Yet, to present there is no satisfactory evidence on these benefits, due to the difficulty in quantitively measuring the potential demand-side benefits.

4.4 ADOPTERS OF AI APPLICATIONS FOR RADIOLOGY

The fourth domain, the adopter system, looks at AI applications for radiology from the perspective of the intended users. It focuses on the reaction of users towards AI applications, and the changes induced by the technology. These aspects will be discussed for the direct adopters, namely the radiologists in 4.4.1, followed by the indirect adopters, namely the referring clinicians in 4.4.2. Section 4.4.3 will show the role of the 'local champion' for the implementation process of AI applications.

4.4.1 Direct adopters of AI applications for radiology: the radiologists

The acceptance of using AI applications in clinical practice differs greatly among individual radiologists. For all hospitals, it was described that some members of the radiology department viewed the technological development more positively than others (IV 1, 2, 7, 11, 12). More generally, the reactions towards using AI technology in clinical radiology range from outright enthusiasm, to curiosity, to skepticism, to fear. These differences in opinion across radiologists were also visible for specific applications, such as BoneXpert. Although the opinion on BoneXpert was predominantly positive across the radiologists in this research (IV 1, 2, 5, 9, 21), some of the interviewees mentioned critical voices among their group of colleagues.

> *"And interestingly enough, even if you've shown that [BoneXpert] works. Because there were multiple studies at the time, showing that it actually does work better and more consistently [than human readers]. […] Still one of my colleagues doesn't use it." (IV 21)*

While in some cases, radiologists refused to use BoneXpert without considering the evidence presented by their colleagues or in scientific publications, others became more critical over time, after having experienced the program to make erroneous assessments.

> *"And there is quite a difference in the age, I noticed that myself, between the BoneXpert and the Greulich and Pyle. At one point we thought, let's just do the BoneXpert [and no manual reading]. That's much easier. But the orthopedics didn't want that. Because they noticed that there is quite a discrepancy." (IV 5)*

This large variation in how radiologists perceive the quality of BoneXpert, leads to different ways of using BoneXpert in clinical practice. This difference was observed across different hospitals, as well as within hospitals. While some radiologists use it as an automated diagnosis, others only use it to double-check their manual reading. This means that the potential benefits of using BoneXpert in clinical practice, namely saving time and more consistent reporting, are de-facto undone. There was no consensus on the origin for these different levels of acceptance among the interviewees. While some believe that there is a generational gap, with more junior radiologist being more open to the use of AI applications than senior ones (IV 11, 12, 14, 21), others contested this idea and argued that acceptance of AI applications rather depends on the personality types, i.e., being more or less open to change (IV 2, 7, 8). *(Variance in acceptance)*

Two cases (TKZ1, UMC2) had experienced active opposition towards AI by radiologists, meaning some radiologists actively blocked efforts by 'local champions' to implement AI applications within the radiology department. This form of negative reaction by radiologists was not based on actual use-experience with AI applications (as it preceded the implementation of the AI applications). Rather, it seemed to emerge from a subjective sentiment towards the technology, resulting from reading and hearing about the potential impact of AI on their profession on congresses and in related publications (IV 1, 11, 13). *(Opposition)*

An important element towards the acceptance of the technology by radiologists seemed to be trust. On the one hand, radiologists need to trust the AI technology enough to accept its output and actually use it in clinical practice (IV 1, 3, 9, 13). On the other hand, it was mentioned that there is a risk that radiologists trust the technology too much, meaning they too easily accept the data generated, failing to double check the results (IV 3, 6, 14, 21). Linked to the question of not trusting AI is also the feeling of fear. Repeatedly, radiologists mentioned that they are afraid of being held accountable (ethically or legally) for mistakes done by AI applications (see also part 4.6.2) (IV 4, 8, 14, 15). *(Trust)*

> *"[My colleagues] don't trust the software [BoneXpert]. They think that we as radiologists can do better than a computer. There's still some anxiety, angst for artificial intelligence. And interestingly my residents trust it blindly. Too blindly sometimes even. They don't think anymore. And that's the risk of AI, that you stop thinking." (IV 21)*

Until present, AI applications have not induced a large change in roles and practices of radiologists. Roles and practices are still determined on the level of the organ-based subgroups. Thus, the sub-group(s) of radiologists targeted by a specific AI application is most central in the process of redefining best-practices. For BoneXpert, this concerned the pediatric and/or musculoskeletal radiologists. Considering that bone age assessments play only a very small part of the total workload of pediatric and musculoskeletal radiologists, the changes induced by BoneXpert were perceived as very small (IV 1, 5, 10, 21). *(Change in practices)*

Repeatedly, the need to create a specialist group on AI or medical imaging informatics was mentioned (IV 3, 5, 9, 16). This group would take the role of scanning for potential applications, advising the adoption process and coordinating the implementation of such applications. Such an 'AI expert group' is present in different forms in five cases (TKZ1, UMC1, UMC2, UMC3, AZ1). This specialist group could potentially even be officialized in the form of a new subspecialty (IV 3, 9). *(Creation of new roles)*

In line with creating AI-experts among the radiologists, several interviewees mentioned the need to reframe their professional identity as a consequence of the arrival of AI applications (IV 3, 6, 7, 22). Following the strategy paper of the Radiological Society of the Netherlands, the radiologists of the future will become an 'imaging consultant' which involves being a more active part in an interdisciplinary patient-focused hospital environment (Nederlandse Vereniging voor Radiologie, 2016). Radiologists' professional identity was also mentioned in a second context: historically, radiologists have been frontrunners in the adoption of novel technology, when compared to other medical specialties. Interviewees reasoned that because radiology was among the first to adopt digital technology or speech recognition, radiologists also need to be leading the adoption of AI within the hospital (IV 2, 5, 13).

However, the topic of professional identity also appeared in a more negative context: the possibility of AI replacing radiologists, and thereby threatening their professional identity, was extensively mentioned in recent radiological publications and repeatedly mentioned in the documents analyzed for this study. Interestingly, none of the interviewees agreed with this idea, and most of them had not heard any of their colleagues explicitly expressing this fear of being replaced by AI. *(Professional identity)*

> *"Sure! There are people who are critical. But I think that the people who are afraid for their job don't know exactly what these types of algorithms can do. So it's a matter of properly educating them." (IV 13)*

While the interviewees did not fear to be replaced by AI, they did mention a more imminent menace coming from AI applications, namely for radiologists to be bypassed by other members of the hospital (IV 1, 4, 7, 11).

> *"One of the things many radiologists are afraid of is that the clinicians themselves are starting the AI collaboration with the companies, for instance, and they just bypass radiology. […] I can imagine that as a physician, you're just asking an exam from the radiologist. And If you have really good software that can do the work on itself, you don't need the radiologists anymore." (IV 7)*

By allowing other clinicians to execute the interpretation of the medical image themselves, AI applications might cause the radiologist to lose part of his or her area of responsibility. *(Loss of responsibility)*

It became evident that radiologists vary strongly in their acceptance of AI applications, in some cases even actively opposing the introduction of AI applications. Several potential causes were identified, such as lacking trust of the applications' output or changes to the professional identity of radiologists caused by AI applications.

4.4.2 Indirect adopters of AI applications for radiology: the referring clinicians

Although, none of the cases in the sample mentioned that referring clinicians were de facto using AI applications for the interpretation of medical images, the referring clinicians are the final recipients of the medical images and the radiologists' reports. Thus, they indirectly adopt AI applications as well, and need to be included in the implementation process. *(Involvement in adoption process)*

> *"And what we do in implementation phase, we mainly have some questionnaires for the users, mainly radiologist, but possibly also laboratory technicians, who make the CT. And also neurologists or other physicians involved." (IV 12)*

Interestingly, in three cases (UMC1, UMC2 and UMC4), it was found that the referring clinicians are not convinced or do not trust the output of the AI application, in this case BoneXpert (IV 3, 10, 11, 21). In fact, in two of these cases, the referring clinicians redo a manual bone age analysis for every scan. *(Acceptance of technology)*

> *"I do really believe in the program. But the endocrinologists of the children hospital, they don't believe in the program. I don't know why. But they don't believe in. They calculate the bone age again, using the atlas of Greulich and Pyle. But I don't care. I spoke to them a few times about it. I said, why don't you believe in it? They have the impression that it didn't work. That's the only thing. I mean they haven't produced any factual mistakes of*

the program." (IV 10)

Thus, just like for the direct adopters, also the indirect adopters (the referring clinicians) showed varying levels of acceptance of AI applications.

4.4.3 Local champion

In all cases studied, the adoption and implementation of AI applications would not happen if it was not for an individual (or more rarely a group) of radiologists, which shows a particularly strong interest in the technological development of AI for radiology and usually has better than average understanding of the technical aspects of the technology: the local champion. Mostly, a local champion was found to start off the adoption process and actively take the lead to implement the application within his or her department. The necessity of having a 'local champion' for the introduction of a new medical technology was also confirmed in the document analysis (Nederlandse Vereniging van Ziekenhuizen, 2018; Nederlandse Vereniging voor Radiologie, 2016). *(Role in adoption process)*

Besides establishing contacts with the technology developers and with other relevant people within the hospital and mobilizing resources, a crucial function of the local champion is to convince his or her skeptical colleagues (IV 1, 2, 3, 8, 14, 21, 22).

> *"I have told them for a couple of years already: ok guys, AI is coming and we need it and we should embrace it. And they asked me, if I was crazy. Then they tried to stop me – saying no AI on this department – and then a couple of weeks later, the same guys go on a congress, and the only thing that they heard there was AI." (IV 1)*

To overcome opposition, two types of strategies to build trust were observed: providing (written) information on AI more broadly or a specific application in particular (in forms of scientific articles, books, presentations) and promoting opportunities for experimentation with an application, e.g. by organizing show-cases or installing a test-version of the application (IV 1, 2, 11). Because test-versions have usually no or low costs involved, they allow to see if the AI application attends the needs and fulfills the expectations of the adopters. *(Trust building)*

In summary, in all cases local champions appeared to perform a crucial role in the implementation process. They not only initiated the introduction of AI applications in their departments or hospitals, but actively pushed the implementation process forward, for example by building trust among their colleagues.

4.5 THE ORGANIZATIONAL CONTEXT OF AI APPLICATIONS FOR RADIOLOGY: THE HOSPITAL AND THE RADIOLOGY DEPARTMENT

Organizational factors undoubtedly play a central role in the implementation of new technologies in the medical field, as proposed in the NASSS framework. For the case of AI applications in radiology, the most relevant organizational units are the radiology department and the hospital.

In the Dutch health system, and particularly within the hospital, the medical specialist (e.g. the radiologist) have a peculiar position: A majority of medical specialists in non-academic hospitals are not formally employed by the hospital, but 'self-employed' service-providers. In most cases, the independent medical specialists of a hospital unite in a 'medical specialist company' (medisch specialistisch bedrijf, MSB). The hospital commissions the MSB to provide the medical specialist care as services, while providing the infrastructure for hospital care (such as rooms and machinery). The MSB therefore functions as a semi-independent company within the hospital. This distinct organizational setting also translates in high autonomy of the departments, for example the radiology department, with regard to personnel decisions and financial investments (Kroneman et al., 2016).

The following section discusses the organizational capacity to innovate regarding AI applications in 4.5.1, as well as the readiness of the hospital and radiology department for AI technology in 4.5.2. Part 4.5.3 explains the nature of the adoption and funding decision concerning AI applications. Next, section 4.5.4 looks into the changes to individual and organizational routines induced by AI applications, followed by the work that is needed to implement the changes induced by AI applications discussed in 4.5.5.

4.5.1 Innovation capacity for AI

At present, with the low volume of AI applications in use in Dutch radiology departments, most of the adoption decisions for implementing AI applications are taken on the level of the radiology departments. From a financial perspective, the radiology department has an assigned budget for technology, which covers mostly short- as well as long-term investments. While the department heads (often a non-medical manager, as well as two or three medical managers) are officially responsible for this type of managerial issues, the bigger decisions (such as the long-term investment strategy) are often taken by the entire group of radiologists in a participatory way (IV 1, 2, 5, 9). This horizontal management and collaborative decision-making appeared to be more strongly articulated in the non-academic hospitals, where radiologists have a different employment relation with the hospital (the MSB). *(Collective decision making)*

> *"But because of those different opinions, it is important to discuss this is in the whole group [with all radiologists]. In a couple of weeks, we have a meeting with everyone, and we will also discuss the plans of our group. And hope to formalize the AI group, because for now it's just a plan. And the whole group has to take a decision, if it wants to start this AI group." (IV 9)*

On the level of the hospital, innovation strategies are developed and implemented by the hospital management (i.e., the board of directors) in collaboration with the innovation specialists (IV 12, 16, 17). Across the cases, the board of directors was found to show different levels of proactivity towards innovation. In three cases (TKZ1, TKZ2, UMC2) the board of directors played a very proactive role in initiating certain innovation projects. *(Innovation leadership)*

The differences in innovation leadership appeared to directly translate to differences in strategic innovation approaches. In four of the cases, a hospital-wide innovation strategy including AI was present (in TKZ1, TKZ2, UMC1 and UMC3). Mostly, their implementation is done through more specific innovation programs (e.g. on a specific technology, e.g. AI), which are related to a specific innovation fund to support the program. *(Hospital-level innovation strategy)*

> *"The added value the board was looking for, was gaining some experience with AI. Not only my department, but we should develop the experience for the whole hospital. […] Because they came from the large congresses, where they heard that AI was going to take over their hospital. And then they said, we need to start somewhere, and decided they would start with radiology. […] The management really pushes for these changes a lot here." (IV 1)*

The innovation strategies and programs are often implemented by designated innovation specialists within the hospital. These can take the form of an innovation manager and/or an interdisciplinary innovation group (IV 1, 12, 16, 17, 18). These individuals or groups containing members of the different departments coordinate and execute innovation programs. They need to have an overview and coordinate the innovation activities within the organization, as well as identify innovation needs and scan for possible (technical) solutions towards these needs. These more centralized innovation manager positions were found in the top-clinical hospitals, but not in the academic hospitals. In TKZ1 and TKZ2, the innovation specialists are actively supporting the radiology department's efforts of implementing AI applications. *(Innovation specialists)*

> *"Yeah, well, the goal of the innovation group is to do quick prototyping of innovations. Basically, if there is an idea, or what I would prefer, if somewhere there is a problem that can be solved by technology, then you look whether that indeed is the case. Whether there is a technological solution that can be easily implemented, prove that it works." (IV 14)*

Regarding the department-level strategy for AI, only TKZ1 had a formalized innovation strategy. However, four more cases (TKZ2, AZ1, UMC2, UMC3) were developing such a strategic approach at the time of this research.

> *"Well, in the [AI-]program we developed, we wanted to do one off-the-shelf product, and that's BoneXpert. We wanted to do a project, which was reachable, like the scaphoid fractures, it's in collaboration. And the lung-nodules is more advanced, it's more a mid-term project." (IV 16)*

It appeared that non-academic radiology departments tend to feel a stronger need to adopt strategic approaches towards the implementation of AI applications in their clinical workflow than academic radiology departments. This trend might be caused by the smaller size of non-academic departments or by the direct relationship of academic radiology departments to universities and research institutes in the field of data science in medical imaging at their respective universities (as is the case for UMC1, UMC2 and UMC3). This means that the innovation activities (and funding) are closely related to the research and development activities of the department, which tend to be on a more fundamental than applied level. *(Department-level AI strategy)*

Summing up, the innovation capacity for the implementation of AI applications in radiology depends on the attitude and leadership of the hospital and to a lesser degree the department management. It is expressed primarily through innovation strategies which cover AI applications and through the presence of innovation specialists within the hospital.

4.5.2 Readiness for AI technology

All cases mentioned the difficulty to deal with increasing workloads, while having to keep the costs capped. Thus, the need for innovations, which can improve efficiency is high. This becomes apparent particularly strongly on hospital-level. While the volume of care increases, hospitals are under large pressure by the government and the insurance companies to keep overall costs constant (Ministerie van Volksgezondheid, Welzijn en Sport (VWS), 2018). At the same time, hospitals have to comply to high quality standards and get scrutinized by the Dutch Health Care inspectorate ('Inspectie Gezondheidszorg en Jeugd') when evidence for lacking quality is found, such as was the case in one of the hospitals, relating to missed diagnoses in radiology.

> *"But on the other hand, we have had some very serious cases of missing large lung tumors every year. And last year, we had three cases and then inspection came to the hospital, who then went to the board. And the board obviously does not like such visits." (IV 1)*

As shown in section 4.3.2, AI applications are thought to have two types of potential demand-side benefits: making better diagnoses and increasing efficiency. The drive to improve their diagnostic quality is an inherent motivation to radiologists, which is expressed more strongly by some individuals than by others (IV 1, 2, 5, 8, 12). *(Tension for change)*

This difference reveals itself in the way the AI technology is accepted by the different individuals in the radiology department (see point 4.4.1), and therefore the level of support towards implementing AI. Two types of dynamics within the radiology departments were observed: collaboration and competition. While in some departments, the different subspecialties collaborate and establish a consensus on how to move forward regarding the implementation of AI applications (IV 2, 9, 22), in other departments there is competition between the subspecialties, such as in TKZ1:

> *"And at the same time, these guys, who said "No AI [for breast radiology] on this department", are now talking with AI-based Neuro-imaging software to implement here. So, something is happening. But it's all politics." (IV 1)*

Interestingly, it appeared that the AI strategy in TKZ1 was not adopted in a collaborative way, including all members of the radiology department, but rather in a top-down decision process by the hospital management, which might have enhanced the opposition to AI within the department. *(Internal dynamics)*

Especially with regard to digitization efforts, hospitals recognize the importance to develop hospital-wide approaches, in order to avoid fragmentation in the digital tools used across the different departments in the hospital (IV 14, 16, 17, 18). The responsibility for the coordination of IT and digitization traditionally lies with the medical physicists, but is increasingly put in the hands of a Chief Medical Information Officer (IV 9, 18). The number of chief medical information officers in Dutch hospitals has increased from 3 to 45 within the last three years (Nederlandse Vereniging van Ziekenhuizen, 2018). *(Technology system-fit)*

All organizational levels of the hospital experience a strong tension for higher efficiency and quality, two of the potential benefits of AI applications. Yet, it appeared that efforts to introduce AI applications were sometimes blocked by members of radiology departments, due to internal conflicts of power.

4.5.3 Nature of adoption and funding decision of AI applications

At present, the adoption decisions for AI applications are mostly taken on department-level and the costs are covered under the existing technology budgets of the radiology department. These technology budgets are generally determined by the hospital board. The adoption decision process can stretch over several months, as it can include lengthy contract negotiations with the supplier and the legal team of the hospital (IV 6, 8, 12). In the case of larger acquisitions, or if a hospital-wide strategy for AI is in place (such as in TKZ1), the adoption decision is taken jointly with the hospital innovation manager, or the board of directors of the hospital (IV 16). The current funding model works in the short term, because the volume of different AI applications is still low and therefore the overall costs for this type of technology are still negligible. As the number of applications used in a hospital increases, these additional costs will have to be absorbed externally or internally, meaning by the hospital as a whole, the radiology department or by the radiologists via the MSB (IV 2).

> *"So, money is now really an issue. And that's why AI is suffering, because if we have to pay even a small percentage of what we receive to AI, we have to cut money somewhere else." (IV 21)*

The uncertainties surrounding the question of funding appeared to be a key concern and were explicitly addressed by 14 interviewees. The main issue is the strong pressure for cost containment within the health system and also within the hospitals. This trickles down to all organizational units. *(Origin funds contested)*

Two scenarios are possible for the medium-term absorption of the extra costs that incur through AI applications. In the first scenario, the hospital can absorb these costs internally, for example by adjusting the internal rates, meaning the price 'paid' to the radiology department for every analysis by other departments of the same hospital. It can be expected that unless the other departments experience clear (financial) benefits for themselves (e.g. if they can reduce the number of biopsies by more detailed diagnoses of lesions), this would create strong opposition (IV 1, 12, 16).

> *"But then the cost savings are not at the radiology departments. And the radiology department is the department that has to pay for it, but it doesn't have then the savings. I think this is the difficult part." (IV 12)*

Alternatively, the radiology department could decrease the rates paid for the radiologists' work, meaning either a de facto pay reduction per person or by hiring less people. It is unlikely that radiologists will accept using a technology that reduces their wages (IV 4, 5, 8). Furthermore, in order to hire less radiologists, AI applications would have to noticeably improve the efficiency of the radiologists by reducing reading time or the number of scans to be read or by fully automating certain reading procedures. As shown in part 3.2.1, this is not yet the case. *(Internal absorption of extra costs)*

The second scenario is to absorb these costs externally, meaning adjusting the rates charged to insurance companies. In theory, hospitals negotiate the price for each treatment with the insurance companies (in the form of 'diagnose-behandelcombinaties', DBC) to which it could technically add the extra costs of the AI applications (IV 1, 4, 7). Yet, hospitals have very little room for negotiation to increase their costs externally because the budget for specialist care is de facto capped (as described more detailed in section 4.6.1). *(external absorption of extra costs)*

In summary, the unclear origin of funds to cover the costs of AI applications for radiology is creating uncertainties for adopting hospitals.

4.5.4 Change to routines induced by AI applications

Overall, the change to work-routines caused by AI applications within the radiology department has so far been limited. As shown in part 4.2.3, there might be small changes in existing workflows, such as additional steps for the radiologists to report the result of the AI application. This depends on how the AI application is used in the workflow. However, even small workflow changes proved to be difficult to implement. This difficulty is even larger if such a change takes extra time from the radiologist, either in the short run while getting familiar with the application, or in the long run, when the application makes diagnostic processes more time-consuming in general.

> *"You already loose half of the radiologists, if you need to open an extra program to determine some extra figures. [It's really hard for my colleagues] to change their routine." (IV 11)*

This reinforces the results from sections 4.2.1 and 4.2.3 that seamless integration of AI applications in the existing workflows is important to avoid non-adoption. Overall, as long as a task is not completely automized by an AI application, it can be stated that routines on the level of the radiology department will not see enormous changes. *(Changes to workflow)*

Also on hospital-level, there have not been large changes to cross-department routines. As was shown for the example of BoneXpert in section 4.1.2, it was found that in many hospitals, communication between the radiology department and the other departments (of the referring clinicians) was not well-established regarding technology related issues. However, interviewees repeatedly mentioned that inter-departmental collaboration for the implementation is important for several reasons (IV 1, 8, 7, 16). From a medical point of view, the referring clinicians need to know how the diagnoses are made. From a managerial point of view, if the other departments are to cover parts of the additional costs, they need to learn about the technology and the potential benefits of using AI applications. *(Cooperation with referring clinicians)*

> *"And now one of the interventional radiologists, took it over, together with one of the neurologists. And they were looking at how they can implement this in our normal daily routine, for every stroke case." (IV 8)*

In line with their professional identity (see 4.4.1), several interviewees mentioned the importance of the radiology department having a leading role in the adoption of AI applications within the hospital. This avoids being bypassed by the other medical disciplines and loosing professional responsibility, while allowing to influence the development and adoption of the technology on the level of the entire hospital (IV 2, 3, 5, 7, 13). *(Leading role)*

Thus, it can be stated that AI applications have so far not induced large changes to organizational routines, neither within the radiology department nor across departments. Yet, cooperating with referring clinicians from other departments, and keeping a leading role in this cooperation, appeared to be of importance to radiologists.

4.5.5 Work needed to implement AI applications

As extensively described in the previous sections, the introduction of radiological AI applications is contested on different organizational levels. In order to overcome this aversion and successfully carry out the change process, several processes are performed, such as building a shared vision and monitoring the change process. An important element in the process of building a shared vision, is creating the 'right' narrative around AI. In order to overcome resistance by radiologists, AI applications are being framed as a 'co-pilot' for radiologists, allowing them to become better doctors, while leaving responsibility in the hands of the radiologist (IV 14). This framing was found to be performed mainly by the local champions and members of the hospital management (IV 1, 3, 16). This element was also repeatedly mentioned in the document analysis (Algra, 2017; Geldermann et al., 2013; Nederlandse Vereniging van Ziekenhuizen, 2018). *(Framing)*

> *"And as a doctor you still can interpret. You are in the responsibility seat. The computer will not take that away, it only helps you. It's like a co-pilot software. That's what we're*

saying to downplay the fuss." (IV 16)

On a more explicit level, creating a shared vision can be done by building innovation strategies in a collaborative way. As mentioned, at this moment such strategies are still relatively uncommon on departmental level, but are in the process of being established. The importance of building these strategies in consensus with all radiologists was repeatedly mentioned (IV 2, 5, 6, 9). Interestingly, in the case of the TKZ1, this strategy was established in a top-down way and the department showed high opposition from radiologists. Creating a shared vision on the level of the hospital is often attempted by creating an organization-wide (innovation) strategy, similarly to what was observed on the level of the radiology department. These strategies are mostly developed in a top-down way. In order to effectively build a shared vision for all members of the organization, they need to successfully trickle down to the lower organizational levels. Attempts for this type of approach were found in two hospitals (TKZ1 and TKZ2) for AI-related technologies. The case of TKZ1 has been described above. In the case of TKZ2, the process is ongoing and includes hiring a new member of staff within the organization, which is charged with developing a vision, building a strategy on how to realize the vision and implement the strategy. *(Vision building)*

The monitoring of the implementation of AI application was found to be rarely done. Only in one of the cases (AZ1), quality control (on a more general level) was mentioned as a concern of the department. For the case of AI, it was found that none of the hospitals had a formal way of monitoring the way that the AI applications were used. Sometimes, there was information on whether the program was being used, but not in what way it was used in the clinical workflow (i.e., as a first or second reader, or completely automated). It is therefore not surprising to observe situations, where different members of a department use the application in a different way, or do not use it at all (as mentioned in section 4.2.3). Interestingly, even though this discrepancy was observed by the respective interviewees, none initiated a formal conversation with his colleagues to 'resolve' this (IV 10, 21). Furthermore, besides technical validation of the algorithms, none of the cases had (successfully) tried to monitor the impact of the AI applications, with regard to potential benefits, such as saving time.

In all cases, the work done to monitor the enacting of existing practices or the impact of the implementation of novel technologies on the level of the hospital is currently limited. Local protocols surrounding implementation specifying how to use the software programs and also how to potentially evaluate it are usually developed with the help of medical physicists or technical clinicians. However, it was found that there was little monitoring if and how the protocols were actually respected within the departments. Moreover, none of the hospitals have a person or department that is formally responsible for issues surrounding the implementation of new interventions or technologies. Nonetheless, several interviewees mentioned that it would be interesting to approach implementation and evaluation of novel techniques in a more strategic way (IV 3, 12, 14). This phenomenon is also recognized on national level and led to the development of a protocol for new interventions in clinical practice, called the 'Leidraad Nieuwe Interventies Klinisch Praktijk' (Orde van Medisch Specialisten, 2014). Although the implementation of AI tools falls under the scope of this protocol, which could provide guidance for the process of implementation in the local setting, the protocol had not been used in any of the cases studied. *(Monitoring)*

The continuous work needed to integrate AI applications in the organization happens in two ways: one the one hand, framing of the technology and vision building is more abstract work, aiming at creating acceptance of AI applications. On the other hand, monitoring of the implementation is more material work, making sure the implementation process proceeds as planned.

4.6 AI APPLICATIONS FOR RADIOLOGY IN THE DUTCH HEALTH CARE SYSTEM

The sixth domain looks into the wider system surrounding the use of AI applications in Dutch radiology departments. Part 4.6.1 discusses the policy aspects as well as the political scenario of the Dutch health care system, followed by the regulatory system for AI applications in 4.6.2. Part 4.6.3 presents the role and position of the professional organizations with regard to AI applications in radiology. Last, part 4.6.4 discusses the socio-cultural context of AI applications in radiology, covering ethical questions and the public opinion.

4.6.1 Dutch health care policy context

The Dutch health system follows a market-based approach, promoting regulated competition among health care providers for health care services (Rosenau & Lako, 2008; Schut & Van de Ven, 2005). As mentioned in section 4.5, this market-based thinking can be found within the organizations, but it is even more strongly manifested between organizations in the health system, e.g. between different hospitals, which are private organizations under Dutch law. This is expressed, for example, in the fact that insurance companies, which are active regionally, will individually negotiate treatment rates (DBC) with each hospital in their region (IV 12, 16). Often insurance companies will send their clients to a particular hospital that offers the lowest rate for a specific DBC (e.g. Bone age assessment), forcing the other hospitals to adjust their rates in order not to lose these patients. De facto, insurance companies often do not negotiate on the level of a DBC, but negotiate a total annual budget for a hospital's total activities. As a consequence, the rates of the DBC and the actual costs of a treatment within the hospital are often disconnected and opaque, making it more complicated to absorb extra costs of technology within these rates, which is strategy many interviewees suggested. *(Market-based thinking)*

One of the key critiques made by interviewees towards the market-based structure, is the difficulty it creates for inter-organizational collaboration between hospitals (IV 15, 16). Nonetheless, hospitals were found to collaborate both informally and formally with each other, with most of these collaborations being of bilateral nature. Repeatedly, the interest of establishing multilateral collaborations was mentioned, particularly in the form of networks, a topic that has recently been taken up by the Radiological Society of the Netherlands (see 4.6.3). Especially collaborations between academic and non-academic hospitals have high potential, as non-academic hospitals have a lot of valuable data, while academic hospitals have resources and knowledge for research and development (IV 1, 8, 9, 16). Interestingly, it was found that all non-academic radiology departments had ongoing collaborations with universities or research institutes, mainly concerning the development of AI based tools. *(Inter-organizational collaboration)*

Furthermore, many hospitals are also collaborating with industry players (UMC1, UMC2, UMC3, TKZ2). As previously mentioned, industry players are interested both in the data of hospitals for the development of their applications, as well as testing and validating their applications in hospitals. Two key issues in this type of collaboration are the division of intellectual property rights and ethical-legal questions related to the use of data and privacy (IV 15). *(Industry-clinic collaboration)*

Besides the hospitals and the insurance companies, the Dutch health care system also counts with a number of other actors, such as patient organizations, the umbrella organization of health insurers (Verbond van Verzekeraars), hospitals (Nederlandse Vereniging van Ziekenhuizen) and medical specialists (Federatie Medisch Specialist) and several (pseudo-)governmental organizations (among others Zorginstituut, ZonMw). Most of these actors have their own focus areas and innovation funds. This makes it particularly hard to direct innovation activities within a particular field, such as AI applications in Radiology (IV 19). *(Fragmented health system)*

Such as all across Europe, the Dutch health care system is confronted with a constant rise in demand for health care, accompanied by strong pressure to stop health care costs from rising further. Accordingly, medical specialist care, including the work of radiologists, which accounts for approximately 30% of total health costs (Volksgezondheidenzorg.info, 2019), has come under particular pressure. In an agreement, signed by all relevant players of the Dutch health system, the growth rate of expenses on medical specialist care was de-facto limited to 0% until 2022 (Ministerie van Volksgezondheid Welzijn en Sport, 2018). This context creates a favorable political scenario for technology that increases the efficiency of health care professionals or even automates certain tasks, with reservations regarding ethical-legal concerns, related to questions of data use and privacy (see part 4.6.4). (*Macro-trends*)

The Dutch health care system is built upon the idea of competition among different actors within the health care system, which complicates inter-organizational collaboration. Due to socio-demographic and macroeconomic trends, these actors are currently submitted to high cost containment pressure.

4.6.2 Regulatory and legal system around AI applications

From a regulatory and legal perspective, AI applications are considered medical devices and have to comply with the European medical device legislation. Under the current regime (European Council Directive 93/42/EEC, in force since 1994), software, such as radiological AI applications, is classified as a type I medical device (low risk). In order to be sold commercially, a product needs to be CE certified, for which a manufacturer simply has to put in place a quality assurance system (Council of the European Communities, 1993). Thus, the CE mark is currently granted without passing by a regulatory body and without any information or proof of the performance of the application (such as is common for medication), or of the added benefit for clinical practice. In comparison to the US-American market (where regulatory approval is granted by the Food and Drug Administration, FDA), the European CE mark is less stringent. However, the existing legislation on European level is going to change with the new Medical Device Regulation (MDR), which will come into effect in May 2020. Under this new MDR, depending on their functionality, software programs will fall into classes IIa, IIb or III. CE certification will then have to be acquired via a notified body, which means a large increase in requirements on quality and safety, as well as post-market surveillance.[10] For a transitional period of three years (until May 2023), existing

[10] Following the Annex VIII, 6.3, Rule 11 of the MDR: “Software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes is classified as class IIa, except if such decisions have an impact that may cause: death or an irreversible deterioration of a person's state of health, in which case it is in class III; or a serious deterioration of a person's state of health or a surgical intervention, in which case it is classified as class IIb.” (European Parliament and the Council of the European Union., 2017)

products can maintain their CE mark without having to fulfill the more stringent requirements, while new products will have to undergo an expensive audit process (IV 15). *(Regulation of medical devices)*

One of the central legal issues is the question of data-security and privacy. Because health care data is perceived as particularly sensitive, this problem is of particular importance for hospitals. In European comparison, the Netherlands has a strict privacy law, which expresses itself in that hospitals want to avoid patient data to 'leave the hospital.' This makes it more difficult to implement AI applications that run on external servers, that is cloud-based services (IV 15, 21). In addition to the national laws, the European General Data Protection Regulation (GDPR), in effect since May 2018, has created new legal rules applicable to the collection and use of personal data in hospitals. While hospitals are interested in using personal data to develop big data solutions (such as radiological AI applications), they are currently struggling with the interpretation of rules on anonymization, established by the GDPR. Because there is no guidance or jurisprudence on how to interpret certain clauses of the GDPR, hospitals do not know how to make use of patient data in compliance with the GDPR, resulting in legal uncertainty (IV 15). *(Privacy)*

Besides the legal uncertainty involving the GDPR, another major uncertainty is related to the question of legal responsibility for damage occurred due to the use of AI applications (e.g. a missed diagnosis by an AI application). This element was one of the key concerns mentioned both by interviewees, as well as in several publications of the NVvR and many newspaper articles. The radiologist is personally responsible for the quality of his diagnoses. Even if a radiologist uses a tool to support his diagnosis, he or she will be fully responsible for the diagnosis. This is why the question of legal responsibility for damage caused by AI applications is very sensitive to many radiologists. Under current Dutch law, the hospital is responsible for any damage that patients suffer, during or due to treatment in the hospital. Hospitals are insured for the case that they, or one of their members such as a radiologist, is responsible for the mistake they have made. For mistakes made by a software program, however, the hospital did nothing wrong. This means that in theory, the existing insurance would not cover this case (IV 15). *(Legal responsibility for damage)*

> *"And of course, we have an insurance, but we are only insured when we are responsible for the mistake we have made. And if we have software or a medical device, which we bought according to the guidelines, what have we done wrong?" (IV 15)*

Medical procedures are strictly regulated in the form of guidelines, which give detailed specifications for medical specialists on how to proceed with diagnosis and treatment, when a patient arrives with a certain condition (IV 19). In the Netherlands, these guidelines are developed by the national specialist organization, for radiology-related procedures, by the Radiological Society of the Netherlands (The Nederlandse Vereniging voor Radiologie, NVvR). These guidelines are created and updated based on the availability of strong empirical evidence. Several interviewees mentioned their expectation that eventually AI applications might become standard clinical practice by being integrated in guidelines (IV 3, 4, 15, 19). Since there is currently only very limited clinical evidence available on the potential clinical benefits of AI applications, it is not surprising that there is currently no process to include AI in guidelines within the NVvR (IV 23). *(Development of guidelines)*

In summary, the regulatory and policy context surrounding AI applications for radiology is characterized by changing and very recent regulation, which leads to high legal uncertainty for radiologists and hospitals.

4.6.3 Professional bodies for Dutch radiologists

Besides their role in developing guidelines, the NVvR represents the interests of radiologists in the Netherlands and is constantly increasing its engagement with AI at several fronts: There is a study group AI within the sub-section 'Technology', which aims at raising awareness among Dutch radiologists. This sub-section serves as an advising body for hospitals to develop an AI strategy and is thinking on how to include AI in the curriculum for future residents (IV 1, 2, 3, 23). This sub-section 'Technology' also recently started to organize meetings open for all the members of the NVvR for knowledge and experience sharing on the topic of AI. The first meeting was held on the 3rd of June 2019 and was attended by approximately 75 radiologists from across the country. Additionally, AI is one of 10 questions on the NVvR's own knowledge agenda 2018-2020 (Nederlandse Vereniging voor Radiologie, 2016), including falling under the scope of the Radiology Research Fund, established in 2017. Several interviewees have mentioned disappointment with the lack of initiative taken by the NVvR, but the AI meeting can be seen as an attempt of the NVvR to develop a more active role (IV 3, 4, 21). *(National professional bodies)*

In addition to the national scientific associations (such as the NVvR), the international societies such as the European Society of Radiology also play an important role in raising awareness on AI for the radiology profession. Annual congresses such as the European Congress of Radiology (organized by the European Society of Radiology) or the annual congress of the Radiological Society of North America act as important sources of information for many radiologists. For the last years, AI has been the most important topic on these congresses, defining the scientific program and the exhibitions of technology providers (IV 1, 3, 4). Additionally, to their congresses, these associations are important opinions makers, for example by publishing white papers (e.g. Neri et al., 2019). *(International professional bodies)*

Overall, it can be said that while the Radiological Society of the Netherlands is no frontrunner on AI for radiology, it appeared to be increasing its engagement with the topic.

4.6.4 Socio-cultural context for AI applications

Through the interviews and mainly through the document analysis, several aspects of the socio-cultural environment were found to be relevant for AI in radiology.

First, interviewees and documents highlighted several ethical concerns. There is a hesitance to the idea of completely automized decisions in the medical field, because the consequence of mistakes could be fatal and it would be difficult to hold a human accountable, accounting to a loss of control (IV 13, 17, 22). Additionally, the black-box nature of AI algorithms makes it as such that doctors would have to proceed with certain decisions without knowing how the algorithm would come to the conclusion and therefore without being able to explain to the patient how and why a decision was taken (Meurs, 2019). This lacking transparency was also mentioned with regard to potential discrimination within algorithms, for example an algorithm that was trained on Caucasian children is used on children with another ethnicity and provides erroneous results (IV 3). Another ethical concern is the question of data ownership and privacy. This issue appeared to be particularly important for the case of medical data, which is perceived as very sensitive and therefore should stay in ownership of the patients (IV 16, Meurs, 2019; Ottes, 2016). Furthermore, the entrance of big technology companies in the medical sector is perceived as dangerous, because of their bad track-record with privacy and data safety and their tendency to concentrate economic power, through ownership of enormous amounts of private data (EY, 2019; Meurs, 2019; Ottes, 2016). (*Ethical concerns)*

Second, it became obvious that the health sector, and automated medical image analysis in particular, are perceived as one of the frontrunner sectors for the large-scale use of AI-based technology in the Dutch economy (AINED, 2018; Nederland Digitaal, 2019; WRR, 2016). Repeatedly, AI in radiology was used as an example for a high-potential solution to curb rising health care costs, through the automation of certain tasks. It was also mentioned that the government should be proactive in providing resourced to the development of a strong AI sector in the Netherlands as well as proactively providing guidance for ethical-legal issues, such as privacy (Meurs, 2019; Ottes, 2016; WRR, 2016). *(Public opinion)*

While the Dutch public recognizes the potential of AI for the challenges in the health care sector, ethical concerns surrounding privacy and lack of transparency are very prominent.

4.7 FUTURE OUTLOOK ON AI APPLICATIONS FOR RADIOLOGY

The last domain touches on the development and implementation of AI applications over time. Part 4.7.1 targets the past and future developments of the technology, while part 4.7.2 looks at the organizational resilience necessary for sustained implementation of AI applications in Dutch clinical radiology.

4.7.1 Development of AI for radiology over time

Although earlier versions of AI applications, such as early CAD systems, have been around for over two decades, the technology is still perceived to be at an early stage of development. As two interviewees put it, the current 'hype' around AI, which results from the development of deep-learning techniques, is creating large expectations on what AI will be able to do for radiology in the future and what its impact will be on the profession (IV 5, 6). Interestingly, half of the radiologists in the sample mentioned that the slow development in the past has led to lower confidence in the performance of AI applications and their potential usefulness for clinical practice and decreased radiologists' expectations for present and current developments (IV 3, 6, 11, 21, 22). This reason also led the manufacturer of BoneXpert, Visiana, not to market BoneXpert as an 'AI application' in the Netherlands.

> *"Most radiologists have been already working here for 20 or 25 years. Yeah, well, what can they expect for the next 10 years? They have heard the voices of computer assisted radiology already for 10 or 15 years. And nothing really changed in the clinical workflow." (IV 11)*

As previously mentioned, in several of the studied cases, some action towards creating groups of AI specialists has happened, showing flexibility and openness to adapt professional scope and responsibilities within the departments (and the hospital at large) with regard to this technological development. Overall, all interviewees are open to try out further AI applications and see the arrival of AI in radiology as inevitable, but do not expect any major changes to their professional practice in the next five years. *(Expectations on future development)*

Thus, although positive, expectations of radiologists on the development of AI applications have been dampened by the slow development in the past.

4.7.2 Organizational resilience concerning AI in radiology

Analyzing the organizational resilience, meaning the organization's ability to reflect and respond to 'critical events', requires a certain time to pass in the implementation process of AI in clinical practice. For BoneXpert, an application that has been in use for a longer period of time, several examples show that this process is not properly working. In three (UMC1, UMC3, UMC4) of the seven cases, it was found that BoneXpert did not satisfy the needs of its (in-)direct users, namely the referring clinicians, which started redoing manual bone age assessments (UMC1, UMC4) or demanded the manual assessments from the radiologists (UMC3). This is a strong sign that the technology is not fulfilling its function and should function as a warning sign to the adopters. Although in all cases the radiologists were aware of the situation, they did not undertake any action to 'resolve' this issue, therefore showing a lacking ability to detect and respond to critical events. Nonetheless, it needs to be stated again that for most AI applications the implementation process is at a very early stage, indicating that organizations have not had sufficient time to engage in collective reflection or continuous adaptation. *(Recognizing critical events)*

It follows that organizations do not have sufficient mechanisms for the recognition of critical events concerning the implementation of AI applications for radiology.

In summary, the seven domains and respective sub-domains of the NASSS framework by Greenhalgh et al., (2017) have been described to understand the facilitating and hindering factors to implementation and adoption of AI applications in Dutch radiology departments. These domains include the clinical condition targeted by AI applications (subchapter 4.1), the technological aspects surrounding AI applications in radiology (subchapter 4.2), the value proposition of AI applications (subchapter 4.3), the adopter system of AI applications in radiology (subchapter 4.4), the organizational context surrounding the implementation of AI applications in radiology (subchapter 4.5), the wider system around AI applications in radiology (subchapter 4.6) and finally the implementation of AI applications in radiology over time (subchapter 4.7). In each domain, concepts were identified that detail the particular facilitating and hindering factors for the implementation of AI applications in clinical radiology. An overview of the results can be found in the final analytical framework in table 4.

TABLE 4: FINAL ANALYTICAL FRAMEWORK

Domain	*Sub-Domain*	*Concept*
1. Targeted Condition	1A. Technical requirements	Desired clinical benefit for condition
		Quality standards required for condition
	1B. Adopters involved	Direct & indirect users
2. Technological Context	2A. Technical features	Technical Performance
		Local Validation
		Empirical evidence
		Integration in Existing IT Infrastructure
		IT support
		Readiness of IT Infrastructure
	2B. Comprehension of output	Understanding of technology
		Technical training
		Dealing with human-machine contradictions
	2C. Use in clinical practice	Guidance on implementation
		Implementation in workflow
	2D. Supply models	Business models
3. Value Proposition	3A. Business case	Different providers
		Development
		Revenue streams
		Cooperation
	3B. Value for clinical practice	Clinical benefits
		Indirect demand-side benefits
		Measuring demand-side value
4. Adopters	4.A. Direct adopters	Variance in acceptance
		Opposition
		Trust
		Change in practices
		Creation of new roles
		Professional identity
		Loss of responsibility
	4.B Indirect adopters	Involvement in adoption process
		Acceptance of technology
	4.C Local champion	Role in adoption process
		Trust building
5. Organizational Context	5A. Innovation capacity	Collective decision making
		Innovation leadership
		Hospital-level innovation strategy
		Innovation specialists
		Department-level innovation strategy
	5B. Readiness for technology	Tension for change
		Internal dynamics
		Technology system-fit
	5C. Nature of adoption/ funding decision	Origin funds contested
		Internal absorption of costs
		External absorption of costs
	5D. Change to routines	Changes to workflow
		Cooperation with referring clinicians
		Leading role
	5E. Work needed to implement change	Framing
		Collective vision building
		Monitoring
6. The Health Care System	6A. Health care policy context	Market-based thinking
		Inter-organizational collaboration
		Industry-clinic collaboration
		Fragmented health system
		Macro-trends
	6B. Regulatory & legal system	Regulation of medical devices
		Privacy
		Legal responsibility for damage
		Development of guidelines
	6C. Professional bodies	National body: NVvR
		International professional bodies
	6D. Socio-cultural context	Ethical concerns
		Public opinion
7. Future Outlook	7A. Development over time	Expectations on future development
	7B. Organizational Resilience	Recognizing critical events

5. ANALYSIS

This chapter puts the individual concepts and sub-domains found in the *Results* chapter in context. It is shown how the identified concepts in the sub-domains and domains interact and thereby impact the (non)adoption and implementation of AI application in clinical radiology.

Targeted Condition

Because different AI applications were studied, it was not possible to single out the role of one specific medical condition on the implementation process. However, it appeared that depending on the condition targeted by a specific AI application, the applications needs to fulfil different technical requirements and different actors are involved in the diagnostic process. Next to the direct users, i.e. the radiologists, the referring clinicians are indirect users of the AI applications outputs, as they will use the findings from the medical images to decide on treatment. The referring clinicians are therefore not only indirect demand-side beneficiaries, but their acceptance of the AI technology is important for successful implementation.

As stated in proposition 1, *the heterogeneity of the diagnostic process across different subspecialties in radiology needs to be taken into account for successful implementation of AI applications.* It was found that the differences across radiological sub-specialties do not materialize in the structure or set-up of the diagnostic processes. Rather, the group of involved direct adopters (the sub-specialty of radiologists) and indirect adopters (the involved referring clinicians) changes. Also, it appeared that depending on the condition, the technical requirements towards AI applications change.

AI Applications for Clinical Radiology

Requirements to the technical performance of AI applications for medical image interpretation were found to be high (in terms of specificity and sensitivity). Interestingly, even a few erroneous results (which on aggregate might be negligible) can lead certain radiologists to abandon the use of an AI application. Next to the performance measures, which become apparent during clinical use, most users (i.e. radiologists and referring clinicians) do not have the necessary technical knowledge to assess the quality and trustworthiness of an AI application, especially prior to gaining use experience. Therefore, they heavily rely on other sources of information, such as scientific literature on the quality of the algorithm, or information on the 'ingredients' of the algorithm, i.e. the data that was used to train it. Such information is scarcely available for most AI applications. It can be assumed that the more use experience radiologists gain, the more they understand about the technical aspects, and the more they have access to external sources of knowledge, the better they will be able to assess the qualities and risks associated with the technology. This in turn can be expected to increase trust in AI applications and their adoption. Furthermore, the smooth integration of AI applications in existing radiological IT infrastructure (i.e. in the picture archiving and communication systems – PACS) appeared to be crucial for adoption by radiologists, as it is the key determinant for perceived user-friendliness.

As the NASSS framework proposes and as was assumed in proposition 2, *user friendliness and high performance, are crucial elements for successful implementation of AI applications in radiology.* It was found that the perception on performance and user-friendliness, are indeed decisive factors for implementation. Furthermore, perceived performance and user-friendliness, as well as the technical knowledge and understanding of AI technology, varies strongly across individual users. In combination with the lack of empirical scientific evidence on the performance and clinical benefits of AI applications, this can cause low acceptance of AI applications by radiologists.

Value Proposition for Developers & Adopters of AI Applications

Because business-models of AI applications for radiology are not consolidated yet, and because there is a low number of (paying) clients for AI applications in clinical radiology, it can be argued that the business case surrounding AI applications in radiology is still underdeveloped. The technological nature of AI applications requires large investments in the development phase. Faced with uncertainty on future revenue streams, many of the smaller developers are led to collaborate with larger medical IT providers and/or the pharma industry. On the one hand, this trend might be beneficial for the adoption rate of AI in clinical radiology, since the functionalities of the algorithms will become automatically integrated in the larger IT systems (such as the picture archiving and communication system, PACS). On the other hand, collaborations between small and large developers might enforce the phenomenon of concentration of resources and market power in the health sector, with potentially negative societal impacts on the cost of the health system.

On the demand-side, a large range of *potential* added-value from using AI applications in clinical practice for the radiologists have been found. These range from better quality of the diagnoses to potential efficiency gains and thereby cost savings. Considering the high (and increasing) workloads of Dutch radiologists and high pressure to curb costs for medical specialist care in the Netherlands, efficiency gains appear particularly promising. However, there is currently very little empirical evidence, if these potential benefits actually materialize in the clinical setting. While the developers focus on providing evidence on the technical performance of the algorithms, the users experience a lack of time and knowledge on how to precisely measure the benefits. The uncertainty about demand-side value leads radiologists to hesitate from engaging in adoption processes of AI applications. Additionally, it complicates the mobilization of financial resources to fund AI applications for radiology within the department or hospital.

Proposition 3, which states that *for successful implementation, AI applications need to provide clear added value for relevant demand- and supply-side stakeholders*, appeared to partially hold. As demonstrated, the added-value for both the demand- and supply-side is uncertain at this point. While on the supply-side, the unclear value proposition does not hinder continuous investments in the development of the technology, uncertainty about demand-side benefits appears to be obstructing the adoption of AI applications in Dutch clinical radiology.

Adopters of AI Applications in Clinical Radiology

The ‘adopter system’ was found to differ between applications. Both the sub-group of radiologists as well as the group of referring clinicians changes depending on the condition targeted by that specific application (e.g. neuro-radiologists and neurologists for stroke detection applications or musculoskeletal/pediatric radiologists and endocrinologists for bone age assessment). Since these two types of adopters work in different departments within the hospital, good communication and alignment between departments on the topic of AI applications for the interpretation of medical images is needed during the adoption decision and implementation process.

From the perspective of the radiology department, it can be assumed that changing adopter systems for each AI application complicates the adoption process. For each application, a distinct set of individuals is involved with different levels of acceptance towards AI technology. As shown, acceptance varies strongly across radiologists, also across members of the same department. Low acceptance can lead to active opposition towards the introduction of AI applications within a radiology department. Low acceptance was shown to originate from several sources: lacking knowledge and trust of AI technology, uncertain added-value to clinical practice, perceived threat to professional responsibilities and potential change in professional identity due to AI. The findings provide evidence in favor of proposition 4, which states that *acceptance of relevant medical professionals is a crucial element for successful implementation of AI applications in radiology*. Besides radiologists as direct adopters, the referring clinicians also play an important role as indirect adopters. This element was not explicitly taken into account by the original NASSS model.

Often one (or more rarely a group of) proactive radiologists, a 'local champion,' actively tried to increase acceptance of skeptical colleagues, by organizing demo sessions of specific AI applications, or courses on the technical aspects of AI and by personal conversations. The local champion also acts as a broker between the radiologists in the department, the referring clinicians, other members of the hospital, (i.e. hospital managers, innovation managers) and external players (i.e. the technology provider), which are part of the adoption process. It can therefore be concluded that, as proposed in proposition 4*, the local champion is a critical element in the adoption system,* because he or she does not only facilitate the adoption process, but is the main driver from start to end.

Additionally, it was found that technical staff and members of the hospital management do not play a direct role in the adoption of radiological AI applications. Rather, these actors assist the radiologists in their adoption and implementation process, by providing resources and technical support. The original NASSS model suggests care-givers and patients as important elements of the adopter system. For the case of AI applications for radiology these actors are not part of the adopter system, since, to present, the interpretation of medical images is the domain of radiologists (and other medical specialists). This means that it does not fall under the scope of care-givers and is done without direct contact to patients, who exert only limited influence on the implementation process from the socio-cultural side.

Hospital & Radiology Department

Because of political and economic forces in the Dutch health care system, Dutch hospitals (and therefore radiology departments) are submitted to high cost containment pressure. This creates favoring conditions for innovation projects that promise efficiency gains. Also, it has led most hospitals to adopt innovation strategies, sometimes explicitly including digitization efforts involving AI applications in radiology. This favors proposition 5a, which states that *an organizational innovation culture, expressed for example through the adoption of innovation strategies, is important for successful implementation of AI applications in radiology*. It was observed that while the non-academic hospitals in the research tended to show a more coordinated strategic approach towards AI compared to the academic hospitals, the latter tend to have more experience with testing and using AI applications. A potential explanation can be that, instead of adopting formalized AI strategies, academic radiology departments decide to adopt AI applications based on ongoing research activities happening within their own departments. These research activities can happen either in collaboration with external developers or are related to the in-house development of new AI applications. In general, it appeared that hospital-wide AI strategies, initiated by the hospital management, create favorable conditions for the implementation of AI applications. Furthermore, collective strategy building, i.e. involving relevant stakeholders such as radiologists and referring clinicians in the strategy building process, was found as a potential way for building a shared vision around AI and appeared to be an important tool to avoid the formation or overcome opposition from radiologists.

The necessity to include radiologists in discussions surround AI appeared even more strongly for the adoption processes for specific AI applications on the level of the radiology department. This appears to support proposition 5b, which mentions *the necessity to include relevant stakeholders in adoption decisions and implementation strategies.* It was already shown that besides the radiologists the referring clinicians are important stakeholders who need to be included in the adoption and implementation process of AI applications. With regard to the implementation of new technologies on organizational level, it was found that the different organizational units – the hospital, the radiology department, the referring clinicians' department and the medisch-specialistisch bedrijf (MSB) – each have their own priorities, strategy and financial resources. These are not automatically and necessarily aligned, and in some cases even opposed to each other. This complicates the adoption process and, in combination with the uncertain clinical added-value, makes the mobilizing of funding more difficult.

Radiology departments appear to lack knowledge on how to organize the implementation phase for AI applications and are failing to develop formal implementation strategies. To present, this does not seem to cause problems for the redesigning of individual and department-level routines, probably due to the fact that these routines have not undergone much change. Thus, it cannot be concluded that, *in order to successfully implement AI applications for radiology, all relevant stakeholders need to be included in (re)designing of routines, on the individual and the organizational level,* as stated in proposition 5c. However, the failure to develop formalized implementation strategies leads to insufficient monitoring of the enactment of the new practices, such as the way an application was used (or not used) in clinical practice within the department. Monitoring can help to 'standardize' the workflow and to identify cases of non-use, as well as to facilitate the measurement of the added-value an application can have for clinical practice. The lack of monitoring can be explained by a short-coming of knowledge on how to monitor and measure such impact, both on departmental as well as hospital level (e.g. in the form of 'implementation managers').

The Dutch Health Care System

Increasing demand for medical imaging and the pressure to contain cost for medical services are exerting pressure on the Dutch health care system as a whole, and hospitals in particular. While this creates demand and favorable political conditions for innovative solutions that promise to increase efficiency, such as AI applications for medical imaging, it also complicates the mobilization of resources to invest in new technologies.

Simultaneously, uncertainties surrounding regulatory and legal questions for AI applications in radiology leads to hesitance of adopting organizations (hospitals) and adopters (radiologists) to take up these technologies. From a regulatory perspective, the new European Medical Device Regulation (MDR) will change the regulatory requirements for AI applications in radiology from 2020 onwards. On the one hand, the new MDR can be expected to increase quality and safety of the products. On the other hand, it creates a barrier to market entry for new products and can, therefore, be expected to raise the prices of existing products, leading to a negative effect on health costs. From a legal perspective, there are two big controversies. The first one regards the unresolved question of who is legally responsible for damage caused to patients resulting from the use of radiological AI applications (e.g. missed diagnoses by AI applications). The second one concerns the uncertainty regarding privacy and use of patient data, related to the lack of jurisprudence on the newly implemented GDPR. Especially the former is a large concern to radiologists and among the reasons for opposition towards the technology. This can be seen as evidence in favor of proposition 6, which assumes that *legal and regulatory certainty on how the technology will be integrated in the existing health system needs to be created and expected institutional opposition needs to be overcome in order to achieve sustained implementation of AI applications in radiology.*

The institutional opposition towards radiological AI applications, as mentioned in proposition 6, did not appear to be a strong factor to this point. Although individual radiologists were found to oppose AI technology in radiology, the professional organization of Dutch radiologists, the Radiological Society of the Netherlands (NVvR), has not expressed any opposition. In fact, the NVvR has increased its activities on the topic of AI in radiology in the last two years and is now actively trying to facilitate knowledge sharing and raise awareness on AI. This can be seen as a sign that acceptance within radiology profession is increasing and will likely accelerate the process of acceptance of individuals and radiology departments across the country.

Future Outlook for AI in Clinical Radiology

Proposition 7 states *the necessity to include moments of reflection and allow for continuous adaption of the implementation process.* It highlights the importance to understand implementation as a continuous and iterative process. This means that radiologists and hospitals need to recognize that the implementation process does not end when an AI application is installed on the local workstation, but that important work follows to assure sustainable implementation of AI applications. Although the implementation processes of most AI applications in Dutch radiology departments too early-stage to assess the long-run implementation aspects, the implementation process of BoneXpert illustrated that monitoring and reflection are currently not given due attention by the people responsible for designing the implementation processes. In fact, it appeared that due to a lack of monitoring of the implementation process and a lack of reflection on the wider adoption system, critical points were not identified, leading to cases of non-adoption and abandon of BoneXpert.

Furthermore, the topic of AI is given increasing importance in the Dutch radiological community, as can be seen from the increased engagement of the NVvR. Although the current impact on the radiologists and the radiology departments is still very small, interviewees assume that the number of applications in use will increase continuously over the next couple of years. This shows an overall openness towards the technology.

An overview of the analysis of the hindering and facilitating factors and the intricate dynamics of the implementation process of AI applications for radiology can be found in figure 3. It shows the domains of the NASSS model adapted for AI applications in radiology, as well as the most important interactions between domains (as arrows), which were found to strongly influence implementation. The interactions are classified as facilitating (green arrows), hindering (red arrows) and ambiguous, meaning the interaction simultaneously has hindering and facilitating features (orange arrows). Because this analysis covers a variety of different AI applications, the first domain 'Targeted Condition' is not clearly specified for one unique medical condition. This means the first domain was not found to directly interact with the other domains, but rather delineates the adopter system and determines the technological requirements towards the AI application (blue arrows).

The green arrows display the facilitating interaction between the Dutch health care system and the technology on the one hand, and the hospital on the other. Macro dynamics in the health care system are increasing the demand for medical images, creating need for quality and efficiency gains in the interpretation of these images, for example through (semi-)automated image analysis done by AI applications for clinical radiology. At the same time, hospitals are experiencing high pressure for cost containment. This creates favorable conditions for the implementation of technology that promises efficiency gains, such as AI applications in radiology.

The orange arrows display interactions that have both hindering and facilitating elements. While AI applications are expected to have a high potential added-value for radiology (by improving the efficiency and/or quality of medical image interpretation), there is currently very little empirical evidence on the de-facto added-value of AI applications in clinical practice. The lacking evidence, related to a lack of knowledge of the technical aspects of AI applications, leads to a lack of trust of the technology by its adopters, namely the radiologists (direct adopters) and the referring clinicians (indirect adopters). However, to present, the AI applications have induced very little change to clinical routines, facilitating a step-by-step familiarization with AI applications by the adopters. A crucial element in this process is assuring a smooth integration of AI applications in the existing hospital-wide IT infrastructure. This process is facilitated by favoring innovation conditions within the hospital, established by innovation strategies that include AI. Finally, the unclear legal and regulatory environment surrounding AI applications for radiology is creating uncertainty for radiologists, leading to hesitance for adopting AI applications. However, the Radiological Society of the Netherlands is increasingly active in providing support and guidance for radiologists to implement AI applications in clinical practice.

Lastly, the red arrows highlight the hindering interactions for the implementation of AI applications. The uncertain added-value for clinical practice is causing varying acceptance of AI applications among radiologists and referring clinicians. Acceptance and trust can be increased by providing access to theoretical and practical knowledge about AI technology. Furthermore, engaging radiologists and referring clinicians in collective decision-making processes (concerning strategy and adoption decisions) allows for the detection of potential reasons for lacking acceptance among adopters. Lacking evidence for the added-value of AI applications also complicates the mobilization of funds to acquire AI applications within the hospital. In order to gain evidence on clinical benefits of AI applications, validation and evaluation should be conducted within the adopting radiology department. In addition, developers of AI applications should develop a methodology to assess clinical benefits and invest in clinical evaluation studies. Funding and adoption decision processes are further complicated by the failure to involve radiologists and referring clinicians in these processes. This is, at least partly, caused by insufficient communication between different organizational departments (i.e. radiology department and department of referring clinicians). Additionally, insufficient attention is given to establish forms of monitoring implementation processes of AI applications. This, in turn, causes critical events to be overlooked such as non-adoption or abandon of AI applications by radiologists or referring clinicians.

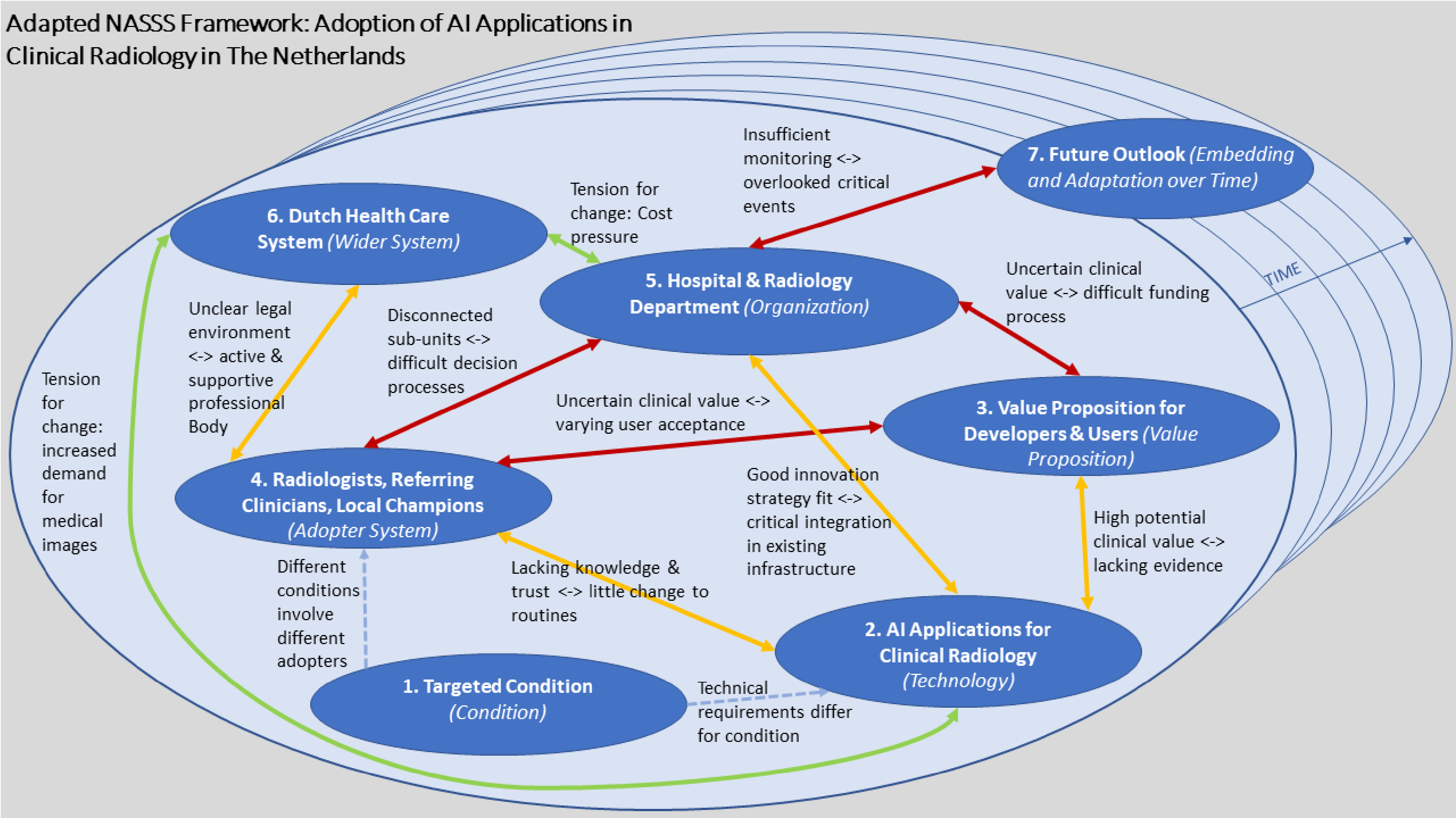

FIGURE 3: INTERACTIONS IN ADAPTED NASSS FRAMEWORK

6. CONCLUSION

“And radiology will change in the years to come. Deep learning is so easy and so fast, and whether people like it or not, it's here. And I like it.” (Interviewee 22)

Clinical radiology in the Netherlands is increasingly being flooded with digital data, mainly in the form of medical images (Obermeyer & Emanuel, 2016). Computerized automated image analysis, in the form of algorithms, is offering an interesting form of dealing with this data. Earlier versions of the technology, commonly referred to as computerized aided diagnosis (CAD) systems, have existed for decades, but failed to reach widespread adoption (van Ginneken et al., 2011). Thanks to recent technological advances in the field of artificial intelligence (AI) such as deep learning, a form of machine learning, these AI algorithms are becoming increasingly powerful and accurate,and perform better than trained radiologists for specific tasks (European Society of Radiology, 2019; He et al., 2019). However, very few AI-based applications are actually in routine clinical use in radiology departments of Dutch hospitals. While technical performance of these algorithms is expected to continuously increase, issues surrounding the practical implementation of this new technology in clinical radiology need to be addressed. Implementing novel technology in the medical field, and especially in hospital settings, has shown to be a complicated process, as it involves a large variety of stakeholders and organizational sub-units, with rigid routines and strong professional identities, as well as strict legal and regulatory standards (Greenhalgh et al., 2004, 2017; Pope et al., 2013). In order to unfold the potential of AI-applications for clinical radiology practice in the Netherlands, the underlying dynamics for successful implementation need to be understood. This has led to the following research question*: What are facilitating and hindering factors for the successful implementation of AI-based applications in radiology departments in Dutch hospitals and how can they be overcome?*

The non-adoption, abandonment, scale-up, spread and sustainability (NASSS) framework by Greenhalgh et al. (2017) aims at detecting the determinants and interacting dynamics of adoption processes of complex technologies in health care. An adapted version of the NASSS framework was used to identify and explain challenges to the sustained implementation of AI applications, which perform image interpretation and/or function as decision support systems, in clinical radiology in the Netherlands. Due to the qualitative and exploratory nature of the research, an embedded multiple case study approach, involving seven cases, was adopted. A total of 24 semi-structured interviews were carried out, complemented by internal and publicly available documents. The cases, seven Dutch hospitals, were selected based on their use of the specific AI application BoneXpert, which appeared to be the only AI-based application for radiology in routine clinical use across several hospitals in the Netherlands.

In order to map barriers and facilitators to implementation and, more importantly, in order to gain understanding of the underlying innovation dynamics between these factors, a two-step research approach was followed. In the first deductive phase, guiding propositions on the facilitating and hindering factors were established, based on the existing NASSS framework (Greenhalgh et al., 2017). In the second, inductive, phase, the concepts identified in the interviews were compared to the original NASSS framework, which was then refined. Additionally, important relations between the concepts were uncovered, resulting in the adapted framework (Figure 4) for the successful implementation of AI-based applications in clinical radiology in the Netherlands. The identified facilitating and hindering factors will be presented in detail.

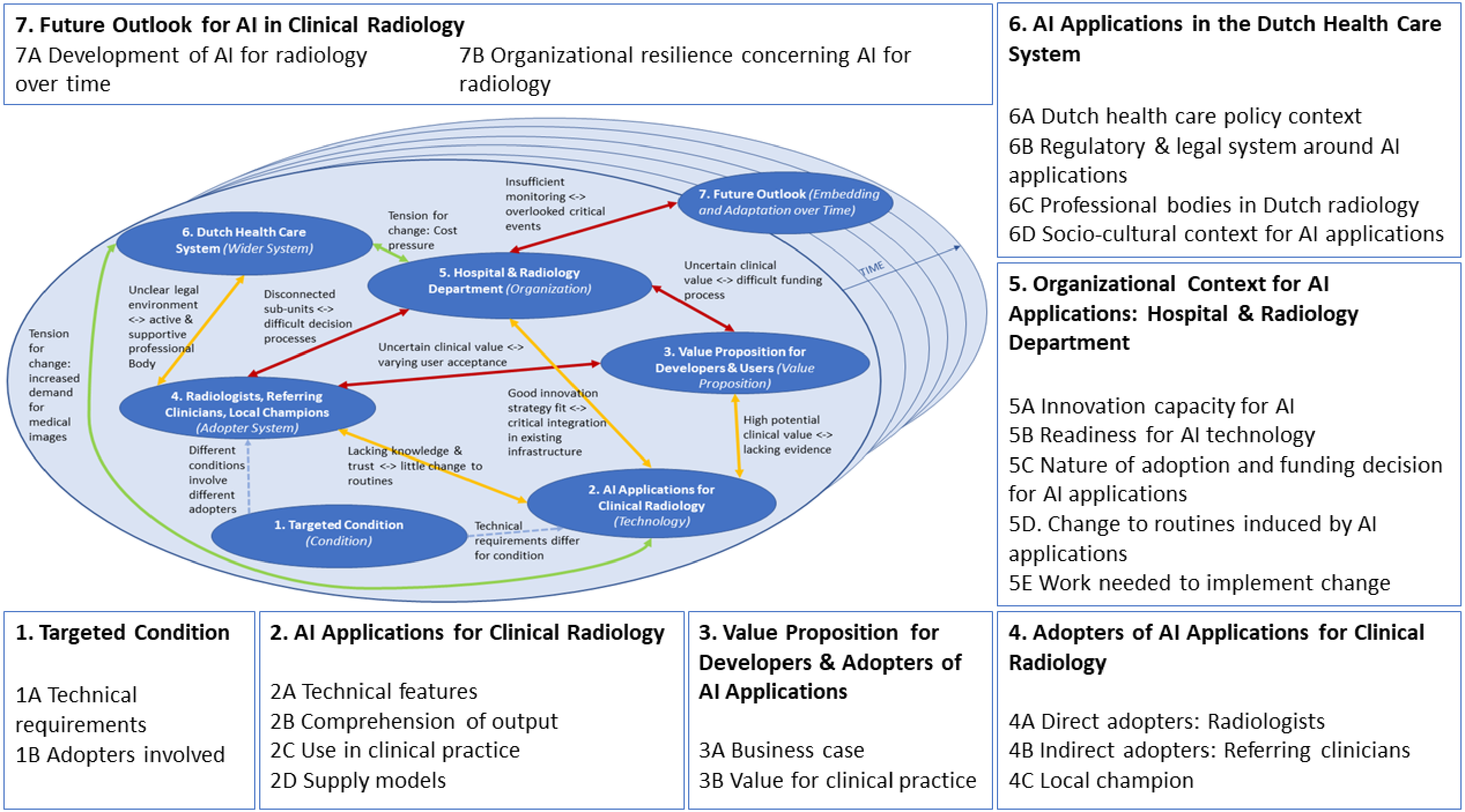

FIGURE 4: NASSS FRAMEWORK FOR AI-BASED APPLICATIONS IN CLINICAL RADIOLOGY IN THE NETHERLANDS

The following facilitating factors for the implementation of AI applications in Dutch clinical radiology were identified: First, macroeconomic and demographic trends in the Netherlands are causing workloads to increase and creating pressure to contain costs, provoking a tension for change. This leads hospitals and radiology departments to develop an openness for the development and adoption of innovative technology, which promises efficiency gains. Second, there are high expectations towards the added-value of using AI applications in radiology. Third, the openness towards AI for radiology is expressed by the adoption of hospital-wide or radiology department specific innovation strategies, which include the use of AI technology. Fourth, adoption and implementation processes are often initiated and pushed forward by a local champion. These local champions are radiologists that show a particularly strong interest in the technological development and usually have better than average understanding of the technical aspects of AI applications. The local champions do vital trust-building work within the entire adopter system, by familiarizing and convincing the other radiologists (direct adopters) and the referring clinicians (indirect adopters) with the AI application. Fifth, the uptake of AI applications by the radiologists is facilitated by implementing the applications without large changes to routines, and by integrating the applications as smoothly as possible in existing IT systems used by radiologists, such as the picture archiving and communication system (PACS). Sixth and last, the Radiological Society of the Netherlands (NVvR) is becoming increasingly active in providing a platform for knowledge-exchange on the topic of AI among its members facilitating the implementation of AI applications.

Likewise, a number of hindering factors was found: First, the technical performance of most AI applications was perceived as inconsistent by the users. In combination with a lack of understanding of the technological aspects of AI, some radiologists are putting the quality and safety of an application in doubt and fail to adopt or abandon AI applications. Second, the added-value of AI applications for clinical radiology practice is uncertain at this point, due to a lack of empirical evidence from scientific studies. Third, the acceptance and trust of AI applications by its direct and indirect adopters (radiologists and referring clinicians) is insufficient. Besides the previously mentioned factors, such as technical factors and the uncertain clinical added-value, perceived changes to professional identity and professional responsibilities were found to influence individual adopters' acceptance. Besides the radiologists, it is also important to achieve acceptance of the indirect adopters, namely the referring clinicians, since they are the final 'customers' of the AI applications' output. Fourth, the adoption and implementation processes of AI applications tend to happen in an unstructured and unguided fashion. This means they do not clearly establish which clinical benefits are hoped to be achieved through the use of the application. Further, they do not involve all relevant stakeholders (the radiologists and referring clinicians). Also, they do not specify how the application is supposed to be integrated in the clinical workflow by its users and fail to include mechanisms to monitor the implementation process. Fifth, the source of funding for AI applications is contested. On the one hand, this is due to the lack of clear evidence on the added-value of AI applications. On the other hand, lacking communication across different hospital sub-units complicates funding decisions. Sixth and last, changes in the regulation and legislation on the use of personal data and medical devices lead to several regulatory and legal uncertainties for AI applications in radiology. Open questions surrounding the use of personal data, the regulatory approval of AI applications and the legal responsibility for AI applications create uneasiness among radiologists and hospitals.

This research examined the opportunities and challenges to the widespread implementation of artificial intelligence-based applications for clinical radiology in the Netherlands. At present, the number of AI applications in routine clinical use in Dutch radiology departments is still very small and the implementation processes of the applications are at a relatively early stage. Yet, AI is gaining rising attention within the radiology profession in particular, and the health care field more broadly. Thus, it can be assumed that an increasing number of Dutch radiology departments will start implementing AI applications in the upcoming years. In order to assure that the use of AI applications for radiology is beneficial to radiologists, to hospitals and to society, the implementation process needs to be as efficient as possible. Reaching this efficiency can hopefully be facilitated through the findings and insights presented above.

7. DISCUSSION

This discussion chapter reviews the theoretical and practical contributions of this research, followed by a critical analysis of the limitations of the research. Finally, some suggestions for further research are given.

Theoretical Contributions

This exploratory research has shown that the success of adoption and implementation of AI applications in Dutch radiology departments depends on the complex interplay of technical, user-related, organizational and systems factors. Previous research on the failed widespread adoption of predecessor technologies of AI application for radiology, such as early CAD systems, considered technical aspects as the most important (if not unique) barrier to adoption. This rather narrow perspective also echoes in current publications on the clinical implementation of AI in radiology in the radiology and data-science literature (He et al., 2019; Liew, 2018; Oakden-Rayner, 2019). By using a refined version of the nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework developed by Greenhalgh et al., (2017), the dynamic nature of the implementation processes of AI technology in clinical radiology became apparent and our understanding of underlying barriers and facilitators was extended. The NASSS framework appeared as a useful tool not only to map hindering and facilitating factors to the implementation of AI applications for clinical radiology, but also to understand the dynamic interactions between these factors.

By targeting cases of unsuccessful adoption, the NASSS framework tries to go beyond the more traditional studies on determinants of adoption. In general, studies on rejection or abandon of innovations tend to be underrepresented in the diffusion of innovation literature, which suffers from a strong bias towards successful innovations (Rogers, 2003). This pro-innovation bias also expresses itself in the form of implicit assumptions inherent to many diffusion and adoption of innovation models, for example the ideas that rejection of innovation is merely temporary and that innovation should be adopted by all potential users as quickly as possible (Bauer, 2017; Rogers, 2003). Especially the last element can also be identified in the NASSS model, showing that pro-innovation bias persists even in models that explicitly target non-adoption and abandon of innovations. Another body of literature has studied the topic of non-adoption from the perspective of the (non-)users of technology (Wyatt, Henwood, Hart, & Smith, 2005; Wyatt, Thomas, & Terranova, 2002). Wyatt et al. (2002) identified four categories of non-users, the resisters, the rejecters, the excluded and the expelled. While this typology contrasts the pro-innovation bias by incorporating the idea that non-use can be an informed and rational decision, it is a rather static view on the phenomenon of non-adoption.

The NASSS framework's focus on the adoption *process* allows for a more dynamic analysis, by covering the implementation phase of an innovation. This phase encompasses the period between the adoption decision and the point of sustainable adoption, i.e. the moment an innovation is used in routine practice after initial adoption efforts are concluded (Greenhalgh et al., 2017). It is in this phase that phenomena like abandonment of the technology, difficulties in scaling up and difficulties to reach sustainable adoption emerge. This became apparent in the case of BoneXpert. While BoneXpert was technically adopted (it was installed and running normally) in all seven cases, sometimes for extended time-periods, several occasions of non-adoption and abandonment by the intended users were observed. Additionally, obstacles to sustainable implementation became apparent, such as inconsistent use in clinical practice across individual users.

Lacking acceptance of adopters was identified as one of the most important causes for non-adoption, abandonment and thus a barrier to successful implementation of AI applications in radiology. The determinants of radiologists' acceptance of AI application found in this study are in line with determinants of clinicians' acceptance of Computerized Decision Support Systems (CDSS) found in previous empirical literature. These include insufficient knowledge (Bulder, 2018; He et al., 2019; Liberati et al., 2017), trust (Liberati et al., 2017; Lugtenberg, Weenink, Van Der Weijden, Westert, & Kool, 2015), change in clinician's professional identity and professional autonomy (Bezemer et al., 2019; Liberati et al., 2017). Furthermore, the role of evidence on innovation adoption has been discussed extensively in the field of evidence-based health care. Scientific evidence was found to be an important determinant of innovation adoption for practitioners in the acute setting, a finding that appears to hold for the case of AI in radiology (Turner et al., 2017; Urquhart et al., 2019). The importance of adopter acceptance confirms a central proposition of Greenhalgh et al. (2017). However, due to the limited number of interviewees, as well as the broad scope of the elements covered by the NASSS framework, it was not possible to single out which of these determinants are the most important ones for determining acceptance from radiologists towards AI applications. This difficulty to go more in-depth on specific elements, caused by the breadth of the framework, highlights one of the key limitations of the NASSS framework for the study of AI applications in radiology.

However, the breadth of the NASSS framework, including different levels of analysis (i.e., the individual users, departments, organizational), proved useful for the detection of examples of abandonment of the technology, as was shown in the example of BoneXpert. Also caused by non-acceptance of AI applications, these examples of abandonment concerned mostly the referring clinicians. The referring clinicians, which do not directly use the technology, can be considered indirect adopters. Because of their indirect relation with AI applications for radiology, they run the risk of getting overlooked or forgotten in adoption and implementation processes that are, at first sight, limited to one organizational unit (e.g. the radiology department). A crucial role in overcoming lacking acceptance of (direct and indirect) adopters is played by the local champion. The importance of having a local champion had previously appeared in research on the adoption of telehealth systems (Wade & Eliott, 2012), as well as on the implementation of CDSS (Liberati et al., 2017). Notably, a recent study on adoption of CDSS in US-American radiology departments, also identified local champions as an important facilitator for adoption (Marcial et al., 2019). Both studies mention the local champions' important role in starting and advancing adoption and implementation processes of CDSS.

The local champions identified in this research were found to engage in vision- and trust-building concerning AI applications for radiology, as well as convincing their peers and mobilizing resources for AI applications. Interestingly, these actions coincide with the activities typically attributed to 'institutional entrepreneurs' (Battilana, Leca, & Boxenbaum, 2009). Institutional entrepreneurs can be defined as "change agents who, whether or not they initially intended to change their institutional environment, initiate, and actively participate in the implementation of, changes that diverge from existing institutions" (Battilana et al., 2009, p. 70). This definition of Battilana et al. (2009) highlights in particular that institutional entrepreneurs not only initiate, but also 'actively participate in the implementation of change'. These were precisely the characteristics attributed to the local champions in this research. Yet, most local champions initiate and implement change first and at foremost locally. This means, a priori, they have only limited direct impact on the institutions surrounding AI applications. An exception is formed by those local champions that are also active outside their organizations, for example within the professional organization. Nonetheless, the identification of the local champion shows the potential overlap between innovation adoption and institutional entrepreneurship literature.

Hence, based on the empirical findings, the NASSS framework was adapted and refined for the case of AI-based applications in radiology. The main adaptation between the original NASSS model and the NASSS for AI in radiology concerns the 4th domain, the 'adopter system'. For the case of AI applications in radiology, it appeared that besides the direct adopters of the technology, the radiologists, the adopter system includes two additional actors: the indirect adopters (most of all the referring clinicians) and the local champion.

Greenhalgh et al. (2017) pays special attention to the role of complexity in determining the success or failure of adoption processes of technology in health care, where complexity is defined by the dynamic, unpredictable nature of the technology, or the fact that a technology is not easily disaggregated into constituent components. It is argued that the higher the complexity of a technology, the more difficult it is to reach sustainable adoption. While determining the level of complexity of AI applications was not an explicit focus of this research, it became evident that several of the elements in the adapted NASSS framework are characterized by rapid change and uncertainty. Thus, AI applications for radiology appear to qualify as complex. This complexity can be related to the fact that AI applications for radiology combine three elements: rapid developments in the field of the analysis of big data, the rigid professional and organizational structures of the health sector and the prospect of automating the work of highly trained individuals. To varying degrees, these three elements also apply to other applications AI in medicine or health care. Therefore, it can be expected that the adapted NASSS framework for AI can provide relevant insights for the implementation of AI applications in other medical or health care domains.

As proposed by Greenhalgh et al. (2017), the NASSS model cannot be applied as a deterministic tool to particular intervention cases, but is intended to be used "in technology implementation projects to address the micro-level challenges of individual adoption, the meso-level challenges of organizational assimilation, and the macro-level challenges of the policy and regulatory environment" (Greenhalgh et al., 2017, p. 15). As such, it proved an excellent tool to map implementation barriers and facilitators on all three levels (micro, meso and macro-level) as well as understanding how these factors interact. The gained insights were condensed in a guide to successful AI implementation, which is discussed in the following part.

Practical Relevance

This research has several practical contributions. The first contribution is capturing an overview of the current state of adoption and implementation of AI applications in Dutch radiology departments. It became evident that, at present, only a very limited number of AI applications are de facto being used in routine clinical practice in radiology departments of Dutch hospitals. On the one hand, this is due to the early phase of the development of the technology, on the other hand it can be explained with the identified barriers for implementation. The systematized mapping of barriers and facilitators, as well as the dynamics between these factors, presents the second, and more important, practical contribution of the study. It provides clinicians and hospital management with important insights on the elements to take into account when planning and executing implementation projects for AI applications. As became evident during the interviews held for the research, this information is of large interest to radiologists, members of hospital management and members of professional organizations in the Netherlands.

Based on the results and analysis of this research, a practical guide to support the planning and implementation process for a specific AI application was created. It turns around four questions: why, who, how and what (as presented in detail in appendix 6). The first question, *why*, should come at the very beginning of the process. It is fundamental to identify the clinical or organization problem, which is aimed to be tackled by the AI application, and determine the desired clinical added-value(s). The second question, *who*, helps to map the relevant stakeholders for the implementation of that particular application. It is important that not only the future users (i.e., the radiologists), but also the indirect adopters (i.e., referring clinicians) as well as supporting staff (management, medical physicists, innovation specialists) are identified. Special attention should also be given to single out individuals, which could act as local champions. The third question, *how,* refers to the nature of how the planning and execution of the implementation process is played out. It is essential to involve the identified stakeholders in strategy building and implementation processes. This should be done through actively promoting discussion sessions and feedback rounds surrounding the implementation process. These activities should be repeated as the implementation advances over time. The fourth question, *what*, targets the elements that should be in the implementation strategy. Besides the more obvious elements, like workflow and PACS integration and funding, special attention should be given to the questions surrounding validation, monitoring and necessary knowledge to use.

The third practical contribution of this research concerns the developers of AI applications in clinical radiology. As was mentioned repeatedly by many of the interviewees, there is a large gap between developers and users of AI applications in radiology. This is expressed both in lacking user-friendliness, as well as poor communication to resolve these problems. The insights that resulted from this research can therefore provide relevant insights also for developers, as was confirmed by the developer of BoneXpert, who has already expressed his interest in the results of the research.

Limitations

Due to its exploratory nature based on qualitative data, several limitations of the research need to be taken into consideration, mainly regarding the internal validity, the measurement validity and generalizability.

Data collection and data analysis were conducted by a single researcher, reducing measurement validity and internal validity of the study. This concern was addressed in three ways: by performing data triangulation, by holding iterative rounds of expert validation and by documenting the empirical and analytical approach in all stages of the research process. First, the highest possible measurement validity was attempted by performing data-triangulation. The rich contextual data gained through a total of 24 semi-structured interviews was complemented with the analysis of a number of documents. The semi-structured nature of the interviews allowed for the emergence of new themes during the conversations, in addition to discussing pre-identified topics. Interviews were conducted until the point of thematical saturation, meaning when no new themes appeared during additional interviews. The concepts and mechanisms identified during the interviews were supported by the internal documents, the publications by relevant stakeholders, the scientific literature and the grey literature examined as part of the document analysis. Second, to compensate for low internal validity, the empirical and analytical approach was closely documented in all stages of the research process. The interview guides, and coding framework can be found in the appendix (appendices 3-5). Transcripts of the interviews can be obtained on request by the researcher. Additionally, both in the beginning and at the end of the research process, experts from the field of study (namely data scientists and radiologists) provided validation for the theoretical framework and propositions, as well as identified concepts and findings. Third, in order to increase the internal validity of the research, repeated discussion rounds with a secondary researcher were held during different phases of the research process, amounting to investigator triangulation. This allowed to validate the identified concepts, their interrelations and the theoretical implications of the empirical findings. Nevertheless, replicability for this type of qualitative research is difficult to the dynamic and constantly changing social nature of the studied social environment.

A further limitation is linked to the generalizability of the research. Overall, it can be said that due to the contextual specificities of the selected cases and interviewees the conclusions drawn from this research might not automatically generalize to a larger population. This is mainly caused by issues concerning the sampling of cases and interviewees. The sampling strategy for the cases was based on the hospitals' use of the software program BoneXpert. As previously mentioned, BoneXpert was chosen as a sampling criterion, because it was identified as the only AI application in routine clinical use across various hospitals. However, BoneXpert influences only a very small part of the day to day work of a majority of radiologists. Therefore, it is questionable whether the results found for the implementation of BoneXpert hold for more complex AI applications, which represent a larger part in radiologists daily work and routines.

Of the identified eight hospitals that use BoneXpert, only seven were included in this research, due to the lack of response from the eighth case. Additionally, the number and position of interviewees varied between these seven cases. The interviewees were sampled based on their knowledge and experience with the topic of AI applications in radiology, with the aim to gain the most insights on the topic from the different organizational levels covered by the theoretical framework. This means that across the cases, individuals with different roles and positions were interviewed, limiting the generalizability. Furthermore, interviewees within cases also varied with regard to their experience and knowledge on the topic, due to the early stage of implementation of AI in clinical radiology practice. On the one hand, this means that parts of the statements from the interviews are based on expectations, rather than actual use experience with AI applications. On the other hand, it can be assumed that the sample of interviewees within cases is biased towards individuals with a particular interest in the technology and its implementation, and that these interviewees have an above-average positive attitude towards the technology.

In addition to the insights of the seven cases, interviews with representatives of professional organizations provided additional macro-level insights and confirmed the findings from the cases. These were identified based on purposive sampling. Interestingly, the interviewees in the sample happened to repeatedly refer to each other for further insights, and it appeared that the most knowledgeable and active individuals in the field of AI in radiology in the Netherlands have been included in the sample, enhancing generalizability to the national context.

The Netherlands were chosen as a case, due to the increase in availability of medical imaging in recent years, combined with high pressure to contain costs for medical specialist care. These factors indeed appeared to play a significant role for the implementation of AI applications in clinical radiology. As previously mentioned, health care systems differ greatly across countries, due to institutional set-up. Thus, generalizability of the present findings to other countries cannot be assured.

Future Research

Considering the theoretical contribution and the limitation of this research, several suggestions for future research can be derived.

Thematically, more in-depth insights are needed concerning the most important barriers and facilitators identified above. First of all, the application at the core of this research, BoneXpert, only presents a very small part of the day to day work of radiologists. In order to achieve better generalizability of these results, an application that presents higher complexity and represents a larger part of the diagnostic work done by radiologists should be studied. Further, the determinants for the acceptance of both the direct and indirect users need to be further studied. It should be investigated, which of the various determinants is most important for determining user acceptance and consequently, through which mechanisms low user acceptance can be increased. In connection herewith, the role played by technical knowledge on AI on user acceptance needs to be better understood. It is important to specifically determine, what kind of knowledge is needed for radiologists, in order to include this in professional training for practicing radiologists and the curriculum of future radiologists. Furthermore, the change of AI on the professional identity of radiologists, its relation to expectations on the technological development and the job replacing effect of the technology need to be better understood. This could be done by comparing the case of AI application in radiology with other applications of AI in the medical field, such as AI applications in pathology or dermatology, two disciplines that are also experiencing the development of different AI applications. Concerning the adopting organizations, the role of innovation culture and strategy on the implementation of AI in radiology needs to be understood better. While the presence of innovation leadership and innovation strategies appeared to be facilitating factors, in-depth comparative case studies can provide further insights on the role of these elements on the sustained implementation of AI in radiology. Finally, this research was limited to one country. Cross-country comparisons can help to gain a better understanding on how the structure and principles of the national health care system influence the implementation process of AI in radiology.

Theoretically, the topic of complexity provides ground for further research. While the NASSS model focuses on complexity as an important determinant for implementation, this aspect was not explicitly considered in the present research. It can be assumed that due to their different scope and functionality, different AI applications present different levels of complexity. Since these applications get adopted in the same contexts, they provide an interesting opportunity to investigate the role of complexity on the success of implementation processes.

Hopefully, the findings of this research in the present report can serve as guidance to the radiologists, members of hospital management and other stakeholders involved in the process of implementing AI applications for radiology. For, as said by a participating interviewee (IV4): "*if it is possible to program a computer to do the boring job of humans, it makes no sense whatsoever to use humans*."

BIBLIOGRAPHY

Abbott, P. A., Foster, J., Marin, H. de F., & Dykes, P. C. (2014). Complexity and the science of implementation in health IT-knowledge gaps and future visions. *International Journal of Medical Informatics*, *83*, e12–e22.

AINED. (2018). *AI voor Nederland*. Retrieved from https://www.vno-ncw.nl/sites/default/files/aivnl_20181106_0.pdf

Algra, P. (2017). Worden we de nieuwe lantaarnopstekers? *MemoRad*, *22*(3), 17–18.

Anderson, N. R., & West, M. A. (1998). Measuring climate for work group innovation: Development and validation of the team climate inventory. *Journal of Organizational Behavior*, *19*, 235–258.

Battilana, J., Leca, B., & Boxenbaum, E. (2009). How Actors Change Institutions: Towards a Theory of Institutional Entrepreneurship. *The Academy of Management Annals*, *3*, 65–107.

Bauer, M. W. (2017). Resistance as a latent factor of innovation. In B. Godin & D. Vinck (Eds.), *Critical Studies of Innovation, Alternative approaches to the pro-innovation bias* (pp. 159–181). Cheltenham: Edward Elgar Publishing.

Bezemer, T., de Groot, M. C., Blasse, E., Ten Berg, M. J., Kappen, T. H., Bredenoord, A. L., … Haitjema, S. (2019). A Human(e) Factor in Clinical Decision Support Systems. *Journal of Medical Internet Research*, *21*, e11732.

Bluemke, D. A. (2018). Radiology in 2018: Are You Working with AI or Being Replaced by AI? *Radiology*, *287*, 365–366.

Bryman, A. (2012). *Social Research Methods* (4th ed.). New York: Oxford University Press.

Bulder, M. (2018). *The Transformation of Radiology from a Knowledge and Learning Perspective (Master's Thesis)*. Retrieved from http://www.ubvu.vu.nl/pub/fulltext/scripties/27_2618614_0.pdf

Chaudoir, S. R., Dugan, A. G., & Barr, C. H. I. (2013). Measuring factors affecting implementation of health innovations: A systematic review of structural, organizational, provider, patient, and innovation level measures. *Implementation Science*, *8*. https://doi.org/10.1186/1748-5908-8-22

Choy, G., Samir, A. E., & Brink, J. A. (2018). Current Applications and Future Impact of Machine Learning in Radiology. *Radiology*, *288*, 318–328.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*, 3–21.

Council of the European Communities. *Council Directive of 14th June 1993 concerning medical devices. 93/42/EEC. 2007*. , (1993).

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., & Medical Research Council Guidance. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ (Clinical Research Ed.)*, *337*, a1655.

Cresswell, K., & Sheikh, A. (2013). Organizational issues in the implementation and adoption of health information technology innovations: An interpretative review. *International Journal of Medical Informatics*, *82*, e73–e86.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C.

(2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science*, *4*, 1–15.

Doi, K. (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, *31*, 198–211.

Edmondson, A. A. C., Bohmer, R. R. M., & Pisano, G. P. (2001). Disrupted routines: Team learning and new technology implementation in hospitals. *Administrative Science Quarterly*, *46*, 685–716.

European Parliament and the Council of the European Union. *Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU*. , (2017).

European Society of Radiology. (2019). What the radiologist should know about artificial intelligence – an ESR white paper. *Insights into Imaging*, *10*, 44.

Eurostat. (2018). Healthcare expenditure statistics. Retrieved December 18, 2018, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Healthcare_expenditure_statistics

EY. (2019). *How will advances in technology put the consumer at the center of health care? The Netherlands 2019*. Retrieved from https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/health/ey-nextwave-health-report-ned-findings.pdf

Fleuren, M., Wiefferink, K., & Paulussen, T. (2004). Examining the Evidence Determinants of innovation within health care organizations Literature review and Delphi study. *International Journal for Quality in Health Care*, *16*, 107–123.

French-Mowat, E., & Burnett, J. (2012). How are medical devices regulated in the European Union? *Journal of the Royal Society of Medicine*, *105*, 22–28.

Geldermann, I., Grouls, C., Kuhl, C., Deserno, T. M., & Spreckelsen, C. (2013). Black box integration of computer-aided diagnosis into PACS deserves a second chance: Results of a usability study concerning bone age assessment. *Journal of Digital Imaging*, *26*, 698–708.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *Milbank Quarterly*, Vol. 82, pp. 581–629.

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., … Shaw, S. (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*, e367.

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., … Shaw, S. (2018). Analysing the role of complexity in explaining the fortunes of technology programmes: Empirical application of the NASSS framework. *BMC Medicine*, *16*, 66.

Gustafson, D. H., Sainfort, F., Eichler, M., Adams, L., Bisognano, M., & Steudel, H. (2003). Developing and Testing a Model to Predict Outcomes of Organizational Change. *Health Services Research*, *38*, 751–776.

Harris, S. (2019). *Funding Analysis of Companies Developing Machine Learning Solutions for Medical Imaging*. Retrieved from https://s3-eu-west-2.amazonaws.com/signifyresearch/app/uploads/2018/06/21115409/21.06.18-Funding-

Analysis-of-Companies-Developing-Machine-Learning-Solutions-for-Medical-Imaging.pdf

Harvey, H. (2018). Why AI will not replace radiologists. Retrieved July 8, 2019, from Towards Data Science website: https://towardsdatascience.com/why-ai-will-not-replace-radiologists-c7736f2c7d80

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

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, *18*, 500–510.

Klein, K. J., & Speer Sorra, J. (1996). The Challenge of Innovation Implementation. *Academy of Management Review*, *21*, 1055–1080.

Kohli, A., & Jha, S. (2018). Why CAD Failed in Mammography. *Journal of the American College of Radiology*, *15*, 535–537.

Kroneman, M., Boerma, W., van den Berg, M., Groenewegen, P., de Jong, J., & Van Ginneken, E. (2016). The Netherlands: health system review. *Health Systems in Transition*, *18*, 1–239.

Kukk, P., Moors, E. H. M., & Hekkert, M. P. (2016). Institutional power play in innovation systems: The case of Herceptin®. *Research Policy*, *45*, 1558–1569.

Lee, C. I., Gupta, S., Sherry, S. J., Chiunda, A., Olson, E., Chokshi, F. H., … Duszak, R. (2018). Translating New Imaging Technologies to Clinical Practice. *Academic Radiology*, *25*, 3–8.

Lee, H., Tajmir, S., Lee, J., Zissen, M., Yeshiwas, B. A., Alkasab, T. K., … Do, S. (2017). Fully Automated Deep Learning System for Bone Age Assessment. *Journal of Digital Imaging*, *30*, 427–441.

Liberati, E. G., Ruggiero, F., Galuppo, L., Gorli, M., González-Lorenzo, M., Maraldi, M., … Moja, L. (2017). What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implementation Science*, *12*, 1–13.

Liew, C. (2018). The future of radiology augmented with Artificial Intelligence: A strategy for success. *European Journal of Radiology*, *102*, 152–156.

Lodwick, G. S. (1966). Computer-aided diagnosis in radiology: A research plan. *Investigative Radiology*, *1*, 72–80.

Lugtenberg, M., Weenink, J. W., Van Der Weijden, T., Westert, G. P., & Kool, R. B. (2015). Implementation of multiple-domain covering computerized decision support systems in primary care: A focus group study on perceived barriers. *BMC Medical Informatics and Decision Making*, *15*, 1–11.

Lupton, D. (2014). Critical Perspectives on Digital Health Technologies. *Sociology Compass*, *8*, 1344–1359.

Lupton, D. (2016). Digitised health, medicine and risk. *Health, Risk and Society*, *17*, 473–476.

Marcial, L. H., Johnston, D. S., Shapiro, M. R., Jacobs, S. R., Blumenfeld, B., & Rojas Smith, L. (2019). A qualitative framework-based evaluation of radiology clinical decision support initiatives: eliciting key factors to physician adoption in implementation. *JAMIA Open*, *2*, 187–196.

Martin, D. D., Deusch, D., Schweizer, R., Binder, G., Thodberg, H. H., & Ranke, M. B. (2009). Clinical application of automated Greulich-Pyle bone age determination in children with short stature. *Pediatric Radiology*, *39*, 598–607.

May, C. (2006). A rational model for assessing and evaluating complex interventions in health care. *BMC Health Services Research*, *6*, 1–11.

May, C., & Finch, T. (2009). Implementing, embedding, and integrating practices: An outline of normalization process theory. *Sociology*, *43*, 535–554.

May, C., Finch, T., Mair, F., Ballini, L., Dowrick, C., Eccles, M., … Heaven, B. (2007). Understanding the implementation of complex interventions in health care: The normalization process model. *BMC Health Services Research*, *7*, 148.

Meurs, P. (2019, February 5). Zorgsector kan kunstmatige intelligentie veel slimmer inzetten. *Het Financieele Dagblad*. Retrieved from https://fd.nl/opinie/1288142/zorgsector-kan-kunstmatige-intelligentie-veel-slimmer-inzetten

Ministerie van Volksgezondheid Welzijn en Sport. *Kamerbrief over aanbieding Bestuurlijk akkoord (Hoofdlijnenakkoord) medisch-specialistische zorg 2019-2022*. , (2018).

Moja, L., Liberati, E. G., Galuppo, L., Gorli, M., Maraldi, M., Nanni, O., … Kwag, K. H. (2014). Barriers and facilitators to the uptake of computerized clinical decision support systems in specialty hospitals: Protocol for a qualitative cross-sectional study. *Implementation Science*, *9*, 1–6.

Morgenstern, M. (2016, June 25). Automation and Anxiety: the impact on jobs. *The Economist - March of the Machines: A Special Report on Artificial Intelligence*. Retrieved from https://www.economist.com/special-report/2016/06/25/automation-and-anxiety

Nawrocki, T., Maldjian, P. D., Slasky, S. E., & Contractor, S. G. (2018). Artificial Intelligence and Radiology: Have Rumors of the Radiologist's Demise Been Greatly Exaggerated? *Academic Radiology*, *25*, 967–972.

Nederland Digitaal. (2019). Resultaten en opbrengsten van de Conferentie Nederland Digitaal 2019. *Conferentie Nederland Digitaal*. Retrieved from https://www.nederlanddigitaal.nl/documenten/publicaties/2019/03/21/opbrengsten-conferentie-nederland-digitaal-2019

Nederlandse Vereniging van Ziekenhuizen. (2018). Ziekenhuiszorg in Cijfers 2018. In *Brancherapport algemene ziekenhuizen 2018*. Retrieved from https://ziekenhuiszorgincijfers.nl/assets/uploads/NVZ-Brancherapport-2018.pdf

Nederlandse Vereniging voor Radiologie. (2016). *De rol van de radioloog in 2020*. Retrieved from https://www.radiologen.nl/system/files/bestanden/documenten/de_rol_van_de_radioloog_in_2020.pdf

Nederlandse Vereniging voor Radiologie. (2019). Secties. Retrieved July 8, 2019, from https://www.radiologen.nl/nvvr/secties

Nishikawa, R. M., & Bae, K. T. (2018). Importance of Better Human-Computer Interaction in the Era of Deep Learning: Mammography Computer-Aided Diagnosis as a Use Case. *Journal of the American College of Radiology*, *15*, 49–52.

Oakden-Rayner, L. (2019). The Rebirth of CAD: How Is Modern AI Different from the CAD We Know? *Radiology: Artificial Intelligence*, *1*, e180089.

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine*, *375*, 1216–1219.

OECD. (2018). Health Resources: Medical Technology. Retrieved December 18, 2018, from https://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT#

Orde van Medisch Specialisten. (2014). *Leidraad Nieuwe Interventies in de klinische praktijk*. Retrieved from https://www.demedischspecialist.nl/sites/default/files/Leidraad Nieuwe interventies in de klinische praktijk def.pdf

Ottes, L. (2016). *Big Data in de zorg* (No. 19). Retrieved from https://www.wrr.nl/publicaties/working-papers/2016/04/28/big-data-in-de-zorg

Pesapane, F., Volonté, C., Codari, M., & Sardanelli, F. (2018). Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. *Insights into Imaging*, 1–9.

Plsek, P. E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. *BMJ: British Medical Journal*, *323*, 625–628.

Pope, C., Halford, S., Turnbull, J., Prichard, J., Calestani, M., & May, C. (2013). Using computer decision support systems in NHS emergency and urgent care: Ethnographic study using normalisation process theory. *BMC Health Services Research*, *13*.

Powell, B. J., Proctor, E. K., Glisson, C. A., Kohl, P. L., Raghavan, R., Brownson, R. C., … Palinkas, L. A. (2013). A mixed methods multiple case study of implementation as usual in children's social service organizations: Study protocol. *Implementation Science*, *8*, 92.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., … Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 65–76.

Robert, G., Greenhalgh, T., MacFarlane, F., & Peacock, R. (2010). Adopting and assimilating new non-pharmaceutical technologies into health care: A systematic review. *Journal of Health Services Research and Policy*, *15*, 243–250.

Rogers, E. M. (1962). *Diffusion of Innovation* (1st ed.). New York: Free Press of Glencoe.

Rogers, E. M. (2003). *Diffusion of Innovations* (4th ed.). New York: Free Press of Glencoe.

Rosenau, P. V., & Lako, C. J. (2008). An Experiment with Regulated Competition and Individual Mandates for Universal Health Care: The New Dutch Health Insurance System. *Journal of Health Politics, Policy and Law*, *33*, 1031–1055.

Schmidhuber, J. (2015). Deep learning in neural networks – An overview. *Neural Networks*, *61*, 85–117.

Scholz, R., & Tietje, O. (2002). *Embedded Case Study Methods*. Thousand Oaks: SAGE Publications, Inc.

Schut, F. T., & Van de Ven, W. P. M. M. (2005). Rationing and competition in the Dutch health-care system. *Health Economics*, *14*, 59–74.

Turner, S., D'Lima, D., Hudson, E., Morris, S., Sheringham, J., Swart, N., & Fulop, N. J. (2017). Evidence use in decision-making on introducing innovations: A systematic scoping review with stakeholder feedback. *Implementation Science*, *12*, 1–12.

Urquhart, R., Kendell, C., Geldenhuys, L., Ross, A., Rajaraman, M., Folkes, A., … Porter, G. A. (2019). The role of scientific evidence in decisions to adopt complex innovations in cancer care settings: A multiple case study in Nova Scotia, Canada. *Implementation Science*, *14*, 1–12.

Van de Ven, A. H., Polley, D. E., Garud, R. A., & Venkataraman, S. (1999). *The Innovation Journey*. Oxford, New York: Oxford University Press.

van Ginneken, B. (2017). Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. *Radiological Physics and Technology*, *10*, 23–32.

van Ginneken, B. (2018). AI and radiologists: a painful divorce? Retrieved June 25, 2019, from Medium website: https://medium.com/@bramvanginneken/ai-and-radiologists-a-painful-divorce-3b39d2c289aa

van Ginneken, B., Schaefer-Prokop, C. M., & Prokop, M. (2011). Computer-aided Diagnosis: How to Move from the Laboratory to the Clinic. *Radiology*, *261*, 719–732.

Van Rijn, R. R., Lequin, M. H., & Thodberg, H. H. (2009). Automatic determination of Greulich and Pyle bone age in healthy Dutch children. *Pediatric Radiology*, *39*, 591–597.

Visiana. (2018). Testimonials and references. Retrieved November 26, 2018, from https://www.bonexpert.com/documentation/testimonials-and-references

Volksgezondheidenzorg.info. (2019). Trend in kosten van ziekenhuiszorg. Retrieved June 25, 2019, from https://www.volksgezondheidenzorg.info/onderwerp/ziekenhuiszorg/kosten/kosten#node-trend-kosten-van-ziekenhuiszorg

Wade, V., & Eliott, J. (2012). The role of the champion in telehealth service development: A qualitative analysis. *Journal of Telemedicine and Telecare*, *18*, 490–492.

WRR. (2016). *Big Data in een vrije en veilige samenleving* (Vol. 95). Retrieved from https://www.wrr.nl/publicaties/rapporten/2016/04/28/big-data-in-een-vrije-en-veilige-samenleving

Wyatt, S., Henwood, F., Hart, A., & Smith, J. (2005). The digital divide, health information and everyday life. *New Media and Society*, *7*, 199–218.

Wyatt, S., Thomas, G., & Terranova, T. (2002). They Came, They Surfed, They Went Back to the Beach, Conceptualising Use and Non-use of the Internet. In S. Woolgar (Ed.), *Virtual Society?* (pp. 23–40). Oxford: Oxford University Press.

Yin, R. K. (2003). *Case Study Research: Design and Methods* (Third Edit). Thousand Oaks: SAGE Publications, Inc.

APPENDICES

Appendix 1: Overview of Interviewees

Interviewees	Role	Seniority	Case Affiliation	Additional Roles	Form of IV	Duration (min)
IV1	Radiologist	Senior	TKZ1	Research collaboration with academic hospital	In Person	75
IV2	Radiologist	Senior	TKZ2	Educational Committee Professional Society	In Person	43
IV3	Chairman of Professional Society	Senior	Professional Organization	Clinical radiologist/ PhD on AI in Medical Imaging	In Person	60
IV4	Data Scientist	Professor	UMC2	Founder of several companies in AI for medical imaging	In Person	45
IV5	Radiologist	Senior	UMC3	Coordinator pediatric radiology + Responsible for residents training	In Person	48
IV6	Radiologist	Senior	UMC3	Management role - Research on Business aspects of AI	In Person	45
IV7	Radiologist	Resident	UMC2	PostDoc on AI in Radiology	In Person	22
IV8	Radiologist	Senior	TKZ2	Medical Manager of Radiology Department	In Person	50
IV9	Radiologist	Senior	AZ1	Medical manager of radiology department	Telephone	37
IV10	Radiologist	Senior	UMC1		In Person	27
IV11	Radiologist	Resident	UMC2	PhD on AI in Radiology	In Person	25
IV12	Technician	Resident	TKZ2	Assesses technology for department	In Person	48
IV13	Radiologist	Senior	UMC1	Research in AI	In Person	25
IV14	Clinical Physicist	Senior	TKZ1		In Person	50

IV15	Legal Advisor	Senior	TKZ1		In Person	54
IV16	Operational Manager Radiology Department	Senior	TKZ1	Also manages other departments	In Person	53
IV17	Advisor Innovation Strategy	Junior	TKZ2		In Person	52
IV18	Innovation and valorization officer	Senior	UMC1		Telephone	40
IV19	Implementation Advisor	Senior	Professional Organization	Member of Dutch Implementation Network	In Person	62
IV20	Innovation Lead	Senior	Developer		In Person	70
IV21	Radiologist	Senior	UMC4	Did Validation Study on BoneXpert	In Person	72
IV22	Radiologist	Senior	UMC1	Board member of Screening Organization	In Person	62
IV23	Managing Director		Professional Organization		Telephone	35
IV24	Data Scientist	Senior	UMC3	Entrepreneur in AI for medical imaging	Telephone	20
Expert IV1	Data Scientist	Assistant Professor	Expert	Research on AI in medical imaging	In Person	50
Expert IV2 & 4	Radiologist	Senior	Expert	Member of AI study group in Professional Organization	In Person	47
Expert IV3	Data Scientist		Expert	Founder of Visiana (developer of BoneXpert)	Telephone	35

Appendix 2: Overview of Cases

Case	Type of Hospital	Name	#inter views	# of Radiologists in department	Time of Use of BoneXpert (in years)
TKZ1	Topklinisch Ziekenhuis	Jeroen Bosch Ziekenhuis - Den Bosch	4	16	< 1
TKZ2	Topklinisch Ziekenhuis	Rijnstate Ziekenhuis - Arnhem	4	18	4
UMC1	UMC	UMC Utrecht (including Wilhelmina Kinderziekenhuis)	4	40(of which 5 pediatric)	7
UMC2	UMC	Radboud UMC Nijmegen	3	25-30	> 5
UMC3	UMC	Erasmus UMC Rotterdam (including Sophia Kinderziekenhuis)	3	30-35	7
UMC4	UMC	AUMC - Amsterdam - locatie AMC	1	35(of which 4 pediatric)	5
AZ1	Algemeen Ziekenhuis	Scheper Ziekenhuis - Emmen	1	14	3

Appendix 3: Interview guide for radiologists

<u>Interview Guide: Radiologist</u>

Personal Information:

- How long have you been working as a radiologist?
- How long have you been working at this hospital?
- Do you have a leadership role?
    - Within the department?
    - Within the hospital?
- Do you have any other professional engagements (e.g. teaching/research)
- Very generally: would you consider yourself a technology savvy person?

General questions on Hospital

- Would you consider this hospital innovative?
- Are the pushers for new technology? Who are they?
- How are innovation decisions normally taken?
    - Management decided alone? Departments are included in decision process?
    - Do you think risk-raking is encouraged by management?
- Do you think, there is a broader innovation strategy in place?

Technology: General questions

- Have you ever worked with Computer Aided Diagnosis systems?
    - If yes: what type of system? Can you explain the type of application/technology behind the system?
        - Which conditions targeted?
    - When did you work with it (time-frame)? Still in use?
        - Could you explain briefly how the process of acquiring and implementing the CAD unfolded (i.e. who came up with the idea of acquiring the CAD, how did the process of implementation go?)
    - Who would use the system?
        - Radiologists / technical staff / other physicians?

Personal Use experience

What was your experience with BoneXpert & Other applications?

- Was it user-friendly, i.e. easy to work with?
- Was there IT support needed and available?
- Was the system easily integrateable in PACs/workflow?
    - How long does it take for the result to show up in the PACS?
    - Have you noticed difference in bone health index, depending on settings used for acquiring of image?

Routines

- How is system used today? (in terms of workflow = first reader/second reader?)
- How did system change the handling and workflows of the targeted cases?
- Who in the team is most familiar with the system (radiologists or technical staff?)

    - Did this change over time?
- Did work-flows across departments change? (e.g. with the departments that request scans?
- Did you undergo specific training for the use of the system (at implementation stage?)
- Do you think there was risk associated with the system?

Dynamics Department

- How big is department?
- How do different people within the radiology Dpt see the surge of this new Tech?
    - Who is more involved, who is less involved?
- Costs:
    - What type of deal with Visana? Do you think that's expensive?
    - How do costs get covered?

Evaluation:

- How was the system supposed to help in general, i.e. what was the improvement the CAD was supposed to bring (*Save time /Make better diagnoses)?*
    - Was this accomplished?
- What were main difficulties with the system?

National Level

- Guidelines NVvR: Do you think BoneXpert should become standard way of doing Bone Age assessment?
    - If yes, who should decide on this and how would this process play out within NVvR?
    - Absorption of costs? Is there room to renegotiate rates with insurances?

Strategy & Collaboration

- Which applications are used besides BoneXpert?
- Is there some sort of strategic approach within the hospital or radiology dpt for implementation of AI?
    - Development of own algorithms?
- Are there any collaborations with other hospitals or industry?
- What would you see as the role of NVvR for AI in general?

Future Outlook

- What are you expecting of this type of technology over the next couple of years?
- Do you think it will induce major changes to the profession of the radiologist?
- How does he imagine the development over next 5-10 years?
- Which countries will adopt quicker?
- Which type of hospitals will adopt quicker?

Appendix 4: Interview Guide Hospital Management

<u>Interview Guide: Member of Hospital Management</u>

Personal Information:

- What is your professional background
- How long have you been working at this hospital?
  - What is your exact role/function?
  - Were you working at other hospitals prior to here?
- Do you have any other professional engagements (e.g. teaching/research)
- Very generally: would you consider yourself a technology savvy person?

Innovation in this Organization:

- How is innovation managed within the organization?
- Would you consider this hospital innovative?
  - Why?
- Do you have a specific department (or person) that is responsible for innovation?
- Is there an official innovation strategy?
  - If yes, could you outline the strategy quickly?
- How are specific innovation needs identified/determined?
- Once a need is identified, how is the appropriate technology chosen?
  - Is department consulted?
    - If yes, only head of department or broader selection of department members?
- Would you say that risk-taking is encouraged by management within the organization?
- Funding
  - Is there a fixed position in the budget for acquiring new technological innovations?
    - Are these resources very controversial among management/staff?
  - Is there also budget for implementation process?

AI in Radiology (only if interviewee has familiarity with the topic, if not, skip to point 4)

- Maybe you can explain about the current discussion/place of AI within the department (Strategy)
- In case of current use of CAD:
  - adoption process of the technology?
    - If yes: who within the department initially pushes for the technology?
    - How does the process unfold?
    - how was the decision process to acquire the innovation taken?
    - Where there any major difficulties/problems?
  - Use of tech:
    - which problem is this technology trying to solve
  - What were some of the major problems identified with the use of the technology?

Responsibility

- Is there a responsible (person or department) for questions/problems surrounding the innovation?

    - Does this person report back to management?
- Have you observed any change processes, following the introduction of the innovation?
    - E.g. updates internal guidelines/best-practices?
- Has there been an evaluation of the innovation?
    - If yes, what was the outcome
    - If no, why not? Is there one planned?
- Are you aware of any additional problems linked to the technology that have come up?
- Is there a strategy for employing more of these types of applications in the future?

Funding

- How much does the technology cost?
- Which department is paying for it? + How is the department paying for it?
- How are these costs covered (handed down to insurers?)?
    - Are these resources very controversial among management/staffs?

External context:

- Is there collaboration with other stakeholders (other hospitals/ insurers) on the topic?

Digitization efforts in this hospital

- Could you give some examples of other digitization efforts happening within the organizations?
- What do you think are the major benefits that digitization can provide for the hospital (and the health system as a whole?)
- What do you think are the biggest challenges associated?
    - Technological
    - Organizational
    - Human/workforce
    - Legal/regulatory

Appendix 5: Coding Framework

Domain	Concept	Description	Example
1. Targeted Condition			
1A. Technical requirements	Desired clinical benefit for condition	Determining the desired clinical benefit(s) for the condition	"Time is one thing. Because, when a stroke patient comes, time means saving brain for the patient."
	Quality standards required for condition	Determining the level of sensitivity and specificity are acceptable for the particular condition targeted by AI application	"And how good an algorithm needs to be depends on the clinical scenario" [IV13]
1B. Adopters involved	Direct & indirect users	Identifying group of users & indirect users involved in that particular condition	"Now, one of our interventional radiologists took it over, together with one of the neurologists. And they were looking at how they can implement this in our normal daily routine, for the every stroke case." (IV8)
2. Technological Context			
2A. Technical features	Technical Performance	Accurateness of algorithm (mostly expressed in sensitivity and specificity of algorithm)	"And they used it [the mammography CAD system] in the beginning, but it's too sensitive, I understood." [IV10]
	Local Validation	Testing the performance of the application in the adopting hospital	"Because you want to validate the software. And they did some tests [on BoneXpert]. But the results showed that the software was better. [One of the radiologists from my department] did a little study and the software performed better.“ [IV16]
	Empirical evidence	Availability and need for evidence from empirical studies on quality and benefit of specific AI application	"When you look at the scientific publications of algorithms. I mean, they all say they're very accurate. But there's a study that has been published recently, only 6% of the algorithms that were developed, have been validated by the users. Only 6%!" [IV3]
	Integration in Existing IT Infrastructure	Interoperability of AI application (and its output) in software packages and workstations used within organization	"You have to make sure that you can use the application on your PACS system. That infrastructure has to be really good. You cannot use a lot of different screens. If there are many different AI applications, they have to be all part of your normal system and pop up if you need it." [IV7]

	IT support	IT support regarding implementation of AI apps necessary, available, helpful	"Yeah, but then you call [the IT support], and they tell you, just restart your computer. And that's a comment that you really don't want to hear."[IV5]
	Readiness of IT Infrastructure	Quality of existing IT infrastructure (hard- & software) with regard to implementation of AI applications	"We've got some complaints about the infrastructure in this hospital. […] You have people complaining about the performance of the PCs, and the communication between the PCs and the PACS and the different software systems that you've got on your PC. It's always too slow. it crashes sometimes." [IV8]
2B. Comprehension of output	Understanding of technology	Knowledge of radiologists (and other users) about technical aspects of AI	"Met een basiskennis van wiskunde, statistiek en algoritmen (ML, DL, NLP) zal een radioloog beter in staat zijn de kwaliteit van de aangeboden software te beoordelen in het geval van een eventuele investering in een AI-applicatie of in het uitbouwen van een toekomstplan voor de afdeling" (NVVR 23.2)
	Technical training	Need for trainings to users on technical aspects of AI	"So, what I did, I asked [a professor for data science] to come over to the hospital to talk with my colleagues. He did a marvelous presentation. He explained, in a simple way, how it [AI/Machine Learning] works." [IV1]
	Dealing with human-machine contradictions	Dealing with situations where human reader and AI application do not agree	"And then you have to defend yourself because you don't agree with the image the AI product gives you, so that gives you more work." [IV11]
2C. Use in clinical practice	Guidance on implementation	Availability of instructions, best-practices on how to implement AI application	"We want to have a more structured implementation, also to be able to do some research on how to do the implementation, and what is good, what is not good in such a process. But it is difficult, because there's nobody who knows how to implement." [IV12]
	Implementation in workflow	Function of output from AI application in clinical workflow (e.g. supporting or automating diagnostic decision)	"At one point we thought, let's just do the BoneXpert [and no manual analysis anymore]. That's much easier. But the orthopedics didn't want that. Because they noticed that there is quite a discrepancy. " [IV5]
2D. Supply models	Business models	Revenue model of AI applications (e.g. pay per view/ license fee)	"Well, there's different business models, of course. You could say: we pay for each case, or there's a

			license. But the business model, as such, has not been defined yet. Everybody's looking for the best way to sell this. Nobody is sure yet, how should we do this.” [IV3]
3. Value Proposition			
3A. Business case	Different providers	Type of software provider of AI applications	“The Siemens and the Philipps of this world are waiting. They are looking, they are there, but they are not doing anything. They are waiting. And there are all these small start-ups coming. And they are still waiting, until some of the smaller companies are winning and then the big ones will buy it.“ [IV1]
	Development	Costs and difficulties for providers in development process of AI application	“And you need money. Because if you want to develop an algorithm, it takes a lot of work hours. Only to prepare the data, that's the most of the work.” [IV3]
	Revenue streams	Current sources of revenue from AI applications	“And when we look at AI applications, we have several agreements, collaboration agreements with AI companies. […] sometimes they want you to pay, but I'm not going to pay for something I do not know if it's going to add value.” [IV6]
	Cooperation	Cooperation models between adopters & users during development process	“I'm not sure whether we're going to pay for it or not, because they [the startup] also wants this hospital for data to improve their own product. So, they're looking to validate their algorithm, their software here." [IV8]
3B. Value for clinical practice	Clinical benefits	Clinical added-value to be gained from use of AI application	“BoneXpert saves time. But it also prevents that one radiologist measures with one method and the other radiologist in a different way.” [IV12]
	Indirect demand-side benefits	Added-value of AI application for indirect adopters	“But yeah, in the end it’s really about helping the patient in the best possible way, and if a computer can do that better, then that’s just it. And you shouldn’t be against that.“ [IV5]
	Measuring demand-side value	Measuring/quantifying added-value	“We also talked about really measuring time, but that's a very difficult part, because then you have to measure how long it takes to see an image. And that's very difficult because, in between, maybe you have a call.” [IV12]

4. Adopters			
4A. Direct adopters	Variance in acceptance	Degree to which users approve the use of AI applications	“Still one of my colleagues doesn't use [BoneXpert]. And one was sort of hesitant, she's now convinced that she can use it. And my other colleague uses it. Says it's easy to use.” [IV22]
	Opposition	Active or passive actions to obstruct the introduction of AI applications	“The guys, who said 'No AI [for breast radiology] on this department', are now talking with AI-based neuro-imaging software developers to implement their solution here." So, something is happening. But it’s all politics.” [IV1]
	Trust	Degree to which users place confidence in the quality and safety of AI applications	“[My colleagues] don't trust the software. They think we as radiologists can do better than a computer. There's still some anxiety, angst for artificial intelligence. And what is interesting, is that my residents, they trust it blindly. Too blindly sometimes even. They don’t think anymore. And that’s the risk of AI, that you stop thinking.” [IV22]
	Change in practices	Degree to which roles and practices of individual adopters change due to AI applications	"We use [BoneXpert] standard for all hand scans. But we also use our old method, which is the comparing in the book." [IV5]
	Creation of new roles	Establishment of new functions related to AI applications within organization	“We really want to have the new AI working group functioning similarly to the other groups we have, of the body parts.” [IV9]
	Professional identity	Change to users' professional identity induced by AI applications	“Not that all radiologists are not needed anymore, but that's the radiologists are working more as a consultant for a specific area and the hospital for instance neurology or orthopedic surgeons.” [IV7]
	Loss of responsibility	Scope of professional practice of radiologists diminished by AI applications	"It could also be the case that non-radiologists will start using it. For example, it already exists that pneumologists are using AI software to analyze lung emphysema, for example, to make a quantification. [...] That would be in the disadvantage of radiologists, of course." [IV3]
4.2 Indirect adopters	Involvement in adoption process	Engagement of indirect users in design and roll out of adoption/implementation process	"And what we do in implementation phase, we mainly have some questionnaires for the users, mainly radiologist, but possibly also laborants, who

			make the CT. And also neurologists or other physicians involved." (IV12)
	Acceptance of technology	Degree to which indirect users approve the use/the output of AI applications	"And I do really believe in the program. But the endocrinologists of the children hospital, they don't believe in the program. They calculate the bone age again, using the atlas of Greulich and Pyle." [IV10]
4.3 Local champion	Role in adoption process	Function in initiation, design and roll-out of AI adoption & implementation process	"And what you need is somebody in the group who is interested in this, who is well aware of what's going on, and also knows what stakeholders should be contacted within the hospital to collaborate" [IV3]
	Trust building	Actions to enhance confidence of other direct & indirect users towards AI application	"You have to prepare them for these innovations. So it's important to at a very early stage to inform everyone on this project. And you have to have a good story. So, convincing them. You have to be a visionary." [IV2]
5. Organizational Context			
5A. Innovation capacity	Collective decision making	Nature of decision-making processes: includes relevant stakeholders	"I think possibly, there are some people, who don't feel the need to be involved in it [the adoption decisions surrounding AI applications]. [...] But I think you have to involve everybody and give everybody a chance to have their say in it. I think that's very important, that people can bring in their opinion, if they want to." [IV12]
	Innovation leadership	Degree of proactivity of management towards AI applications	"The 'Voorzitter van de Raad van Bestuur', the boss of the hospital, he is also involved [in these innovation projects]. He also wants to innovate." [IV12]
	Hospital-level innovation strategy	Presence of a (formalized) innovation strategy on hospital-level which covers AI applications for radiology	"Well, the five-year strategy is the basis of a lot of strategic decisions from the board, as well as the managing layer. The fact that I'm here, and that I can develop [the innovation program on robotics] is based on that strategic plan, where they want to really develop the robotic expertise. " [IV17]
	Innovation specialists	Presence and of individuals within hospital that initiate and support innovation processes	"Yeah, well, the goal of the innovation group is to do quick prototyping of innovations. Basically, if there is an idea, or what I would prefer, if somewhere there is a problem that can be solved by technology, then

			you look whether that indeed is the case. Whether there is a technological solution that can be easily implemented, prove that it works." [IV14]
	Department-level innovation strategy	Presence of a (formalized) innovation strategy on radiology department-level which covers AI applications	"Well, in the [AI-]program we developed, we wanted to do one off-the-shelf product, and that's BoneXpert. We wanted to do a project, which was reachable, like the scaphoid fractures, it's in collaboration. And the lung-nodules is more advanced, it's more a mid-term project." [IV16]
5B. Readiness for technology	Tension for change	Presence of organizational conditions on department- and hospital-level that demand benefits promised by AI applications for radiology	"But on the other hand, we have had some very serious cases of missing large lung tumors every year. And last year, we had three cases and then inspection came to the hospital, who then went to the board. And the board obviously does not like such visits." [IV1]
	Internal dynamics	Nature of interactions between individuals within radiology department	"I guess we have a group that's very cooperative. So with the right investment of knowledge and discussion of how we're going to improve, we all have a consensus on where to go." [IV2]
	Technology system-fit	Degree to which AI applications fit in, or are integrated in, existing organizational digitization/IT strategy	"And I think from a hospital view, you should have more of these lines together. If you look at the patient dossiers, you will never have every discipline have its own software. You will structure that from the hospital and everybody will have to work with it. Then you can have special fields for a specific discipline, but the core is the same." [IV19]
5C. Nature of adoption/ funding decision	Origin funds contested	Source of funding for AI applications disputed	"This is the hurdle that needs to be taken now. Who has to invest? Is it the radiologists? Is it the patient? Is it the hospital? Is it the State?" [IV3]
	Internal absorption of costs	Mobilization of internal resources to fund of AI applications	"But then the cost savings are not at the radiology departments. And the radiology department is the department that has to pay for it, but it doesn't have then the savings. I think this is the difficult part." [IV12]
	External absorption of costs	Mobilization of external resources to fund of AI applications	"So I said to the clinicians, find a way to get the product paid here in the hospital, go discuss it with the insurance companies." [IV1]

5D. Change to routines	Changes to workflow	Degree to which workflow routines of users changes due to AI application	“You already loose half of the radiologists, if you need to open an extra program for determining some extra figures. [It’s really hard for my colleagues] to change their routine.” [IV11]
	Cooperation with referring clinicians	Nature of collaboration between radiologists and referring clinicians concerning AI applications	“And now one of the interventional radiologists, took it over, together with one of the neurologists. And they were looking at how they can implement this in our normal daily routine, for every stroke case.” [IV8]
	Leading role	Role of radiology department concerning AI with regard to the rest of the hospital	“And it’s better to keep it in our own hands. We make the scans, we have the data, and how we best analyze the scans is something we also should decide.” [IV5]
5E. Work needed to implement change	Framing	Contextualization of AI applications for the radiology profession, or medicine more broadly	“And as a doctor you still can interpret. You are in the responsibility seat. The computer will not take that away, it only helps you. It’s like a co-pilot software. That's what we're saying to downplay the fuss.” [IV16]
	Collective vision building	Nature of vision building process concerning AI applications for radiology: including or excluding relevant stakeholders	"In a couple of weeks, we have a meeting with everyone, and we will also discuss the plans of our AI group, hopefully formalize it. [...] And with the whole group we have to take a decision if we want to continue with the AI group. So we’ll make plans for the next two years. Because, if there is no support, it is better to change the plans, or wait a bit more, or do other projects" [IV9]
	Monitoring	Tracking of implementation process of AI applications	"In our department, we have several quality control projects. And we evaluate them: we usually take measure before and after the projects. And then we do a small research project about the project." [IV9]
6. The Health Care System			
6A. Health care policy context	Market-based thinking	Structural mechanisms of competition between health care providers	“I don't know whether that's because of the free market thinking in health care, that they're supposed to be competitors, and not collaborators. And so sometimes they [the hospitals] just sit on their own innovation, because that's something that distinguish them from the rest." [IV19]

	Inter-organizational collaboration	Mechanisms of collaboration between health care providers	“Yes, a formalization by contract is something we are working on. We have different sorts of contracts. We have a cooperation agreement with our radiology department and a department at an academic hospital. And we have a project agreement for scaphoid fractures. And I think there's another contract for the exchange of data.” [IV15]
	Industry-clinic collaboration	Mechanisms of collaboration between health care providers and developers/suppliers of AI applications	“But also the PACS companies cannot develop everything themselves. Of course, they also need the data. And the hospitals have the data. The academic centers have the data. And the PACS suppliers? Well, they make agreements with hospitals to develop this. “ [IV3]
	Fragmented health system	Range of players in the health care system indirectly involved with implementation of AI applications	“And everybody has his own pot of money.[...] And let's say on something that integrates everyone within a hospital: the nurses don't have their money, the GPs have a little pot, and the medical specialists have a pot, and the patients got a pot of money since a couple of years, but you have to bring it all together. But everybody has his own priorities.” [IV19]
	Macro-trends	Macro-level socio-economic context of health care system	“Especially with a new decision that by 2020, we cannot have an increase in the cost for hospital care. So, the costs need to stay at the same level, although the number of patients is increasing.” [IV21]
6B. Regulatory & legal system	Regulation of medical devices	Regulatory context for AI applications for radiology as medical devices	"And the MDR is being implemented at the moment. And of course, the question is, how is that going to affect artificial intelligence, deep learning, machine learning? I also talked to the hospital association and they said apps are going to be part of it as well. So yeah, everybody's struggling with it." [IV19]
	Privacy	Regulatory context for protection, and use, of personal data	“And with GDPR, it becomes more and more difficult to sustain software systems running in the cloud. And honestly, I never thought about GDPR with this software [BoneXpert], because it is CE

			marked. It's within our own cloud services, so it should be fine. And are all our cloud-services are been checking now for GDPR compliance.“ [IV21]
	Legal responsibility for damage	Legal context of responsibility for potential injury caused by the use of AI applications	“And of course, we have an insurance, but we are only insured when we are responsible for the mistake we have made. And if we have software or a medical device, which we bought according to the guidelines, what have we done wrong?” [IV15]
	Development of guidelines	Progress of development and modification of medical guidelines in response to AI applications in radiology	“Or you could think that the National Society of radiologists in the Netherlands would say “we recommend” - because in medicine it often goes like that, right, with guidelines. [...] There's nothing like that yet for this software, which is actually, I think, something that will be nice if that would evolve.” [IV4]
6C. Professional bodies	National body: NVvR	Role and position of Radiological Society of the Netherlands (NVvR) with regard to AI applications	"They [the NVvR] are setting up working groups that are focused on AI to look at it more from the national way. For instance, for identifying software that is really clinically applicable, meaning that you can use in the clinic. Because not that many applications are CE and FDA approved right now." [IV7]
	International professional bodies	Role and position of international professional organizations with regard to AI applications	“Well, if international societies have an opinion, then usually they can put more weight into the scale. Their influence is larger. And then usually, what you see is that the national, the local societies are using that kind of guidelines to create their own guidelines. " [IV3]
6D. Socio-cultural context	Ethical concerns	Perceived threats from use of AI applications in medical field to societal values and standards of conduct	"Ethische afwegingen zijn ook een reden waarom kunstmatige intelligentie met voorzichtigheid een rol in de zorg krijgt. Kunstmatige intelligentie is namelijk niet neutraal. Ieder ontwerp is waarden-geladen. Soms gebeurt dat expliciet, maar veel vaker is dat impliciet en worden de gevolgen van de morele keuzes pas zichtbaar bij de concrete toepassingen." (Meurs, 2019) *(Ethical considerations are also among the reasons why artificial intelligence is only*

			slowly becoming relevant in health care. In fact, artificial intelligence is not neutral, but incorporates values. Sometimes that is explicit, but more often these values get incorporated implicitly. In these cases, the consequences of the moral choices only become visible in concrete applications)
	Public opinion	Attitude of society towards use of AI applications in medical field	"Een vlotte acceptatie van kunstmatige intelligentie door arts en patiënt van een door computers gegenereerde diagnose valt nog te bezien" (Algra, 2017) *(A quick acceptation of artificial intelligence by doctors and patients, for example in the form of computer-generated diagnosis, still needs to manifest itself.)*
7. Future Outlook			
7A. Development over time	Expectations on future development	Anticipation of development and rate of adoption of AI applications in the future	"Most radiologists have been already working here for 20 or 25 years. Yeah, well, what can they expect for the next 10 years? They have heard the voices of computer assisted radiology already for 10 or 15 years. And nothing really changed in the clinical workflow." [IV11]
7B. Organizational Resilience	Recognizing critical events	Ability to identify problems and evaluate implementation process of AI applications within organization	"So we need to do an evaluation of the past year. And I think that will mainly focus on how the radiologists have experienced the technology, but also the referral physicians, the pediatricians. How they have experienced the use of boneXpert." [IV14]

Why, who, how & what: practical questions for implementation of AI applications in clinical radiology

Why: Determine the clinical or organizational problem to be targeted

- What is the desired clinical added-value?

Who: Identify relevant stakeholders

- Which radiologists should be involved (sub-group of radiologists or entire department)?
- Who are the relevant referring clinicians?
- Who is(are) the local champion(s)?
- Which members of management (from department level and hospital level) should be involved?
- Who are the innovation experts (e.g. innovation managers, innovation group, implementation specialists)?
- Which technical staff should be involved (e.g. medical physicists, IT department)?

How: involve all relevant stakeholders

- Which mechanisms allow for stakeholders to be included in the decision-making process (e.g. discussion & feedback rounds)?
- How often should stakeholders be consulted?

What: Design implementation strategy

- Knowledge
  - Which technical knowledge is needed for executing the implementation?
  - Where and how can the necessary technical knowledge for executing the implementation be acquired?
  - Which technical knowledge is needed for using the AI application?
  - Where and how can the necessary technical knowledge for using the AI application be acquired?
- Technical Elements:
  - Is the existing infrastructure sufficient for the integration of the AI application?
  - How can the AI application best be integrated in the existing IT infrastructure (i.e. PACS)?
  - How can the AI application best be integrated in the existing workflow?
  - How should the AI application be used in the clinical workflow (second-reader, first reader)?
  - Who is responsible for the technical elements of the implementation process?
- Validation & Monitoring
  - How can the AI application be validated locally?
  - How can the expected clinical added-value be measured?
  - How can the ‘correct’ use of the AI application be monitored?

Description:

When planning to adopt and implement a specific AI application in a department, the following four questions should be answered: why, who, how and what. The first question, *why*, should come at the very beginning of the process. It is fundamental to identify the clinical or organization problem, which is aimed to be tackled by the AI application. More specifically, the desired clinical added-value(s) should be singled out (e.g. saving time, improving diagnostic accuracy). The second question, *who*, helps to map the relevant stakeholders for the implementation of that particular application. It is important that not only the future users (i.e. radiologists), but also the indirect users (i.e. referring clinicians) as well as supporting staff (management, medical physicists, innovation specialists) are identified. Special attention should also be given to single out individuals, which could act as local champions. By determining the group of individuals involved in the implementation ex ante, mis-alignments and lacking communication across different units within the hospital can be prevented. The third question, *how,* refers to the nature of how the planning and execution of the implementation process is played out. It is essential to involve the identified stakeholders in strategy building and adoption process. This should be done through actively promoting discussion sessions and feedback rounds surrounding the adoption process. These activities should be repeated as the implementation advances over time. This collaborative approach enables the detection of potential reasons for opposition among adopters or other stakeholders and avoids mis-alignment. It can also be expected to increases the internal legitimacy of the project and allows for interested stakeholders to take ownership of the project. The fourth question, *what*, targets the elements that should be in the implementation strategy. Besides the more obvious elements, like workflow and PACS integration and funding, special attention should be given to the questions surrounding validation and monitoring and required knowledge. Answering these questions allows for the evaluation of the implementation process, including determining if the expected clinical added-value (the *why* question) is de facto materializing.