LUND UNIVERSITY
School of Economics and Management

Department of Informatics

Acceptance of AI technology for lung cancer screening diagnosis

Unpacking the factors affecting the radiologists' acceptance of AI

Master thesis 15 HEC, course INFM10 in Information Systems

Authors: Jan Niklas Harder
Joachim Felix Maßberg

Supervisor: Gemza Ademaj

Grading Teachers: Saonee Sarker
Osama Mansour

Acceptance of AI Technology for Lung Cancer Screening Diagnosis: Unpacking the Factors Affecting the Radiologists' Acceptancee of AI

AUTHORS: Jan Niklas Harder and Joachim Felix Maßberg

PUBLISHER: Department of Informatics, Lund School of Economics and Management, Lund University

PRESENTED: May, 2022

DOCUMENT TYPE: Master Thesis

FORMAL EXAMINER: Osama Mansour, PhD

NUMBER OF PAGES: 258

KEY WORDS: User Perspective, Health Information Technology, Technology Acceptance Model, Task-Technology Fit, Lung Cancer Screening, Artificial Intelligence

ABSTRACT (MAX. 200 WORDS): Lung cancer is the leading cause of cancer death. Much benefit can be found in national screening programs to diagnose cancer early and reduce mortality and such is about to be introduced in Germany. However, the process of screening diagnosis is very time-intensive for radiologists and the workload is not feasible without further support. Artificial Intelligence was assessed to be able to support radiologists and make lung cancer screening feasible. But many integrations of AI in healthcare fail due to lack of technology acceptance of physicians. Therefore, this research takes a qualitative approach to investigate the technology acceptance of radiologists. Based on the Technology Acceptance Model and the Task-Technology Fit model a framework was developed to investigate the influential factors for radiologists' acceptance of the AI. Five semi-structured interviews were conducted with radiologists who work with the AI for lung cancer screening in currently ongoing German trials and five factors were identified as influential: Perceived Usefulness, Perceived Ease of Use/Training, Locatability, Relationship with Users, and Result Demonstrability.

Acknowledgements

We must stress that this work would not have been possible without the support from our supervisor Gemza Ademaj. She provided valuable criticism, feedback and advice and guided us with this to keep us on the right path with our research and for this we would like to express our biggest heartfelt gratitude. Thank you so much.

Content

Figures

Tables

Abbreviations

AI	Artificial Intelligence
AIS	Association for Information Systems
CAD	Computer Aided Detection
CNN	Convolutional Neural Network
CT	Computer Tomography
DL	Deep Learning
E	Experience
HIT	Health Information Technologies
IS	Information System
IT	Information Technology
JR	Job Relevance
LDCT	Low-Dose Computed Tomography
ML	Machine Learning
NLST	National Lung Screening Trial
NN	Neural Networks
L	Locatability
OQ	Output Quality
PEU	Perceived Ease of Use
PEUT	Perceived Ease of Use/ Training
PT	Production Timeliness
PU	Perceived Usefulness
RD	Result Demonstrability
RU	Relationship with Users
SR	Systems Reliability
SV	Subjective Voluntariness
TAM	Technology Acceptance Model

TPC	Technology-to-Performance Chain
TTF	Task Technology Fit

1 Introduction

The idea of Artificial Intelligence (AI), a system that can perform human tasks better than humans themselves, has been around since the 1950s (Buchanan, 2005). The data-intensive nature of healthcare makes it one of the areas where AI applications are most promising (Rajkomar, Dean & Kohane, 2019; Yu, Beam & Kohane, 2018). However, research on AI integrations lacks, and when it is integrated it often fails to deliver the promised outcomes, and little benefit in clinical outcomes is observed (Brocklehurst et al., 2017; Nunes et al., 2017; Wolff et al., 2021).

Many problems in the healthcare sector, such as staff shortages, high patient numbers, or incorrect diagnoses, can be alleviated by the efficiency and high performance that AI solutions offer (Topol, 2019). Health Information Technologies (HIT) which use AI can support healthcare institutions in many routine tasks like bed management, patient care, or diagnosis of diseases (Kalis, Collier & Fu, 2018). But also complicated tasks like surgeries by robotic technologies or the research of new drugs by machine learning technologies can be solved by AI (Garbuio & Lin, 2019).

Lung cancer is the most dangerous cancer worldwide in terms of deaths (Tanoue et al., 2015). Often lung cancer is detected rather late such that most patients have only a few years to live after diagnosis (Tanoue et al., 2015). Therefore, early diagnosis is very important to be able to treat cancer at an early stage. In order to detect lung cancer at an early stage, preventive lung cancer screenings have been introduced in China and the United States to screen high-risk patients (European Lung Foundation, 2021). Studies are currently underway in several countries in Europe to investigate the effectiveness of lung cancer screening and whether it can be introduced on a large scale in these countries (Becker et al., 2020). Germany is also currently testing this introduction of such a lung cancer screening program in two trials, namely the HANSE study and the 4-IN-THE-LUNG-RUN study (BREATH, 2021; European Commission, 2021; Siegmund-Schultze, 2019).

Several studies in the United States and Europe have shown that regular screening can reduce lung cancer mortality in high-risk patients by 25% (de Koning et al., 2020). Lung cancer screening thus not only saves the lives of many people but also saves the healthcare system large amounts of money spent on expensive lung cancer therapies.

The screening involves taking multiple images of the patient's lungs in a low-dose computed tomography scan using an x-ray machine, which are then scanned by radiologists for specific disease patterns in order to make a diagnosis (de Koning et al., 2020). This work is very time-consuming for the radiologists due to the number of images and the number of patients invited for lung screening (Fischer et al., 2020; Liu et al., 2021). Almost 2.7 million people in Germany are estimated to be invited for lung cancer screening and it is expected that 50% of these people will participate in the program (Stang et al., 2015). Therefore, lung cancer screening programs must be tailored to the limited capacity of radiologists and healthcare institutions and be designed efficiently (Deutsches Krebsforschungszentrum, n.d.).

Diagnosis supported by a computer-aided detection system can increase performance and take less time (Awai et al., 2004; Brown et al., 2019; Ma et al., 2020). In recent years huge advances in the low-dose computed tomography (LDCT) diagnosis support have been made with AI, mostly by using convolutional neural networks (Ma et al., 2020). Existing academic literature has laid a good foundation by developing various helpful AI systems which can support lung cancer diagnosis in some way (Espinoza & Dong, 2020). However, it is unclear how these different systems can be used together to support diagnosis in an efficient way and the design of comprehensive end-to-end AI solutions lacks (Guo et al., 2020). Those systems need to be integrated into the daily workflows of practitioners to achieve meaningful use (Agarwal et al., 2010).

In Germany, lung cancer screening is currently at a critical point. Recently the German Institute for Quality and Efficiency in Health Care (2020) has confirmed the benefits of a LDCT screening for high-risk patients, and currently, it is being looked into implementation (Klein, 2021).

The workload for radiologists can be expected to increase significantly by this and with existing processes, this amount appears unfeasible (Fischer et al., 2020; Liu et al., 2021). Lung cancer screening using LDCT scans will create a workload that is not manageable for radiologists unless some part is automated or efficiency is increased (Nasrullah et al., 2019) and AI can solve the workload problem (Mathew, David & Mathew, 2020; Nasrullah et al., 2019).

Furthermore, AI also holds much potential in not only increasing efficiency but can also increase performance at the same time (Liu et al., 2019b). In the case of AI to support radiologists in the diagnosis for lung cancer screening it was estimated that it could reduce the workload by up to 86,7% (Lancaster et al., 2022). However, there is still a large gap between the benefits of AI in theory and implementation in practice (Seneviratne, Shah & Chu, 2020; Strohm et al., 2019; Wolff et al., 2021). In particular, the perspective of radiologists has often been neglected, making user resistance a crucial point for the failure of AI projects in radiology (Agarwal et al., 2010; Buck, Hennrich & Kauffmann, 2021; Kauffmann et al., 2022; Khanijahani et al., 2022).

Consequently, lung cancer screening is expected to bring many benefits and a nationwide implementation is expected soon. However, implementation requires more efficient processes in the diagnosis of lung cancer and computer-aided detection (CAD) using AI is found as one way that this can be achieved. However, to achieve the promised efficiency gains, these systems must be adopted and accepted successfully and this implementation often fails.

1.1 Problem

The success of lung cancer screening in Germany depends on the successful adoption and acceptance of AI systems, as the workload cannot be handled by radiologists alone (Klein, 2021). However, problems must be expected with AI acceptance, as AI applications in radiology have rarely made it into widespread use in the past (Strohm et al., 2019). Generally, a large gap between the extensive AI research and the lack of successful integration in healthcare institutions exists (Seneviratne, Shah & Chu, 2020; Strohm et al., 2019; Wolff et al., 2021). Existing research mostly focused on the development and validation of AI algorithms and the radiologists' perspective has been mostly neglected (Buck, Hennrich &

Kauffmann, 2021; Kauffmann et al., 2022), leading to a lack of consideration of the concerns of radiologists in the development of AI software (Kauffmann et al., 2022). Without taking into account the needs of radiologists, there may be high user resistance to AI use, which is one of the main reasons for the failure of technology implementations (Agarwal et al., 2010; Davis, 1989; Goldfarb & Teodoridi, 2022; Jorritsma, Cnossen & van Ooijen, 2015; Khanijahani et al., 2022; Park et al., 2020; Reis et al., 2020; Strohm et al., 2019; van der Aalst, Ten Haaf & de Koning, 2021). The reasons for the user resistance are many. A bad fit between technology and task (Goodhue & Thompson, 1995) or lacking consideration of how the technology can be integrated into the radiologists' workflows can cause user resistance (Hwang & Park, 2020). The disruptive nature of AI applications further contributes to user resistance (Reis et al., 2020). The understanding of which factors are responsible for user resistance and how this can be counteracted is not yet sufficiently developed in information systems (IS) research (Kauffmann et al., 2022). In addition, the lung cancer screening trials in Germany only started last year, so there are no scientific publications on this yet (BREATH, 2021; European Commission, 2021).

1.2 Purpose

The importance of appropriate AI integration which considers the user's needs has been highlighted (Al Mohammad et al., 2019; Buck, Hennrich & Kauffmann, 2021; Lee et al., 2017; Nasrullah et al., 2019). Furthermore, various challenges and concerns in that process have been hinted at by the IS literature (Huisman et al., 2021; Maier, Jussupow & Heinzl, 2019; Pumplun et al., 2021; Yin, Ngiam & Teo, 2020), but these depend on various factors which cannot be generalized for different cases and therefore, more research is needed (Buck, Hennrich & Kauffmann, 2021). These findings are supported by medical research (Gaube et al., 2021; Laï, Brian & Mamzer, 2020) and also appear to be applicable to the integration of AI for lung cancer screening (Guo et al., 2020; Hwang & Park, 2020; Mathew, David & Mathew, 2020; Nasrullah et al., 2019). Therefore, an objective of this thesis is to take a closer look at the user acceptance of AI systems for lung cancer screening. Using factors from the relevant technology acceptance models from the IS literature, a qualitative data research approach is used to investigate the most influential factors of the radiologist's acceptance of AI for lung cancer screening diagnosis.

Based on this, we formulate the following research question for this thesis:

What are the most influential factors affecting the radiologist's acceptance of AI supporting lung cancer screening diagnosis?

The research purpose is to develop an understanding of the most impactful factors which affect radiologist's technology acceptance for lung cancer screening. This understanding contributes to the academic knowledge about the practitioner's perspective, needs and acceptance of AI technology and helps to close the gap between AI research and the integration of AI in a healthcare setting.

From this, system design decisions and technology selection criteria can be derived to achieve better user acceptance and also the way is leveled for more AI integration projects in healthcare which are accepted by practitioners.

1.3 Delimitation

We concentrate on the technology acceptance of radiologists working with AI for lung cancer screening in scope of a trial in Germany. The IS literature has highlighted country specific technology acceptance factors and as this research is motivated by a lung cancer screening implementation in the near future in Germany, focusing on locations which are in Germany was considered most appropriate.

The trials have been running since August 2021 and this research only takes into account their current technology acceptance in the trial, so the radiologist's perspective in the trial pre- or post-integration or outside of a trial is not looked into.

Existing IS research has identified various influential factors for user acceptance and has developed several models to present relationships between these factors. This study has focused on the Technology Acceptance Model and the Task-Technology Fit Model. Furthermore, due to time and scope restraints of this research, the factors of both models were limited and integrated into a selection of relevant factors which were then used to guide the research.

2 Theoretical Background

2.1 AI in Healthcare

Since the 1950s, there has been the idea of artificial intelligence, systems that surpass human capabilities in human tasks (Buchanan, 2005). John McCarthy, a pioneer of AI, coined the term Artificial Intelligence as “the science and engineering of making intelligent machines” (Hamet & Tremblay, 2017, p.36). While AI was initially only able to solve simple tasks, today AI can also solve very complex problems due to the significantly higher computing power and new advanced algorithms (Lee, 2020). Moreover, due to ever-increasing amounts of data and data mining and processing technologies, problems can be solved by AI that would be unsolvable for humans due to the large amounts of data, creating many new opportunities (Hamet & Tremblay, 2017).

In order to examine the use of artificial intelligence technologies in the healthcare sector, an updated definition of AI must first be provided. The High-Level Expert Group on Artificial Intelligence of the European Commission defines AI as "systems that display intelligent behavior by analyzing their environment and taking actions - with some degree of autonomy - to achieve specific goals" (Pekka et al., 2018, p.1). With this definition, Pekka et al. (2018) emphasize the analysis of the environment, in addition to the autonomous derivation of actions, as necessary for successful AI systems. Russell and Norvig (2020) have elaborated as commonalities among the various definitions of AI the motives to think and act like humans and to make rational decisions.

One of the most important areas of AI is Machine Learning (ML) (Yu, Beam, & Kohane, 2018). The key advantage of machine learning is that it does not require the design of predefined rules for the input data, but that ML can autonomously find patterns in the data (Deo, 2015; Murphy, 2012; Yu, Beam, & Kohane, 2018). This not only saves time but also finds new correlations that would have been overlooked by fixed rules. To analyze correlations in the data, deep learning can be used, which uses multilayer neural networks to examine very large data sets for similar features. Deep learning can find these features without first labeling the data. This process is therefore called unsupervised learning (Deo, 2015). This allows large datasets to be used, as labeling the data would otherwise take a lot of time, as well as requires expert knowledge depending on the type of data. These ML models can then be trained using labeled data (supervised learning) and used for predictions (Yu, Beam, & Kohane, 2018).

Healthcare institutions can use these AI technologies in many areas to make their processes more efficient and increase the quality of medical treatments (Raghupathi & Raghupathi, 2014; Topol, 2019). McKinsey & Company and EIT Health (2020) name the use of AI as the way to address the growing healthcare challenges of the near future, such as the rapidly aging population or increasing demands, while keeping processes sustainable. Already today, AI is making a significant contribution to improving healthcare with numerous e-health applications that, for example, reduce information asymmetries between doctors, healthcare institutions, and patients through digital solutions (Garbuio & Lin, 2019). In particular, administrative tasks such as arranging appointments, prescribing medications, or automatically reordering supplies or drugs are already being handled by AI, allowing doctors more time to actually treat patients (Kalis, Collier & Fu, 2018). In some areas, deep learning approaches are already

being used successfully in diagnostics and early detection of diseases (Kassania et al., 2021). Further potential can be seen in clinical decision-support tools, triage tools or robots helping assist tasks such as bringing medical equipment (Davenport & Kalakota, 2019; McKinsey & Company & EIT Health, 2020).

The healthcare sector is also particularly well suited for the application of Artificial Intelligence, as enormous amounts of data are available through patient records, X-ray images, or sensors (Topol, 2019). This provides sufficient data for training a wide variety of ML models. At the same time, such large amounts of data also require the automation of processes, as they can no longer be processed manually.

Wolff et al. (2021) found that there is still a large gap between the extensive AI research and the lack of successful implementation in healthcare. This is not mainly due to a lack of technological progress but to structural problems and concerns. One challenge is that in everyday clinical practice, a lot of unstructured data is still generated, such as findings in free text form, which can only be used for training AI algorithms at great expense (Jiang et al., 2017). In addition, databases must first be built that can store large amounts of data in a structured way (Panch, Mattie & Celi, 2019). Ethical concerns also play a major role in the introduction of AI in healthcare (Bærøe, Jansen & Kerasidou, 2020; Chen & Verghese, 2020; Morley et al., 2020; Parikh, Teeple & Navathe, 2019; World Health Organization, 2021). AI algorithms may contain biases that favor certain populations or genders or have better accuracy for certain groups when trained on biased data (Morley et al., 2020; Parikh, Teeple & Navathe, 2019). It is also important to ensure that patients' privacy is protected and that their data is not accessible to third parties. Therefore, it must be ensured that AI decisions are transparent, traceable, and understandable (Lockey et al., 2021; Reis et al., 2020). In addition, the healthcare sector is a highly regulated market with extensive security and privacy laws that make it difficult to adopt AI applications, as they must go through lengthy approval processes and data processing requires consent from affected patients (Jiang et al., 2017).

In Germany, AI is already being tested in the healthcare sector in many areas, but it has not yet been implemented across the board (Krumm, & Dwertmann, 2019; Mumm et al., 2021). However, the Covid-19 pandemic has accelerated this process, as the potential of digital solutions has become even more evident (Mumm et al., 2021). Krumm and Dwertmann (2019) predicted in 2019 that in the next 1-5 years, deep learning technologies will be deployed in Germany that can perform diagnoses by analyzing image data. Even robots that perform surgeries using AI are expected by Krumm and Dwertmann (2019) to be introduced in Germany in the next 5-10 years. Radiologists are among the first in Germany to use AI in everyday medical practice (Krumm, & Dwertmann, 2019; Mumm et al., 2021). For instance, AI solutions help them by automatically detecting abnormalities in images, which the radiologists then evaluate (Krumm, & Dwertmann, 2019). The use of AI in German healthcare facilities often fails due to the strict approval criteria for medical products (Dünnebeil et al., 2012). The process is particularly time-consuming for deep learning algorithms, as they have to be continuously developed and trained on new data, which means that a single approval for the release version of the product is not sufficient (Krumm, & Dwertmann, 2019). The approval processes in addition to the corresponding studies are very time-consuming and cost-intensive, so not every software vendor can afford them, which slows down the implementation (Hufnagl, 2020). Strict data protection laws in Germany also make it difficult to obtain approval and find suitable data sets for training AI (Eitle & Buxmann, 2020; Krumm, & Dwertmann, 2019). The German Ethics Council is already calling for an adjustment of data protection laws for the use of AI, as current laws are not designed for the collection of patient data

on a large scale in the context of Big Data and for the evaluation of this data by AI (Deutscher Ethikrat, 2017). Another problem in Germany is the reluctant acceptance of AI applications in medicine among the population. Only 61% are satisfied with a diagnosis made by a doctor with the help of diagnostic software. Furthermore, only 12% would trust a diagnosis made by an AI system alone (Krumm & Dwertmann, 2019).

2.1.1 Lung Cancer Screening

Lung cancer is the world's most dangerous type of cancer in terms of deaths (Tanoue et al., 2015). Only 18% of people in the U.S. diagnosed with lung cancer live more than five years after their diagnosis. This makes lung cancer the third leading cause of death in the United States (Tanoue et al., 2015). Medical treatments for lung cancer patients resulted in an estimated direct cost of $12.1 billion to the healthcare system in the United States in 2010 (Wender et al., 2013). A study from Germany, where about 500,000 people develop cancer each year, showed that 37.4% of cancers are preventable in age groups between 35 and 84 (Siegmund-Schultze, 2019). Furthermore, with the ageing population, cancer cases are expected to continue to increase. Currently, lung cancer is often diagnosed at the advanced stage, when treatment options are minimal. These patients then usually have a life expectancy of only one to two years (Tanoue et al., 2015).

Therefore, in order to counteract the consequences of lung cancer, early detection is very important. However, examinations for lung cancer usually only take place in the case of typical symptoms, and if a doctor prescribes it. For other types of cancer, large-scale preventive examinations are already organized in some countries. For example, preventive examination programs for breast and cervical cancer are in place in many countries, such as Germany and the USA (Weller et al., 2009; World Health Organization, 2020a; World Health Organization, 2020b). Such preventive examinations are called screenings and are carried out for certain groups of patients either once or at regular intervals (Morabia & Zhang, 2004). The tests performed do not have to be used as a final diagnosis, but as an indication to refer a patient to an appropriate specialist (Morabia & Zhang, 2004). The US Preventive Services Task Force (2021) recommends lung cancer screening for people aged 50 to 80 years who are smokers or have quit smoking in the past 15 years. Smokers include those with a 20 pack-year history, where one pack is the equivalent of 20 cigarettes per day per year (US Preventive Services Task Force, 2021). Research agrees that the most promising technique for implementation is low-dose computed tomography, also called a low-dose CT scan, which is reflected in the recommendations of relevant organizations such as the US Preventive Services Task Force (Wender et al., 2013). For the screenings, an X-ray machine is used to produce images of the lungs, which are examined by radiologists who scan these images for patterns of disease they know of and then make a diagnosis.

Studies from several countries show that lung cancer screenings can actually reduce the risks of smokers dying from lung cancer. A study examining lung cancer screening trials in Belgium and the Netherlands (NELSON) concluded that the risk of death from lung cancer for high-risk patients decreased by 25% (de Koning et al., 2020). Trials in Germany, Denmark, and Italy have also been studied and positive results in lung cancer mortality have been reported (Becker et al., 2020). Becker et al. (2020) also found a 20% decrease in lung cancer mortality for patients who underwent LDCT screening instead of conventional X-ray

screening during the National Lung Screening Trial (NLST) in the United States. The decreased mortality from lung cancer is confirmed by other studies and their evaluation by other researchers (Aberle, Abtin & Brown, 2013; Bundesamt für Strahlenschutz, 2021; Duffy & Field, 2020; US Preventive Services Task Force, 2021). Another advantage of screening is that the CT images of patients can be examined for other diseases such as coronary artery disease without the patient having to be examined more often, which increases the efficiency (Hecht et al., 2014). In addition, it reaches many smokers who can be encouraged to participate in smoking cessation programs, which is an important part of lung cancer screening, as it can detect lung cancer early but not prevent it (US Preventive Services Task Force, 2021).

For lung cancer screening to be approved, it must be possible to prove that the medical benefits outweigh the risks and that the costs are commensurate with the benefits. The German Institute for Quality and Efficiency in Health Care (2020) has investigated this and concluded that LDCT screening for people with increased risk reduces mortality from lung cancer and the side effects are low enough not to outweigh this benefit. Side effects include irradiation of patients and overdosing of patients who have been falsely diagnosed with lung cancer (false positives) and are subjected to treatments that are not necessary (Institute for Quality and Efficiency in Health Care, 2020).

2.1.2 Lung Cancer Screening with LDCT

Most benefits of lung cancer screening can be found in the early detection of lung cancer and in the past few years, many developments have enabled new or improved ways to detect and catch lung cancer at an early stage (Mathew, David & Mathew, 2020).

One of the most promising innovations and feasible implementations comes from the detection of lung nodules in LDCT images (Mathew, David & Mathew, 2020; Nasrullah et al., 2019). A significant number of lung nodules that are found represent early and curable forms of lung cancer (Cruickshank, Stieler & Ameer, 2019). However, depending upon slice thickness, a single CT scan can consist of up to 500 images (Bogoni et al., 2012). Therefore, searching through a CT scan for the existence of a lung nodule means an immense workload for a radiologist (Bogoni et al., 2012). Consequently, implementing lung cancer screening and repeating this process for a significant share of the population is an unfeasible task unless radiologists' processes become more efficient or automated in some way (Lancaster et al., 2022; Nasrullah et al., 2019). Mathew, David, and Mathew (2020) point out how well AI can support lung nodule identification and that it enables radiologists to efficiently manage the new data load. They explain that AI can surpass the performance of radiologists, however, this does not mean that AI can replace radiologists, which is hardly feasible due to task complexity. Instead, the most practical and efficient system is reached, when AI is used to support radiologists.

The process of analyzing a CT scan to detect lung cancer consists of two core steps (Espinoza & Dong, 2020; Guo et al., 2020). First, the CT scans must be looked through for the existence of lung nodules and secondly, the long nodules must be judged on their malignancy. Both steps can be supported by AI.

2.1.2.1 Lung Nodule Detection

As highlighted, the examination consists of the radiologists quickly scanning through hundreds of images. To catch cancer early it is important to detect small nodules, however, those can also easily be missed by radiologists (Becker et al., 2012; Fischer et al., 2020). This makes AI a powerful asset to detect small abnormalities and enable cancer to be detected very early (Becker et al., 2012; Fischer et al., 2020). In the past few years, many CAD techniques have been shown to improve both work performance and efficiency of radiologists in the detection of lung nodules (Goncalves, Fong & Blokhina, 2022; Guo et al., 2020; Svoboda, 2020). The techniques involved mostly ML with hand-crafted features (Guo et al., 2020) which comes with some risks (Lebovitz, Levina & Lifshitz-Assaf, 2021). Data labels are assigned by experts, but their opinions can vary. This also leads to high variance in model accuracy (Guo et al., 2020; Lebovitz, Levina & Lifshitz-Assaf, 2021), and furthermore, even though accuracy on a test set was high, models can disappoint in practice (Lebovitz, Levina & Lifshitz-Assaf, 2021). In recent years neural networks have been popularised which are mapping inputs to outputs and by this, are able to bypass a conceptualization of an expert's tacit knowledge (Lebovitz, Levina & Lifshitz-Assaf, 2021). Various studies have looked into those AI systems and their performance in detecting lung nodules, but various issues to compare studies exist. First, how well long nodules are detected depends on various factors as image resolution (Liu et al., 2021) and slice thickness (Guo et al., 2020). Furthermore, size of the nodules and presence of various forms and their location in the lung affect how well lung nodules are detected (Liu et al., 2021; Setio et al., 2017). Therefore, reported lung nodule detection accuracy varies for different combinations of those factors and size of the nodules from about 50% to 99% (Guo et al., 2020; Liu et al, 2019a; 2021; Pinsky et al., 2013). Setio et al. (2017) recognized the need for a comparable study and set up a challenge to compare the performance of various AI algorithms on a large representable data set. Sensitivity of AI systems varied between 31.8% to 92,9% and the top three AI systems made use of CNN. Furthermore, a combination of different deep learning systems was able to outperform any individual system with a sensitivity of 95% and a relatively low rate of false-positives. Lastly, they also found that the AI systems were finding nodules which were missed even by expert radiologists, emphasizing the value of AI to support lung nodule detection.

2.1.2.2 Lung Nodule Classification

While most lung nodules are benign, a small but considerable number are malign and do indicate lung cancer (Cruickshank, Stieler & Ameer, 2019). Accurate classification is very desirable as false-negative classifications defeat the purpose of detecting lung cancer early on and false positives may lead to unnecessary, expensive, and physically harmful treatments (Institute for Quality and Efficiency in Health Care, 2020). Therefore, the classification of benign/malignant lung nodules can be considered the most crucial step in the early detection of lung cancer (Xie, Zhang & Xia, 2019). Many ML techniques are well suited for image classification and various have been tried out and evaluated for lung cancer screening.

Although validation of malignancy is difficult as it can often only be known for sure after operation (Cruickshank, Stieler & Ameer, 2019) and therefore, data sets are often limited, the classification is more straightforward and reported performances do not vary as much. Xie, Zhang and Xia (2019) used semi-supervised adversarial classification on LDCT scans and reached an accuracy of 92,53% with a specificity of 96,28%. This accuracy and specificity was matched by further studies using various neural networks (Bonavita et al., 2020; Guo et al., 2020; Mathew, David and Mathew, 2020) and can be considered a satisfactory performance in medicine (Guo et al., 2020).

2.1.2.3 Comprehensive AI Solutions

Lung nodule detection and classification are the core two steps in the detection of lung cancer and what most studies have focused on (Guo et al., 2020). However, various concepts and ideas of different AI models which address further important aspects in and around those steps exist and have been developed (Espinoza & Dong, 2020). In addition to the core functions described above, this includes also narrowing down a smaller high-risk group with a higher likelihood of having cancer (Huang et al., 2019), estimating the likelihood of developing lung cancer (Huang et al., 2019), detecting regions of interest in CT scans (Ardila et al., 2019) and reducing the false-positive rate by integration of multilevel contextual information (Dou et al., 2017). While all solutions can generate additional value to some extent, it remains unclear how a comprehensive end-to-end solution can be designed which integrates and combines different AI models along the way to support the radiologists optimally (Ma et al., 2020). Furthermore, an integration of various lung cancer indicators into the AI models analyzing LDCT images can provide further benefits. Shaffie et al. (2021) found that the integration of breath bio-markers into an image-based CAD system resulted in superior performance. Even a combination and integration of different AI models with similar functionality can result in systems which are superior to individual systems (Setio et al., 2017) and consequently, good individual systems for lung cancer detection exist but currently much potential is unused due to lack of comprehensive solutions (Ma et al., 2020).

2.1.3 Lung Cancer Screening Status

While many studies and research institutions recommend the introduction of lung cancer screenings, they are currently only performed nationwide in China and the USA at no cost to patients (European Lung Foundation, 2021). However, trials are also underway in many European countries for the widespread implementation of lung cancer screening (Becker et al., 2020). The government in Germany plans to improve early diagnostics and cancer prevention (Siegmund-Schultze, 2019) and has positive feedback from the Institute for Quality and Efficiency in Health Care (2020) that lung cancer screenings contribute positively to the health of high-risk patients. In addition, Germany already has experience with other cancer screening programs (BREATH, 2021; World Health Organization, 2020a; World Health Organization, 2020b), so expertise is already available for implementation. Currently, the Federal Joint Committee of German public health agencies is reviewing the approval of lung cancer screening (Klein, 2021). Subsequently, this could be introduced in Germany. In northern Germany, the HANSE study, a test program with 5000 participants is already underway to demonstrate the implementation of nationwide lung cancer screening in Germany (BREATH, 2021). In parallel, the German Cancer Research Center and the University Hospital Essen are participating in the Lung Cancer Trial of the research project 4-IN THE LUNG RUN of the European Union (European Commission, 2021). The 4-IN THE LUNG RUN study (2021) will also target 5000 patients in Germany and 24000 patients across Europe. Stang et al. (2015) calculated that based on the criteria for participants in the NLST study, nearly 2.7 million people in Germany would be invited for lung cancer screenings. The criteria for the NLST study that could be adopted in Germany were to be a current smoker or to have quit smoking in the last 15 years, to have a pack-year history of 30 years, which means smoking an average of one pack of cigarettes per day, and to be between 55 and 74 years of age (Pinsky & Berg, 2012). With an expected participation rate of 50%, this would be 1.35 million participants, who would all visit a radiology facility annually for lung cancer screening (Stang et al., 2015). This high number of attendees once again highlights the importance of effective processes for radiologists as well as support from computer systems. Klein (2021) emphasizes that for the first

implementation tests should be performed in lung centers that have been certified by the German Cancer Society to guarantee standardized processes and software.

Tang et al. (2018) studied the performance of an ML model trained on data from the U.S. National Lung Screening Trials and then used in a lung cancer trial at Oxford University Hospitals. They found that despite high performance on the training data, the ML model performed less well on patients in England. This indicates that ML models need to be adapted for use in other environments, as both the patients and the use of AI applications in healthcare institutions differ. This emphasizes the need to conduct research in Germany to achieve relevance for the German lung cancer screening implementation.

2.2 User Resistance in HIT

Although AI promises various benefits for lung cancer screenings and the healthcare sector in general, previous research has observed the failure of AI deployment in healthcare institutions (Abouzahra, 2011; Lee & Yoon, 2021). Only 53% of planned AI projects are completed and used afterward in practice (Gartner, 2020). Kitzmiller, Hunt and Sproat (2006) estimate that 80% of healthcare IT projects fail and that almost all projects require twice their original budget and fail to meet their schedule. The reasons for this failure are many. Projects in the healthcare sector face some risks and challenges due to strict regulations (Jiang et al., 2017). But especially in the field of artificial intelligence, these challenges become even more complex.

An important factor that researchers see as a major barrier to the adoption and success of AI applications is user resistance (Agarwal et al., 2010; Goldfarb & Teodoridi, 2022; Khanijahani et al., 2022; Park et al., 2020; Reis et al., 2020; Strohm et al., 2019). Managing this user resistance is especially important since clinical applications are critical to patient health (Topol, 2019). Reis et al. (2020) attribute the user resistance mainly to the disruptive character of AI applications, as they can fundamentally change the processes and routines of practitioners. They, therefore, observe a general mistrust of practitioners in these AI systems. Lockey et al. (2021) define different AI trust challenges that need to be overcome, like transparency and explainability as well as accuracy and reliability. Radiologists' concerns about AI have not yet been well researched and are therefore not usually taken into account in the development of AI systems (Buck, Hennrich & Kauffmann, 2021; Kauffmann et al., 2022; Strohm et al., 2019). Radiologists are considered pioneers of new technologies among physicians because their work has long been supported by computer systems (Thrall et al., 2018). Nevertheless, it is important that AI applications are easy to use and learn in order to avoid errors in medical decisions (Park et al., 2020).

For healthcare practitioners, the introduction of AI systems primarily means additional work, although they are supposed to have the opposite effect (Buck, Hennrich & Kauffmann, 2021; van der Aalst, Ten Haaf & de Koning, 2021; Yu, Beam & Kohane, 2018). Buck, Hennrich, and Kauffmann (2021) also list the loss of control, additional effort, job loss, loss of autonomy, and unclear responsibilities as additional reasons for practitioners' aversion to AI systems. The loss of control can be attributed to non-transparent AI systems that present medical decisions to physicians without explaining to them in an understandable way the criteria according to which the AI system made this decision (Reis et al., 2020). In addition, Reis et al. (2020) found that physicians found AI tools obstructive rather than supportive because they

interfered with their work without being able to communicate with them to understand or change their diagnoses.

To address these issues, practitioners' workflows in particular need to be examined in order to tailor AI applications to them, so that practitioners see them as a help rather than an intrusion into their workflow (Agarwal et al., 2010; Yu, Beam & Kohane 2018). Otherwise, the benefits of AI systems cannot be exhausted. Yu, Beam, and Kohane (2018) noted that integrating AI systems into practioners workflows is not easy, for example, due to the widely divergent processes and information systems. Pumplun et al. (2021) also emphasize the importance of integrating ML into clinical workflows, noting that clinics often fail in this task and that more research is needed on this topic to overcome this. These findings are supported by Huisman et al. (2021), who also identify workflow integration as one of the barriers to AI applications in radiology. Shibl, Lawley and Debuse (2013) found that users will not use a new technology if its adoption requires major changes to their workflows. The integration of AI systems into users' workflows is therefore an important factor for the acceptance of new technologies (Khanijahani et al., 2022; Shibl, Lawley & Debuse, 2013). Especially in the healthcare sector, AI applications need to be thoroughly adapted to the complex workflows of physicians in order to use them efficiently and reduce error rates (Hosny et al., 2018; Park et al., 2020). By adapting AI applications to routines and workflows, additional steps for radiologists while using the AI application can be reduced, thus eliminating additional work and making the application more user-friendly (Strohm et al., 2020).

There is also a need to understand the needs of radiologists, in which situations they find support from AI systems helpful, and in which situations they find it obstructive. Lebovitz, Levina, and Lifshitz-Assaf (2021) suggest that in areas that require physicians' expert knowledge and are even less researched, physicians should continue to be used as decision-makers, and in areas that are very well researched, supportive AI should be increasingly used that meets high-performance requirements and is designed together with physicians and adapted to their knowledge. Nishikawa and Bae (2018) name adjusting the AI Systems to the radiologists' behavior as an important factor for the success of clinical AI Systems. Furthermore, radiologists should be included in the training and evaluation process of AI systems to further increase radiologists' confidence in these systems and thereby decrease their resistance to using them (Rubin, 2019).

2.3 Frameworks

To develop our understanding of influential factors and also to structure our research two relevant frameworks from the IS literature were selected. The selection process is described in Section 3.3.

Core concern of our research is technology acceptance and therefore first, an appropriate framework of technology acceptance was selected. Various technology acceptance models exist in the IS literature which have their origins in different academic fields and also contain vastly different assumptions about the environment in which the system is deployed (Taherdoost, 2018). The models are applicable to the early stages of an IS and are therefore appropriate to our research as we are investigating a novel technology that has not seen much use before (Lai, 2017).

The Technology Acceptance Model (TAM) is one of the most frequently used acceptance models in IS research (Liao, Palvia & Chen, 2009; Marangunić & Granić, 2015). It offers multiple dimensions and factors to analyze the acceptance of new technologies and was found to be a good fit to our research. The TAM focuses mainly on the perceived ease of use and the perceived usefulness of a system, to predict its acceptance. We will discuss the TAM further in Section 2.3.1.

Similar to most other technology acceptance models, cognitive processes and emotions are at the core of the TAM and mostly attitudes, beliefs and perceptions are used as predictors of technology acceptance (Taherdoost, 2018). Therefore, the Task Technology Fit model (TTF), which focuses on characteristics of task and technology, is used in addition as it complements the TAM well (Dishaw & Strong, 1999; Goodhue & Thompson, 1995) and IS research has also proven that TAM and TTF can be integrated to create an even more comprehensive acceptance model (Dishaw & Strong, 1999; Klopping & McKinney, 2004; Pagani, 2006). The TTF model is briefly described in Section 2.3.2.

Section 2.3.3 explains an integrated model of TAM and TTF and the theoretical framework behind it. It shows how the factors of the TAM and TTF model relate and enables a more comprehensive understanding of the factors.

In Section 2.3.4, the relevance of the factors of the TAM and TTF model are discussed to restrict the scope of the research to allow enough time for factors that are deemed more relevant to be investigated in the interviews. In Section 2.3.5 the relevant factors are then integrated into our research framework.

2.3.1 Technology Acceptance Model

The Technology Acceptance Model was developed in 1986 by Fred Davis inspired by the theory of reasoned action and the theory of planned behavior (Davis, 1989). The purpose of the TAM is to investigate the factors that influence the acceptance of users of the technology (Marangunić & Granić, 2015). This makes the TAM applicable even before extensive system use and especially useful for systems in an early development stage (Davis, 1989; Liao, Palvia & Chen, 2009). Therefore it can be used to estimate the acceptance and the success of a technology before it is introduced on the market.

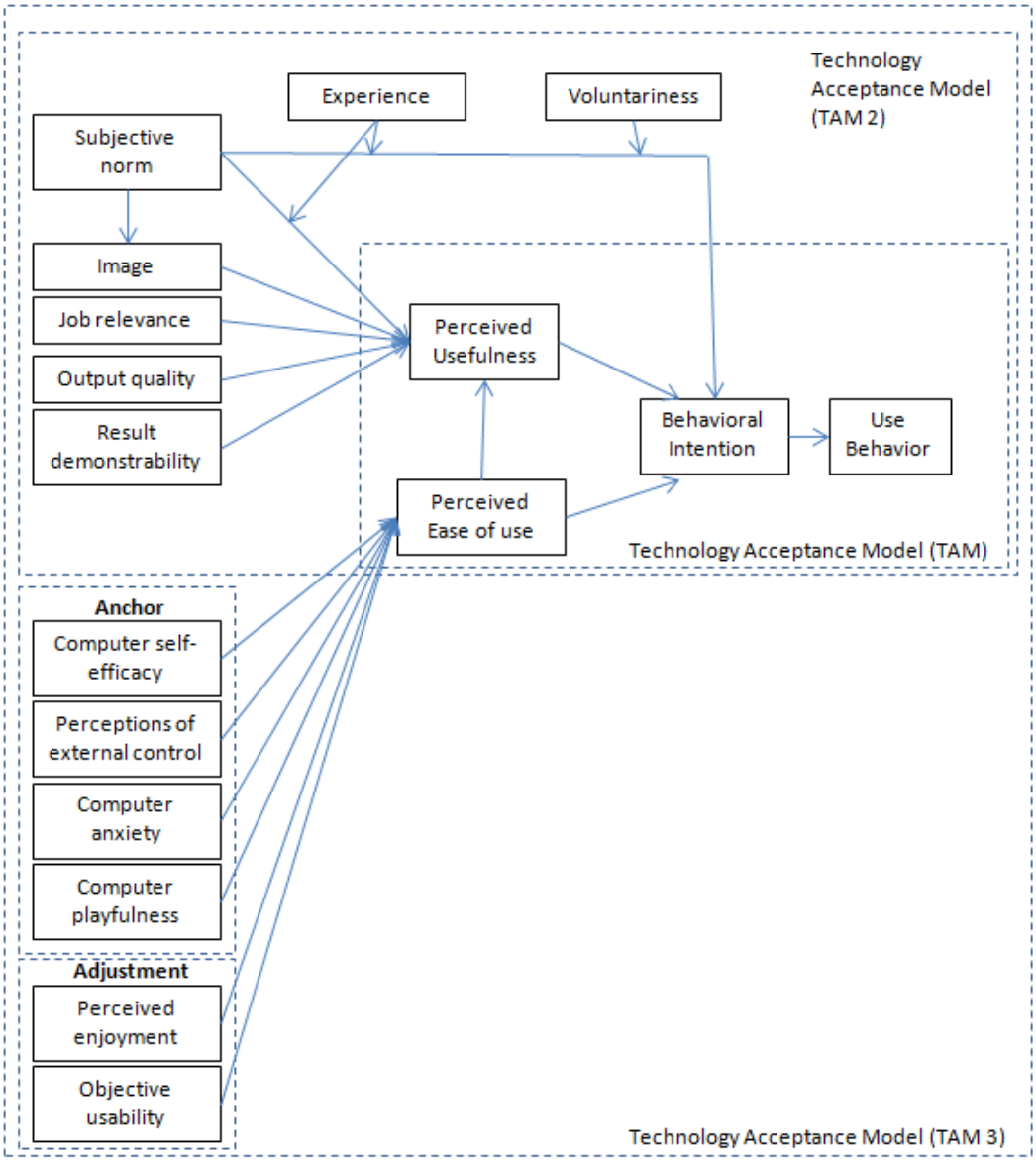

Figure 1: Technology acceptance models (TAM, TAM2 vs. TAM3) (Boughzala, 2014, p.169)

In the first version of the TAM, the Perceived Ease of Use and the Perceived Usefulness, as well as influences of external factors on these, are examined in order to analyze the actual use of the user and thus its acceptance opposite the technology (Davis, 1989). Figure 1 shows these factors included in the model as well as their relationships represented by the arrows. According to Davis (1989), the intention to use a system is mainly influenced by Perceived Ease of Use and Perceived Usefulness. Thus, if a user does not understand the purpose of the system or has difficulty using it, the user is likely to reject it and not use it. Therefore, Behavioral Intention also directly influences the Actual Use of the system (Davis, 1989).

Venkatesh and Davis (2000) presented the TAM2, a revised version of the TAM that more closely examines the factors that influence the Perceived Usefulness of the system, as Perceived Usefulness significantly influences acceptance of the system (Marangunić & Granić, 2015). These factors are Subjective Norm, Image, Job Relevance, Output Quality, Result, and Demonstrability (Venkatesh & Davis, 2000). In addition, two moderating factors were introduced that influence the Subjective Norm, namely Experience and Voluntariness (Marangunić

& Granić, 2015). With these additional factors, Venkatesh and Davis (2000) highlight the influence of social factors such as Image or Subjective Norm for the acceptance of an information system. This is because the expectations of others, such as supervisors, can influence the Perceived Usefulness of the system for the user, according to Venkatesh and Davis (2000). In addition to social factors, they also highlight cognitive instrumental factors such as the relevance of the information system to the job or the quality of the output. A system may be well suited for a task, but if the quality is not sufficient it will still not be perceived as useful, or a system with a high-quality output without relevance to the job is not useful either.

Venkatesh and Bala (2008) developed the TAM3 as a further improvement of the TAM, which builds on and combines the TAM2 and the model of the determinants of perceived ease of use by Venkatesh (2000). They introduced new factors that influence the Perceived Ease of Use and explored further interrelationships between the different factors. In particular, they introduced factors that describe users' feelings when using information systems. These include Computer Self-efficacy, Perceptions of External Control, Computer Anxiety, and Computer Playfulness. According to Venkatesh and Bala (2008), these factors are significantly influenced by user experience, as users can then, for example, reduce anxiety or improve perceptions of external control, i.e. the belief that resources exist within the company that are helpful in using the information system (Venkatesh et al., 2003). In addition, Venkatesh and Bala (2008) introduced two system-specific factors. Objective Usability, the measured effort required to use the system, and Perceived Enjoyment, the pleasure derived from using the system regardless of the quality of the results. These factors are especially important after prolonged use of the system, as increased experience has then reduced biases or problems in using the system and objective values become more important (Venkatesh, 2000). The Objective Usability also represents a major change in the TAM, since previously only subjective values were included in the determination of system acceptance, since only the perception of the users was analyzed, but not the system itself (Taherdoost, 2018).

The TAM is one of the most widely used models in IS research to better understand the acceptance of information systems (Liao, Palvia & Chen, 2009; Marangunić & Granić, 2015). However, the TAM also has some limitations. Lee, Kozar, and Larsen (2003) criticize that only the subjective opinion of the users is taken into account but not the actual quality of the system. This has been improved somewhat by TAM3, but the model still uses almost exclusively subjective factors. In addition, the lack of actionable guidance to practitioners on how to encourage users to better accept and use the systems is criticized (Faqih & Jaradat, 2015). Nevertheless, the Technology Acceptance Model seems to be promising for our research, as it provides us with many different relevant factors on which we can develop appropriate questions regarding user resistance and integration of AI systems. Current research has shown that in studies of health IT system acceptance, the TAM model has almost always been adapted and additional factors were added to fit the socio-technical context of healthcare facilities (Harst, Lantzsch, & Scheibe, 2019, Heinsch et al. 2021). The TAM model should therefore not be understood as a final framework to be applied but as a basis for research on the acceptance of information systems, which also made it suitable for this research.

The TAM, as well as the TAM2 and TAM3, were originally designed for quantitative research methods in the form of questionnaires (Davis, 1989; Venkatesh & Davis, 2000). However, this only allows researchers to determine which factors have a high influence on technology acceptance and not exactly why technology acceptance is low or how this could be improved (Vogelsang, Steinhüser & Hoppe, 2013). Therefore, IS and medical research is also increasingly focusing on qualitative approaches when applying the Technology Acceptance

Model (Holden & Karsh, 2010; Karsh et al., 2006; Wu, 2011), which can also provide many informative insights (Holden & Karsh, 2010). Venkatesh, who was involved in the development of TAM2 and TAM3, along with Bala (TAM3) also calls for more qualitative approaches in IS research (Venkatesh, Brown & Bala, 2013). Therefore, despite its quantitative origin, the TAM is appropriate with our qualitative approach.

To investigate the lack of acceptance of AI systems among healthcare professionals (Agarwal et al., 2010; Goldfarb & Teodoridi, 2022; Khanijahani et al., 2022; Park et al., 2020; Reis et al., 2020), the TAM is helpful to investigate the different factors that affect this acceptance. We decided to use only the factors of the TAM2 in order to be able to examine the Perceived Usefulness of the AI applications in great detail and to be able to devote a large part of the interview time on this. Especially in the introductory phase of the systems, factors such as Output Quality or Job Relevance are more important than the individual factors of Perceived Ease of Use (Venkatesh & Davis, 2000). The decision for the TAM2 and the focus on Perceived Usefulness also follows the finding of Venkatesh and Davis (2000), which identified Perceived Usefulness as the most important factor for Behavioral Intention and thus for Actual Use, since a system is only accepted if it is perceived as useful.

2.3.2 Task-Technology Fit Model

A considerable amount of IS research has investigated relations between technology and performance and it suggests that IS only has a positive impact on performance when technology functionality and task requirements of users fit (Goodhue & Thompson, 1995). Furthermore, two separate, but related streams of research, investigated the effects of utilization and task-technology fit on performance (Goodhue & Thompson, 1995). Combining and integrating ideas from both streams, Goodhue and Thompson (1995) came up with the Technology-to-Performance Chain (TPC). The TTF model which is of interest in this research is an integral part of the TPC and shown in Figure 2. It is based on three assumptions. First, task and technology characteristics affect the task-technology fit. Secondly, a good task-technology fit will also increase the utilization of technology. And lastly, a good task-technology fit and also utilization of the IS will positively affect performance outcomes. Goodhue and Thompson (1995) recognize that utilization is only addressed very generically in their model. They argue that the model can benefit and be complemented by models which address utilization as a more complex construct. However, the novelty and focus of Goodhue and Thompson (1995) is on the construct of TTF. TTF is understood as "the degree to which a technology assists an individual in performing his or her portfolio of tasks" (Goodhue & Thompson, 1995, p.216).

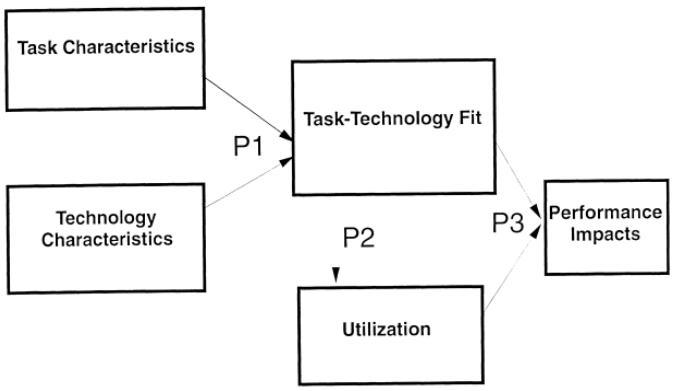

Figure 2: Task-technology fit model (Goodhue & Thompson, 1995)

However, in a more precise definition, they describe it as “the correspondence between task requirements, individual abilities, and the functionality of the technology” (Goodhue & Thompson, 1995, p.217). This second understanding much more emphasizes the importance of not only considering technology and task characteristics but also taking a closer look at the individuals performing the tasks, their abilities, and their needs. By a total of 16 dimensions, summarized in 8 factors, they describe the TTF: Quality, Locatability, Authorization, Compatibility, Ease of Use/Training, Production Timeliness, Systems Reliability, and Relationship with Users (Goodhue & Thompson, 1995). The factors and their relevance to our research are assessed in Section 2.3.4.2.

2.3.3 Integration of TAM/TTF Model

The highlighted complementarity of the TAM and TTF model is not novel and Dishaw and Strong (1999) integrated both models (Figure 3) and tested its performance (Figure 4). The integration outperformed both individual models and therefore, the integrated view of both models can provide a more comprehensive understanding of the relation of the different factors which determine user acceptance and actual system use (Dishaw & Strong, 1999). Six paths connect the two models as can be seen in Figure 3.

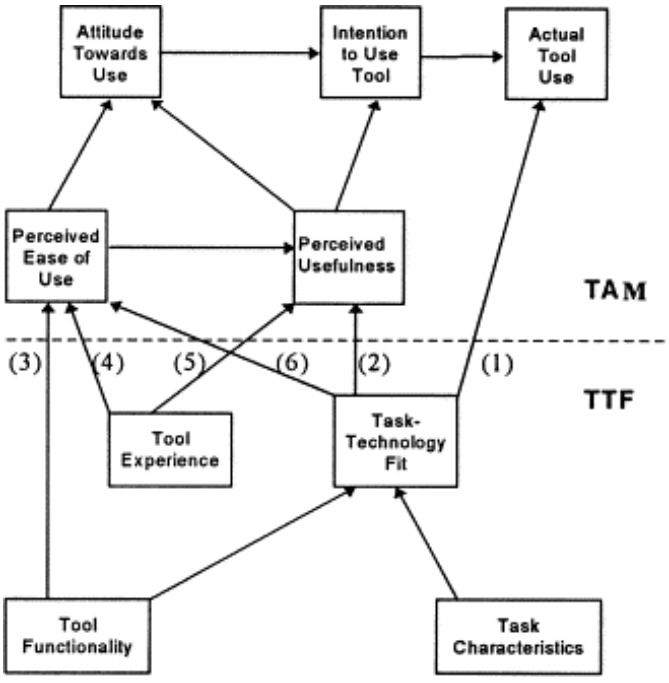

Figure 3: Integrated TAM/TTF model (Dishaw & Strong, 1999)

Both models explain actual use (utilization) so path 1 is the natural first relation of both models. If both models were just connected by path 1 their factors could be regarded as exclusive and while much complimentary value were gained, an integrated view would not be sensible (Dishaw & Strong, 1999). However, five other paths signal the interrelatedness of their factors. Path 3 captures the idea that a tool with more functionalities is also more complex which negatively affects the Perceived Ease of Use (PEU). The other four paths incorporate the argument that both more tool experience and a good TTF will positively affect the user’s

perception of ease of use and usefulness (Dishaw & Strong, 1999). Dishaw and Strong (1999) find that the data does not support the idea that a good TTF would increase Perceived Usefulness (PU), but it is argued that this effect is likely mediated indirectly by the effect of TTF on PEU and PEU on PU. Generally, their findings (Figure 4) support their theoretical construct (Figure 3), which consequently provides a novel and better understanding of how factors relate and affect system use. The theoretical construct of Dishaw and Strong (1999) helps to develop a more comprehensive understanding of the factors of TAM and TTF and how they relate and affect user acceptance and utilization.

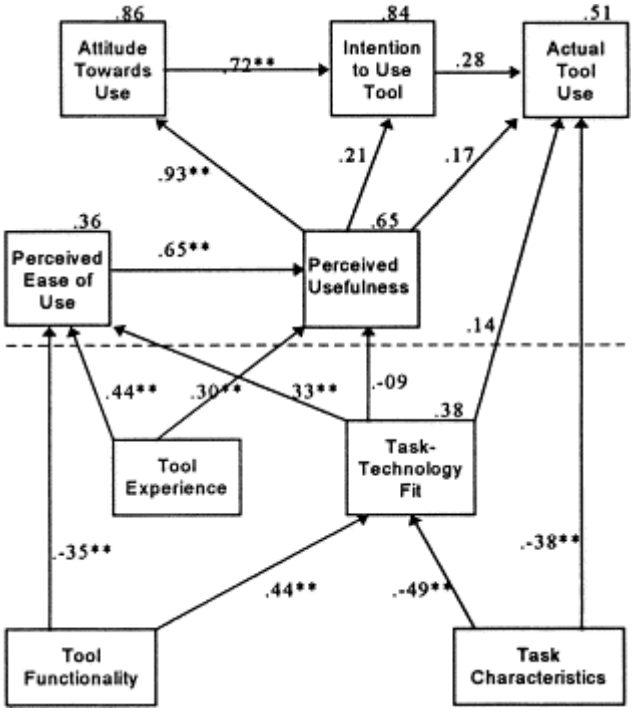

Figure 4: Integrated TAM/TTF model tested and adjusted to data (Dishaw & Strong, 1999)

2.3.4 Relevant TAM/TTF Factors

In Section 2.3 the relevance of TAM and TTF was explained and both models were presented in Section 2.3.1 and 2.3.2. This section looks closely at the variables of both models and assesses their relevance to our research. In Section 2.3.5, the research practices of Dishaw and Strong (1999), Klopping and McKinney (2004), and Pagani (2006) are followed and the relevant variables from both models are integrated in our research framework. We then adjust these factors to fit our research topic (Gagnon et al., 2012; Holden & Karsh, 2010).

2.3.4.1 Technology Acceptance Model

The TAM2 consists of nine factors that influence the behavioral intention of a user to use a technology system.

Perceived Usefulness is a core construct of the TAM which can be defined as "the extent to which a person believes that using the system will enhance his or her job performance" (Venkatesh & Davis, 2000, p.187). It is an understandable and central, but also complex and broad concept (Venkatesh & Davis, 2000). Venkatesh and Davis (2000) see Perceived Usefulness as the most important factor that positively influences Behavioral Intention and thus Actual Use,

which can also be observed in technology acceptance by radiologists (Aldosari, 2012). If the radiologists do not consider the AI system useful, they will not adopt it, which would mean that the potential benefits from its use would be lost. Therefore, Perceived Usefulness was assessed to be an important factor for our research framework.

Perceived Ease of Use is one of the two factors from the original TAM which positively influence the behavioral intention of the user directly (Venkatesh & Davis, 2000). The factor Perceived Ease of Use describes "[the] extent to which a person believes that using the system will be free of effort" (Venkatesh & Davis, 2000, p.187). Not only the behavioral intention is positively influenced by the Perceived Ease of Use but also the Perceived Usefulness as a system is more useful for them if they have no difficulty using it (Venkatesh & Davis, 2000, p.187). Especially in the healthcare sector, it is important that the computer systems are easy to use because the work of the healthcare professionals is already very complicated and time-consuming and this complexity should not be further extended by the computer systems (Tubaishat, 2018). In addition, healthcare professionals have less time to familiarize themselves with the AI systems, which makes intuitive operation and good usability important. Since a very high workload is expected for radiologists in the context of lung cancer screening (Fischer et al., 2020; Liu et al., 2021) and they, therefore, need to focus on their work per se and not on familiarizing themselves with new AI systems, Perceived Ease of Use is highly relevant to this study and was therefore included in the research framework.

Subjective Norm is one of the three social influence factors introduced in TAM2 that influences Perceived Usefulness. Venkatesh and Davis (2000) adopted the definition of Fishbein and Ajzen (1975, p. 302) and described the Subjective Norm as a "person's perception that most people who are important to him think he should or should not perform the behavior in question". Venkatesh and Davis (2000) explained the influence on Perceived Usefulness by the fact that users could also use a system if someone important to them uses it or finds it useful, even if they do not find it useful themselves. The use of Artificial Intelligence is still new in many areas, which means people lack experience and have to rely more on the opinions of others (Li, Hess & Valacich, 2008). In the field of radiology, there have been many attempts to integrate AI into the workflows, but this is often not yet implemented across the board, so the Subjective Norm remains an important factor and was included in the research framework.

Voluntariness is another social influence factor introduced in TAM2. This mediating factor only affects the influence of the Subjective Norm on the Behavioral Intention and not on the Perceived Usefulness. TAM2 uses the definition of Moore and Benbasat (1991, p.195) that Voluntariness means "the degree to which use of the innovation is perceived as being voluntary, or of free will". According to Venkatesh and Davis (2000), Subjective Norm affects Behavioral Intention only when the user is required to use the system. Since medicine, and in particular, a screening program, is highly regulated (Jiang et al., 2017), it is necessary to investigate whether radiologists are required to use AI and, if so, whether this affects the acceptance of the AI systems. Voluntariness was therefore considered in the research framework.

Image is the third social influence factor in the TAM2 and influences the Perceived Usefulness, itself being mediated by the Subjective Norm (Venkatesh & David, 2000). The TAM2 defines Image, following Moore and Benbasat's (1991, p.195) definition, as "the degree to which use of an innovation is perceived to enhance one's ... status in one's social system". Venkatesh and Davis (2000) explain the influence of Subjective Norms on Image by the fact that people try to imitate the behavior of people with higher status in order to consolidate their

group membership and thus improve their own status, which in turn leads them to expect higher productivity due to more available resources resulting from the higher status. Since the acceptance of AI technology in lung cancer screening trials is being studied in the setting of research studies without a competitive environment that would reinforce the pursuit of status, Image was found not to be relevant and was not included in the research framework

Experience is a mediating factor for the influence of the Subjective Norm on Perceived Usefulness and Behavioural Intention (Venkatesh & Davis, 2000). Venkatesh and David (2000) found that the influence of the Subjective Norm decreases with growing Experience with the system, as users can then form their own opinions about the system and are less dependent on others. Since new AI systems are being introduced in the context of lung cancer screening in Germany and the relevant trials have only started in the last year, the systems are at a stage where the user experience is just growing and influences the acceptance, which is why Experience was taken into account in the research framework.

Job Relevance is another TAM2 factor that positively influences Perceived Usefulness and is about "an individual's perception regarding the degree to which the target system is applicable to his or her job" (Venkatesh & Davis, 2000, p.191). Job Relevance is high when the tasks which the system is supporting are important for the user's job (Venkatesh & Davis, 2000). Job Relevance has a great impact on Perceived Usefulness, as the output must be relevant to the task in order to appear useful (Venkatesh & Davis, 2000). There are many different AI systems, each promising many different benefits, but for these to be accepted, these benefits must fit the work of the radiologists, and therefore Job Relevance was investigated in our research framework.

Output Quality was defined by Venkatesh and Davis (2000, p.191) as the perception " how well the system performs those tasks". In contrast to job relevance, which is an exclusion criterion for the choice of software if the software is not relevant to the job, output quality is a criterion after which a choice is made from a selection of software by selecting the one with the highest output quality (Venkatesh & Davis, 2000). Since data quality is particularly important in the health sector, since medical decisions are made on the basis of the data, or the risk of serious illnesses increases considerably if lung nodules are not detected (Tanoue et al., 2015), output quality is very important for the acceptance of the software by radiologists and was therefore taken into account in the research framework.

Result Demonstrability was introduced in TAM2 (Venkatesh & Davis, 2000) and describes "the tangibility of the results of using the innovation" (Moore & Benbasat, 1991, p.203). In other words, a system is more likely to be adopted when its results can easily be demonstrated and its advantages are visible (Zaltman, Duncan & Holbek, 1973). Venkatesh and Davis (2000) highlight the importance of Result Demonstrability by arguing that an effective system won't be accepted if this efficiency is not noticeable. Also, systems with high Job Relevance won't be perceived as useful if this Job Relevance is not apparent to the users (Venkatesh & Davis, 2000). Therefore, Result Demonstrability has a positive effect on Perceived Usefulness. Especially during the implementation of AI systems for radiologists, it is important that the quality of the results is evident so that it can be evaluated whether the systems facilitate the work of radiologists. Result Demonstrability was therefore included in the research framework to explore the effects.

2.3.4.2 Task-Technology Fit Model

There are eight factors that make up the TTF (Goodhue & Thompson, 1995).

Quality describes that data must be sufficient in currency, completeness, and level of detail (Goodhue & Thompson, 1995). It refers to data maintained by an organization and its suitability for the user's needs (Goodhue & Thompson, 1995). Therefore, in a healthcare setting, it is most applicable to data on patients. This data is not integrated into AI systems and therefore not relevant to our research. However, Quality is applicable in the way that the output of the AI system can be treated as data which is then used by the radiologists. The currency of the data is then not an issue. Nevertheless, the AI system might provide insufficient information or even miss critical information altogether. Therefore, applied to the output of the AI, the factor Quality was deemed relevant.

Locatability is made up of the two dimensions, locatability and meaning (Goodhue & Thompson, 1995). The Locatability dimension is defined as the "[e]ase of determining what data is available and where" (Goodhue & Thompson, 1995, p.234). This dimension refers to data maintained by the organization which was deemed irrelevant in the research context which focuses on the use of AI. However, the dimension meaning is applicable to the AI system output. Meaning describes the "[e]ase of determining what a data element on a report or file means, or what is excluded or included in calculating it" (Goodhue & Thompson, 1995, p.234). Therefore, for Locatability to be present, the meaning and definition of the AI output must be either obvious or easy to find out (Goodhue & Thompson, 1995). Addressing the dimension of meaning, Locatability was also included in the research framework.

Authorization describes the issue that difficulties with data authorization can get in the way of a person to execute their job (Goodhue & Thompson, 1995). Data authorization problems appear unlikely since radiologists were expected to have everything necessary for diagnosis available. Data authorization was not found to be relevant enough to be included in the research framework.

Compatibility concerns that "[d]ata from different sources can be consolidated or compared without inconsistencies" (Goodhue & Thompson, 1995, p.234). If multiple AI with different functionalities would be applied for lung cancer screening, this factor would concern these differences. However, adoptions restrict themselves to one AI system and consequently, Compatibility is not further considered.

Production Timeliness refers to "IS [that] meets pre-defined production turnaround schedules" (Goodhue & Thompson, 1995, p.235) and it investigates whether IS functions/tasks are completed by IS in time. The integration of AI into the lung cancer screening process is necessary to increase process efficiency and enable lung cancer screening on a national scale (Fischer et al., 2020; Liu et al., 2021). Therefore, *Production Timeliness* is likely a concerning factor for radiologists. Furthermore, it appeared to be closely related to the relevant factor *Relationship with Users* which made inclusion in the framework further useful to investigate codependencies. Thus, it was included in the framework.

Systems Reliability is about the "[d]ependability and consistency of access and uptime of systems" (Goodhue & Thompson, 1995, p.235). In our research context, this means that the AI system must be available when it is needed, does not have inconvenient downtimes, and does not crash or have other problems which negatively impact its use (Goodhue & Thompson, 1995). As this factor is very applicable to our research it was also included in the research framework.

Ease of Use/ Training consists of two dimensions (Goodhue & Thompson, 1995). First is, ease of use of hardware & software, which describes in our research the perceived ease of use for the radiologists to use AI to support them (Goodhue & Thompson, 1995). Secondly, the training dimension questions whether the necessary training is available for radiologists so they can use the AI system effectively (Goodhue & Thompson, 1995). Both dimensions apply to our research so the entire construct of Ease of Use/ Training was integrated into the research framework.

Lastly, **Relationship with Users** is defined broadly and spans five dimensions. Those five dimensions generally describe the relationship of the system user with the IS staff. This includes whether IS keeps its agreements, understands and takes business problems seriously, and responds timely with an appropriate answer (Goodhue & Thompson, 1995). Furthermore, this factor also spans the reaction of the provider to issues, concerns, and improvements which are raised by the radiologists and their clinics (Goodhue & Thompson, 1995). Software solutions in hospitals are increasingly implemented as software as a service (Lucas, 2021), which prevents the internal IT department from introducing new features into the software or fixing bugs, since the healthcare institution only buys licenses for the software instead of programming it itself. Therefore, when radiologists have problems with the software, they have to contact the software vendor's support, which is more time-consuming or less convenient for the radiologists than contacting the in-house IT support, which could investigate the problem directly on site. Therefore, providing easily accessible support to radiologists and involving them in the planning of the software is even more important, which is why the Relationship with Users factor was included in the research framework.

2.3.5 Research Framework

In Sections 2.3.4.1 and 2.3.4.2 various relevant dimensions of the TAM and TTF model which are applicable to radiologists' technology acceptance in lung cancer screening have been highlighted. The theoretical foundation of both models overlaps (Dishaw & Strong, 1999) and therefore also common practice was followed and the factors were grouped together and adopted where need was seen (Harst, Lantzsch & Scheibe, 2019). The resulting factors were sorted into one of three themes: System Output, Usability Experience, and Operationality (Table 1). The themes serve as a structure of the individual factors allowing the analysis of the findings to be clearly arranged.

The factors *Result Demonstrability, Job Relevance, Locatability, Perceived Usefulness, Experience, Relationship with Users, Production Timeliness,* and *Systems Reliability* were considered to be unique and were sorted into the research framework unchanged. Overlap was found with the ideas of the other factors, and the three factors *Output Quality, Subjective Voluntariness,* and *Perceived Ease of Use/ Training* were newly developed or changed and are explained below.

System Output describes how good the AI system is in terms of its results and output. *Output Quality* merges the ideas from *Output Quality* from TAM2 and *Quality* from the TTF measures. The factor of the TAM2 emphasizes that the technology must perform well at its task (Venkatesh & Davis, 2000), which is in our research to provide output that supports the radiologists with their tasks. The TTF factor addresses this output and describes that the output should be complete and consist of all necessary information (Goodhue & Thompson, 1995). Both definitions are summed up in our new definition for *Output Quality:* The system

performs its job well and generates complete and relevant output. Furthermore, the factors *Result Demonstrability, Job Relevance, Locatability,* and *Perceived Usefulness* were included in the theme System Output. The theme is relevant to the adoption of AI in the context of lung cancer screening because it links the quality of the software output, which must be highly accurate in a medical setting, to the relevance of that output, since an accurate output that is not relevant would likely hinder the acceptance of AI technologies.

Usability Experience addresses different aspects of the experience with the use of the system. *Subjective Voluntariness* merges two factors of the TAM2 (Venkatesh & Davis, 2000). *Subjective Norm* and *Voluntariness* are closely related concepts in the TAM2 already and the idea of *Subjective Norm* which is relevant to our research intersects with *Voluntariness*. *Subjective Norm* refers to social pressure a user feels to use the system and *Voluntariness* is to which degree the use of the system is perceived as voluntary. *Subjective Voluntariness* captures the joint idea of both factors which we define as the pressure a person feels to use a system. *Experience* and *Relationship with Users* are also to the theme of Usability Experience. *Perceived Ease of Use* is defined in the TAM2 as “the extent to which a person believes that using the system will be free of effort" (Venkatesh & Davis, 2000, p.187). The TTF model describes with *Ease of Use/ Training* a very similar concept that adds the training dimension which looks at whether necessary training to use the system is available (Goodhue & Thompson, 1995). Both ideas are included in the new framework factor *Perceived Ease of Use/ Training* which is the ease with which the system can be used effectively and the necessary training completed. *Usability Experience* takes into account the usability of AI applications which is important for radiologists due to their time constraints and complex tasks. The dimension also considers the radiologists' relationship with the software vendors and employers to examine the socio-technical aspects of healthcare facilities. And thirdly, **Operationality** involves aspects of the AI which directly affect the operational utilization. To investigate this *Operationality*, *Production Timeliness* and *Systems Reliability* were grouped together to study the quality of systems independent of the quality of outputs and usability. *Operationality* also helped to examine how the availability and timeliness of AI systems affect radiologists and their reliance on AI systems.

Themes	Factors	Definition	References
System output	Output Quality	The system performs its job well and generates complete and relevant output	Goodhue & Thompson, 1995; Marangunić & Granić, 2015; Venkatesh & Davis, 2000
	Result Demonstrability	"[T]he tangibility of the results of using the innovation" (Moore & Benbasat, 1991, p.203)	Marangunić & Granić, 2015; Moore & Benbasat, 1991; Venkatesh & Davis, 2000; Zaltman, Duncan & Holbek, 1973
	Job Relevance	"[The] individual's perception regarding the degree to which the target system is applicable to his or her job” (Venkatesh & Davis, 2000, p.191)	Marangunić & Granić, 2015; Venkatesh & Davis, 2000

Themes	Factors	Definition	References
	Locatability	“Ease of determining what a data element on a report or file means, or what is excluded or included in calculating” (Goodhue & Thompson, 1995, p.234)	Goodhue & Thompson, 1995
	Perceived Usefulness	"The extent to which a person believes that using the system will enhance his or her job performance" (Venkatesh & Davis, 2000, p.187)	Davis, 1989; Marangunić & Granić, 2015; Venkatesh & Davis, 2000
Usability Experience	Subjective Voluntariness	The pressure a person feels to use a system.	Marangunić & Granić, 2015; Venkatesh & Davis, 2000
	Perceived Ease of Use/ Training	The ease with which the system can be used effectively and the necessary training completed	Davis, 1989; Marangunić & Granić, 2015; Goodhue & Thompson, 1995; Venkatesh & Davis, 2000
	Experience	“[D]irect experience with a system over time” (Venkatesh & Davis, 2000, p. 190)	Marangunić & Granić, 2015; Venkatesh & Davis, 2000
	Relationship with Users	The relationship of the information system’s department with the users. Including IS Understanding of Business, IS Interest and Dedication, Responsiveness, Consulting, and IS Performance.	El-Gayar, Deokar, & Wills, 2010; Goodhue & Thompson, 1995
Operationality	Production Timeliness	"IS meets pre-defined production turnaround schedules" (Goodhue & Thompson, 1995, p.235)	Goodhue & Thompson, 1995
	Systems Reliability	“Dependability and consistency of access and uptime of systems” (Goodhue & Thompson, 1995, p.235)	Goodhue & Thompson, 1995

Table 1: Research framework

3 Methodology

This chapter explains and motivates the choices for the research design. In Section 3.1 the philosophical considerations are presented. In Section 3.2 the research approach is described and motivated and in Section 3.3 the data collection process is discussed. In Section 3.4 the interview design is shown and Section 3.5 covers the methods of data analysis. Finally, Section 3.6 addresses ethical considerations before the means to achieve scientific quality are disclosed.

3.1 Research Philosophy

Research philosophy is about "the development of knowledge and the nature of that knowledge" (Saunders, Lewis & Thornhill, 2007, p.101). Therefore, it is at the core of research and must be addressed to ensure the quality and insightfulness value of the findings (Hassan, Mingers & Stahl, 2018). The research philosophy consists of different dimensions with different implications for practice (Creswell, 2007), and the elements considered most important in our research are discussed below.

In our research, we take an interpretative stance and a subjectivist perspective (Saunders, Lewis & Thornhill, 2007; Walsham, 1995). This means that we do not believe that we can observe reality directly, but instead, we try to make sense of it by social means (Walsham, 1995). We respect that in our research we will interview people from different healthcare institutions who have had their individual experiences and that all have a unique perspective on the matter (Schultze & Avital, 2011; Walsham, 1995).

The development of our understanding started with our first thoughts on the topic and was then further developed with the literature research. When a sound understanding of the topic and context was developed, an appropriate research framework was developed (Table 1) to structure our approach to assess the research question.

Through interviews, we were given insights into the different perspectives of the interviewees and after, our interpretations, analysis, and discourse regarding the interviews led to our final understanding of the topic which is discussed in Section 5.

This research process is highly subjective and thus it is important to openly and transparently disclose the means of how the research was conducted and the understanding of the topic developed (Creswell, 2007).

Therefore, this thesis is written in a narrative style as this is best suited to convey the story of how we conducted our research and developed our understanding (Creswell, 2007). In Section 4 we disclose our perceptions of the findings and the statements from the interviews which were most influential in the development of our understanding. And then, in Section 5, the findings are related with the literature and our thoughts and expectations on the findings are given.

The interpretative stance works well and is appropriate and often applied in exploratory research (Patton, 2015; Recker, 2013). Exploratory research implies that not much on a topic is

known and an interpretive stance allows the development of a rich understanding (Gregor, 2006; Recker, 2013). This case also applies here as there are only a few locations in Germany where novel AI systems are tested for lung cancer screening and extracting rich information from the few available sources was critical to develop a sound understanding. As there is no streamlined lung cancer screening process yet, AI integrations differ across the different locations and so did the experiences which were made with the technology. The interpretative stance allowed us to consider the context and environment of the experiences which were made to find common themes and develop a sound understanding of the challenges that occur with AI integration (Saunders, Lewis & Thornhill, 2007; Schultze & Avital, 2011).

3.2 Research Approach

The research approach describes the "plans and procedures for conducting research" (Creswell, 2013, p.3). It can also be described as the "strategy of investigation to answer a particular research question" (Recker, 2013, p.36). This research will follow a qualitative approach. This means, using the data we get from interviews, we are the instrument of analysis to arrive at our findings (Patton, 2015). A qualitative approach with our interpretive stance is common as both emphasize similar core elements such as context, subjectivity, and the necessary involvement of the researcher in the results (Creswell, 2007; Patton, 2015).

To structure our research, we have integrated topic-relevant elements of the TAM and TTF model into a framework (Table 1) to develop a comprehensive list of themes that address the relevant factors which determine the radiologist's acceptance of AI for lung cancer screening diagnosis. The framework provides an overview of the factors which affect radiologist's technology acceptance and "[u]nderstanding the influence of social and organizational context on systems use" is named by Kaplan and Maxwell (2005, p.33) also as one of the main reasons for using qualitative methods in evaluating computer IS. Factors of the TAM and TTF model are commonly tested and measured in surveys (Davis, 1989; Goodhue & Thompson, 1995), however, the applicability and usefulness of a qualitative approach in research that uses semi-structured interviews to assess the factors has also been shown in the IS literature (Bandyopadhyay & Zadeh, 2014; Hou & Gao, 2017; Sturm & Peters, 2020; Xiao, Meredith & Gao, 2017) and is adopted here.

Shortcomings of a qualitative approach include lack of controllability, deductibility, repeatability, and generalizability which also threatens validity (Recker, 2013). However, this can be mediated by credibility (Creswell, 2007), which was pursued utilizing various strategies in our research (Section 3.7). Arguments in favor of a qualitative approach are its strength in discovering phenomena, generating rich and complex knowledge, and developing a comprehensive understanding (Recker, 2013; Saunders, Lewis & Thornhill, 2007). As apparent from the literature review, the relevant context of our research question is rather complex, and therefore, the arguments in favor of a qualitative approach prevail.

3.3 Data Collection Methods

When conducting the research we knew we lacked knowledge on lung cancer screening, the role of AI for lung cancer screening and the specifics of technology acceptance in the

healthcare context to assess the research question. Therefore, those three topics were the focus of our initial literature research. The Association for Information Systems eLibrary is a large database with IS articles and was selected as a database for our search. Furthermore, Google Scholar was selected as well to extend the scope to healthcare, technology and other related literature, which was also deemed relevant to us (Webster & Watson, 2002).

The databases were searched for a selection of the keywords “Artificial Intelligence”, “Lung Cancer Screening”, “Health Information Technology” and “Healthcare” in combination with a selection of the keywords “Radiology”, “Technology Acceptance” and “User Resistance”.

This yielded a solid understanding of the topic and only the understanding of the AI systems and the workflow of radiologists in lung cancer screening was found to be important and lacking.

Therefore, Google Scholar was searched with the keywords “Lung Cancer Screening”, “Artificial Intelligence” and “Pulmonary Nodule” in combination with “Detection” or “Identification” or “Classification”.

Information and knowledge from this literature research was categorized (Webster & Watson, 2002) and discussed in Section 2.1 and Section 2.2.

The literature research already included literature where the TAM was applied and further research suggested that the TAM was appropriate in the healthcare context (Section 2.3.1). However, it was also pointed out that the TAM was lacking in certain aspects (Section 2.3) and a search for “Technology Acceptance TAM Framework” in combination with “Extension” or “Complement” revealed the TTF to us after some time which was then assessed to be an appropriate complement (Section 2.3).

After preparing the research with a literature review (Webster & Watson, 2002), interviews were found to be the most suitable primary data collection method. In a timeframe of nine days we conducted five semi-structured interviews (Kaplan & Maxwell, 2005) using the video call platform Zoom (Zoom Video Communications, 2022). Interpersonal interviews have the advantage that one can “ask open-ended questions and probe for in-depth responses about people’s experiences, perceptions, opinions, feelings, and knowledge” (Patton, 2015, p.87). Targeted questions can be asked to generate rich, insightful, and relevant knowledge (Recker, 2013). In particular, the incomplete interview script of semi-structured interviews allows for emerging questions to be discussed which is ideal for exploratory research and makes them very suitable for this research (Myers & Newman, 2007; Saunders, Lewis & Thornhill, 2007).

It is important to find relevant and knowledgeable interview partners to generate quality data from the interviews (Recker, 2013). A Google search with the keywords “Lungenkrebs Früherkennung Deutschland” (lung Cancer Screening Germany) and “Lungenkrebs Früherkennung Deutschland Implementierung” (Lung Cancer Screening Germany Implementation) was conducted which revealed two relevant active trials, the HANSE trial (Medizinische Hochschule Hannover, 2021) and 4-IN THE LUNG RUN (Deutsches Krebsforschungszentrum, n.d.). Suitable interview partners were sorted into one of two categories. First, clinic directors and department heads which manage and oversee the trial and are informed about the current findings in the trials as well as the general acceptance of AI by radiologists for lung cancer screening at their location were considered suitable interview candidates. R1, R3 and R4 belong to this category (Table 2). Secondly, radiologists who work with

AI for lung cancer screening in those trials were considered suitable complimentary interview partners as they have more personal experience with the AI system and from this the technology acceptance factors can be better examined and assessed. R2 and R5 belong to this category (Table 2).

The radiologists in more senior positions (R1, R3 & R4) were very busy and were only available for limited periods of time which limited the benefits of the semi-structured interview design (Saunders, Lewis & Thornhill, 2007; Section 3.4). The interview design was adjusted to mediate the shortcomings (Section 3.4), nevertheless, the interview with R4 was only 18 minutes and too short for an extensive discussion about all aspects which were of interest to us. The radiologists who were working with the AI software in their clinical routine regularly were in less senior positions and were flexible with the duration of the interview and available for longer interviews (R2 & R5). It is noteworthy that the duration of the interview with R2 stands out as many personal experiences, details and honest opinions were shared by them which were very insightful. However, this might have led to an overrepresentation of their personal stance in the findings with few methods of statement validation due to the limited number of interviews with comparable interview partners. New insights were still gained from the final fifth interview, suggesting that this research could have benefited from more interviews which was not possible due to time constraints.

All interviews were conducted in German, transcribed, and then translated to English. Both German and English transcripts are attached in Appendices D through H.

Respondent ID	Role	Organization type	Date	Duration	Platform
R1	Director of a radiology clinic, professor of radiology	University hospital	April 26, 2022	25 minutes	Zoom
R2	Head senior radiology physician	Lung center	April 28, 2022	69 minutes	Zoom
R3	Director of a radiology clinic	University hospital	May 3, 2022	37 minutes	Zoom
R4	Head of a lung imaging lab, professor of radiology	University hospital	May 3, 2022	18 minutes	Zoom
R5	Director of a radiology department, head senior radiology physician	Lung center	May 5, 2022	39 minutes	Zoom

Table 2: Respondents

3.4 Interview Design

Interviews can be a powerful data gathering technique when executed correctly (Myers & Newman, 2007). However, various problems and pitfalls exist in the context of the interview and this emphasizes the need to carefully plan ahead (Myers & Newman, 2007; Recker, 2013). One of the strengths of the semi-structured interview are the unscripted questions which emerge during the interview, however, to guide and lay a foundation for the interview some structural elements should be considered (Myers & Newman, 2007; Recker, 2013).

In Appendix A, a translation of our interview guide is attached. For further reference here, the questions in the interview guide were numbered.

As suggested by Myers and Newman (2007), the interview starts with a short introduction of ourselves and the purpose of the interview. In addition, we quickly talk about the consent form (Appendix B) that interviewees have signed regarding the processing of their personal data and ask them to participate in member checking (Appendix C) by validating our preliminary analysis of the interviews. We proceed with four questions (1-4) to understand the relationship of our interviewee to lung cancer screening and the relevance of lung cancer screening at their institution.

It is advisable to proceed with general questions on the topic which can then serve as a foundation to develop the interview in a conversational manner to more specific topics (Recker, 2013). Therefore, when more time was available for an interview, questions 5 to 9 were used to get a conversation on the topic started in which the themes from the research framework and questions 10 to 16 were covered. However, when little time was available for the interview, the structure of the interview guide was more rigid, and questions 5 to 9 helped to build the understanding of the role of AI in the lung cancer screening process at their institution to dive deeper into the topic in the following questions.

Questions 10 to 16 were designed to address the themes from our research framework (Table 1) which were not already covered in our prior questions. In Table 3 the mapping of factors to questions in the interview guide (Appendix A) can be seen.

Dimension	Factor	Questions
System output	Output Quality	10, 11, 12, 13
	Result Demonstrability	12
	Job Relevance	10, 11, 12
	Locatability	12
	Perceived Usefulness	12, 13
Usability experience	Subjective Voluntariness	8
	Perceived Ease of Use/ Training	12, 14
	Experience	13, 14
	Relationship with Users	14, 15, 16
Operationality	Production Timeliness	13, 14, 15, 16
	Systems Reliability	14, 15, 16

Table 3: Interview design

Questions 17 to 22 are rather specific and difficult questions which mostly served as reminders as they are topic-wise related to questions 5 to 9 and were planned to be asked in a conversational manner as follow-up. However, in shorter interviews where it was important to ensure that all topics were covered, not much conversation was developed in questions 5 to 16 and questions 17 to 22 were then asked to use the remaining time to dive deeper into aspects which were considered insufficiently addressed or most relevant. Finally, questions 23 and 24 served to prepare for closure by allowing the interviewees to address further aspects and ask questions (Myers & Newman, 2007).

3.5 Data Analysis Methods

All interviews were transcribed and then translated into English. To analyze the data, the interviews were then coded. Coding refers to grouping pieces of the data and assigning labels to "facilitate insight, comparison, and the development of theory" (Kaplan & Maxwell, 2005, p.42) and is generally regarded as an essential step to prepare data for analysis (Patton, 2015). The analytic process and the development of the findings can be understood as an interplay between us and the data and coding is one of the procedures to bring standardization and rigor to that process (Strauss & Corbin, 1998). We used the factors of our framework (Table 4) as themes and followed a deductive coding approach (Patton, 2015). This did not restrict novel ideas to be found and presented, but it limited the potential for novel findings (Saunders, Lewis & Thornhill, 2007). However, at the same time, the findings were aligned with prior research and could be integrated well with existing findings (Saunders, Lewis & Thornhill, 2007; Section 5).

The interviews were reviewed separately by both researchers, and the sections in which a relevant factor was addressed were marked with codes (Table 4), as seen in Appendices D through H. The relevant statements were then sorted below their respective factors and grouped into similar aspects. In joint discussion both researchers then agreed on core findings visible as bold subheadings in Section 4.

Dimension	Factor	Code
System Output	Output Quality	OQ
	Result demonstrability	RD
	Job Relevance	JR
	Locatability	L
	Perceived Usefulness	PU
Usability experience	Subjective Voluntariness	SV
	Perceived Ease of Use/ Training	PEUT
	Experience	E
	Relationship with Users	RU
Operationality	Production Timeliness	PT
	Systems Reliability	SR

Table 4: Coding themes

Member checking was done in which general questions about statements were asked to confirm our understanding and the feedback of respondents was requested on two findings which we deemed most controversial (Appendix C). Four of the five interviewees responded. R1, R2, R3 and R5 agreed with the first finding. R1 and R2 agreed also with the second controversial finding, while R3 said that bigger studies are needed to investigate this and R5 simply did not comment.

3.6 Ethical Considerations

In our research we took a deontological view that ends cannot justify the means and that research must be ethical under any circumstances (Saunders, Lewis & Thornhill, 2007). We follow the recommendations of our university to follow the European Code of Conduct for Research Integrity (ALLEA, 2017; Lund University, 2021) which includes to “comply with codes and regulations relevant to [our] discipline” (ALLEA, 2017, p.6). The Association for Information Systems (AIS) Code of Ethics and Professional Conduct (AIS, 2019) is set by the professional community in IS and is considered and chosen as an appropriate code of ethics to

follow for IS research (Recker, 2013). Below we address our most relevant ethical considerations.

We have conducted interviews and with that comes "the responsibility to secure the actual permission and interest of all those involved in the study" (Recker, 2013, p.143). All participants received the personal data processing consent form (Appendix B) before the interviews and were again informed before the recording of the interview about the personal data processing, research purpose, scope of the research, voluntariness of answering our questions and the planned member checking (Appendix A). Two respondents expressed concern that important data from the study might be leaked before they made their disclosures, and accordingly we complied with their requests and anonymized all data that might allow inferences about the AI software used.

One central ethical principle is that this research must contribute to society, do good and avoid harm (AIS, 2019). Generally, the ethics of AI in healthcare have been discussed extensively (Bærøe, Jansen & Kerasidou, 2020; Chen & Verghese, 2020; Morley et al., 2020; Parikh, Teeple & Navathe, 2019; World Health Organization, 2021). The aforementioned studies conclude that the ethics of AI in healthcare depend on a variety of factors and require individualized judgment. This exploratory research developed an understanding about the radiologist's acceptance of AI systems and from this design decisions might be derived. Care has been taken to transparently disclose all relevant data from this research, how we arrived at our findings and their limitations so that inappropriate actions are not derived. This research might also facilitate the implementation of lung cancer screening in Germany and after extensive and careful considerations we agree with the prevailing assessment that lung cancer screening and our research is ethical (Institute for Quality and Efficiency in Health Care, 2020).

Honesty and trustworthiness are also of the highest importance, and we met this with precise and transparent documentation of our research and furthermore, all research documents will be stored for five years after the research has concluded to allow for external reinvestigation (AIS, 2019; Recker, 2013). Both researchers take full responsibility for the integrity of the research (ALLEA, 2017) and both separately ensure the storage of all documents in a secure way. We do not want to discriminate and have investigated our research carefully for subconscious or indirect discrimination to ensure fairness. Lastly, adequate academic writing criteria are known and followed and our research reflects our own work and ideas. Both researchers also declare no conflict of interest.

3.7 Scientific Quality

While the importance of high scientific quality is commonly agreed upon, clear assessment is difficult in practice (Patton, 2015). No consensus has emerged about universal criteria and suitable criteria vary on the nature and approach of the research (Creswell, 2007; Patton, 2015). For our interpretive qualitative research we found considerations of relevance, validity, and reliability to be exceptionally important (Creswell, 2007; Patton, 2015). Relevance is discussed in Section 1, so below we focus on our considerations regarding validity and reliability.

One popular understanding of validity in interpretive research comes from Angen (Creswell, 2007) who argues that validity in interpretive research is a question of morality (Angen, 2000). Angen (2000) explains that validity can be seen as a continuous process that consists of the judgment of the trustworthiness of a piece of research by other researchers. She distinguishes between ethical and substantive validation (Angen, 2000). Ethical validation describes how "scientific endeavors should contribute to our ability to carry on in an ethical way" (Angen, 2000, p.388). It concerns the morality of assumptions and implications of the research (Angen, 2000) which are discussed in Section 3.6. Substantive validation requires self-reflexivity and transparency (Angen, 2000). Research is of substantive validity when the development of the understanding can be comprehended (Creswell, 2007). For this, all interviews were recorded, transcribed, translated and are attached both in German and English to this thesis in Appendices A through H. Reliability in qualitative is often measured as "the stability of responses to multiple coders of data sets" (Creswell, 2007, p.210) and was addressed by us by conducting observer triangulation in coding (Creswell, 2007). However, an intercoder process often includes many vaguely defined and dynamic steps and this can raise questions about how researchers arrived at findings (Creswell, 2007). Therefore, the coding process was described in detail in Section 3.5 and findings in Section 4 were documented in a way that the development of ideas from the interviews can be comprehended (Section 3.5; Section 4). An effort was made to write down our research in a transparent and narrative way to make the development of our understanding clear (Angen, 2000; Creswell, 2007; Recker, 2013) and further increase validity and reliability. Finally, member checking was conducted (Section 3.5) which is also considered an excellent way to increase the validity of findings (Creswell, 2007).

4 Findings

The interviews were coded according to our research framework (Table 1), and the results for each factor are presented in their respective themes in Section 4.2. An understanding of the research context is important to understand the relevance of the interviewees' statements and the findings and is briefly presented in Section 4.1.

4.1 AI Integration

The five respondents came from four different locations. All five were radiologists, however only R2 and R5 worked most of their time in a clinical routine while R1, R3 and R4 were more engaged with the management tasks as directors of radiology facilities. Nevertheless, R1,R3 and R4 also had insights into the clinical routine as and have worked with the AI software as senior radiologists.

R2 and R5 were colleagues and worked at the same location. Although both were experienced senior radiologists, they worked with the AI software for lung cancer screening regularly as part of their daily routine.

Furthermore, R2, R3, R4 and R5 were all working at clinics that were involved in the HANSE lung cancer screening trial and R1 was a director of a radiology clinic, which was involved in the 4-IN THE LUNG RUN screening trial in Germany.

The involvement in those trials led to certain aspects, characteristics, and processes which were similar across all locations. Most notable is that the AI software for screening at all locations was the same.

In the scope of the lung cancer screening trials, the focus was on three different diagnoses, all of which were supported by the deployed AI software. The first and primary goal was to find long nodules and classify their malignancy (1.29; 3.14). This function is directly related to detecting lung cancer. Two more secondary diagnoses were looked into. The software also calculates a coronary calcium score (1.29; 3.14) as the smokers which are the focus group of lung cancer screening are also more likely to have a high coronary calcium score which indicates risk of a heart attack (3.10). And finally, it was looked for lung emphysema, which means that the severity of the damage of the lung tissue is calculated (1.29; 3.14).

However, hardware and integration of the software into processes were different at all locations. At the clinic of R1, the AI software was integrated but there was no process defined on how the system is to be used (1.32). At the location of R2 and R5 the AI only ran when the radiologists started their diagnosis. This took one minute per function and was considered as dead time (2.65; 5.38). Also, the AI was integrated into the image display system so the AI had to run to look at the images and the output was always already displayed/overlayed on the images (2.59). At the locations of R3 and R4, the AI took between 5 to 20 minutes to run over the images (3.98; 4.10). However, this happened before the images were investigated by radiologists, so no time was lost (3.98; 4.10). However, the image evaluation at the location of R4 was web-based which means there was a waiting time of about a minute when accessing and saving the images (4.18). The overall time of the process of the diagnosis with AI was

depending on the severity of the case but usually ranged from 5 to 20 minutes, with an average of about 10 minutes (2.211; 3.73; 5.84).

4.2 Factors of User Acceptance

4.2.1 System Output

The System Output theme spans the factors *Output Quality*, *Results Demonstrability*, *Job Relevance*, *Locatability* and *Perceived Usefulness.*

4.2.1.1 Output Quality

The *Output Quality* factor addresses how system performance is evaluated and whether it generates complete and relevant output.

Radiologists perceive system performance to be good

The trials have not been running for long and validation of system output is difficult (1.42; 2.154). And while the system is flawed in some known ways (1.42; 2.73; 2.81; 2.89; 2.191; 3.38; 3.56; 5.54; 5.62), the radiologists have clearly commented positively on the performance of the system.

"Most radiologists take this relatively positively for such a new application as well as a clinical routine." (1.68)

"Yes, yes, it is better. Actually, it's a good system that has, it has its weaknesses" (2.89)

"Of course, there are still things where you might be able to develop more modern user interfaces, but basically, that works well. It's fast, it's available everywhere. So we're on a super good way there." (3.46)

Positive perception of system performance can partly be attributed to low expectations

While objectively the system performance appeared also reasonably good (4.12), it felt as if that did not quite justify the positive statements of the radiologists regarding system performance. This was attributed to worse prior experiences which might have lowered the expectations. R1 said that they do not even care so much for high accuracy anymore and that they just care for the system to function robustly (1.50). Also, they said that they are now very satisfied with the system for which they did not have too high requirements, but they also had bad experiences with other systems prior to that.

"So we now have a new system from [AI software, anonymized], which we tested in a clinical context and with which we were very satisfied. And that makes the interfaces the way we want them to be. The system we had before was more of an experimental one. That didn't make the interfaces the way we wanted them to be. … But that's what I meant earlier, it doesn't really matter how good the system is. So as far as measurement is concerned, not when it comes to that, … you shouldn't mark too many false positive findings. So we used to have a [AI software, anonymized] system that made dots everywhere. Of course, you can't work with that." (1.86)

R2 also mentioned that they had and have other systems which are not actually used in clinical routine because it is cumbersome to use them.

"In the past, there were also these measuring systems from [software, anonymized], but that was so cumbersome that you really had to sit down with everyone on an extra workstation and do that. We did that sometimes, but mostly not. I have to be honest." (2.87)

"So, in [city, anonymized] I don't think anyone uses it [AI], to be honest … So that's actually more work and you usually don't just do that." (2.160)

Therefore, the AI software for lung cancer screening was already outstanding in that way, that it actually facilitates the work and saves time (2.164).

Discrepancy between what is supposed to be done and what is supported by the software versus what is actually done by the radiologists
Additionally, the interviewees were in agreement that the system output is complete since the software does everything it is supposed to do.

"Nope. In fact, they can now deliver anything that interests us" (3.54)

"Right now, it's doing what we need it to do." (4.24)

"No, the software provides everything necessary, that's practically the content of the study, so that's what it covers. The content of the study is covered." (5.12)

However, there appeared to be a discrepancy between what is supposed to be done and what is done. Three functionalities are currently investigated in the lung cancer trials and the AI software is supporting those (1.29; 3.14). However, there are also more things which can be diagnosed using the CT scans and those are also looked for.

"So, I actually look at the bones, the soft tissues, that is, and there is information in there, for example, I look to see if the patient has a biliary obstruction if I see that, or I look to see if the patient, if the stomach, if he has a sliding stomach. That is also something that is relatively common, which people often don't know. This you really only see in such an examination and that would be stupid not to write it down." (2.221)

This is not within scope of the lung cancer screening and AI is not used in supporting these functions simply because R2 does not know whether the system can do that (2.221).

4.2.1.2 Result Demonstrability

This factor addresses how visible the advantages of system use are to the radiologists.

Benefits regarding System Output are not very visible
Benefits of system use included making results more objective (1.64; 1.66; 2.172) and decreasing inter-reader variability, however this is an aspect which was only observed when investigating multiple diagnoses over time and was not very visible in clinical routine. Furthermore, it is also questionable how much the increases in *Output Quality* (Section 4.2.1.1) are visible.

In lung cancer detection, whether a nodule was correctly identified will sometimes only be clear when patients come again after a few months and therefore in the clinical routine it is also not very visible how good an assessment of the AI is (2.146; 2.154). Therefore, advantages of the many additional small nodules which are shown of which some are correct, which are important to be identified for screening benefits, were not clearly visible when the radiologist worked (2.146; 2.154). Instead, more apparent to them was that many more small irrelevant nodules are recognized, which meant more work for them (2.146; 2.154). And additionally, the benefit of detecting big malign nodules which indicate cancer was far more apparent, but in detecting those the system was not so good and instead those could easily be found by the radiologists (2.150).

Time savings and work relief are clearly visible advantages
System Output was not perfect and therefore, sometimes, the system complicated work.

"But with every tenth patient, and with the bypasses or with the stents, which are then incorrectly detected, that also happens with every patient, that's perhaps so, yes, I guess on average with every tenth patient I have more work." (2.229)

However, the AI system was generally doing a large chunk of the radiologist's job for them and overall the radiologists also voiced that they know that the system saves time and also really helps and facilitates work (2.156; 2.164; 3.42; 4.10; 5.42; 5.62; 5.122).

"... and when the radiologist comes, he sees the analysis of the computer and he just has to check if that's right I would say." (4.10)

"No, I think everyone has now realized this, especially this topic of pulmonary artery detection. That is simply good. As I said, you have to sort out the false positives. But it is a real support and shortens the reporting process. There is absolutely no question about that. And other things, i.e. lung emphysema quantification, to do that manually, no one will even get the idea. Because that will take an hour to paint its contours somehow. This is only possible with automatic segmentation, which now always works with AI support." (3.42)

"So the software, let's say, the image data too, it's really excellent, and the image processing and the structure of the system, that's really excellent" (5.62)

4.2.1.3 Job Relevance

Job Relevance is about the perceived importance of the system for the radiologist's job. As relevant output and system relevance for the supported job are closely related, so is *Job Relevance* and *System Output* as defined in the research framework.

Radiologists have different preferences in their needs and system requirements
Whether the *System Output* was applicable to the radiologist's job was more difficult to assess as there were different opinions on what kind of output is needed.

"Ah, that's where the opinions of radiologists differ. So there are already things that are displayed that I don't need. But usually…" (1.52)

"You can still imagine various things that such a system can do. And there will be radiologists who want that and others who don't need it." (1.56)

This was also observed when talking about helpful functionalities which can support radiologists in general. R1 suggested a useful functionality where AI would incorporate more patient data into the diagnosis (1.10). R2 agreed that this was a useful integration (2.107), however R3 did not think it would make much sense or really help (3.108).

All important functionalities are included in the AI system

The AI software for lung cancer screening was assessed to have found a good way in-between. The respondents voiced various different ideas about further functionalities which could be included, but the importance of the missed functionalities was not assessed as critical but rather as nice-to-have.

"*No, actually I think that's pretty good. So that's kind of. Well, we have this Agatston score for calcium. It's relatively good. I find it a little bit vague … You could stratify it even more because the system actually knows whether it's an upper lobe or a lower lobe, and these are actually different clinical pictures, or they have different relevance for the patient. This could certainly be refined a bit more, but I would say that for a basic patient risk assessment, I think it's quite okay.*" (2.111)

"Yes, I think what you would still wish for is that you, that is, we still live very much with these textual findings. That is the way we communicate. And, um, that's usually free text. It is a great challenge to immediately make a structured text finding out of these AI findings, but that will help significantly further." (3.80)

"I think that's acceptable. So the software, let's say, the image data too, it's really excellent, and the image processing and the structure of the system, that's really excellent, but I think that there's still, for the future you wish maybe even sharper or possibilities of verification or control." (5.62)

Furthermore, the System Output was relevant and supported the full scope of the radiologist's job well during the lung cancer screening trials as discussed in Section 4.2.1.1 (3.54; 4.24; 5.12). R5 mentioned that they had a problem once where a critical functionality was missed, but this was then quickly integrated by the system provider (5.74). This highlights that the good *Relationship with Users* (Section 4.2.2.4) also positively affected the *Job Relevance*.

Impact of system weaknesses is perceived low

The system did not yet work perfectly and had some known systematic weaknesses (2.73; 2.81; 2.89; 2.146; 5.54; 5.58; 5.62), however, the system was designed in a way that the impact of its weaknesses was minimized. R2 described how the system was still very applicable even when nodule measurements were incorrect. They could use alternative suggested long nodule shapes suggested by the system and alternatively, the system could also be used to do the long nodule measurements without automation.

"There is a nodule measurement. You can make @Nodule and then you can measure it and you can also change it. I can always change the shape and change the size." (2.199)

"Exactly. I can go all the way out and remeasure it. I do that sometimes when I get all desperate. But I can also take the measurement it gives me and then in moderation, I can change that. It gives me plus/minus and I can do that in one direction twice in direction three times or so, I can then always play around like that, and then he gives me different shapes …" (2.203)

Furthermore, R5 described that some criteria which affect the likelihood of nodule malignancy were not well recognized by the AI (5.16). However, those criteria could also be assessed and added or changed by the radiologist (5.14; 5.18) and were then included in the malignancy assessment of the system (5.20; 5.22). Therefore, working with a system which had some weaknesses was not considered a problem by the radiologists.

System is seen in a supportive function and is not blamed for bad performance
There also appeared to be little to no blame on the system for its weaknesses. It appeared that this was at least partly due to the role which the radiologists attribute to the system. The system was seen merely as a supportive system which was applicable to the radiologist's job not by doing it for them, but by supporting it and the radiologists still took full responsibility for the results.

"That is very well accepted, because it is the right workflow because the radiologist is still the one who has to take responsibility for that and then also nodules, if he says, oh that is not recorded correctly or whatever and that is not a solid nodule. That wouldn't bother at all, the category of the nodule is wrong. Then he can rather change that manually …" (4.10)

It was also a known issue that the system has the tendency to detect too many lung nodules and there is the risk that for the final diagnosis of the radiologist this increases the false positive rate (1.86; 3.38; 3.42). Nevertheless, R2 took the responsibility to sort out the false positives (2.150) and R5 did not even consider this an issue of the system as the final interpretation is their thing.

"That is my interpretation and so to that extent this is called false-positive, false-negative, that is not relevant for this software now." (5.124)

This interpretation is in line with the statement of R1 who did not think that system accuracy is an important criterion.

"I'm always of the opinion that the systems don't have to be super accurate, but the systems have to be robust" (1.50)

System enables combination of the strengths of system and radiologist
And while the shortcomings of the system appeared not to negatively affect the degree of Job Relevance, some strong arguments for the importance of the system's benefits exist. The system decreased inter-reader variability (1.66) and the diagnosis became more objective (1.64). In the scope of lung cancer screening the same patient will be looked at multiple times over the years by different radiologists and the findings become more standardized and can be compared better when the same AI is used for the diagnosis.

"So, I think for this screening program it's very important that you use something standardized. Somehow, if, if then you should use the same program for every patient who comes … also always remembering that yes not always the same radiologist is available the next time when the patient comes" (2.172)

"I believe that it will be important to have a system that is as uniform as possible in Germany. As I said earlier, if you have four different systems in Germany, different things come out.

Which is not good, especially if the participants change their place of residence and do their examination." (1.83)

High degree of Job Relevance of the AI system is limited to the precise purpose of lung cancer screening

It is also interesting and important to highlight that the high degree of applicability of the system applies to lung cancer screening and does not necessarily cover lung cancer detection in general. R2 stated that they would not normally use AI for lung cancer detection but they are likely going to use it for lung cancer screening.

"[I]f I get a patient like that and the question is, a focus was seen and he coughs too much or something and look at it to see if he has something, then I don't use AI." (2.158)

"[I]f now people would come, just for that [lung cancer screening], then I think I would do that [use the system], because that also really facilitates the work." (2.164)

This highlights that the benefits of the AI system which are only applicable to lung cancer screening are of high significance in the assessment of whether AI is applicable to the radiologist's job and should be used. Alongside standardization, system strengths named by R2, which are mostly only useful in lung cancer screening, are the detection of small nodules and time savings in finding those (2.148; 2.156). And with this interpretation, the statement above is no surprise. In lung cancer detection it is more important to pay attention to rather larger nodules, which the AI was actually not very good at (2.148). As pointed out earlier, this shortcoming is not relevant to the applicability of the system for lung cancer screening as those large nodules are rarely overlooked by the radiologists anyway (2.148). However, it just does not make the system very useful for them in lung cancer detection.

"I don't even know if I would have missed a single really bad finding so far." (2.150)

The importance of the time savings was also highlighted by R5 who initially stated that they think that to *"link [lung cancer screening] to AI … is not absolutely necessary"* (5.8). However, later they responded to the question whether they thought that AI was necessary to deal with the workload (5.131), that this depends on how many people would have to be screened.

"They have a long-term smoker and then simply send him to a low-dose CT. Take a look at him whether, it is already being practiced, that is not something that is now being kicked off in one stroke. Um, then the other thing, of course, if you institutionalize this correctly … [a]nd that's exactly what you have to prepare for, that's important. Yes." (5.132)

4.2.1.4 Locatability

The *Locatability* factor is about the ease of understanding the system output.

There are minor Locatability issues and room for improvement exists

The system had a variety of functions and outputs (3.60) and therefore, it is not surprising that when various functions were discussed in detail not all system outputs were found to be completely understandable to the radiologists. One concern was raised by R2 who mentioned that they sometimes have to add information to the automated report on why possibly the AI system assessed the lung emphysema incorrectly, which hints at a lack of understandability of the output (2.209). This is similarly assessed by R5.

"Exactly, that's exactly what I can do with emphysema when calculating emphysema, because we had the impression that emphysema is somewhat underestimated. That there is a systematic mistake in it, but proving that now is not easy at all." (5.60)

Furthermore, R5 also addressed that they cannot assess how accurate the key figures of the system output regarding the coronary calcium calculation are and that they are dependent on the machine there.

"And in the other two fields, where this is the coronary calcium calculation, the machine only provides me with figures, I accept them and how good the software is, is very difficult for me to assess. I don't know that, of course, but I'm basically dependent on the machine." (5.58)

Overall the system is very transparent and Locatability is high
However, aside from those two specific cases, no other concerns were raised regarding understandability and instead it became clear that the system is showing clearly how it is doing its measurements and arriving at key findings (1.62; 2.191; 5.58).

"And then you look at what it has measured there and sometimes you have to correct it a little bit because it measures some caesurae or something adjacent … And then, um, and recently it also gives us histograms, which show how dense the focus is, so to speak, or how much soft tissue, how much calcium, so how much away from the visual aspect? Is it a dense focus? Is there calcium in there? So is that natural, if it's really calcium in it, then it indicates that, and then it's automatically judged a 1 focal, so it's judged benign. But if it's a little bit bright, this is something you can see, if it's more of a necrotic focus, if it's more of a dense focus somehow, that's shown by the histogram." (2.81)

And this also highlighted what has already been discussed in Section 4.2.1.3; the system output was flawed sometimes, but the transparency of the system enabled error detection and correction (2.199; 5.14). R5 also summed everything up very well.
"I think that's acceptable. So the software, let's say, the image data too, it's really excellent, and the image processing and the structure of the system, that's really excellent, but I think that there's still, for the future you wish maybe even sharper or possibilities of verification or control." (5.62)

4.2.1.5 Perceived Usefulness

Perceived Usefulness covers the perception of the radiologists whether system use will enhance their job performance. Two core components of job performance were mentioned in the interviews: Output quality and efficiency.

Radiologists perceive the output quality to improve by system use
The radiologists named many quality aspects which are improved by system use. First, their diagnosis became more objective (1.64) which also decreased inter-reader variability (1.66; 2.172).

Additionally, it enabled more sophisticated diagnoses by making complex aspects of the diagnosis automated and easier.

"[O]nly 20% of radiologists do volume-based evaluations and I have to say quite honestly, we in the clinic also do not do volume-based evaluations, when it comes to the … tumor follow-up, for example, we only do two, three parameters, that's it. But since we know

scientifically that volume is much more accurate, because differences are better recognized than with the diameter and so it improves the quality." (4.14)

The diagnosis process also incorporated various measurements which needed to be done and the computer was found to be very good at those measurements.

"But actually, it is so, that I think a few times with these emphysema bubbles somehow, okay, if one had estimated that, so I would have estimated that maybe differently, so if I had seen that now in the clinical context, but if one measure that exactly, it can be that it [the AI system] is right." (2.91)

"I think that's good in principle because no one likes to evaluate scores or count lung nodules or anything, or measure things. To be quite honest, it's all a bit of a dork's job, to put it kindly, somehow, that's, the computer can do that better and it should" (2.188)

Furthermore, a difficult aspect for the radiologists was also finding small nodules and they also recognized that the AI is better at finding those.

"So it sees more lung nodules than I do because I just click away a lot of lung nodules in my subconscious somehow" (2.148)

R3 also mentioned that compared to radiologists the AI system is *"much more sensitive, but not super specific yet"* (3.38), meaning that the system finds more small nodules, but also finds more things, which are actually not nodules. And it appeared that the radiologists perceived this system to mostly only enhance their performance as they did not think that the false positives which the system also finds many of, would affect their diagnosis.
"[T]he machine recognizes the nodule but the nodule interpretation is my job." (5.126)

"No, so we would sort out the false-positives when checking the findings of AI. In other words, this does not actually result in any disadvantage for patients." (3.76)

Quality could be improved even further
While many benefits of the system are recognized already which improve the quality of the diagnosis, it should also be noted that there is still some good potential for further improvements.

"It could always be better." (1.42)

The system accuracy in lung nodule detection was about 92% (4.12), lung emphysema was possibly systematically underestimated (5.60), inclusion and evaluation of additional data could improve risk profiles and diagnosis (1.10; 2.111) and with some more system flexibility for the findings report, the usefulness of that report would also increase (3.80).

AI also saves a lot of time and increases efficiency
There was also strong agreement among the radiologists that the process which involved AI became much quicker (1.64; 2.229; 3.32; 4.10; 5.122).

"So I think it's going to be a lot faster" (1.64)

"Yes, I think that then [without AI] you wouldn't have to think about whether you can somehow get towards 100 CTs. So that's only possible with IT support." (3.32)

"Yes, I mean, considering how much information this software provides, it's definitely a time saver." (5.122)

Most interesting about the agreement about the time saved is that the five radiologists span four different AI integrations. R2 and R5 also strongly voiced that AI is a time-saver (2.229; 5.122) although, as shortly discussed in Section 4.1, compared to the other AI integrations in the HANSE trial of R3 and R4 (3.98; 4.10; 4.18), a lot of time is still unnecessarily lost in the process at their location.

"[W]hen this lung screen program then opens, that takes at least a minute, maybe even two, that there you sit and wait." (2.65)

"No, no. The loading times, that is, that matters. That costs time." (5.32)
"I would say about a minute." (5.36)
"[P]er function." (5.38)
"Yes that is, that is unused waiting time" (5.40)
"So everyone is a bit annoyed about the loading times." (5.86)

Therefore, even not so good integrations still save a lot of time as the AI software appears to be a very good time saver.

The entire diagnosis process also took on average ten minutes at the location of R2 and R5 (2.211; 5.84) and apparently multiple minutes were currently just spent waiting on the loading times. This signals that the loading times of the integration at the location of R2 and R5 made up a significant proportion of the overall diagnosis time and that consequently, process efficiency with a better AI integration could be much higher.

Time savings are more important to the radiologists than increases in quality
The importance of a time-efficient integration is further stressed, considering that the radiologists appeared to value the time savings more than the gains in quality. This evaluation is based on a couple of remarks.

First, R1 voiced their opinion *"that the systems don't have to be super accurate"* (1.50), hinting that not too much value is attached to quality. They also added that however *"the systems have to be robust"* (1.50), showing that more value is attached to operationality, which is closely connected to the time savings, as it also regarded that with lung cancer screening the workload on radiologists will highly increase (1.79; 2.10; 4.6) and AI is needed to deal with that (1.77; 5.132).

Furthermore, R2 also mentioned that normally they and other radiologists do not use AI for their diagnosis.

"So, in [city, anonymized] I don't think anyone uses it [AI], to be honest." (2.160)

And the reason for why AI is not used is that it is simply more time-consuming (2.161; 2.162). R2 was asked for reasons which motivate them to use AI in the scope of lung cancer screening and they said that it saves them work (2.158). After, R2 was asked whether they think they would actually use AI when lung cancer screening is coming.

"[I]f now people would come, just for that [lung cancer screening], then I think I would do that, because that also really facilitates the work." (2.164)

It is notable that quality improvements were not mentioned as motivation to use AI and instead the time savings were highlighted.

AI makes the radiologists' job easier and interesting which is an important motivational factor
Another understanding from the statements of R2 we developed was that the work which needs to be done when working together with AI was generally less exhausting and easier. This second understanding was also in line with another comment of R2 shortly after.

"[N]o one likes to evaluate scores or count lung nodules or anything, or measure things. To be quite honest, it's all a bit of a dork's job, to put it kindly, somehow, that's, the computer can do that better and it should" (2.188)

When R2 was asked in member checking what they meant by AI *"saves them work"* (2.158) and *"facilitates the work"* (2.164), they confirmed that they meant not only time savings, but also that they consider AI valuable since it takes over the monotonous tasks.

4.2.2 Usability Experience

The theme *Usability Experience* includes the factors *Subjective Voluntarines, Perceived Ease of Use/ Training, Experience, Relationship with Users.*

4.2.2.1 Subjective Voluntariness
The factor *Subjective Voluntariness* includes the voluntariness of using the introduced technology. Therefore, we investigated whether the use of AI systems is mandatory for radiologists in lung cancer screening and whether they perceive pressure to use the AI system and whether there are specific guidelines for their use. *Subjective Voluntariness* also includes the user's perception of what people they value think about the technology and the resulting social pressure. Therefore, it was also investigated whether radiologists use AI to satisfy others or rely on their opinion on the use of AI.

Despite objective pressure to use the system, system use is perceived voluntarily
R1 (1.32) mentioned that there is *"no finished process of what you are allowed to and what you are not allowed to do"*. R3 (3.40) agreed that they can decide on their own whether they use the AI system or not. R2 (2.158) emphasized that they use the AI system voluntarily because it saves time and not because it is prescribed and that they think they would also use it when the program officially starts (2.164). However, R2 (2.168; 2.170) also stated that the software is mandatory in the HANSE study, to which R5 (5.100) agreed, adding that the use of the AI system is part of the HANSE study and is mandatory at all sites involved in the study. As this seemed to contradict the first statement of R2 and the statement of R3, who is also part of the HANSE study, R2 was asked during the member checking about the voluntariness of the use of AI and confirmed that there was a prescribed process of how to use the AI and that the LDCT images could not be viewed without using the AI software. Therefore, based on the member checking and the statements of R1 (1.32), R2 (2.158), and R3 (3.40), it was concluded that although the use of the AI software is mandatory according to the study guidelines, the radiologists still perceive voluntariness to use the systems according to their preferences. This can be attributed to the supporting function of AI which only suggests diagnosis and interpretation is still fully considered by radiologists as their task (5.124).

R2 (2.158; 2.160) said radiologists at their location typically do not use AI unless a referring doctor requests it. Thus, in the context of radiologists, this pressure seemed to come particularly from other physicians demanding diagnosis by AI.

"Then the pneumologists always want a machine evaluation … but I believe that otherwise, so if I get a patient like that … then I don't use AI." (2.158)

There are referrers who attach importance to that [AI diagnosis]. (2.160)

Radiologists R1 (1.68), R4 (4.16), and R5 (5.102) felt that their colleagues are well accepting and satisfied with the AI software. However, R1 (1.70) also emphasized that the opinions of the radiologists about AI vary.

"Most radiologists therefore take this relatively positively for such a new application as well as a clinical routine. Yes." (1.68)

"And otherwise the IT affinity of radiologists is, of course, different and some have more fun than others dealing with such topics and then kneeling in and working their way through them." (1.70)

Therefore, despite the mandated use of the AI software and requests from other physicians to use AI, the radiologists felt little pressure to use the system.

Radiologists demand standardized processes which mandate the use of AI
R1 was convinced that radiologists must want to use AI on their own to get along well with it. R3 (3.40) stated that the prescribed use of an AI would be something that did not exist before. However, the voluntariness of use was not demanded by the radiologists R1 (1.34) and R2 (2.172; 2.174). Instead, they argued in favor of standardized processes that precisely regulate use as this is relevant to the success of lung cancer screening (Section 4.2.1.3).

"Yes, I think we need an SOP [standard operating procedure] that says what to use in what order and how." (1.34)

"[I]t should just be then that when a patient comes in for screening, they should use the AI, and the next time they come in, the same AI should be used." (2.174)

4.2.2.2 Perceived Ease of Use/ Training

The factor *Perceived Ease of Use/ Training* combines the perception of how easy a system can be used effectively with the training provided for this system. It was investigated whether the radiologists found it difficult to use the AI systems and whether they received training on how to use the AI systems and how helpful this was.

Radiologists' limited time requires intuitive AI systems
The AI system was generally perceived as easy to learn and intuitive by the radiologists R2 (2.239; 2.241) and R4 (4.16). R2 also emphasized that it is important for them that the system is intuitive to use, as they do not have time to spend a long time learning how to use the system (2.237). However, there seemed to be differences between the ease of use of basic functions of the AI software and additional features (5.106).

No, you learn that within a week, faster I think (2.241)

"So it's those details, details that you actually have to know, how to document something again, or, in that respect, and you then have to know that, you can probably acquire the basics relatively quickly" (5.106)

The AI software was designed for being used worldwide by different programs and therefore had a lot of different functions which makes it difficult to use comprehensively and as a result, not all aspects of the AI software were perceived as intuitive (3.60). Therefore, radiologists found training on how to properly use the software necessary (3.60; 5.106).

"I actually always use the same and for me personally, it would be enough if the software could do this one way, and then it would probably be much more intuitive and easier to use. But because it offers so many options, most IT solutions have too many buttons, too many options, and that's why you simply need an introduction." (3.60)

Continuous training maintains Perceived Usefulness and Ease of Use

It was reported that training was provided at the beginning, but this was a one-time course (2.123; 2.233; 3.62; 5.104). This initial training was found to be helpful and radiologists were shown how to adapt the AI software to their own workflows (3.62). R2 did not seem to be satisfied with this approach of having only one training, because it does not allow them to remember everything about the use of AI (2.123). They also had to summarize their training on their own and had to write their own manual (2.237).

"[T]he system was introduced, you got a password and then the basic features were explained, I guess. Of course, that happened once before and then never again …" (2.233)

"There is also such an operating manual, but I must honestly say, I had that at the beginning once, we have all sat down ourselves. We threw everything we remembered into it and tried it out a bit" (2.237)

It has been assumed that when lung cancer screening officially starts in Germany, training in the use of AI software will be a necessary part of the certification required by radiologists who will perform the screening (1.79).

Unexpected behavior and non-customizability hinder the adoption

Questions regarding the factor *Perceived Ease of Use/ Training* also helped us to analyze how the radiologists feel about the usability of the AI software. Ease of use, in general, was found to be good (2.89; 3.46; 4.16; 5.98), especially compared to previous systems the radiologists have used (2.88; 2.89). However, the radiologists were also not yet completely convinced and pointed out several minor problems with the user-friendliness of the AI systems.

"Of course, there are still things where you might be able to develop more modern user interfaces, but basically, that works well. It's fast, it's available everywhere. So we're on a super good way there." (3.46)

"[A]ctually, it's a good system that has, it has its weaknesses, somehow …" (2.89)

It was mentioned that the user interface could be optimized (3.46; 3.48). R2 stated that *"sometimes you have to fight the system a bit"* (2.97). For example, the measuring function was perceived as complicated to use (2.119). The loading times of the software as well as the switching between different applications and workstations were found to be annoying for the

radiologists and therefore hindered the Perceived Ease of Use (2.18; 2.69; 2.137; 2.143; 3.98; 4.18; 5.32; 5.40; 5.86).

"So everyone is a bit annoyed about the loading times." (5.86)

Also, radiologists were bothered by settings in AI systems that they would not expect (2.69; 2.73). For example, the system expected the radiologist to submit a diagnosis after the handling of the first function, when the radiologist is not done yet and the system then opens a window that interferes and must be clicked away before the system can be used further (2.69). Also, some keys were unusually assigned, requiring radiologists to adjust between using the AI system and other systems, which costs time and leads to operator errors (2.119; 2.140; 5.88; 5.90).

However, there were also some setting options for radiologists to adapt the AI systems to their needs and expectations (4.14). This can preserve the autonomy of the radiologists. Furthermore, it can increase the *Job Relevance* if the radiologists can adjust the software to their needs (Section 4.2.1.3). In order for these setting options to be understood and used by all, it would probably require more extensive training for the software.

"You can adjust the threshold, and you'll notice that it's also very interactive. You can set the threshold a bit higher, manually, if the radiologist doesn't like it. And therefore, the radiologist has full control." (4.14)

Fun from using AI increases happiness but is inhibited by AI errors
In the discussions with the radiologists, it had become apparent that some radiologists enjoyed using the software, which we attributed to the *Perceived Ease of Use/Training* factor, as enjoyment can facilitate the use of the AI software. R2 (2.195) and R5 (5.102) mentioned that AI software is fun to use. R2 (2.178) "*always find[s] that kind of thing exciting*" which is an interesting factor for the early adoption of AI software.

However, opinions differed and not everyone had so much fun using AI (1.70). Due to many errors in the software that had to be tediously corrected manually, the fun can also quickly disappear (2.227).

Fun seemed to be a relevant factor for the radiologists as they wanted to keep the tasks they enjoy and the tedious tasks could be taken over by the AI (2.180).

4.2.2.3 Experience

The *Experience* factor was used to investigate whether radiologists who have already had contact with similar AI systems have fewer problems using the software and are more likely to accept it. Furthermore, it was investigated whether the use over time affects the other technology acceptance factors.

Missing technology Experience needs to be compensated by training
The *Experience* of radiologists in terms of computer affinity and AI usage differed (2.121; 2.87; 5.2), so it can be expected that when lung cancer screening is rolled out nationwide, some radiologists will do well with the software and others less well.

"I just click on everything once and see what happens. But I'll say when people, just older people, maybe you see, what am I saying, older people. But then there are actually people

who have even less experience with computers. For them, it's perhaps not quite so intuitive." (2.121)

The missing *Experience* of some radiologists must be compensated by extensive training, both in the use of AI software and in the use of technology in general. R1 (1.72) expected that the systems will be taught in a school-like setting through explanation and training, which could allow overcoming the lack of *Experience*.

Previous experience cannot be fully exploited by unusual configurations
Experience with similar systems was no guarantee for *Perceived Ease of Use* as the radiologists have reported that the new AI system differed from previous systems (2.69; 2.73; 2.119; 2.137; 3.60; 5.88; 5.90). Previous *Experience* with similar systems had a positive impact on Ease of Use, but because these systems were handled differently, radiologists will need to get used to using AI software for lung cancer screening (2.137). Also, there were various ways in the software to get a specific solution of which only one way is used as it is more convenient (3.60). As a result, not all of the systems' functions were explored. This underscored the need for extensive training to educate radiologists on all the functions of the system so that they know the most appropriate and time-efficient function for each task.

4.2.2.4 Relationship with Users

The factor *Relationship with Users* includes the dedication of interest of the IS department to provide the user of the software with everything they need. To investigate this factor the radiologists were asked whether there was helpful technical support available when they had issues using the software and whether their needs are taken into account for improving the AI software further.

Technical support is good but it can be hard to communicate problems
The radiologists agreed that there was technical support, which responded quickly and solved their problems well (1.70; 2.47; 2.125; 5.110). However, there were still problems with the support for questions that go beyond the technical and concern the use of the software in the daily work of radiologists. R5 (5.112) stated that "*it's not that easy to communicate the problem and that it's not really understood*".

"All right, there is a point of contact for technical problems. I would say that works very well and in the application, these are questions of the application, I would say it works medium. I would say medium." (5.110)

"Yes, yes, there is, there is. And you can also ask questions. If there's something we didn't know, there's an IT person who's responsible for us, you can just write to him and ask. Well, he's very nice." (2.125)

It was also emphasized that radiologists must accept the software, or they will not use the system properly despite the support, showing that *Relationship with Users* alone was not a guarantee of acceptance of the software.

"*So first of all, I'm sure the radiologists must want that, yes. So no matter which IT system you use for diagnosis, you must want to make friends with the new system. Then you can understand the help you're getting.*" (1.70)

Continuous improvements of AI have favored the acceptance
The AI software for lung cancer screening has been continuously developed. New features have been introduced as well as bugs have been fixed (2.79; 2.235; 2.237; 3.68; 5.72). These updates have also been communicated to radiologists so that they are aware of the new features of the software and how to use them (2.237). The radiologists were very satisfied with the communication with the software provider, as they were responsive to their problems and requests and quickly implemented new features or bug fixes (3.68; 5.72). The introduction of features that were requested by the radiologists and were critical in their view, positively influenced the radiologists *Perceived Usefulness* and *Job Relevance* (Section 4.2.1.3). This was something special for them, that during the study the software was continuously adjusted and improved to their needs (5.72; 5.74).

"A nice lady came along … because then the system was introduced, you got a password and then the basic features were explained … and then this woman [name, anonymized] who trained us, she came a few more times" (2.233)

"To any, if there were new features, for example, there were a few new features added at the beginning, which she explained again." (2.235)

"I did it more when there were problems because again something didn't work … and that [IT support] has always worked very well. There is a gentleman who takes care of it." (2.129)

"Yes. Yes. Yes. Totally. Yes, that is really great and I think that [functionality of the software] is very good. So this is another interesting point that we have actually still been working on and improving since the introduction of this software. We kept reporting to [country, anonymized] that we would like something else, and can you change that again? And that the software has actually been continuously improved and edited. I think that's something, actually something so special, that during a study, so to speak, it was still being worked on." (5.72)

4.2.3 Operationality

The *Operationality* theme covers the factors *Production Timelines* and *System Reliability*.

4.2.3.1 Production Timeliness
The factor *Production Timeliness* investigated whether IS could meet the operational time and efficiency expectations.

Production Timeliness is decent
There were no requirements voiced when a scan needs to be evaluated. R3 said that they were trying to evaluate the scans the same day and it appeared they were mostly successful with that (3.20; 3.22; 3.26). R2 also said that evaluating scans on a daily basis is important, but it appeared that this is not something that is really done or even attempted at their location.

"[S]ometimes we really do send out the findings almost on a daily basis. I think that's also relatively important." (2.47)
Exactly that will be evaluated later. This is then run during the day and if it's good, then I might also sit on it online. If that's the case, I only work on Thursday and Friday, so mostly on Thursday, when I work on Thursday and Friday, I evaluate at least the results from Wednesday, Thursday, Friday, and sometimes even the results from Tuesday." (2.45)

Production Timeliness is not concerning
Also, achieving satisfiable production turnover did not appear to be very problematic or concerning. R3 said that they do *"about 25 such CTs per day for the patients"* (3.20) *"[a]nd that costs a good half day of work, at least"* (3.22). One radiologist is then doing this full time that day and *"they're moderately busy then"* (3.26).

R2 also explained they are not under time pressure and they can just work off the workload on any two days.

"Most of the time I'm there for two full days. Of course, what do I know, if I work on Thursday, and Friday, then I can work on Saturday and Sunday as well, it's not like that. That doesn't go away" (2.14)

Furthermore, R5 described how they initially had technical problems once in a while (5.116) which got solved within two days (5.118) and that this is fine since it is a study after all (5.120).

4.2.3.2 Systems Reliability
This section looks at the reliability of the AI software for operationality.

Systems Reliability requirements are not very demanding
The importance of system reliability was highlighted by R1 who said that *"systems don't have to be super accurate, but the systems have to be robust"* (1.50) and it is important that *"they can survive in the real world"* (1.50). However, the requirements to achieve satisfactory *System Reliability* were not too high as there are no emergency cases in screening that need to be dealt with (4.20) and scans can also be looked at at a later day so it is not an issue if the system is not available for some time (2.14). However, it also must be noted that the lower requirements regarding system reliability can also be partly attributed to the study setting. R5 stated that IT issues were solved within two days (5.118) and when asked whether this is acceptable in a lung cancer screening setting they said *"[i]t is a study after all. Yes."* (5.120).

Initial system issues were overcome and the system works reliably now
How severe a system downtime affects operationality also depends on the integration of the AI software. While lung cancer screening diagnosis could also be done in reasonable time by the radiologists without AI in case of AI system failure (5.8), the AI software might be integrated in a way which makes this impossible as AI analysis is a required upstream process (2.55; 3.96; 4.10).

"So, you don't have access to the [AI software, anonymized] and … you couldn't diagnose, which was a bit painful." (2.51)

Anyway, no location appeared to struggle with *System Reliability*. All kinds of problems were had in the beginning when the AI integration was novel and some of them were also severe (2.47; 3.64). However, despite some trouble in the beginning no significant problems were faced anymore, and also, it was made clear that IT support works very well in case there is a system issue (2.125; 2.135; 3.68; 4.26; 5.118).

"And yes, there were … such a few hick-ups, but on the whole, it worked." (3.64)

"I have to say, the first cycle was terrible. Nothing worked at all. The system broke down all the time and we had big server problems actually. That was stupid because then we couldn't

notify the patients online at all. That was a bit difficult. But since then, it's actually working well" (2.47)

"If there's something we didn't know, there's an IT person who's responsible for us, you can just write to him and ask. Well, he's very nice." (2.125)

5 Discussion

5.1 Factors of User Acceptance

In the following sections, we discuss our results and compare them with our expectations that we developed based on the literature on user acceptance, the TAM, and the TTF model.

5.1.1 System Output

The *System Output* theme consisted of various factors. The theme covered the quality and completeness of the output (Section 4.2.1.1), whether it was understandable (Section 4.2.1.4) and whether the advantages were visible to the radiologist (Section 4.2.1.2). Furthermore, the perceived relevance of the system for the radiologist's job and the *Perceived Usefulness* (Section 4.2.1.5) were investigated.

As suggested by the literature, a lack of comprehensibility of the AI system in use was found and support for detection of lung cancer was restricted to the two core functions discussed in Section 2.1.2.1 and Section 2.1.2.2. It is interesting to note that two radiologists wished for integration of further AI functionality to improve patient's risk profile (Section 4.2.1.3) which is a function which has also been developed and looked into (Huang et al., 2019). However, another radiologist stated that they did not think that such functionality would help (Section 4.2.1.3). It became apparent that radiologists had different preferences in what was needed for their job and that this would also make it more difficult for the *System Output* to be evaluated as relevant and complete by all radiologists (Section 4.2.1.3). Nevertheless, and at least partly due to a good *Relationship with Users* (Section 4.2.1.3; Section 4.2.2.4), all important functions which were needed by the radiologists were covered by the system (Section 4.2.1.3). However, in one case, the *System Output* was not complete and did not provide all necessary information because a discrepancy between what was supposed to be done, and what was actually done by the radiologist, was found (Section 4.2.1.1). R2 described how they performed additional diagnosis steps and that they also did not use AI for this because they simply did not know whether the system had functions which would support that (2.221).

This finding was presented to our interviewees in member checking and they all agreed that this is how radiologists behave and do their diagnoses. This discrepancy reduces the relevance of AI systems to radiologist's jobs, reduces the positive impact systems can have and presents a critical issue. This also raises the relevance of the recommendation of R1 who called for a strict standard operating procedure for the lung cancer screening process (1.34) which is also found in the literature to be useful before implementation of advanced IT systems (Dünnebeil et al., 2012).

In recent years the literature has described many improvements of AI for lung cancer detection with CNN and that good and satisfactory results are reached (Section 2.1.2.1; Section 2.1.2.2). Although not too much information was provided on system performance in the interviews, it appeared that the AI software achieves results similar to those described in the literature (Section 2.1.2.1; 4.12). Nevertheless, it was unclear how this would be perceived by radiologists and the literature suggested that although system performance was better than theirs, radiologists may not be very tolerant of mistakes of the AI (Reis et al., 2020).

However, radiologists were very happy with system performance, and it appeared that actually radiologists did not have that high expectations towards system performance (Section 4.2.1.1). These lower expectations seemed to be rooted in prior experiences where systems did not perform well, and system integration was not successful (Section 4.2.1.1). Since the interviewees were all radiologists with many years of experience, they all had experienced their fair share of bad system integrations. This also suggests that by the majority of radiologists, who are less senior than the interviewees and have had less experience with previous unsuccessful AI system integrations, the perceived system performance might not be evaluated as positively as by the interviewees. And less senior radiologists might also develop an aversion to AI for another reason. Although due to task complexity, AI is unlikely to replace radiologists and instead it will change the workflows (Lebovitz, Levina & Lifshitz-Assaf, 2021; Mathew, David & Mathew, 2020), this is perceived differently by the radiologists (2.184; 2.186), especially in the case of lung cancer screening (3.104). This aversion was not observed in our research, but it needs to be noted that radiologists were senior, and they believe that this is not applicable to their remaining working time (2.184). Anyway, younger radiologists may perceive the AI systems as a threat to their job and display a different attitude. Further research should also cover more junior radiologists to investigate this more closely.

While the *System Output* is perceived as good, it is definitely not perfect and has some known weaknesses (Section 4.2.1.1; Section 4.2.1.5). However, the system is very transparent and the output is very well understood (Section 4.2.1.4). Various benefits of high *Locatability* are mentioned in the literature, as it is is important for a good TTF, mitigates challenges with AI trust (Lockey et al., 2021), can be expected to generally facilitate IT acceptance (Dishaw & Strong, 1999) and also lead to better performance (Goodhue & Thompson, 1995). In our findings we were able to identify further positive effects which were at least partly attributable to the *Locatability*. The high *Locatability* makes it easy to detect and correct mistakes and so the impact of the systems' weaknesses can be reduced (Section 4.2.1.3). Also, the radiologists do not blame the AI software and instead take responsibility for the interpretation and the final output (Section 4.2.1.3). This behavior is likely at least enabled by the high *Locatability* as Reis et al. (2020) argued that without the ability to understand or change the diagnosis suggested by the AI system, the systems can be considered rather obstructive than actually supportive and furthermore, radiologists can feel a loss of control which can lead to system aversion (Kauffmann, 2021, Khanijahani et al., 2022; Reis et al., 2020).

The AI system also allows additional useful information to be added which is also taken into consideration for the System Output, so the system always has a high degree of applicability to the radiologist’s job (Section 4.2.1.3). Additionally, the system has strengths in important functionalities which cannot be easily done by the radiologists without system support (Section 4.2.1.3). Therefore, the AI system and radiologists are able to play to their strengths and make up a strong complement (Section 4.2.1.3). It needs to be noted that those strengths have limited applicability to lung cancer detection in general and it is unclear whether the AI software would be used in a daily clinical context for that (Section 4.2.1.3). Anyhow, the system enables the detection of small nodules, objectivity, standardization and efficiency, which are all essential to make lung cancer screening feasible and also useful in Germany and consequently, *Job Relevance* of the system for lung cancer screening is also high (Section 4.2.1.3). This is in line with the expectations which were set by the literature which clearly emphasizes the importance of AI to make lung cancer screening feasible (Becker et al., 2012; Bogoni et al., 2012; Bonavita et al., 2020; Fischer et al., 2020; Liu et al., 2021). It could have been expected that the *Perceived Usefulness* of the system is found to be equally high as normally performance expectancy is found to have a significant effect on the technology acceptance

(Khanijahani et al., 2022). However, those factors were not found in the responses of radiologists when asked what motivates them to use AI (4.2.1.5). Instead, radiologists named time savings and easing of work, the two benefits which were most visible and useful to them (Section 4.2.1.2; Section 4.2.1.5). A similar relationship is also described by Venkatesh and Davis (2000) regarding *Result Demonstrability* that "if a system produces effective job-relevant results desired by a user, but does so in an obscure fashion, users of the system are unlikely to understand how useful such a system really is" (Venkatesh & Davis, 2000, p.192). Therefore, because the increases in *Output Quality*, objectivity and standardization, are not very visible in the daily clinical routine (Section 4.2.1.2) they are also not as influential for *Perceived Usefulness*. More visible in the daily routine are the system benefits which facilitate work and lower the time needed for diagnosis (Section 4.2.1.2). It is therefore also no surprise that it was found that those two very visible advantages also emerged as the driving factors of *Perceived Usefulness* (Khanijahani et al., 2022; Section 4.2.1.5). This suggests that *Result Demonstrability* is a central factor which influences *Perceived Usefulness* which was named as motivation by the radiologists to use AI.

The low *Perceived Usefulness* of the less visible factors which are also deemed essential for lung cancer screening signals a discrepancy that the system's usefulness is perceived lower than it actually is. *Perceived Usefulness* is considered by Venkatesh and Davis (2000) the most important factor that positively affects Behavioural Intention and therefore Actual Use and therefore, a lack of *Perceived Usefulness* is quite impactful and concerning. Venkatesh and Davis (2000) state that "even effective systems can fail to garner user acceptance if people have difficulty attributing gains in their job performance specifically to their use of the system" (Venkatesh & Davis, 2000, p.192). The radiologists were presented with this finding in member checking and R1 and R2 agreed with this finding, while R3 called for more research into this. R4 did not respond and R5 did not comment on this finding.

5.1.2 Usability Experience

The *Usability Experience* covered the factors *Subjective Voluntariness*, *Perceived Ease of Use/Training*, *Experience* and *Relationship with Users*. This covers perceived pressure felt by radiologists to use the system, how easy the system can be used, how experience over time affects the radiologists and how the relation between radiologists and IS provider affects technology acceptance.

It was found that radiologists wanted to use the AI software voluntarily as they saw various benefits in its use. (Section 4.2.2.1). They also did not feel pressured by others to use AI technology. Therefore, the *Subjective Voluntariness* did not significantly influence the *Intention to Use* AI (Section 4.2.2.1). This finding was not expected for a mandatory software as Venkatesh and Davis (2000, p.188) found that "subjective norm had a significant effect on intention in mandatory settings but not in voluntary settings". However, this can be explained since the radiologists perceived the use of the software as voluntary. Previous studies concerning the AI acceptance of health care professionals would have suggested different findings regarding the impact of social factors on AI acceptance (Fan et al. 2020; Ye et al., 2019). Since both studies were conducted in China, the results may not be comparable due to cultural differences, as Eitle and Buxmann (2020) found that social factors and cultural norms have an impact on the acceptance of AI technology. However, as both studies were conducted in the same country and resulted in differences in the effect of social factors on the AI acceptance, it can be expected that the impact of social norms depends on the setting of the project and the

pressure from the supervisors. Also, if participants come at their own request and are not prescribed lung cancer screening by another physician who might require an AI diagnosis (1.34; 2.158; 2.160), the *Subjective Voluntariness* factor becomes more negligible. However, the positive perception of the radiologists about voluntariness should be taken into account, so that although the use of AI is mandatory, the radiologists still have some leeway in its use. This could reduce the fear of losing autonomy, which Buck, Hennrich, and Kauffmann (2021) and Kauffmann et al. (2022) found to be a hindering factor in a study on AI acceptance by radiologists. Especially among physicians, they see this fear high, since the freedom of physicians to make their independent decisions on how to treat patients is guaranteed by law in Germany (Buck, Hennrich & Kauffmann, 2021; Kauffmann et al., 2022; Reis et al., 2020). While the fear of losing autonomy was not directly mentioned by the radiologists, it was mentioned by R4 (4.14) that the physician retaining full control was a selection criterion for the chosen AI software, suggesting that this fear is already taken into account.

The radiologists called for intuitive AI systems as they only have limited time for learning how to use the systems (Section 4.2.2.2). This concern that the learning and understanding of the AI system mean additional effort is in line with previous research (Buck, Hennrich & Kauffmann, 2021; Kauffmann et al., 2022; Yu, Beam & Kohane, 2018). This additional work caused by the AI must be minimized as radiology already have a tight working schedule (Fischer et al., 2020; Liu et al., 2021). This emphasizes the importance of intuitive systems that can be easily learned by radiologists to achieve high technology acceptance. The radiologists were in general satisfied with the *Ease of Use* of the AI software but there were also some minor usability issues that annoyed them (Section 4.2.2.2). However, it has to be noted that all sites that were investigated used the same AI software in their research trials and that it is not certain that the same software will be used in a nationwide rollout. Therefore, the observed opinions of the radiologists regarding the importance of easy-to-use systems should be used to underline the importance of usability for radiologists instead of evaluating this specific AI software.

Radiologists also mentioned that the AI software was configured in ways they did not expect. These included that the keys were not programmed to industry standards or that the software's behavior annoyed them, such as windows that opened for no reason and had to be closed each time, costing additional time. Such problems interrupt radiologists' workflow and lead to errors that can be critical in a medical environment and should be therefore avoided (Park et al, 2020). However, there were also some customization options for the software, allowing radiologists to customize it to their liking, making it easy for them to use. In addition, this can limit the fear of loss of autonomy identified by Buck, Hennrich, and Kauffmann (2021) and Kauffmann et al. (2022) by giving radiologists many options to use the software as they see fit. At the same time, it is important to note that benefits gained by standardized processes, such as minimizing errors or ensuring objectivity, may be lost. The findings of Venkatesh and Davids (2000) that *Perceived Ease of Use* positively increases the *Perceived Usefulness*, Intention to Use as well as the job performance in general, could not be directly confirmed or disproved since the radiologists perceived the *Ease of Use* as positive, but due to the qualitative methodology of the study no direct correlations could be concluded and the radiologists did not explicitly mention that the good usability positively influenced the job performance or *Perceived Usefulnes*s. However, due to the positive attitude toward the *Perceived Ease of Use*, it can be assumed that the opinions of the radiologists at least have a positive influence on the acceptance in general and are thus similar to the basic idea of Venkatesh and Davis (2000).

The radiologists not only found the AI application easy to use, but they also enjoyed using it and found this exciting. This also indicates a positive influence on the *Perceived Ease of Use* without this being explicitly mentioned. Fun seemed to be a general important point for the radiologists, as they suggested leaving tasks that can be solved without cognitive effort and that come up again and again to the AI and to keep tasks that challenge them and are interesting for themselves (Section 4.2.2.2). Similar findings were presented by Chen et al. (2021, p.5), which found by interviewing radiologists on AI adoption that they "saw AI innovations as potentially helping to eliminate the more routine aspects of their work, allowing them to focus on more interesting and challenging tasks". Buck, Hennrich, and Kauffmann (2021) and Lebovitz, Levina, and Lifshitz-Assaf (2021) also found that radiologists would like to hand tedious or boring work over to AI. This could further increase the *Perceived Usefulness* of radiologists if they are involved in the process of deciding which tasks should and should not be performed by AI.

In the trials on lung cancer screening, training was often offered only once at the beginning (Section 4.2.2.2). The radiologists criticized the fact that only one training session took place. Since the training can help to increase the *Perceived Ease of Use* and it was reported that the *Experiences* of the radiologists in terms of technology use vary greatly, which will be even more extreme in a nationwide program, regular training is very important (Section 4.2.2.3). As the AI will receive updates for bug fixes and new features over time, it is important that the training is regularly adjusted so that radiologists can learn the new features, otherwise, the *Perceived Ease of Use and Perceived Usefulness* may decrease over time if radiologists do not understand the new features. Extensive training should be seen as a chance to comprehend the missing *Experience* with AI, especially since AI has so far only very rarely been implemented in the health care sector in Germany outside of research projects and must therefore be assumed to have low *Experience* of AI (Krumm, & Dwertmann, 2019; Mumm et al., 2021). The importance of extensive *Training* for AI adoption has also been highlighted by previous literature. Khanijahani et al. (2022) identified user *Training* as one of the main factors for user acceptance of AI in healthcare and BenMessaoud, Kharrazi & MacDorman (2011) found that technical *Training* has a positive effect on the adoption of surgery robots. Health care professionals interviewed by Hercheui and Mech (2021) also expected that extensive *Training* for doctors is a prerequisite for AI adoption.

The radiologists valued the technical support that was available and the continuous improvements of the AI software according to their needs (Section 4.2.2.4). Therefore the *Relationship with Users* was perceived as high. Support for content-related questions about AI use was experienced as less accessible, as it was not always perceived as easy to communicate one's problems (Section 4.2.2.4). Without this easy communication of issues to the support, the application's *Perceived Ease of Use* suffered. Therefore, it must be ensured that in addition to enough *Training*, there is also documentation that provides guidance on how to use the software for lung cancer screening, such as which function can be used for which medical purpose. Also, the support staff must be trained to understand the problems of the radiologists well. The well working technical support that was observed in the trials does not provide insights into the future implementation of the nationwide lung cancer screening program but it highlights the importance of the technical support as it can help the radiologists with technical problems with the software and can so compensate for lacking *Experience* or *Training*.

Another aspect that resulted in high *Relationship with Users* was the continuous improvements of the AI software according to the needs of the radiologists (Section 4.2.2.4). The radiologists had the opportunity to report issues or missing features in the software which was

well accepted and regularly used. The software provider reacted quickly to those needs. It was mentioned that the flexibility to respond to the radiologists' needs was one criterion to choose the software provider for the trials (Section 4.2.2.4). *Relationship with Users* seemed to positively influence *Perceived Usefulness* and *Job Relevance*, as the features of the AI software were better understood by the radiologists as well as considered relevant as they participated in their selection. Also, the radiologists feel more included when the *Relationship with Users* is high, which could lead to a high level of satisfaction and therefore to higher AI acceptance (Karantanas & Efremidis, 2022). However, it was noted by the radiologists that these considerations of their needs were very special in the trials and were nothing that can be expected in regular health care programs. The high impact of *Relationship with Users* is in line with the research of Rubin (2019) who found the inclusion of the radiologists and their need in the AI development an important factor for the acceptance of AI.

5.1.3 Operationality

Within *Operationality* it was looked at the dependability and consistency of access and uptime of systems as well as whether AI can meet operational expectations and requirements.

Although a large proportion of AI healthcare projects fail, problems to operate the technology is not much named in this regards in the literature (Davenport & Kalakota, 2019; Kitzmiller, Hunt & Sproat, 2006; Park et al., 2020; Strohm et al., 2020) and consequently, it was not very surprising that *System Reliability* was also not raised as problematic by the interviewees.

On the other hand, *Production Timeliness* was expected as a crucial factor as the system efficiency was promised to enable lung cancer screening and failure to meet user's needs is known in the literature as a common factor which leads to adoption failure (Kitzmiller, Hunt & Sproat, 2006; Khanijahani et al., 2022; Strohm et al., 2020).

However, it appeared that *Production Timeliness* was not very concerning to the interviewees and only a moderate attempt was made at achieving wishful production turnover results (Section 4.2.3.1).

One possible explanation is found in the known visible strength of the system to increase efficiency to an extent where efficiency is not concerning anymore. AI might be crucial to achieve necessary process efficiency for lung cancer screening, however, then with AI involved, it might not be difficult to achieve that necessary process efficiency.

Other factors also played at least partially a significant role in the few *Operationality* concerns. After an initial struggle, few more issues with *Operationality* occured, *Operationality* demands and requirements were not difficult to meet and radiologists knew that good and timely support was available (Section 4.2.3.1; Section 4.2.3.2).

Furthermore, the study setting also impacted *Operationality* and lowered the bar for acceptable *Operationality* (Section 4.2.3.1; Section 4.2.3.2). It must also be noted that the artificiality of the study setting limits the implications of the findings (Black, 1996) especially since it was voiced by the interviewees that the work process of radiologist's in the trial includes additional tasks (3.26) and that "[t]he time that [was] now invested in the study is certainly higher than … afterward in the clinical routine" (3.28).

5.2 Summary and Research Implications

The factors which were investigated in the radiologist's acceptance of AI for lung cancer screening diagnosis were grouped into the three themes *System Output, Usability Experience,* and *Operationality.*

In *System Output* it was described how good the AI systems output was, how relevant it was to the radiologist's job and how it was perceived. The AI system promised a good and satisfactory performance which it was able to deliver. Also affected by negative previous experiences with AI integrations, the radiologists were satisfied with the system performance. However, even though the AI supported all functions and tasks which were supposed to be done, the output could not be considered complete in all cases. Radiologists had different opinions on what is important for a diagnosis and also voiced different system requirements. A discrepancy was found between - *What tasks were supported by AI and supposed to be done by the radiologists* and - *What tasks were done by the radiologist*. Additional examination steps which were not supported by AI were conducted, which limited the relevance and completeness of the *System Output* to the radiologist's job.

Nevertheless, it was found to have a good impact on the quality of the output which was also known by the radiologists. Furthermore, the radiologists also voiced that the AI was very important for lung cancer screening and that the objectivity and decreased inter-rater variability would lead to standardizable diagnoses which are of significant value. However, those aspects were not named by radiologists when they were asked for their intention and motives to use AI for lung cancer screening.

Instead, it became clear that they were referring to aspects which affected their *Perceived Usefulness* of the system. Advantages of the AI which were perceived as useful by the radiologists were its efficiency and that it facilitates work. Those factors set themselves apart by having a high *Result Demonstrability*. Therefore, it became apparent that while the system's advantages included quality and job-relevant aspects, it was the visibility of a system advantage which made it influential for *Perceived Usefulness* and for radiologists' technology acceptance.

Under the *Locatability* factor the impact on technology acceptance which depends on the transparency of the system was investigated. Since the system was found to be very transparent, it was not possible to investigate potential technology acceptance problems that the literature suggested to occur with low transparency. Nevertheless, the high transparency was found to enable a strong complement between radiologists and the system, where they could compensate for their weaknesses and leverage their strengths. This was considered crucial to the radiologist's assessment that they found the AI supportive and enjoyed working with it. Although the extent of the effect of locatability was not comprehensively clear, it was judged to be influential on radiologists' technology acceptance.

The factor *Subjective Voluntariness* investigated the radiologist perceived pressure to use the AI system. Although the use of AI was mandatory in the trial studies, the use of AI was perceived as voluntary by all radiologists. The missing pressure to use the AI software allowed the radiologists to use the software in the way they prefered which was observed in the usage of different functions for the same task. Nevertheless, they called for standardized processes for the use of AI, suggesting that objectivity and decreased inter-rater variability are more

important to them and *Subjective Voluntariness* was not an influential factor to their technology acceptance.

Perceived Ease of Use/Training was an important technology acceptance factor for the radiologists as they called for intuitive systems due to already high workloads and differences in the radiologists' experience in the use of technology. Therefore it was observed that the *Perceived Ease of Use/Training* had a positive influence on the acceptance of the radiologists to use the AI system as it was suggested by the literature. This became clear when unexpected system behaviors and key assignments annoyed them because they had to get used to them, which resulted in operating errors.

Extensive training and user manuals of the software were also very important to the radiologists, especially since in the trials only a one-time training was offered, but the software was constantly developed which can lead to a decrease in acceptance if the purpose and use of new features are not explained. The training and the high usability can compensate for the lack of experience with similar systems. The *Experience* itself was therefore not considered an influential factor for AI acceptance, as the intuitive systems showed that even new radiologists who did not know the software quickly became proficient with it.

The technical support which helped the radiologists quickly to solve their issues with the AI software together with the continuous improvements and adjustments of AI software according to the radiologists' needs led to a good *Relationship with Users*. This was very much appreciated by the radiologists and helped to mitigate their concerns about the AI software. Moreover, the features had been designed according to the needs and wishes of the radiologists, so that these features were tailored to their work and were found useful and job-relevant by them. Thus, a good relationship with users was considered an impactful factor for AI acceptance.

The *Operationality* theme addressed the reliability of the AI system and whether it was able to meet the operational productivity expectations. Only a few concerns were voiced and this was partly attributed to a few operational issues which had occurred in the trials. Also, the AI integration was investigated in a study setting and it was voiced by the interviewees that this lowered their expectations. Furthermore, a good relationship to the AI provider likely mediated concerns that IS difficulties would lead to significant problems. Nevertheless, the few operational issues and concerns in this area were also in line with expectations expressed in the existing literature on technology adoption in healthcare, and furthermore, the requirements for achieving satisfactory *Operationality* seemed manageable. Therefore, operational factors were not found to be impactful.

This research is contributing in various ways. It contributes to IS research by helping to close the research gap between IS research and practice which result in many HIT projects to fail. The research also added to the IS literature which uses TAM and TTF in a qualitative approach and has demonstrated the applicability and usefulness of such approach. This approach also allowed us to contribute to the relevant literature by unpacking *Locatability* and *Relationship with Users* as two influential factors in technology acceptance, which are both normally not investigated. Furthermore, the investigation into influential factors which affect radiologist’s acceptance of AI for lung cancer screening is novel and the developed understanding contributes as well in different ways.

The findings can help managers in the health care context to understand radiologists' acceptance of AI in their clinics. The understanding can also be used to prepare measures to improve acceptance and provides also relevant criteria which should be considered when a system is selected for adoption so high technology acceptance is achieved. Lastly, the findings also signal the importance to both healthcare managers and providers to maintain a good relationship and act upon radiologists' needs to achieve technology acceptance.

5.3 Limitations

The findings and research implications pointed out above need to be viewed with care as they are limited in respect to the aspects below.

First, the research was restricted to Germany and the literature highlighted that aspects which are found influential to technology acceptance can be country specific. Therefore, generalizability for other countries needs to be viewed with care and instead results should be used as the foundation upon which similar research in other countries can emerge.

Furthermore, due to the novelty of AI for lung cancer screening we were only able to investigate radiologist's technology acceptance in trial settings. The trial setting is artificial and it was also found that it restricted the emergence of operational concerns. It is unclear how else the artificiality might have affected the results and therefore further research will be needed to investigate applicability of results in settings outside of trials.

The scope of the research also had to be limited in the investigation of influential factors to a selection of TAM and TTF factors which were deemed relevant and sufficiently comprehensive. Nevertheless, in technology acceptance research various frameworks with different factors have been developed for research and the list of potentially influential factors is long. Also, time constraints limited us to only five interviews and we found that novel insights were still gained from the fifth interview and findings could likely have been enriched by further interviews. Consequently, the list of potentially influential factors which was investigated in this research and also the list of factors which were found to be impactful for radiologist's acceptance of AI for lung cancer screening does not claim to be complete.

It must also be noted that the research took place nine months after trial start and introduction of AI and its first use. Venkatesh and Davis (2000) pointed out that *Experience* and *Perceived Ease of Use/ Training* change over time and so does their influence on radiologist's technology acceptance. What potential change in factors took place and whether the initial influential factors of technology acceptance at trial begin might have differed was not further explored in this research.

Also, in the trials only one AI software and one respective provider were relevant and this led generally to rather one-sided expressions and perceptions in the themes and factors. This limits its comprehensibility of the understanding of the factors developed in this research as e.g. without having observed radiologist's perception of a system with low transparency, only little can be said about how this may have affected their technology acceptance.

6 Conclusion

6.1 Purpose

Lung cancer screening holds a lot of potential for society, and a rollout in Germany is planned soon. This introduction would overwhelm radiologists with existing processes, but AI can increase efficiency and enable screening. However, technology acceptance in healthcare is often problematic, and it is unclear whether radiologists will accept AI for lung cancer screening and what this depends on. Therefore, this research aimed to answer the question: *What are the most influential factors affecting the radiologist's acceptance of AI supporting lung cancer screening diagnosis*

6.2 Methodology

An interpretive qualitative research approach was used in this exploratory study. A framework based on the TAM and TTF model was developed and used to structure a comprehensive investigation of factors influencing radiologists' acceptance of AI. Five semi-structured interviews were conducted with senior radiologists, and deductive coding based on the research framework was performed independently to obtain reliable results. Member checking was performed to confirm unclear statements made by the interviewees and to get their opinion on findings to assess their validity.

6.3 Key Findings

This research unpacked the influential factors in radiologists' acceptance of AI for lung cancer screening.

In the evaluation of the importance of AI for lung cancer screening, radiologists highlighted among others the improvements in diagnosis quality and objectivity. Nevertheless, those factors were not named in the radiologists' considerations of whether they would use AI. The impact of *Result Demonstrability* was detected when instead time savings and easing of work were named which was attributed to the visibility of these benefits to radiologists in their daily work.

Furthermore, the impact of *Perceived Ease of Use/Training* became apparent. Radiologists work with different systems under time pressure in their daily routine and this emphasizes the importance of intuitive systems. Otherwise, operational mistakes occur and this leads to unnecessary time losses and frustration which negatively affects technology acceptance.

Locatability, which refers to the transparency of the system, was highlighted as another influential factor on radiologists' acceptance of the AI. It was crucial to enable a strong radiologists-AI complement with which the work with the AI system was enjoyed and the AI was considered to be supportive.

Finally, *Relationship with Users* was also found to make a significant multi-faceted impact on radiologists' acceptance. The radiologists had a good relationship with their provider and the trust in the provider's ability to handle operational problems was deemed to have hindered the emergence of operational concerns which could have negatively affected the technology acceptance. Furthermore, the radiologists were able to request system functionalities which were then implemented. This increased the relevance of the system to their job and their satisfaction with the system. And this was assessed to further increase their perceived ease of use and perceived usefulness of the system and with that also their acceptance of the AI system.

6.4 Future Research

It was found that radiologists do not always do what is supposed to be done in their examinations and that this negatively affects TTF and endangers a successful lung cancer screening implementation. Further research into this discrepancy is needed and was also called for by the radiologists when this finding was presented to them in member checking.

Also, the interviewees were senior radiologists and some findings indicated that seniority affects radiologists' technology acceptance. Therefore, the meaningfulness of the findings for the relevant factors which determine system acceptance of junior radiologists is limited and should also be addressed in further research.

Finally, the integration of AI for lung cancer screening in Germany, although soon to come, has not been researched before much and the relevant trials have only been running since August 2021. Further relevant factors might have been missed due to the limited scope of this research and others might emerge over time. This calls for later research into the validation of our findings and into the emergence of new relevant factors.

Since the acceptance of AI depends on cultural differences and other countries have different laws and guidelines for the use of AI in healthcare, the findings from Germany cannot be directly transferred to other countries, requiring IS research to investigate AI acceptance in other countries as well.

Generally, with lung cancer screening this qualitative study investigated a novel context of technology acceptance and arrived at interesting findings. A quantitative study to generalize, complement and support the findings would also be very valuable.

Appendix A - Interview Guide

Introduction:

- Thank the interviewee for taking the time
- Introduce ourselves and our research project in 2-3 sentences
- Explain the process to the interviewee. They should introduce themselves and then we will ask questions about lung cancer screening and their work with AI in lung cancer screening
- Inform the interviewee that he/she does not have to answer any question
- Offer to the interviewee to ask questions at the end of the interview
- Ask the interviewee for member checking
- Ask for permission to record the meeting

General questions about the person:

1. Can you tell us in 2-3 sentences who you are and in what context you have been involved with lung cancer screening so far?

General questions about lung cancer screening:

2. How do you think the introduction of lung cancer screening will be carried out in Germany? Where will it be offered and who will be responsible for this additional work?
3. Can you tell us about the process of screening a patient in the lung cancer screening program?
4. Is this process working as planned or are there difficulties?

Questions about radiologists' work with AI:

5. When and how does the radiologist evaluate the CT scan?
6. What functions do the systems have that are to be integrated?
7. When and how are these systems used?
8. Is there a prescribed process for how the AI supports the radiologist or how the system should be used?
9. Are there any touchpoints where you would say that AI does not yet optimally support radiologists' work processes?

Questions about the constructs from the frameworks/ properties of the system:

10. Do the systems provide you with all the information you need?
11. How good is the output of the system in terms of relevance and detail?
12. Is the output of the system understandable and can it be used well?
13. How does the use of the AI system change the outcome/output of the radiologists?
14. Do the radiologists all get along well with the system or are there any difficulties?
15. What do you consider to be the biggest problems when radiologists use these systems?
16. Do radiologists get support quickly to solve these problems?

Specific questions about radiologists' work with AI:

17. Do you see potential at any point that the work process could function even more efficiently?
18. Can you imagine functions with which AI could support the work processes of radiologists even better?
19. What else is costing radiologists a lot of time in the current process?
20. What do you think the radiologists' work process would look like if no AI was brought in - Would it still be the same?
21. What are the general feelings about the use of AI in lung cancer screening?
22. Do you anticipate any hurdles or issues that you will face before successfully using the system in your lung cancer screening clinic?

Ending the interview:

23. Do you have any comments or are you aware of any important aspects of lung cancer screening that have not been mentioned or addressed so far?
24. Do you have any questions?

Thank you for the interview and have a nice day

Appendix B - Consent Form

LUNDS UNIVERSITET

Consent to take part in an interview

- I voluntarily agree to participate in this research study.
- I understand that even if I agree to participate now, I can withdraw at any time or refuse to answer any question without any consequences of any kind.
- I have had the purpose and nature of the study explained to me in writing and I have had the opportunity to ask questions about the study.
- I understand that participation involves participating in an interview about the challenges encountered when integrating artificial intelligence into radiologists' workflows for lung cancer screening.
- I understand that I will not benefit directly from participating in this research.
- I agree to my interview being audio-recorded.

Consent for personal data processing

I consent to my personal data in the form of **audio and video recordings** processed by Lund University for the following purpose:
The research project aims to improve understanding of the challenges encountered when integrating artificial intelligence into radiologists' workflows for lung cancer screening. Audio and video recordings help us to evaluate the interviews after conducting them.

Information

The personal data will be processed in the following way:
The audio and video recordings are stored in a password-protected environment to which only Jan Harder and Joachim Maßberg have access. Upon request, the recordings must be disclosed to the supervisor of the master's thesis as well as to the examination board but will not be made public. The transcripts of the interviews will be published freely accessible in anonymized form as part of the master thesis without identifiable personal information such as the name of the interviewee.

The data will be used for the above purpose and in accordance with this form. The legal basis for the processing of your personal data is that you have given your voluntary consent. We do not share your personal data with third parties.

Lund University, Box 117, 221 00 Lund, corporate identity number 202100-3211 is the personal data controller. You can find information about the processing of personal data at Lund University at www.lunduniversity.lu.se.

The consent is valid until further notice. You have the right to withdraw your consent at any time. You do this by contacting Jan Harder [ja5751ha-s@student.lu.se] or Joachim Maßberg [jo1055ma-s@student.lu.se] or registrator@lu.se. We will in this case cease to process personal data that we have collected based on this consent. However, data included in results that have already been obtained will not be affected by the withdrawal of your consent. Certain data may also be archived in accordance with Swedish law.

You have the right of access to information about the personal data we process about you. You also have the right to have incorrect personal data about you corrected. If you have a complaint about our processing of your personal data, you can contact our data protection officer via dataskyddsombud@lu.se. You also have the right to submit a complaint to the supervisory authority (Swedish Data Protection Agency) if you think that we process your personal data incorrectly.

I agree to all terms listed under “Consent to take part in an interview” and I consent to Lund University processing personal data about me in accordance with the above.

Town/City	Signature
Date	Name in block letters

Appendix C - Member Checking

English

Dear [person, anonymized],

In one week we will hand in our master thesis. We have now analyzed all the data and two controversial findings in particular are currently attracting our attention. We would like to briefly present these two findings to you here and ask for your comments, which may then be integrated as a further data point in the interpretation of the results of our master's thesis. We would like to thank you again for your participation and considerable contribution to our master thesis and hope that the findings are also interesting for you.

1. A discrepancy was partially found between - *what is supposed to be examined in lung cancer screening* and - *what is actually looked at and examined*. If further examination steps are performed by radiologists, they may not be supported by the AI or the radiologists may not have been trained to use the AI for these purposes. The importance of AI to the radiologists' job decreases with the wider this discrepancy

2. AI is considered to have a very high importance for lung cancer screening. On the other hand, some benefits are not very visible. These include aspects of diagnostic quality such as higher standardizability, objectivity, and reduction of inter-rater variability. These factors were mentioned in the benefits of AI for lung cancer screening. However, the factors did not come up in questions about whether radiologists themselves want/will use AI systems for screening. This is attributed to the low visibility of the benefits to radiologists in their daily practice. Visible benefits are instead that AI makes work easier and leads to time savings. These aspects are therefore presumably more influential on the behavior of radiologists in daily clinical practice.

In addition, we also would like to check our understanding of some statements of yours.

[personalized questions]

Thank you very much!

Joachim Maßberg & Jan Harder

German

Sehr geehrte/r [person, anonymized],

In einer Woche ist es so weit und wir geben unsere Masterarbeit ab. Wir haben jetzt alle Daten analysiert und insbesondere zwei kontroverse Findings ziehen zurzeit unser Augenmerk auf sich. Wir möchten Ihnen gerne diese zwei Findings hier kurz präsentieren und einen Kommentar von Ihnen erbitten, welcher dann eventuell auch so als weiterer Datenpunkt in die Interpretation der Ergebnisse unsere Masterarbeit integriert wird. Wir wollten Ihnen nochmals danken für Ihre Teilnahme und erheblichen Beitrag zu unserer Masterarbeit und hoffen, dass die Findings auch für Sie spannend sind.

1. Es wurde teilweise eine Diskrepanz festgestellt zwischen - *Was soll in der Lungenkrebs Früherkennung untersucht werden* und - *Was wird tatsächlich alles angeguckt und untersucht*. Wenn weitergehende Untersuchungsschritte von Radiologen durchgeführt werden, sind diese eventuell nicht von der AI unterstützt oder die Radiologen wurden nicht ausgebildet die AI für diese Zwecke zu benutzen. Die Bedeutung der AI für den Job der Radiologen sinkt mit dieser Diskrepanz

2. AI wird eine sehr hohe Bedeutung für die Lungenkrebs Früherkennung zugesprochen. Andererseits sind einige Vorteile nicht sehr sichtbar. Dazu gehören unter anderem Aspekte der Qualität der Diagnose wie eine höhere Standardisierbarkeit, Objektivität und Reduzierung der Inter-Rater Variabilität. Diese Faktoren wurden bei den Vorteilen der AI für Lungenkrebs Früherkennung genannt. Die Faktoren kamen aber nicht bei Fragen auf ob Radiologen selber die AI Systeme benutzen wollen/werden für die Früherkennung. Das wird auf die geringe Sichtbarkeit der Vorteile für die Radiologen im täglichen Alltag zurückgeführt. Sichtbare Vorteile sind stattdessen, dass AI die Arbeit erleichtert und zu Zeiteinsparungen führt. Diese Aspekte sind deswegen vermutlich auch einflussreicher auf das Verhalten der Radiologen im klinischen Alltag.

Zusätzlich wollen wir bei Ihnen auch noch unser Verständnis von einigen Aussagen überprüfen.

[personalized questions]

Vielen herzlichen Dank!

Joachim Maßberg & Jan Harder

Appendix D - Interview 1

Duration: 25 min

Date: April 26, 2022

Interviewee: R1

English

Row	Speaker	Transcribed text	Codes
1.1	JH	Exactly. yeah. Then let's start right away. Can you tell us in two to three sentences who you are and in what context you have come into contact with lung cancer screening so far?	
1.2	R1	So my name is [name, anonymized]. I am the Medical Director of Diagnostic Interventional Radiology at the University Hospital in [city, anonymized] and my focus has always been on thoracic radiology. In this capacity, I am also on the board of the German Center for Pulmonary Research and have been working on the subject of lung cancer screening for ten years and am on various specialist associations in various committees for this purpose.	
1.3	JM	Um great, then I would like to come directly to the general questions that I would like to ask at the beginning. Namely, that is exactly what you are already watching at many locations at the moment, how can lung cancer screening be implemented? How do you think this introduction is going to happen in the end? Where and in which locations, such as university hospitals or large practices? And exactly who would deal with this?	
1.4	R1	So I think that this is mainly triggered via the certified lung cancer centers. Some of them are university, some are not university and, in order to reach a broad, i.e. geographical latitude, they will then agree in a contract with established partners that not all participants then have to travel 50 or 80 kilometers somewhere, but that they can be closer.	

1.5	JM	And how would you imagine the process of how lung cancer screening would then be implemented exactly, i.e. now in terms of how the patient is examined, i.e. what would be done there?	
1.6	R1	So the general opinion in Germany at the moment is that it is through registered specialists who have patients who are at risk rather than their patients or serve as a point of contact for others and that they are informed there about this offer. And if they agree, then participation in this program will begin, which will certainly result in smoking cessation and CT. And then there are the various algorithms that happen when the CT is inconspicuous. And then the participant goes back, so to speak, to his lung specialist, receives the message from him that it is inconspicuous and is then asked again by him that he should come back in a year.	
1.7	JM	Could you be a bit more precise about the part exactly as it happens in the clinic. So, from CT scan to communication of the diagnosis for the patient.	
1.8	R1	Yes, it's still a bit blurred. So my idea is that when the CT scan is performed, that the radiologist does it according to certain guidelines, that will be relatively strict. Finds that this finding, if it is indisputable, goes back to this lung specialist and that he then reveals this to the participant.	
1.9	JM	Um, exactly, where would, I say, the system or the AI work? Then the AI would simply say in advance is something conspicuous here, is there something not conspicuous here or how?	
1.10	R1	So AI can engage in a variety of places. So one is that AI is effective in radiation dose reduction. The next is the AI helps in the detection of the nodules. Then the AI probably intervenes in the assignment as to whether a nodule that was found, what is the probability that it is malicious or not. That is probably the most exciting topic. And then it would be clever if we also had AI to feed these risk models there or to evaluate them even better. Because individual participants have more risk profile data based on age, gender, and history than just their smoking history. Ehm, and if you then create an individualized profile for the individual nodule that you find or can create for the individual person, then you can probably use it to control participation in this program in a more individualized way.	

1.11	JM	I've been thinking about it too. In theory, it would also make sense if, I say, previous patient data were also included in the diagnosis. But that's too complicated now to somehow connect it to the system, isn't it? Or is that done?	
1.12	R1	In Germany, this is difficult, if not at all, possible.	
1.13	JM	Because of data privacy?	
1.14	R1	So in Germany, this would be done by collecting the medical history again. And what the participant tells in this context is recorded, but it is not possible in Germany to access the patient's previous history, so to speak, via health insurance data.	
1.15	JM	Okay, thank you very much. Just like the process you've just described and the way it may have been tested on you before. Did that work so well or were there any difficulties?	
1.16	R1	Once that is settled in, then there is no difficulty.	
1.17	JM	An, an, okay, yeah, well, we'll go into more detail about the difficulty later, because okay, but well.	
1.18	JH	Yes, then we would go back to, uh, the more detailed questions about radiologists' workflows with this AI, as you have already said at which points the CT scans are evaluated. Um and what exactly are the functions of these systems that support radiologists? So you will probably then be shown the, yes, pre-recognition by AI. But are there more?	
1.19	R1	So formally, it is unclear whether it will be allowed for the AI to look first and then the radiologist. In terms of content, however, I am completely involved in this, the AI has to look ahead. And then the program that analyzed the images displays the findings and then the radiologist looks at both the displayed findings and the complete examination again, and the point will be that the system first finds the nodule, that it automatically measures the nodule or volumitizes it, and that it then does a gives an assessment that can either go according to this Lung-RADS assessment or according to any other, um, tree, based on which it wrote.	

1.20	JH	Exactly. The question was which functions exactly support radiologists in their workflows, i.e. how the AI systems also act as second or first reader, what functions they have to support radiologists.	
1.21	R1	So, they have to have the functions that they find the nodules, that they measure the nodules and that they determine the probability of malignancy of these nodules. And what else goes with that, I forgot that earlier, is that it of course compares preliminary tests, that is, nodules that were already there the first time, automatically finds them again and determines the change, if there is one. Or just make sure whether a nodule is new or whether a nodule has disappeared. And then a structured report has to be generated, where these whole measures are included accordingly.	
1.22	JM	Do you know what the system you're using is called?	
1.23	R1	Ah, different one can do that differently well. So there's this from Siemens, there's this from Philips, there are at least five different smaller companies. [AI software, anonymized] is currently hot in the HANSE study, but there are also plenty of others who are trying to do this or something similar. With AI for malignancy assessment. They're all still in their infancy, but some already have AI in there.	
1.24	JM	Exactly, uh, for example, the [AI software, anonymized] which is also used in the HANSE study. This also calculates coronary calcium, for example. Do you think that makes sense or why aren't you doing it?	
1.25	R1	That makes sense. But this is actually an extension of the program, which, at the moment, I have the feeling that it makes sense in terms of content, but politically, if we want to push through the program, it only confuses, because you have to put other SOPs and other workflows behind it. And in some cases, there is not enough evidence on this.	
1.26	JM	This means you think that in order to really get lung cancer screening off the ground, you should first focus on the function of lung nodule detection and classification.	
1.27	R1	Yes.	
1.28	JM	Okay, yes.	

1.29	R1	So what many of the manufacturers are doing now, that they are already offering a comprehensive lung or thorax package, so to speak, where not only are the nodules in it, but there is also airways, emphysema and calcification. So if we think from the perspective of software tools, then of course that makes sense, that they offered them as modules side by side. And then everyone can pick out the things they want to use for that now or decide which ones they want to do without.	
1.30	JM	Yes, it makes sense, yes.	
1.31	JH	Oky thanks. Is there a prescribed process for how radiologists should use AI? Or does everyone use it as they like?	
1.32	R1	At the moment, there is no finished process of what you are allowed to and what you are not allowed to do.	SV
1.33	JH	Okay. And, um, what would your opinion be? If lung cancer screening is then used nationwide?	
1.34	R1	Yes, I think we need an SOP that says what to use in what order and how. So, in the case of a national programme, it would of course be nice if everyone used the same system, because the results also vary between the systems. However, I find that difficult to do in Germany and under competition law.	SV
1.35	JM	yeah. As I have now understood it again, to refer back to it once again. Did you think that the system looks first, and then the radiologist, the, would then use the output of the system, so at the moment, as you imagine it, or how you do that, is AI as a first reader, right?	
1.36	R1	That's how I imagine it, yes.	
1.37	JM	Why? Or what do you think of it, why not AI as a second reader? Or what I've seen often, I think how it's done with [AI software, anonymized], is viewing it on two screens in parallel.	
1.38	R1	So on two screens in parallel, uh, probably makes sense. But with First Reader, I meant that the system must pre-process the images automatically when they look out of the scanner.	

1.39	JM	Yes, that's true.	
1.40	R1	So I don't think that I, as a radiologist, have to look first without this annotation, because I don't think that, firstly, it takes longer, and I don't think that this supposed independence will yield any better results.	
1.41	JM	Yes, um, and the way AI works right now, I'll say, or that's it evaluates the image beforehand. Are you happy with that or could you or would you say the AI could actually maybe, it would be ideal if it could do a little bit more for exactly the intended process of lung nodule detection, it could help even better or would you say it does everything great?	
1.42	R1	It could always be better. And the other thing with this probability is whether something is malignant or not. That is incredibly difficult in validation. Yes, because. Um, depending on what the preselection bias is from the cohort which one developed this on, the performance will be different. Yes, there are still unlimited possibilities for improvements. They say.	OQ
1.43	JM	Yes, okay. Simply to describe a specific possibility, how and now not only depending on the system being able to work better, but functionally, whether it could still do something else. Can you think of anything right there that the system could still do? Anything specific?	
1.44	R1	So I believe that with the Lung-RADS classification or the classification for which you want to work, it automatically makes the determination, so to speak, that you no longer have to look it up.	OQ
1.45	JM	yeah.	
1.46	R1	Um, that she, that it automatically calculates and enters these whole measurement figures, i.e. also this volume doubling time. And the other is then more IT connectivity, I think. This is not AI then, but it must, as it were, be via problem-free interfaces into the structured findings and then to this resident lung specialist who included this person, must this then go through automatically. Yes, not that we have to generate PDFs that we then just start faxing and stuff like that, yes.	OQ

1.47	JM	We would like to talk more specifically about the system again and we have a few questions to test your opinion on the system again. We would start with, um, okay, we've just touched on the point of whether the system makes all the information available for the function it is supposed to fulfill. You said, right there, it could actually make the detection better, or exactly the classification according to the Lung RADS 1.1, right?	OQ
1.48	R1	Yes.	OQ
1.49	JM	Yes, then we would go straight ahead and ask, what would you say is the output of the system in terms of quantity or relevance and also accuracy? So a very broad question.	
1.50	R1	So I believe that the system must of course manage with the volume and I'm always of the opinion that the systems don't have to be super accurate, but the systems have to be robust. Precisely because they can survive in the real world, so to speak, they simply have to function and produce objective, repeatable results. It always optimizes whether or not we always have a measurement error of one millimeter, that doesn't matter at the end of the day. As long as there is only one, some sort of offset from the truth, so to speak, the proportionality will hopefully be the same and that is actually quite enough. So I find that sometimes these things are done and supervised too much by physicists who have a completely different approach to accuracy and error margins and stuff like that. I don't think we need to be hyper specific there.	SR OQ
1.51	JM	Yes, but even if the production group not the ones, well, but they already understand what the system is able to do. So is the output of the system completely relevant or does it also indicate something that you don't even need?	
1.52	R1	Ah, that's where the opinions of radiologists differ. So there are already things that are displayed that I don't need. But usually...	OQ JR
1.53	JM	Which other radiologists then need or who are generally not needed?	

1.54	R1	Opinions differ on this. So there are radiologists who say they love that now that they have some information and others don't need it. So that's where it starts, for example, the texture parameters, to what extent I then play them out in detail or whether he makes an anatomical assignment to me in which lobe or in which segment the nodule is.	OQ JR
1.55	JM	Yes.	
1.56	R1	You can still imagine various things that such a system can do. And there will be radiologists who want that and others who don't need it.	
1.57	JM	Um, short side question again do you actually have to go halfway away at a point or would it be okay if five minutes overrun?	
1.58	R1	I actually have to go to the other video conference.	
1.59	JM	Okay, then we have to get on with the last few questions.	
1.60	R1	I can come in four minutes later, I think.	
1.61	JH	Um, that's right. So the output of the system is presented well and clearly for radiologists, so do the systems have good usability?	RD L PEUT
1.62	R1	Yes, I think so.	RD L PEUT
1.63	JM	Continuing directly, how would you say, does the use of the system change radiologists' output? I think you've already said it might be a bit more specific. Will work become more time-efficient?	
1.64	R1	So I think it's going to be a lot faster and it's getting more objective.	JR PU
1.65	JM	Okay, yes.	

1.66	R1	Well, the system certainly works. So if we have radiologists doing initial assesment and then look at inter-reader variability, then it may be relatively high. If we have the system first and then the radiologist looks at it, the inter-reader variability will certainly be significantly reduced.	JR PU
1.67	JH	Yes. Are the radiologists all getting along well with the AI system or are there still any difficulties that should be addressed?	
1.68	R1	Most radiologists take this relatively positively for such a new application as well as a clinical routine. Yes.	PEUT
1.69	JM	Are there any problems in use anyway and can radiologists also get help if they have problems, or are there any technical difficulties?	
1.70	R1	So first of all, I'm sure the radiologists must want that, yes. So no matter which IT system you use for diagnosis, you must want to make friends with the new system. Then you can understand the help you're getting. And otherwise the IT affinity of radiologists is, of course, different and some have more fun than others dealing with such topics and then kneeling in and working their way through them.	SV PEUT RU
1.71	JM	Yes.	
1.72	R1	But I believe that in such a lung cancer screening approach, that whole thing will be so school-like that it doesn't matter.	SV PEUT E RU
1.73	JM	Okay, yes. Uh, can you imagine any potential where the radiologists' work process could be made even more efficient at the moment? Or is there a place where you say that there is still an unnecessary amount of time spent on something here?	

1.74	R1	So the time goes before and after in the diagnosis process, if it is supported accordingly and if the workflow is described accordingly, is standardized, I don't think you can save much more. Efficiency is then created in communication. Or what happens if any nodule is found that lies at some limit in the rating? Yes, but then we planned that a second opinion would then be obtained, or that it would be presented in a kind of tumor board, as how one actually deal with the result you have now received.	PU PT
1.75	JM	Yes, yes.	
1.76	JH	What do you think the radiologists' work process would look like if AI were not involved now? Would it still be the same or would lung cancer screening itself not be feasible at all? Without this AI support?	
1.77	R1	Then it would look like mammography screening. And in my opinion, it would no longer be feasible these days, that we have so many radiologists doing a first report, a second report and then an adjudication report, that is not feasible in my opinion.	JR PU
1.78	JM	Who exactly would use AI now? Is it the classic radiologists who are already employed by you anyway? Or would you say that it will be a new full-time job, where new radiologists are actually hired...	
1.79	R1	So if that really goes off and there are participants, then it will be quite a full-time job and, in addition, a certain quality control certification would be provided in training. Um, that means that there will be a special group that does that.	PEUT
1.80	JM	Okay, yes, then we're on to the last question. Do you still see any hurdles or problems that would be in front of successful use of the system in your clinic now or that could come your way?	
1.81	R1	So the most difficult thing is that this topic is even accepted with this initial diagnosis by AI. And the other is, of course, the approval of the overall system, of course, of the overall program. It is always a question of the extent to which you get involved in such a black box, because at the end of the day nobody understands exactly what the system actually does to get there. And, um then, what else did I want to say? I believe that it will be important to have a system that is as uniform as possible in Germany. As I	

		said earlier, if you have four different systems in Germany, different things come out. Which is not good, especially if the participants change their place of residence and do their examination.	
1.82	JH	Okay.	
1.83	JM	Oh, can I ask another last, second question then? Why do you use the system that is currently in use and not [AI software, anonymized], for example? Or what motivated the choice of the system for you?	
1.84	R1	So most of the time, you have money at some point in time and you can procure something, and then you take what is that at the time that you're best at. That is a very practical answer, but it often plays a role.	
1.85	JM	Since when do you actually have the system that you are currently using?	
1.86	R1	So we now have a new system from [AI software, anonymized], which we tested in a clinical context and with which we were very satisfied. And that makes the interfaces the way we want them to be. The system we had before was more of an experimental one. That didn't make the interfaces the way we wanted them to be. Yes, the pictures were not properly stored in PACS. And so on. Yes, it is often decided on such things. But that's what I meant earlier, it doesn't really matter how good the system is. So as far as measurement is concerned, not when it comes to that, of course, you shouldn't mark too many false positive findings. So we used to have a [AI software, anonymized] system that made dots everywhere. Of course, you can't work with that.	OQ
1.87	JM	Great, thank you.	
1.88	JH	Yes, thank you very much. Um, do you have any more comments or are there any important aspects that you think are still missing out that you would like to explain?	
1.89	R1	No, I think I've told you everything. You had a nice list of questions.	
1.90	JH	Okay. Do you have any other questions for us?	

1.91	R1	No, I think everything is clear for now. I think you put it together and then send it over again. And if you think of something later or I think of something, we can also exchange it again via email.	
1.92	JM	Great, great.	
1.93	JH	Yes, thank you again for taking the time.	
1.94	R1	All right, I'd love to, bye.	
1.95	JM	Have a nice day.	
1.96	JH	For you too. Thank you, bye.	

German

Row	Speaker	Transcribed text	Codes
1.1	JH	Genau. Ja. Dann fangen wir direkt mal an. Können Sie uns in zwei bis drei Sätzen sagen, wer Sie sind und in welchem Zusammenhang Sie bis jetzt mit dem Lungenkrebs Screening in Berührung gekommen sind.	
1.2	R1	Also mein Name ist [name, anonymized]. Ich bin der Ärztliche Direktor der diagnostischen Interventionalen Radiologie am Universitätsklinikum in [city, anonymized] und mein Fokus lag schon immer in der Thorax Radiologie. Ich bin in dieser Eigenschaft auch im Deutschen Zentrum für Lungenorschung im Vorstand und befasse mich seit zehn Jahren mit dem Thema Lungenkrebs Screening und bin dafür diverse Fachgesellschaften in verschiedensten Gremien unterwegs.	
1.3	JM	Ehm super, dann würde ich jetzt mal direkt zu den allgemeinen Fragen erst mal kommen, die noch gerne am Anfang stellen würden. Nämlich genau das wird ja momentan auch schon an vielen Standorten auch bei Ihnen ja geguckt, wie kann Lungenkrebs Screening umgesetzt werden? Wie denken Sie, wird schlussendlich diese Einführung passieren? Wo und in was für Standorten, zum Beispiel Unikliniken oder in Großpraxen? Und genau wer würde sich damit beschäftigen?	

1.4	R1	Also ich glaube, dass das hauptsächlich über die zertifizierten Lungenkrebs Zentren getriggert wird. Die sind teils universitär, teils nicht universitär und die werden sich dann, um eine breite, also geographische Breite zu erreichen, damit niedergelassenen Partnern noch vertraglich einigen, dass nicht alle Teilnehmer dann 50 oder 80 Kilometer irgendwo hinfahren müssen, sondern dass man näher dran sein kann.	
1.5	JM	Und wie würden Sie den Prozess sich vorstellen, wie Lungencancer Screening dann genau umgesetzt werden würde, also jetzt bezogen auf, wie der Patient untersucht wird, also was da alles so gemacht würde?	
1.6	R1	Also die allgemeine Meinung in Deutschland ist im Moment, dass es über die niedergelassenen Fachärzte geht, die Patienten, die im Risiko sind, eher als ihre Patienten haben bzw für andere als Anlaufstation dienen und dass die dort über dieses Angebot aufgeklärt werden. Und wenn die dort einwilligen, dann beginnt die Teilnahme an diesem Programm, was dann sicherlich eine Raucherentwöhnung und ein CT nach sich zieht. Und dann gibt es die verschiedenen Algorithmen, die dann passieren, wenn das CT unauffällig ist. Und dann geht der Teilnehmer sozusagen zu seinem Lungenfacharzt zurück, kriegt von dem die Mitteilung, dass es unauffällig ist und wird dann von dem wieder angesprochen, dass er in nem Jahr wieder kommen soll.	
1.7	JM	Könnten Sie das noch mal ein bisschen präziser ausführen den Teil genau wie das in der Klinik passiert. Also vom CT Scan bis zu der Kommunikation der Diagnose für den Patienten.	
1.8	R1	Ja, das ist ein bisschen unscharf noch. Also meine Vorstellung ist, dass wenn der CT Scan durchgeführt wird, dass der Radiologe das nach bestimmten Vorgaben, das wird relativ streng sein. Befundet, dass dieser Befund dann, wenn er unstrittig ist, an diesen Lungenfacharzt zurückgeht und dass der das dann dem Teilnehmer eröffnet.	
1.9	JM	Ehm genau, wo würde da, sage ich mal, das System oder die AI greifen? Dann würde einfach nur im Vorfeld sagen ist hier was auffälliges, ist hier was nicht auffälliges oder wie?	
1.10	R1	Also AI kann eine an verschiedensten Stellen greifen. Also das eine ist, dass AI bei der Strahlendosis Reduktion greift. Das nächste ist das AI bei der Detektion, der Rundherde greift. Dann greift die AI wahrscheinlich in der Zuordnung, ob ein Herd, der gefunden wurde, wie hoch die Wahrscheinlichkeit ist, dass der bösartig ist oder nicht. Das ist das spannendste Thema wahrscheinlich. Und dann wäre es schlau, wenn wir auch AI hätten, um diese Risikomodelle da zu füttern oder die noch besser auszuwerten. Weil die einzelnen Teilnehmer haben ja durch Alter,	

		Geschlecht, Vorgeschichte ja mehr Risikoprofildaten als nur ihre Raucheranamnese. Ehm, und wenn man dann ein individualisiertes Profil für den einzelnen Herd den man findet oder für die einzelne Person erstellen kann, dann kann man darüber wahrscheinlich auch individualisierter die Teilnahme an diesem Programm steuern.	
1.11	JM	Ich habe mir auch überlegt. Theoretisch wäre es auch sinnvoll, wenn man, sage ich mal, frühere Patientendaten auch einbeziehen würde in Diagnose. Aber das ist jetzt auch zu kompliziert, das irgendwie an das System anzubinden, oder? Oder wird das gemacht?	
1.12	R1	In Deutschland ist das schwer bis gar nicht möglich.	
1.13	JM	Wegen Datenschutz?	
1.14	R1	Also in Deutschland würde das so laufen, dass die Anamnese noch mal erhoben wird. Und was der Teilnehmer in diesem Zusammenhang erzählt, das wird erfasst., aber dass sozusagen über Krankenkassendaten auf die Vorgeschichte des Patienten zugegriffen wird, das ist in Deutschland nicht möglich.	
1.15	JM	Okay, vielen Dank. So wie der Prozess, wie Sie ihn gerade beschrieben haben und so wie der auch vielleicht schon mal bei Ihnen getestet wurde. Hat das so gut funktioniert oder gab es da Schwierigkeiten?	
1.16	JM	Also wenn sich das eingelaufen hat, dann gibt es da keine Schwierigkeiten.	
1.17	JM	An, an, okay, ja gut, wir werden wir später nochmal präziser auf die Schwierigkeit eventuell eingehen, weil okay, aber naja.	
1.18	JH	Ja dann würden wir noch mal zu, äh ja, den genaueren Fragen zu den Workflows der Radiologen ja genau mit dieser AI eingehen, Sie haben ja schon gesagt, an welchen Stellen die CT Scans ausgewertet werden. Ähm und was für Funktionen genau haben diese Systeme, die die Radiologen unterstützen? Also Ihnen werden dann wahrscheinlich die, ja halt so, Vorerkennung durch die AI angezeigt. Aber gibt es noch weitere?	

1.19	R1	Also formal ist es unklar, ob es erlaubt sein wird, dass die AI zuerst guckt und dann der Radiologe. Inhaltlich bin ich da aber völlig dabei, die AI muss vor gucken. Und dann zeigt das Programm, das die Bilder analysiert hat, die Befunde an und dann guckt der Radiologe sowohl die angezeigten Befunde als auch die komplette Untersuchung noch mal an und dort wird es darum gehen, dass das System erst mal die Herde findet, dass sie die Herde automatisch vermisst oder wo volumitriert und dass sie dann eine Einschätzung abgibt, die entweder nach diesem Lung-RADS Einschätzung gehen kann oder nach irgendeinem anderen, ähm, Baum, nachdem sozusagen der schrieb.	
1.20	JH	Genau. Die Frage war dann, welche Funktionen, die die Radiologen genau in ihren Workflows unterstützen, also wie die AI Systeme sozusagen auch als Second oder First Reader, äh, also was für Funktionen die haben um die Radiologen zu unterstützen.	
1.21	R1	Also, die sollen die Funktionen haben, dass sie die Herde finden, dass sie die Herde vermessen und dass sie die Malignitätswahrscheinlichkeit dieser Herde bestimmen. Und was da noch dazugehört, das habe ich auch vorhin vergessen, ist das er natürlich Voruntersuchungen vergleicht, das heißt Herde, die beim ersten Mal schon da waren, automatisch wiederfindet und die Veränderung, wenn es eine gibt, bestimmt. Oder dann halt auch sicherstellt, ob ein Herd neu ist oder ein Herd verschwunden ist. Und dann muss da ein strukturierter Befund dazu erzeugt werden, wo diese ganzen Maßzahlen entsprechend drinstehen.	
1.22	JM	Wissen Sie, wie das System heißt, das Sie da benutzen?	
1.23	R1	Ah, das können Verschiedene unterschiedlich gut. Also das gibt es von Siemens, das gibts von Philips, gibt es ja mindestens fünf verschiedene kleinere Firmen. [AI software, anonymized] ist im Moment in der HANSE Studie angesagt, aber es gibt auch genug zahlreiche andere, die dieses oder ähnlich versuchen zu machen. Mit der AI zur Maliginitätsabschätzung. Da sind die alle noch in den Kinderschuhen, aber einige haben schon AI da drin.	
1.24	JM	Genau, äh zum Beispiel, das [AI software, anonymized] was auch bei der HANSE Studie benutzt wird. Das berechnet zum Beispiel auch noch den Koronarkalk. Denken Sie, das ist sinnvoll oder warum wird es bei Ihnen nicht gemacht?	
1.25	R1	Das ist sinnvoll. Aber das ist eigentlich eine Erweiterung des Programms, die, wo ich im Moment das Gefühl habe, dass es inhaltlich sinnvoll, aber politisch, wenn wir erst mal das Programm durchsetzen wollen, verwirrt das nur, weil da musst du wieder andere SOPs und andere Workflows dahinter schalten. Und dazu gibt es teilweise nicht genug Evidenz.	

1.26	JM	Das heißt Sie denken, um erstmal wirklich Lung Cancer Screening auf den Weg zu bringen, sollte man sich erstmal auf die Funktion der Lungenknötchen Erkennung und der Klassifizierung reduzieren.	
1.27	R1	Ja.	
1.28	JM	Okay, ja.	
1.29	R1	Also was viele von den Herstellern jetzt machen, dass sie dann schon sozusagen ein umfassendes Lungen oder Thorax Paket anbieten, wo eben nicht nur die Knötchen drin sind, sondern da sind auch Atemwege, Emphysem und die Verkalkung drin. Also wenn wir aus der Brille der Software Tools denken, dann macht das natürlich Sinn, dass man das als Module nebeneinander angeboten hat. Und dann kann jeder sich die Sachen raussuchen, die er jetzt dazu nutzen will oder entscheiden, auf welche er verzichtet.	
1.30	JM	Ja, macht Sinn, ja.	
1.31	JH	Oky danke. Gibt es denn einen vorgeschriebenen Prozess, wie die Radiologen die AI einsetzen sollen? Oder benutzt das jeder so wie ihm das lieb ist?	
1.32	R1	Im Moment gibt es keinen fertigen Prozess, was man darf und was man nicht darf.	SV
1.33	JH	Okay. Und, ähm, was wäre Ihre Meinung? Wenn dann das Lungenkrebs Screening flächendeckend eingesetzt wird?	
1.34	R1	Ja, ich glaube schon wir brauchen ein SOP, wo drinsteht, was man in welcher Reihenfolge wie zu benutzen hat. Also, es wäre bei einem nationalen Programm natürlich schön, wenn alle das gleiche System verwenden würden, weil die Ergebnisse auch zwischen den Systemen variieren. Das halte ich aber in Deutschland und wettbewerbsrechtlich für schwer möglich.	SV
1.35	JM	Ja. So wie ich das jetzt auch noch mal verstanden hat, um noch einmal darauf zurückzugreifen. Meinten Sie, dass sich das System, das ja erst anguckt, und dann der Radiologe, die, den Output des Systems dann benutzen würde, also momentan, wie Sie sich das auch vorstellen, oder wie das bei Ihnen gemacht wird, ist AI als First Reader, oder?	
1.36	R1	So stell ich mir das vor, ja.	
1.37	JM	Warum? Oder was halten Sie davon, wieso nicht AI nicht als Second Reader? Oder was ich auch oft gesehen habe, ich glaube wie es auch bei [AI software, anonymized] gemacht wird ist parallel sich das anzugucken auf zwei Screens.	

1.38	R1	Also parallel auf zwei Screens, äh, macht wahrscheinlich Sinn. Aber mit First Reader meinte ich das, das System muss die Bilder automatisch, wenn die aus dem Scanner rausgucken, äh, vorprozessieren.	
1.39	JM	Ja, das stimmt ja.	
1.40	R1	Also ich halte nichts davon, dass ich als Radiologe, die zuerst ohne diese Annotation gucken muss, weil ich glaube nicht, dass, dass es erstens dauert es länger, und ich glaube nicht, dass das diese angebliche Unabhängigkeit bessere Ergebnisse bringt.	
1.41	JM	Ja, ähm, und so wie die AI momentan, sage ich mal, funktioniert oder das vorher auswertet das Bild. Sind Sie damit zufrieden oder könnten Sie oder würden Sie sagen, die AI könnte eigentlich vielleicht, es wäre ideal, wenn sie noch ein bisschen mehr können würde für genau den vorgesehenen Prozess der der Lungenknötchenerkennung, könnte es noch besser helfen oder sagen Sie es macht alles top.	
1.42	R1	Also besser, geht immer. Und das andere ist mit dieser Wahrscheinlichkeit, ob etwas maligne ist oder nicht. Das ist in der Validierung ja unglaublich schwierig. Ja, weil. Ehm, jenachdem wie der Preselektionsbias von der Kohorte ist, auf denen ich das entwickelt habe, werden, wird die Leistungsfähigkeit unterschiedlich sein. Ja, also Möglichkeiten für Verbesserungen gibt es noch beliebig. Soll es heißen.	OQ
1.43	JM	Ja okay. Einfach nur um eine konkrete Möglichkeit mal zu beschreiben, wie und jetzt nicht nur abhängig davon, dass das System besser funktionieren könnte, sondern funktional, ob es noch etwas tun könnte. Fällt Ihnen direkt da irgendwas ein, was das System vielleicht noch machen könnte? Irgendwas Konkretes?	
1.44	R1	Also ich glaube, dass es automatisch dann mit dem Lung-RADS Klassifikation oder der Klassifikation, für die man, nachder man arbeiten will, dass er die Bestimmung sozusagen automatisch macht, dass man das nicht mehr nachgucken muss.	OQ
1.45	JM	Ja.	
1.46	R1	Ähm, dass sie, dass sie diese ganzen Maßzahlen, also auch diese Volumenverdopplungszeit automatisch berechnet und einträgt. Und das andere ist dann mehr IT Konnektivität, glaube ich. Das ist ja nicht dann nicht AI, aber es muss sozusagen möglichst über problemfreie Schnittstellen in den strukturierten Befund und dann eben zu diesem niedergelassenen Lungenfacharzt, der diese Person eingeschlossen hat, muss das dann automatisch durchgehen.	OQ

		Ja, nicht, dass wir PDFs erzeugen müssen, die wir dann nur anfangen zu faxen und so ein Kram, ja.	
1.47	JM	Wir würden noch mal ganz gerne konkreter auf das System zu sprechen kommen und da haben wir ein paar Fragen nochmal Ihre Meinung zum System zu testen. Da würden wir anfangen mit, ehm, ja okay haben wir gerade schon berührt den Punkt genau, ob das System alle Informationen, die für die Funktion die es erfüllen sollen, auch zur Verfügung stellt. Da haben sie gesagt, genau da, es könnte eigentlich noch genau die Erkennung besser machen oder genau die Klassifizierung nach dem Lung-RADS 1.1, ne?	OQ
1.48	R1	Ja.	OQ
1.49	JM	Ja, dann würde wir direkt weitergehen und fragen, wie würden Sie sagen, ist der Output des Systems in Bezug auf die Menge oder Relevanz und auch die Genauigkeit? Also ganz breite Frage.	
1.50	R1	Also ich glaube, dass das System muss natürlich mit dem Volumen auskommen und ich bin immer der Meinung, dass die Systeme nicht super genau sein müssen, sondern die Systeme müssen robust sein. Eben weil damit die sozusagen in der Real World überleben können, müssen die einfach funktionieren und objektive wiederholbare Ergebnisse bringen. Er optimiert immer, ob wir immer nen Messfehler von einem Millimeter haben oder nicht, das ist am Ende des Tages völlig egal. Solange sozusagen nur ein, ein irgendwie gestalteten Offset haben von der Wahrheit ist die Verhältnismäßigkeit ja hoffentlich die gleiche und das reicht eigentlich völlig aus. Also ich finde, manchmal werden diese Dinge zu sehr von Physikern gemacht und betreut, die ganz andere Herangehensweise an Genauigkeit und Fehlerbreiten und so was haben. Ich glaube nicht, dass wir dort hyper genau sein müssen.	SR OQ
1.51	JM	Ja, aber auch wenn das jetzt sag ich mal die, die Produktionsgruppe jetzt nicht die, na ja, aber die, die verstehen schon was das System können muss. Also der Output des Systems auch komplett relevant oder zeigt auch irgendwas an, was sie gar nicht brauchen?	
1.52	R1	Ah da gehen die Meinungen der Radiologen auseinander. Also da gibt es schon Dinge, die angezeigt werden, die ich nicht brauch. Aber normalerweise...	OQ JR
1.53	JM	Die andere Radiologen dann brauchen oder die allgemein nicht gebraucht werden?	

1.54	R1	Da gehen die Meinungen auseinander. Also es gibt Radiologen, die sagen, sie finden das jetzt toll, dass sie irgendeine Information haben und andere brauchen das nicht. Also da fängt es zum Beispiel an, die Texturparameter, inwieweit ich die dann detailliert ausspiele oder ob der mir eine anatomische Zuordnung macht, in welchem Lappen oder in welchem Segment der Herd ist.	OQ JR
1.55	JM	Ja.	
1.56	R1	Da kann man sich ja noch diverse Dinge vorstellen, die so ein System kann. Und da wird es Radiologen geben, die das wollen und andere, die das nicht brauchen.	
1.57	JM	Ähm, kurze Nebenfrage noch mal müssen Sie eigentlich um Punkt halb weg oder wäre es okay, wenn fünf Minuten überziehen?	
1.58	R1	Ich muss eigentlich in die andere Videokonferenz.	
1.59	JM	Okay, dann müssen wir uns mit den letzten Fragen jetzt mal ranhalten.	
1.60	R1	Ich kann auch vier Minuten später reinkommen, glaube ich.	
1.61	JH	Ähm, genau. Also wird der Output des Systems für die Radiologen denn gut dargestellt und übersichtlich, also haben die Systeme eine gute Usability?	RD L PEUT
1.62	R1	Ja, finde ich schon.	RD L PEUT
1.63	JH	Direkt weiter, wie würden Sie sagen, verändert die Nutzung des Systems den Output der Radiologen? Ich glaube, Sie haben bereits gesagt, es wird vielleicht ein bisschen genauer. Werden die Arbeiten zeiteffizienter?	
1.64	R1	Also ich glaube, es wird viel schneller und es wird objektiver.	PU
1.65	JH	Okay, ja.	
1.66	R1	Also, das System arbeitet sicherlich. Also wenn wir Radiologen als Erstbefunder haben und dann die Inter-Reader-Variabilität angucken, dann ist die vielleicht relativ hoch. Wenn wir das System zuerst haben und dann guckt der Radiologe drauf, wird dadurch die Inter-Reader-Variabilität sicherlich deutlich kleiner.	PU

1.67	R1	Ja. Kommen die Radiologen denn alle gut klar mit dem AI-System oder gibt es da noch irgendwelche Schwierigkeiten an denen, die angegangen werden sollten?	
1.68	R1	Also die meisten Radiologen nehmen das sowohl für so eine neue Applikation als auch eine klinische Routine relativ positiv auf. Ja.	PEUT
1.69	R1	Gibt es denn trotzdem mal irgendwie Probleme in der Benutzung und können dann auch die Radiologen Hilfe bekommen, wenn sie da Probleme haben oder gibt es auch mal technische Schwierigkeiten?	
1.70	R1	Also erstens, ich bin sicher, dass Radiologen, die müssen das wollen, ja. Also egal welches IT-System man zur Befundung verwendet, man muss sich mit dem neuen System anfreunden wollen. Dann kann man die Hilfe, die man bekommt, auch verstehen. Und ansonsten die IT Affinität der Radiologen ist natürlich unterschiedlich und die einen haben mehr Spaß dran als andere, sich mit solchen Themen zu befassen und sich dann auch da richtig rein zu knien und sich reinzuarbeiten.	SV PEUT RU
1.71	JM	Ja.	
1.72	R1	Ich glaube aber, dass in so einem Lungenkrebs Screening Ansatz, dass ganze so verschult sein wird, dass das keine Rolle spielt.	SV PEUT E RU
1.73	JM	Okay, ja. Äh, können Sie sich an irgendeiner Stelle auch noch Potenzial vorstellen wo momentan der Arbeitsprozess der Radiologen effizienter noch gestaltet werden könnte? Oder gibt es irgendwo eine Stelle, wo Sie sagen, hier wird noch unnötig viel Zeit für irgendwas verwendet?	
1.74	R1	Also die Zeit geht vorher und nachher drauf in dem Befundungsprozess, wenn der entsprechend unterstützt wird und wenn der Workflow entsprechend beschrieben ist, standardisiert ist, glaube ich, kann man nicht mehr viel einsparen. Die Effizienz entsteht dann in der Kommunikation. Oder was passiert, wenn irgendein Herd gefunden wird, der in der Bewertung halt an irgendeiner Grenze liegt? Ja, aber dann haben wir ja doch vorgesehen, dass dann eine Zweitmeinung eingeholt wird, oder dass das in so einer Art Tumor Board vorgestellt wird, wie man jetzt mit dem Ergebnis, das man bekommen hat, tatsächlich umgeht.	PU PT
1.75	JM	Ja, ja.	

1.76	JH	Wie denken Sie denn, würde der Arbeitsprozess der Radiologen aussehen, wenn jetzt keine AI dazu gezogen werden würden? Wäre der noch derselbe oder wäre das Lungen Cancer Screening an sich gar nicht umsetzbar? Ohne diese AI Unterstützung?	
1.77	R1	Dann würde das so aussehen wie beim Mammografie Screening. Und meines Erachtens wäre das heutzutage nicht mehr umsetzbar, dass wir so viele Radiologen hätten, die ne Erstbefundung, ne Zweitbefundung und dann Adjudication-Befundung machen, das ist meines Erachtens nicht umsetzbar.	JR PU
1.78	JM	Wer genau würde jetzt eigentlich die AI benutzen? Sind es die klassischen Radiologen, die da auch schon jetzt bei Ihnen sowieso angestellt sind? Oder würden Sie sagen, es wird ein neuer Full-timejob, wo eigentlich auch neue Radiologen einstellt...	
1.79	R1	Also wenn das richtig abhebt und da Teilnehmer kommen, dann wird das ein ziemlicher Fulltimejob und außerdem würde eine bestimmte Zertifizierung Qualitätskontrolle in der Ausbildung vorgesehen. Ähm, das heißt, es wird da schon eine spezielle Gruppe geben, die das macht.	PEUT
1.80	JM	Okay, ja, dann kommen wir schon zur letzten Frage. Sehen Sie noch irgendwelche Hürden oder Probleme, die vor einer erfolgreichen Nutzung des Systems in Ihrer Klinik jetzt irgendwie stehen würden oder die auf Sie zukommen könnten?	
1.81	R1	Also das Schwierigste ist erst mal, dass das Thema mit dieser Erstbefundung durch die AI überhaupt angenommen wird. Und das andere ist natürlich die Zulassung von dem Gesamtsystem, ist klar, von dem Gesamtprogramm. Es ist immer die Frage, inwieweit man sich auf so eine Black Box einlässt, weil am Ende des Tages versteht ja keiner, was das System eigentlich genau macht, um dahin zu kommen. Und, ähm dann, was wollte ich noch sagen? Ich glaube, dass es wichtig sein wird, dass man ein möglichst einheitliches System hat in Deutschland. Also ich hatte es ja vorhin schon gesagt, wenn man vier verschiedene Systeme in Deutschland hat, kommen unterschiedliche Sachen raus. Was vor allem dann, wenn die Teilnehmer den Wohnort wechseln und ihre Untersuchung machen, dann nicht gut.	
1.82	JH	Okay.	
1.83	JM	Oh, darf ich dann noch eine weitere letzte, zweitletzte Frage stellen? Warum wird denn bei Ihnen das System verwendet, das gerade verwendet wird und zum Beispiel nicht [AI software, anonymized]? Oder was hat für Sie die die Wahl des Systems motiviert?	

1.84	R1	Also meistens ist es so, dass man zu irgendeinem bestimmten Zeitpunkt Geld hat und irgendwas beschafft, beschaffen kann, und dann nimmt man das, was zu dem Zeitpunkt das ist, mit dem man am besten klarkommt. Das ist eine ganz praktische Antwort, aber die spielt da häufig ne Rolle.	
1.85	R1	Seit wann haben Sie das System, was Sie gerade benutzen, eigentlich?	
1.86	R1	Also wir haben jetzt neu ein System von [AI software, anonymized], was wir im klinischen Kontext getestet haben und mit dem wir sehr zufrieden waren. Und das macht die Schnittstellen so, wie wir sie haben wollen. Das System, das wir vorher hatten, war mehr ein experimentelles. Das hat die Schnittstellen nicht gemacht, wie wir sie haben wollten. Ja wurden die Bilder nicht im PACS vernünftig abgespeichert. Und so weiter. Ja, also es entscheidet sich dann häufig an solchen Dingen. Aber das ist das, was ich vorhin meinte, es ist eigentlich relativ egal, wie gut das System ist. Also was die Messung betrifft, nicht wenn es darum geht, man darf natürlich nicht zu viele falsch positive Befunde anmarkern. Also früher hatten wir ein [AI software, anoynmized] System, das hat überall Pünktchen gemacht. Damit kannst du natürlich nicht arbeiten.	OQ
1.87	JM	Super, danke.	
1.88	JH	Ja, vielen Dank. Ähm, haben Sie noch irgendwelche Anmerkungen oder gibt es irgendwelche wichtigen Aspekte, die Ihrer Meinung nach noch zu kurz gekommen sind, die Sie gern erläutern würden?	
1.89	R1	Nein, ich glaube, ich habe alles erzählt. Sie haben einen schönen Fragenkatalog gehabt.	
1.90	JH	Okay. Haben Sie sonst noch irgendwelche Fragen an uns?	
1.91	R1	Nein, ich glaube so weit erst mal alles klar. Ich denke, Sie stellen das mal zusammen und dann schicken sie es ja nochmal vorbei. Und wenn Ihnen nachher noch was einfällt oder mir, dann können wir es per E-Mail auch nochmal austauschen.	
1.92	JM	Super, klasse.	
1.93	JH	Ja dann noch einmal vielen, vielen Dank, dass Sie sich die Zeit genommen haben.	
1.94	R1	Alles klar, gerne, tschüss.	

1.95	JM	Schönen Tag noch.	
1.96	JH	Ihnen auch. Dankeschön, tschüss.	

Appendix E - Interview 2

Duration: 69 minutes

Date: April 28, 2022

Interviewee: R2

English

Row	Speaker	Transcribed text	Codes
2.1	JM	Can you tell us in two to three sentences who you are and in what context you have been involved with lung cancer screening so far?	
2.2	R2	So I'm a radiologist. I've been in the job for a long time, a specialist in radiology, also have quite a lot of experience with lung imaging or with thorax imaging, worked in different centers, first in [city, anonymized] and [city, anonymized] and now here in [city, anonymized]. Yes, and since August of last year, I've been co-managing a study where we do lung screening with an AI, where we screen smokers for the presence of lung cancer, just all the people who smoked once or who still smoke from the age of 55 to I think 79, are then allowed to present and come by. And that is evaluated in this study and I participate in that and for that, I had to do a couple of certification courses, from the ST, from the Thorax Society, I did that of course somehow, that you then kind of know the pitfalls of lung cancer screening, what to look out for, that was just such a course with several lectures from very renowned people and that was also very interesting.	
2.3	JH	Was that for the process of lung cancer screening, that is, for the image evaluation, or also already with the handling of the AI? Were there also courses on this?	

2.4	R2	Yes, yes, yes. The one…that… it was a bit integrated. This was assumed a bit, that what was talked about was done with AI. That wasn't, that wasn't a specific course on AI, but it was just like that, the working course in dealing with the patient, with the AI, with the knowledge that you had from other larger studies, screenings like that.	
2.5	JM	This was done in advance before the whole thing was introduced also now in [city, anonymized] before the study started?	
2.6	R2	Yes, that is right.	
2.7	JM	Okay. And again, so I've understood correctly. This was generally about working with AI or what?	
2.8	R2	No, it was actually more about the epidemiology behind these studies. So that was then said, which studies have already been conducted, the NELSON study, this study, the large European study. What were the results there? What do we have then in this study? What should we pay attention to? What should be investigated additionally? Like this.	
2.9	JM	So, let's move on to the first question, which is a bit more difficult. If you had to think now about how lung cancer screening would be introduced in Germany, where can you imagine? Or where do you think it would be introduced? Where would it be offered and for whom would the additional work be? Would you say, for example, certain practices do it, all hospitals do it, university hospitals do it? Or how can you imagine this being implemented?	

2.10	R2	I could imagine that being a bit like mammography screening. There would have to be very specific centers, it would need certification and then people would come there especially for that. I think that this would then be part of the normal everyday routine, I don't know, it could simply run completely parallel to the everyday work of the radiology practices, so to speak. It is basically a bit like that already. I actually don't think it's so bad that people sit down specifically just for this, and that people come specifically just for this and that this runs as a separate program, so as examinations, which are also the case in addition to the normal clinical workload, they are actually not patients in the sense. They are healthy people. Theoretically healthy people.	PT
2.11	JH	That sounds like it would be a lot of extra work, do you think the capacities would be sufficient at the moment? Or would it be necessary to hire additional radiologists who are explicitly commissioned for this, so to speak?	
2.12	R2	It's a little bit difficult to say, theoretically, if you're going to do this on a larger scale or if you always want to do this. I mean, basically, we are radiologists who are specially assigned for this. It becomes a part of my working time or is compensated for it somehow. I'm just working on the study at the moment. I don't just look at the knee again and then look at the lung again, but I really sit down and spend a few hours just doing that.	
2.13	JM	Yeah, but I mean, you do that then maybe one day a week or two days a week, right?	
2.14	R2	It always depends, the truck is only with us every six weeks, and then it's there for two weeks and we are three people who share it. Four, actually four, three, four people, and um, ye,s and everyone is on it, so to speak. Most of the time I'm there for two full days. Of course, what do I know, if I work on Thursday, and Friday, then I can work on Saturday and Sunday as well, it's not like that. That doesn't go away, but it's actually the case that I actually only do that during that time, and that also makes sense, because once I open the interface and create this working environment for myself and so on, then it actually makes sense to stay with it for the time being and to go away with a certain amount of patients.	
2.15	JM	Because you familiarize yourself a little bit, or how?	

2.16	R2	Yes, that too. And above all, you first have to open it with 1,000 passwords, as is always the case, and then I imagine that if I were to do that, it would be a certain effort to switch back and forth, if I were to have three people with the knee beforehand, I would now make two more lungs from a completely different surface and then I would go back to the knee again.	PEUT PT
2.17	JM	Yes. So, in order to relate this to our topic, which is also about the workflow of the radiologist, to classify it more precisely somehow. So you say it's a lot of work, I'll say, to diagnose the lungs with AI, if you've done something else before, because you have to call up all these programs first. And that's also why it makes sense to do a lot of this on the back end so that you don't always have this switch between the systems.	
2.18	R2	Exactly. Exactly. So actually, that's what this is. Of course, that's really only three minutes or so, but that's of course, that accumulates over the course of the day, and I do, I do five, 5 to 6 of them per hour. So that means at the moment, that is, if I… Exactly. So, if I switch back and forth so often, then you lose something like 30 minutes or so, that's not that little.	PT
2.19	JM	But then overall you're always working that away for two days when it comes up again. But that is now also not super much overall more work just for the lung clinic and you would say that in general, even if it is implemented overall in Germany, you would also manage to get that done for the most part, right?	
2.20	R2	You are frozen right now. I hope my internet is enough here.	
2.21	JM	I think that's up to us, unfortunately.	
2.22	R2	But as far as I can still hear you, all is well.	
2.23	JM	Exactly. Yes.	
2.24	R2	What did I want to say? What was the question again? Well, it's a completely different kind of work. In the [institution, anonymized], people come to me with a clinical problem. There I look at it with a completely different eye, this screening eye.	

2.25	JM	Yes	
2.26	R2	That is already so, that is time, which I use only for this program, it would be now in such a way, if now a patient would come and would have something completely different and would have then nevertheless still this lung cancer screening, which I would let run then so, I would say, I would have that in a program in it. Maybe you can imagine that I look at the knee first, and then I switch to the lung screen and then I go through it.	
2.27	JM	Yes.	
2.28	R2	Then it would probably cost me an extra ten minutes.	PT
2.29	JM	Okay. Yes. Yes.	
2.30	R2	That would help, but of course, it's time. As I said, I can about five to six per hour. That is, so ten minutes per patient is realistic.	
2.31	JM	Yes, then of course ten extra minutes are critical again.	
2.32	R2	But that is then extra time. So it's not that the, well it's also that we then look at a whole lot of other, so we also always look at the calcium score in the program and we also always look at the emphysema. If I were to look purely at Tumor now, it might just, and the AI helps me, then it might only cost me five minutes.	
2.33	JM	Yes.	
2.34	R2	So, I think.	
2.35	JH	Will that also be part of, or do you think it will also be part of the official lung cancer screening program that you then look at other things independent of lung cancer because you're already in the process anyway?	

2.36	R2	That's a good question. Would make sense, of course, because all smokers also have this tendency to get the calcium on the coronary arteries somehow. And if you examine that anyway, then you can easily look at that. I think that's a sensible thing to do. And emphysema they all have anyway. So, I mean, quantifying that. We're just trying to find out in this study how much quantification, so once you're told you have emphysema, whether that changes your behavior, that's part of the study, I kind of think that some, it's randomized, some get told, some don't.	
2.37	JH	Yes, do you mean smoking behavior or what behavior?	
2.38	R2	And whether smoking behavior changes. Exactly, whether the people who are told ui, ui, ui, there are already 17% no longer there, whether they then rather say then even more or whether they then say now no more.	
2.39	JM	Can you then explain again in more detail the whole process of screening a patient in lung cancer screening? So, you said there are three different functions and sometimes it's told to the patient, sometimes it's not told to the patient, can you kind of try to explain exactly the overall process there?	
2.40	R2	So, it's more complex, they also come for a clinical examination beforehand and only then get their CT. There they are questioned and, of course, there are also inclusion and exclusion criteria. They must not have had any tumors in the last ten years, I believe. They don't have to know about the tumor themselves, and otherwise, they are asked about things. So it is always listed whether they use statins. But that's I think as part of this calcium score then whether they that's a drug that lowers blood levels somehow. And then it asks if they smoke naturally or not or if they smoked. And then, or I think still, whether they're an active smoker is also the question. And then they ask if you have a family history of lung cancer.	
2.41	JM	Yeah, okay.	
2.42	R2	Those are the three things that I can always look at in the study, which, so the other things I could, so there's an extra portal, but that's a different portal where the clinical data is collected. But I usually don't look at that because it actually biases me when I look at that, so to speak. And then the patients get their CT, and then	

		they go home again, and they are told that they will get the results sent to their family doctor in the next few weeks.	
2.43	R2	Sometimes also the patients themselves. I think they can specify that if they want to.	
2.44	JM	Then separately, later, will this CT scan be evaluated?	
2.45	R2	Exactly that will be evaluated later. This is then run during the day and if it's good, then I might also sit on it online. If that's the case, I only work on Thursday and Friday, so mostly on Thursday, when I work on Thursday and Friday, I evaluate at least the results from Wednesday, Thursday, Friday, and sometimes even the results from Tuesday. That's how it is from the workload. The patients are examined every quarter of an hour and on one day between 15 and 20 patients are examined.	PT
2.46	JM	Yes. Yes.	
2.47	R2	And then I always work 3 to 4 days away but depending on what's going on. And sometimes my colleagues are also online. Then I don't have so much to do and sometimes I work a little bit after. And the first time, I have to say, the first cycle was terrible. Nothing worked at all. The system broke down all the time and we had big server problems actually. That was stupid because then we couldn't notify the patients online at all. That was a bit difficult. But since then, it's actually working well and sometimes we really do send out the findings almost on a daily basis. I think that's also relatively important.	RUS PEUT
2.48	JM	What exactly were those server problems again?	
2.49	R2	I don't even remember what exactly was broken there, that the IT at that time, it was constantly like that, so I opened the program and the connection was gone.	SR
2.50	JM	On the AI program or on which program or the program to inform the patient or the AI program. Sorry, I didn't quite understand that.	

2.51	R2	No, but my own program to evaluate the software. So, the [AI software, anonymized] program worked, remained then black. So, you don't have access to the [AI software, anonymized] and then that was and of course you couldn't diagnose, which was a bit painful.	SR
2.52	JM	[AI software, anonymized], so that's the AI that's supportive of the findings, right?	
2.53	R2	That's the AI from [country, anonymized] that we're working with, exactly.	
2.54	JM	But theoretically, you can also find that without the AI. Or would that take too long and that is why it is not done?	
2.55	R2	But then I don't have the images at all. So, the images were at that time from the CT directly to this, this, this [country, anonymized] server played.	
2.56	JM	I thought they will be displayed in the PACS reader anyway again in parallel without the AI.	
2.57	R2	But no, that's actually the case that they were not transferred to general PACS in theory.	
2.58	JM	Okay, that means they will be displayed directly with the AI only.	
2.59	R2	Exactly, they are displayed directly with the [AI software, anonymized], they are only transferred to the PACS, if I order that manually because the patient has a problem and will come back, then I can send that directly via DICOM to the PACS, so to speak. But I do that, so theoretically they are in the normal PACS, they are not there, they are only on this one server and from there the data is somehow now collected in [city anonymized], I mean, is the procedure now. But that was not the case at the beginning, that came later.	
2.60	JM	Can you take us through the process of how you diagnose it? So, when you sit down in front of the picture, or the initial situation is apparently...	

2.61	R2	Exactly, I can do that. I open the screen, there is a list of patients.	
2.62	JM	So you get on it and there it was already processed by the AI right?	
2.63	R2	I let it process it at that moment. So, I open the patient and then I say, and then I have a menu on the side, which I click on. And then it's like I open the patient and then I say, I would like to have the LAA now, so the emphysema evaluation. And I click on it and then it takes a moment to run through it.	PT
2.64	JM	Okay, then the AI runs over it, okay.	
2.65	R2	First, I have to segment it. That's always the case, I must check it or at least the segmentation. It then shows me a segmentation, I then look at it to see if it's correct. So middle lobe, upper lobe, lower lobe, just these lung segments and release those. And then I can determine the emphysema from there. And then it also always tells me how much emphysema is in which lobe. That's why it's important that that's correct. And that runs but at the moment, anyway, I always have to wait a little bit while that… while that the surface, until the surface opens up. I think it's only in the moment that it runs through, anyway, for me. Somehow.	PT
2.66	JM	How long does that take, approximately?	
2.67	JM	One, one minute like this.	PT
2.68	JM	Oh, is that a lost minute where you can just sit there, right?	
2.69	R2	Yes, it goes relatively quickly, so the first goes relatively quickly. Opening the patient takes a little bit, then the surface opens, then the lung segmentation goes relatively quickly, then the emphysema evaluation takes a little bit of a moment. Then I switch to the calcium score evaluation. That's also that's all displayed. CIC it says then I click on that and then another screen opens up. So, everything is now so far in the soft kernel, always in the soft kernel. So, I have to switch once at the beginning, I mean. I have to once… It opens... The problem is that once, how is that, I open that and say I want this lung segmentation and then it always	PEUT

		already wants that, that I release the patient, because it somehow understands, I'm already done, but somehow I have to click that away then. Then I say no, I'm not done yet. Then it opens this lung segmentation for me. Then I go to the, then I'm in the soft kernel and we look at the emphysema in the soft kernel, the calcium score in the soft kernel. Then I go to the lung screen, so the cancer evaluation. Then it opens a screen where I can choose the kernel, so to speak. And then I must click on the other kernel, the high-resolution one, the lung kernel, which is high-resolution.	
2.70	JM	Right.	
2.71	R2	And it has another signal to noise and that takes, that takes then really so when this lung screen program then opens, that takes at least a minute, maybe even two, that there you sit and wait.	PT
2.72	JH	Okay.	
2.73	R2	Yes, exactly. And then I actually always do it in such a way that I, I could then already, I can then click on this CAD, but I always look at the lungs once myself. Just screen it once to see, because sometimes there are things that it doesn't recognize, whether there are any things that are conspicuous to me, so to speak, then I go through it once myself. And then I do the CAD once and see what it gives me. And sometimes it also gives nonsense. I then click that away. Then I can take that out on the CAD and then, or it recognizes very large herds, there he has the tendency not to recognize, for example. So, I once had a really, well two, two times we had really large foci, which he then did not even recognize as a pulmonary focus. Because it was not a pulmonary focus, in the sense of a pulmonary focus recognition, it was somehow so bizarrely configured and was not nice.	OQ PEUT
2.74	JH	I would imagine that with the big ones then it's especially critical if you don't recognize them, right?	
2.75	R2	Yes, exactly, of course. But on the other hand, you don't need to use CAD to recognize them.	OQ
2.76	JM	But still, well, then you also ask yourself how good the system is really what I'm using here if it does not recognize something so big, right?	

2.77	R2	Yeah, but I think that's so, that's so big that it thinks that's an organ or so I think that thinks that's a gastric hernia or that's a, so that's so, if the thing, don't know how big the part was, like six or seven centimeters, then it thinks it's probably wrong. So that's how that's entered.	
2.78	JM	What I wonder is, the system only identifies and displays the lung nodules there, or? Or does it display...	
2.79	R2	No, no, no, it also shows a histogram in the meantime, that's new. It shows me the lungs, it shows me on there if they contain fat, it shows me the lung nodules and it then also specifies them right away. So, when I click on the individual ones, it tells me that's from the volume and so there's this Lung-RADS, that's then 1 to 4 and then there's 4A and B and X and exactly. And it always tells me the same thing, so when I click on the individual nodes, it always tells me what kind of nodule it is and I just look, is that right, so to speak.	RU OQ
2.80	JM	Oh, that classifies that also directly according to this Lung-RADS 1.1, or what is the current?	
2.81	R2	It does it, it does it, and it does it, so it classifies the individual lung nodules right away, and then I just look at them, and otherwise, there are cases where you say that this can never be a nodule of size three. It is simply not. And then you look at what it has measured there and sometimes you have to correct it a little bit because it measures some caesurae or something adjacent. Any vessels that border there or so, and then, so you always have to look over it once, that also learns a bit of the system. And then, um, and recently it also gives us histograms, which show how dense the focus is, so to speak, or how much soft tissue, how much calcium, so how much away from the visual aspect? Is it a dense focus? Is there calcium in there? So is that natural, if it's really calcium in it, then it indicates that, and then it's automatically judged a 1 focal, so it's judged benign. But if it's a little bit bright, this is something you can see, if it's more of a necrotic focus, if it's more of a dense focus somehow, that's shown by the histogram.	L
2.82	JH	So, the histogram already helps then.	
2.83	R2	That helps, too, yes.	

2.84	JM	I can imagine that especially if you don't agree with the AI now, for example, about which the classifications of the size of the lung node now take more time, then take a closer look at it in order to somehow change the measurement of the AI system. Or does that still take time off?	
2.85	R2	No, it's easy, you change it already. It takes more time. It's just clear, the more nodes, the more time. If a patient has two foci, then it's done quickly. If a patient somehow, unfortunately, has an underlying inflammatory disease and has 20 foci, then you still have to look at each one and then and then so well, I do that because I know, yes, the patient will come back and if I don't do that, then I'll do it next time. I don't know how I would have that now in my daily routine. Really, whether I wouldn't look over it briefly and say, oh well, it's a bit big now, it's a bit small, but on the whole it's right. It may be that I would do that somehow. If someone comes with 20 lung nodules and I look and say, hmm is that really a two, hmm is that really a one, but that I would just look over it and say, well, two is okay.	
2.86	JH	But have you ever done something like lung cancer screening before these AI tools? So would you have a comparison there, whether that changed the workflow or the time as well? Or were the systems there from the beginning?	
2.87	R2	In the past, there were also these measuring systems from [software, anonymized], but that was so cumbersome that you really had to sit down with everyone on an extra workstation and do that. We did that sometimes, but mostly not. I have to be honest.	PEUTE
2.88	JH	It sounds as if you now find the user-friendliness of the current [AI software, anonymized] tool much better.	PEUT
2.89	R2	Yes, yes, it is better. Actually, it's a good system that has, it has its weaknesses, somehow and you have to somehow still, that's why we try that as much as possible. So, a little bit there in-depth to make it as accurate as possible, so that the system really learns and you look, where are the sources of error? I think, for example, that emphysema assessment can also be significantly improved. Now it's so that you also often misjudge visually, as the AI is sometimes superior to you somehow. You must not forget that.	PEUTL

2.90	JH	Have you already had a few times, for example, oh no that's absolutely not that big, and have you looked at it more closely and then you were like, oh yes, that has completely correctly assessed? And at first, I thought I had misjudged it.	
2.91	R2	No, actually not. But actually, it is so, that I think a few times with these emphysema bubbles somehow, okay, if one had estimated that, so I would have estimated that maybe differently, so if I had seen that now in the clinical context, but if one measure that exactly, it can be that it is right. So, I thought that once, that one rather went back a little bit and said okay, it can be somehow.	PUL
2.92	JM	Yes, I think emphysema is also about the volume calculation, right?	
2.93	R2	There it is also about the volume calculation. But that's more with the stove the problem somehow.	
2.94	JM	What I still wondered when you just talked about the, the round foci also, is and where they said that every now and then you go through the images the see very large round foci that the system has not even recognized. If that really just shows where it thinks lung nodules are and doesn't show any other abnormalities, wouldn't it be helpful to now show all consolidations in general so that they can be safe? Oh, it saw that for example, but didn't identify it as a lung nodule, did it?	
2.95	R2	But that actually does that, actually, it does that. That is then really in such a way that that must have taken that somehow for something completely different. It's no longer lung parenchyma or something.	
2.96	JM	Oh, so this actually shows general abnormalities in the consolidations?	
2.97	R2	That is the problem. You see 20 lung nodules and have to go through them and say, that's nothing. That is somehow a cut, that for example. A relatively large number of older people have had tuberculosis at one time and there are these and the tuberculosis causes scarring of the lung tips. However, these scarrings are not smooth, but they are always a little bit with such small extensions and the system recognizes this very, very gladly as lung nodules	OQ PEUT

		and that is, I say, with every fourth patient you have to go through it. Now, it is the case that tumors also like to settle in this area, and therefore it is quite stupid if you are told that such a lung nodule is present and then say no, but that is nothing. You have to be sure somehow. And that's it. And then sometimes you have to fight with the system a bit.	
2.98	JM	Yes.	
2.99	JH	I think we should get on with our actual questions before we digress too much. We had another question about lung cancer screening in general, so is the process working well there, or are there still any issues?	
2.100	R2	I think it actually works well. There is always the problem of what to do with the sensitivity. How often do you want to worry the patient and call them? They may have something and have to come back. And you have to have a certain degree of not saying too often that it could be something bad and then call the patient in and, let's say, your own need for security, if you see something, couldn't it be something? Or when do I want to see the patient again? Is it enough if he comes back in a year? Does he have to come back in six months?	OQ
2.101	JH	Hmm.	
2.102	R2	But that otherwise, so actually it works well. And this, this, this assessment, we now have a fairly standardized protocol and I think it works quite well. In individual cases, we then discuss it with the clinical team.	
2.103	JH	Then we would move on to questions about the...	
2.104	JM	Actually, I have one more question, because our last interview partner also pointed out to us where AI could be used additionally well. For example, in suggestions on when a patient should come back again or simply in the risk classification of patients. Or? Would you agree with that? Or how do you see that?	
2.105	R2	Yes, it is, because that's basically how we do it, so we have this, this Lung-RADS classification. And actually, it's like this, from stage three they have to come back in six months and then there's	

		4A and 4B and 4A comes back in three months and um and 4B is then immediately clinically called in.	
2.106	JM	I mean now for example also like that also age or how much the patient has smoked has also to be considered in the decision. When should this patient come back or how likely is it really that he has developed lung cancer, that this could somehow be integrated into the previous diagnosis?	
2.107	R2	That would be nice. Yes. Yes.	
2.108	JH	That's probably the question of how well that then reconciles with the lung cancer screening framework.	
2.109	R2	Basically, you know the age of the patient, even if they're randomized. I see the age somehow and basically, I know man or woman. And I mean theoretically, the system knows, well it's just not stratified yet, how much did he really smoke, how long? I don't think that's showing up at the moment.	OQ
2.110	JM	Would you say that for the three functions that are now to be covered by this, i.e., coronary calcification in general, pulmonary nodule identification and segmentation, and also the determination of emphysema? Are you getting all the information that you need for the diagnosis overall from that system? Or would you say if it was still giving me that, that would be great?	
2.111	R2	No, actually I think that's pretty good. So that's kind of. Well, we have this Agatston score for calcium. It's relatively good. I find it a little bit vague because it's positive from, I would say it's positive from 15 or so. And actually, I have the values and then I say between 35 and 1000, so if they are positive, you could maybe segment a little bit finer, I think, but I don't know, because I'm not a cardiologist. I don't know that much about it. I don't usually do the calcium evaluation. I think emphysema is actually quite good. You could stratify it even more because the system actually knows whether it's an upper lobe or a lower lobe, and these are actually different clinical pictures, or they have different relevance for the patient. This could certainly be refined a bit more, but I would say that for a basic patient risk assessment, I think it's quite okay. So, it's not like that, these are three values that are important, that we also have.	OQ

2.112	JM	Yes. So, all the data that the system also throws out is also relevant to them or throws out the irrelevant data that they don't even use, where you say. Then that just makes it more unclear or confusing.	
2.113	R2	Nope, nope, it doesn't.	JR
2.114	JM	Also, let's say the level of detail, is actually all pretty good, or would you say somewhere would be, let's say, more general better, or somewhere is not precise enough yet, or it could indicate more, more precise anything?	
2.115	R2	I know that, as I said with this calcium, I don't usually do this CP cardiac evaluation, so I don't know that for sure. But this is such a subject that I have neglected a little bit. But whether you can still a little bit, a little bit more precise, which vessel or I know which vessel is affected, I can see all that, but whether you could still go a little bit finer there in the evaluation the risk profile. But basically, I think it's very detailed. As I said, the lung is almost sometimes a little bit too detailed. Whether you really have to classify every tiny focus, I don't know.	OQ JR
2.116	JH	Do you find that the output of the AI system is presented in an understandable way?	
2.117	R2	So, our output is very basal I say, that we deliver to the patient, somehow. That is, thank you very much for participating. You have no risk at all, minimal risk and in one year, the control is enough. You have a probably benign finding and should in six months, should be checked in six months. Then there is another probably good, so also will also be reassured, but you should present again in three months and then there is this please come, so that is. It is important to know that even these 4A findings, where we call them back in three months, often only have a real risk of disease malignancy of 2 to 4%. So that is not the case that they really have cancer.	
2.118	JM	And with regard to the output that the system now offers you in image analysis. Would you say that everything is understandable and that you can handle it well and that it can also be used well?	

2.119	R2	It's a tiny bit cumbersome, the measurement functions are cumbersome, I always have to click on those and there are no shortcuts. Of course, you could probably just set that up, somehow. That doesn't exist now because I, because they don't have that set up. I guess that wouldn't be a problem actually, but you just always have to, measurements are a bit cumbersome in the system, both diameter measurements and dense measurements. So also, and yeah, then there's little things like that. You scroll with a key that you don't scroll anything else with and so, but maybe that's also common in [country, anonymized]. I don't know that for sure.	PEUTE RU
2.120	JM	Yes. Do you know if it's just like this for you? Or have you ever talked about it with your colleagues?	
2.121	R2	I know that it bothered everybody a little bit in the beginning that you kind of, when you, when you scroll, that you take the other mouse button there, kind of. And it's, um, I'll say, not for everybody, I'm like that now, I just click on everything once and see what happens. But I'll say when people, just older people, maybe you see, what am I saying, older people. But then there are actually people who have even less experience with computers. For them, it's perhaps not quite so intuitive. Operable, I would say.	
2.122	JH	So, is there something in the tool...	
2.123	R2	So, there was a little induction that I barely remember and then that somehow didn't happen anymore.	PEUT
2.124	JH	But is there documentation that you can look up. Or is there any help in the tool?	
2.125	R2	Yes, yes, there is, there is. And you can also ask questions. If there's something we didn't know, there's an IT person who's responsible for us, you can just write to him and ask. Well, he's very nice.	RU
2.126	JM	And have you ever done that, does it work quickly and does everything also worked well?	
2.127	R2	Yes. Yes.	

2.128	JM	Perfect. Yes.	
2.129	R2	Mostly, more, I did it more when there were problems because again something didn't work, or something didn't move, or I couldn't send a patient. Or sometimes I also, they are sent from the CT into the system and then they arrive incomprehensible or somehow have a bug in there and that has always worked very well. There is a gentleman who takes care of it.	RU
2.130	JM	But in general, the system is very buggy, or do you have problems with it regularly?	
2.131	R2	Nah, nah, nah.	
2.132	JM	These are then exceptional cases where the system does not work?	
2.133	R2	Yes, for example, once a patient had a pacemaker that produced very large artifacts. It completely collapsed and then had to first, so this automatic evaluation did not work at all, because all these, because these evaluation artifacts, I think all completely overwhelmed the evaluation system.	SR
2.134	JM	Yes.	
2.135	R2	That was a problem. Then the other day one was somehow read out incorrectly, I don't know where that was. That was somehow, I think that in the CT the TA had set wrong. This was where the center of the image was, somehow. And there was then in the back quasi the back was cut off, somehow and that had to be sent again. But otherwise, nope. So, since this, this, this server connection, or this connection problem is fixed, it actually, that was what was worst at the beginning, that it had to be restarted again and again the system. And there was an IT person who was on call for a whole weekend, so to speak, and he booted up the system again and again so that I could continue working.	

2.136	JM	Let's go back to what you said before because I found it quite interesting. You said that it was a bit annoying that you had to scroll with the other mouse button. I was just thinking, that's not because the system is badly designed per se, but that's because you normally scroll with the other mouse button with another system that you use, isn't it?	
2.137	R2	Yes, exactly. The functions are simply assigned a bit differently. I mean, you get used to it after the third patient. It's not difficult, but it is, but the other key has other functions. So that is then somehow…	
2.138	JM	That's another change of habit, yes.	
2.139	JH	There should actually also be default assigned keys.	
2.140	R2	Exactly, but that's, that's ridiculous. But of course, it's like that at that moment, the first two times, or for example when you measure the nodules yourself. You always have to; you measure a nodule and then you have to click that function away again and again. And that's something, as stupid as that sounds, that you often forget, so it's really every time again and then, then, how is that. And then you scroll on with the same mouse button, I think, and when you forget that you suddenly draw some nodules in the system that you don't want to draw. That's a little bit, that happened to me a couple of times, so somehow, that's then and then you have to delete them again. That is of course now somehow, but I say that it is simply, if you consider, you would like to make 30 of this patient or so. Hello, but now I don't see anything.	PEUT
2.141	JM	I, I was just thinking, maybe once I turn off the video and turn it back on if then that is smooth again. Yes, it is perfect. Because it is annoying if you look all the time at this frozen image.	
2.142	JH	Then one more question. Um, what do you consider to be the biggest problems when using the system? You've already listed a few, but what bothers you the most about the system?	

2.143	R2	The time delay, well there's probably not that much you can do. I still find that a bit annoying. I, no it's more like, it's more like user-unfriendly things, like I said, like this, for example, so that actually bothers me. For one thing, it bothers me that I still have to measure so many lung nodules, but I think that is because it is still in its infancy. It's supposed to do that all by itself at some point and not be corrected. But me...	PEUT
2.144	JM	You would also say that the system is simply not good enough yet, wouldn't you?	
2.145	JH	Or it just doesn't do that yet?	
2.146	R2	No, I would say that it can, that it can't do that well enough yet somehow. That's the, that, so we always say so, or it recognizes too many lung nodules. You always have to remember that people come back in a year, and they have 20 lung nodules, some of which are tiny, which the system stores and then controls in the course of time, and of these tiny lung nodules, if a lung nodule is two millimeters in size, the patient only has to breathe a little differently. Suddenly it is three millimeters in size and has a doubling... and is indicated as a malignant long nodule. So, I can still see that we have now had the first three-month checks. That was a bit difficult in part, somehow, i.e., with the comparability. And you really have to be careful that you don't take lung nodules that are too small, because then it's really very sensitive with this, where maybe just a little bit of lung parenchyma continues to stick together or something, but that's really a very sensitive increase in size, in the increase in size, that's really very sensitive and that's somehow, that's of course also meant to see better than the human eye, but that's then at that moment, you have to be very careful that you don't exaggerate with the very small lung nodules, that's difficult, I think.	OQ JR PU
2.147	JM	Yeah okay. How would you say does using this system change your final diagnosis in general or how does that change quasi, or how if you were to think about if you were to do it all without the AI system? Now how much would quasi the final diagnosis result, how does that look different than quasi that if you were doing it without AI?	
2.148	R2	So it sees more lung nodules than I do because I just click away a lot of lung nodules in my subconscious somehow and that's so, I scroll through there and if there is such a fissure, such a small	OQ PU

		thickening has or so, that, I look no minute longer on it somehow and it gets me all out and makes me but again aware and quite that happens rarely, but sometimes it actually sees small lung nodules, but perhaps have potential that I overlook. Just so, ones that are just so between two and three, I say so between lung nodule but not so important and could be what somehow and there I think it is better than me. I also see larger lung nodules and it sees more unimportant lung nodules than I do, that has to be said.	
2.149	JM	Okay, would you say that this increases, well, the number of false positives, or how do you think? So maybe it's hard to say so far, but how do you think that affects the overall output?	
2.150	R2	Without control it would certainly increase the number of false positives, I'm quite sure. But since we work through everything again and then somehow also partly, if you change such a lung nodule, then you always give it to someone else to look at it, actually, that you are also so sure that it really is nothing somehow, that you actually, so I think there would be more false positives if we would not look at it. It would not, not significantly, so I don't even know if I would have missed a single really bad finding so far.	PU
2.151	JM	Okay, yeah.	
2.152	R2	And the other thing is and there you have to wait, the control in half a year or in the year, whether the small herds that have been classified as a little bit suspicious…	PU
2.153	JM	Are there any patients that have come in for a second time where they have an assessment of whether overall the performance is generally very good, along with the system, or can you not say anything about that yet.	
2.154	R2	So now we're just getting started, it started in August, now we're just getting the six-monthly histories in, they're just coming in and unfortunately I haven't seen that much, because we had divided it up a bit, that, of course, these are significantly less than the total mass of patients and a colleague who doesn't have quite as much time, she prefers to do these controls, so to speak, because it's also, she has a lot of experience. That is [name, anonymized], with whom you will also speak, she has a lot of experience and especially in these somewhat critical cases, that is always, I think, very good when she does it. But of course, the rest of us learn	

		from that as well. It's not that we share it completely, that we give it away completely. But it is, it is then, for example, that I sit down for a day and only do that and she actually, more in her workflow, then in the afternoon, if there is still time, looks at the difficult cases.	
2.155	JM	Yes, yes, and um, but would you say that overall time is saved with the AI system? Or how much faster do you think you work together with the AI system than without the AI system?	
2.156	R2	But in any case, if I would only do that now, and so at least for the pulmonary nodules, if I would only do that now, click on the thing and, and that gives me the nodule and I then really only check the nodule. It also depends a little bit, as I said, if the patient had 20 foci, of which I would have clicked off maybe ten or not at all, not at all, it would be more work, of course, but I think, so on the whole it makes, it makes my work easier, because then I only look through what it shows me. And of course, since I go over it once over the lungs anyway, things like that that it doesn't see, I'd probably notice. So that.	OQ PU
2.157	JH	Is there actually a prescribed process, like for you radiologists, that you have to use the AI? Is it mandated that they use it at all? Or do they use it because it makes their job easier.	
2.158	R2	I do that because it saves me the work, I don't have to do that at all, so in the normal workflow, so apart from that, but also the mammographers, no we don't have to do it at all, so for example for the emphysema evaluation, is running right now in [city, anonymized], [name, anonymized] knows that better for sure, because she does that more. Then the pneumologists always want a machine evaluation, so to speak. So, because it quantifies it better, but I believe that otherwise, so if I get a patient like that and the question is, a focus was seen and he coughs too much or something and look at it to see if he has something, then I don't use AI.	PU JR SV
2.159	JM	Yes. Also, related to other radiologists do they also know whether they also all use the AI similar to you and also like to use it? Or do they see it more critically in part, or?	

2.160	R2	So, in [city, anonymized] I don't think anyone uses it, to be honest. I know it always depends on where the device is and where it's available. So that's how it is and that's often also in consultation with the referrer. There are referrers who attach importance to that. I think in [hospital, anonymized] we have a syngo.via where some things are actually looked at on syngo.via. But then, as I said, it is always an extra work step. Then you sit down there once, then you call up the patient, then it is usually read out somehow and then you have to look through it again. So that's actually more work and you usually don't just do that.	PU JR SV PT
2.161	JH	And that's probably why others don't like to use it because it's more time-consuming.	
2.162	R2	Exactly.	PT
2.163	JM	But now with lung cancer screening, that's much better integrated, isn't it? And where you work everything away, most probably use that already, right?	
2.164	R2	But this lung cancer screening is only about that at the moment, which is basically a difficult thing to say. But if I now so, if now people would come, just for that, then I think I would do that, because that also really facilitates the work. And in 80% of the cases, I would say so.	SV PU
2.165	JH	Do you see any potential for making this even better, this AI Tool, where it can be integrated even better into the work process?	
2.166	R2	You would just have to, so if you have your normal interface, basically it would have to run in the background of the interface. So that's then you call the patient and then you kind of look at something else and then the moment you need the AI. So that is always difficult. I don't just look at lungs, it's somehow. It would have to be able to call basically and on the same user interface, and quickly with one click and now somehow, I click here and then the AI opens, and then I do it quickly like this. And then I continue in my normal program, on the same screen, on the same user interface, without somehow having to do anything else and switch back and forth or somehow click something away.	PU

2.167	JM	Again, because I'm not sure if I fully understood your answer to the previous question. So now, when the other radiologists who evaluate the CT-Scans as part of the lung cancer screening, they all use [AI software, anonymized] now, right? And they look at using this tool?	
2.168	R2	But because with this system, that is the study. So, there is no other, no other system that is integrated there yes. We always take that. So that is more or less the default.	SV
2.169	JM	But very briefly. You also said because this is directly integrated with it, that is, you actually have to use it as part of the study?	
2.170	R2	Yes, there is no other way. The study is that in the, in the applications and in everything is prescribed this software.	SV
2.171	JM	Yes, but even if that's introduced at some point and AI is not mandatory, would you say that it would still be used because it saves time somehow? Or do you think rather not?	
2.172	R2	So, I think for this screening program it's very important that you use something standardized. Somehow, if, if then you should use the same program for every patient who comes. And to that extent, I think that's the case. And if they come for screening, then you should also do the screening. So that is then actually somehow, then it's just not enough if the radiologist looks over there just once briefly, also always remembering that yes not always the same radiologist is available the next time when the patient comes and that is yes, mostly that is supposed to be yes, such a lung screening is only meaningful if that is repeated every few years somehow.	JR
2.173	JH	If you mean that this [AI software, anonymized] program is mandated by the study, does mandated mean that every practice that participates is going to purchase that program? But it is not mandatory for the radiologists to use it. But if they use it, they should use it?	

2.174	R2	It's like, for our study that we're doing right now, it's mandatory for the lung cancer screening itself, I don't think it really matters what they use. It should just be then, it should just be then that when a patient comes in for screening, they should use the AI, and the next time they come in, the same AI should be used.	SV
2.175	JM	Yes.	
2.176	R2	That's important, that and so that it becomes a bit, let's say so, the patient is now somehow in his mid-50s and then comes every few years, he now becomes happy 85 years old or so somehow and then is just then at some point that's no longer the same radiologist sitting there no longer, because I then also already put my feet up.	
2.177	JM	Do you generally look forward to the use of AI in the context of lung cancer screening? So, do you think it's nice to work with AI, like, because it makes it more efficient? Or are you saying like, everything used to be better when everything was done manually? That's the real radiologist work.	
2.178	R2	No, not at all, I always find that kind of thing exciting. I'm not quite, well, not quite convinced of it. Of course, the best thing would be to let the people come. Then the program runs over it and spits out something, a few numbers at the end and you can rely on it one hundred percent. Unfortunately, that's not the way it is, and I don't know if we'll ever get there, but if that's the goal, it makes a lot of sense, because it saves time and manpower and all kinds of things. And then I think that's great. I mean, I'm rationing myself away a little bit right now, but that's because then you can spend time on other things. I mean, I'll just say that counting lung nodules is also a job that needs to be done a bit.	SV
2.179	JH	So, you're not afraid that this will replace your job in general, but just the tedious work rather, and then you can also focus on more meaningful things?	
2.180	R2	It would be nice if not, if you keep the things where you have a little bit of fun or where you have to think a little bit or something, if you keep those and things where you count points, then the AI takes over.	SV

2.181	JM	As long as that is then straight in such a support function, that's great, right?	
2.182	R2	Yes, so I find.	SV
2.183	JM	But would you then be afraid that at some point it would actually be able to do that? That it will make the diagnosis itself? Does that scare you?	
2.184	R2	Oh, I won't be working until then.	
2.185	JM	How are your colleagues feeling about this topic?	
2.186	R2	Yes, exactly. In fact, that's how it is now. I also have children who want to study, they also want to study at some point, and my daughter probably also wants to study medicine. And then, of course, you have to think about whether I can recommend that she become a radiologist? Or do I assume that in 30 years this will no longer exist as a separate specialty or whether it will not exist on a very, very different scale, whether there will be much fewer of them and they will have a much more interdisciplinary field of activity?	
2.187	JH	You probably just have to make sure that you also have some technical subjects with which you can also understand and operate this AI. This is probably an important competence in this area.	
2.188	R2	I think that's good in principle because no one likes to evaluate scores or count lung nodules or anything, or measure things. To be quite honest, it's all a bit of a dork's job, to put it kindly, somehow, that's, the computer can do that better and it should. So that, exactly, and otherwise I don't know exactly where it's going somehow. That remains exciting.	
2.189	JH	Yes. Is it actually in the program that if the AI shows anything you can track that? So, is there any explanation? Or when you see that it's displayed, then you're also like ah yeah okay, that makes sense that that was somehow flagged as a concern?	
2.190	JM	For example, in the calculation of the coronary calcium score or something.	

2.191	R2	Yes. Nope, actually, it has, so that makes sense. That's somehow, the problem, the problem at the moment is actually that it always measures all the stents, where I don't know yet exactly why that's not implemented, that that, so it takes some out, some not and I actually correct the whole time only the stents out or delete there the whole time only the stents out of the program. But that', of course, that's, of course, that's super useful and that you then, when you go for such a CT right away, that then they say, you have a risk of getting a heart attack, are you already taking medication? If not, you should, that's maybe a sensible thing to do.	OQ L RU
2.192	JM	But for example, even if it says, so the lung nodule is so and so big, then you can sort of see the outlines that it's identified there exactly, right?	
2.193	R2	Yes. Yes. Yes. I can look at it in three levels. It's actually, I can probably look at it in three dimensions if I click on three buttons. I just don't really need that, actually, you work… The program is, I think, much broader than I really use it somehow. So, it's not that I play with all, with all features that it shows me above. I don't do every measurement and so or so, of course, I use the things that make sense.	L PEUT
2.194	JH	Does that also bring a bit of fun to play around with the tool?	SV PEUT
2.195	R2	Yeah, yeah, sure.	SV PEUT
2.196	R2	What I'm still wondering is, for example, if you don't agree to a measurement now or say, no, the pulmonary focus is not that big. What exactly would you do then? Do you then go into the system and distort the outlines in the system, or do you measure it again yourself?	
2.197	R2	Exactly, I measure it, I actually measure it again myself, it's like this, I first look at it and think to myself, what's the reason that the pulmonary focus is incorrectly detected somehow, because for me it's clearly, clearly not a high-risk pulmonary focus, and then I look like this, why does it assess it incorrectly? Is it really the size? Maybe I'm wrong and then I usually measure it again and if it is then and if it is then, so if I then see, okay it's really above the spot, then I look again, are there any vessels in there somehow. Why do I think that it's not, sometimes I'm also wrong, and then I	L

		have to admit that, and sometimes, but completely it just measures any neighboring structures. And I have to look at it in three layers, then I see well, it goes along there, it has an extension to the fissure or it measures this pulmonary vessel that runs directly past there or so and you just try to cut it out, that's...	
2.198	JM	But then use, so then you do not use the system for measurement, but then rather manually, you look…	
2.199	R2	There is a nodule measurement. You can make @Nodule and then you can measure it and you can also change it. I can always change the shape and change the size.	
2.200	JM	So, you're already using the system for that?	
2.201	R2	Yes, I use the system for that.	
2.202	JH	So manually then, so to speak?	
2.203	R2	Exactly. I can go all the way out and remeasure it. I do that sometimes when I get all desperate. But I can also take the measurement it gives me and then in moderation, I can change that. It gives me plus/minus and I can do that in one direction twice in direction three times or so, I can then always play around like that, and then he gives me different shapes, for example, because he says well then that lies [inaudible] and then he gives me a star or an ellipse or something and then I say comes there.	JR
2.204	JM	Yes.	
2.205	R2	Something like that.	
2.206	JH	Does the program then actually only help with recognition, or does it also generate a report at the end, something in writing, so to speak, which you would otherwise have to do manually yourself?	

2.207	R2	No, exactly, there is a report, which is kept very short. Then there is really only, Dear Mr. So-and-So, we have found out this and that during the evaluation and then randomized is added and then a picture is always printed out and a sentence about the emphysema and then it is said, a significant emphysema is present from a threshold value of 7% and then the emphysema value is spat out or it is written that there is no emphysema and then it gets a calcium value and then it is said that there is a significant risk of heart attack from a value of so and so. Or there is no emphysema, and this is randomized and means that the negative findings are also informed or not.	
2.208	JH	That's quite a bit of information. If you didn't have the tool, would you have to write it all yourself? Well, it sounds like that would save quite a bit of time.	
2.209	R2	So actually it is so that the software, there are such things, there are such basic diseases, where the software does not recognize the emphysema correctly and if I find and I see that and I find, the emphysema is clearly incorrectly assessed, then I write and he really has a pronounced emphysema, then I write it manually underneath and then I write that underneath that for this and that reasons which are probably not recognized by the software it is a relevant emphysema. I then write that down like that. So, I can add additional findings at the end, I can do that. Or if I see that the patient has a huge thyroid gland, for example. Then that also comes in at the bottom, the system does that, writes something like that, of course, the system doesn't recognize that and doesn't write it in itself. And that is then again, a little bit of the clinician work on it, but so about the lungs, which the lung focus information is always in there in the findings. It's always in there, the rest is randomized.	
2.210	JM	I think you meant that takes about ten minutes per patient, right?	
2.211	R2	Yes, I can do it in five minutes if there really, if he really has nothing and then the next patient comes who has 20 foci, which takes fifteen minutes. So, on average it takes ten.	
2.212	JM	That's what I wanted to address. What are all these ten minutes or five minutes, or 20 minutes used for? So just looking through all the images again, checking all the pulmonary nodules? What exactly is the most time-consuming thing there?	

2.213	R2	Well, I look, I look, once the emphysema, so that's actually pretty quick, look at the segments, if there's nothing, it really only takes a few seconds. Then I look at the emphysema, and it always depends, if there's not much, it's pretty quick and then I wait until it changes. Then I look at the calcium score and that's a bit time-consuming because I always look at all the organs. That takes maybe three or four minutes, I would say, and then I jump to these...	
2.214	JM	What exactly do you have to look at, sorry?	
2.215	R2	The liver, whether there is any water in the lungs. Then I look at the upper abdominal organs that are shown whether they have any pathologies. Then I look at the thyroid gland and take a look at the bones.	
2.216	JM	That means you go through all the images for irregularities, right?	
2.217	R2	Yes, exactly. So, the vertebral body fractures and so on, the system just doesn't do that. But of course, that's information that I have, and I can't just overlook it somehow - that's not possible.	OQ
2.218	JM	So, you do that, they do that, but the system doesn't do that yet.	
2.219	R2	The system does not do that yet. The system only measures the calcium score in this image.	OQ
2.220	JM	That is, that would actually be a function that you would still like to have, you could say that, right?	
2.221	R2	To be honest, I don't know if the system can do that. It's a very complex thing. It has to look at several organ systems. So, I actually look at the bones, the soft tissues, that is, and there is information in there, for example, I look to see if the patient has a biliary obstruction if I see that, or I look to see if the patient, if the stomach, if he has a sliding stomach. That is also something that is relatively common, which people often don't know. This you really only see in such an examination and that would be stupid not to write it down. I look at such things. It takes three or four minutes, then I switch back and it takes one or two minutes and	OQ

		then I take a look at the lungs and if they're okay, then they're super fit.	
2.222	JM	By switching, what else do you mean by switching?	
2.223	R2	I always have to switch from the soft kernel to the hard kernel to view the lungs. I always have to look at the lungs in such a high-resolution mode and this switching takes time on the system for some reason.	
2.224	JM	And that time is really just waiting, right? Where you can't do anything else?	
2.225	R2	Exactly and then, if everything is good there, I need at most another two minutes, I'll say, and if that's the case, if there's something there and I have to change something somehow, then it depends on that. But if I just look up once, look down once, click on the other side and the CAD and it doesn't find anything, then I'm also through and then I still have to look at, isn't there something else and I still have to confirm that, the readings are. It always tells me emphysema and calcium score, I have to EXIST again or not. I always have to confirm that again. Then it sees CONFIRMED CONFIRMED, otherwise, it grumbles, and then I can [inaudible] the patient, then I say please make a report. And actually, in our center it's like this, this report, we don't print it directly, but I save it first between and they are then printed in a big batch. That is then again, I then make such a download simply into a folder on the computer and collect those there under the date, mostly under which I that, under which they have been studied. And then it's all printed out by a study nurse, afterward. That's not our job anymore.	
2.226	JM	You meant, now in a complex case, it takes 20 minutes. That just takes because you, there the time is more on the lungs…	
2.227	R2	Because I have to look at all these lung foci, because I might have to take out all the stents in the calcium score, because I don't have the lung segments right, because it doesn't recognize any lobes and I have to draw them in manually, which is also a lot of fun, somehow. So, I probably wouldn't have to do that, but I mean, you want to apply it the way it's applied so that everything is correct.	PEUT

2.228	JH	How often does it happen that you feel like things are taking longer than they should because of bad decisions made by AI?	
2.229	R2	Maybe for every tenth patient. So, it always depends a little bit. So, the lung nodules, that's easy, the AI is supposed to recognize them, so to speak. That's just more work for me, but this lung segmentation, if that goes wrong, that's relatively much work. But with every tenth patient, and with the bypasses or with the stents, which are then incorrectly detected, that also happens with every patient, that's perhaps so, yes, I guess on average with every tenth patient I have more work.	PU PT
2.230	JM	What I'm also just wondering again. You meant, you had a preparation course at that time before the whole study, but it was generally about what studies were already there.	
2.231	R2	Yes, what kind of studies there were. Also, so a little bit the pitfalls, also so a little bit on what do you have to look out for, for example, if you see a certain type of lung nodule, what is there a sign that it could be malignant? Which, so there is also this classification 4X. That means that you have a focus, which um, which in and of itself also worries you because of the size, but which above all has clear malignancy criteria, which are then explained again. So, these were very different lectures. One of them was about the fact that patients should not be excessively worried. This is then also very often backed up with data, somehow, where I then sometimes switch off a bit.	
2.232	JM	Were you also prepared in some way to use the system or were you just thrown into it, then one day?	
2.233	R2	No, no, the use of the system. A nice lady came along somehow. So there then came also, from [country, anonymized] were some people there. So that was a kind of small welcome meeting, because then the system was introduced, you got a password and then the basic features were explained, I guess. Of course, that happened once before and then never again and then this woman [name, anonymized] who trained us, she came a few more times.	RU PEUT
2.234	JM	On site then to you?	

2.235	R2	To any, if there were new features, for example, there were a few new features added at the beginning, which she explained again.	RU
2.236	JH	So, when there are updates, you will be informed about them as well?	
2.237	R2	Then we were informed again. There is also such an operating manual, but I must honestly say, I had that at the beginning once, we have all sat down ourselves. We threw everything we remembered into it and tried it out a bit, and basically, yes, it's important that something like this can be operated intuitively, I'd say, you can't do it every time.	PEUT
2.238	JM	But you would say, in general, it's intuitive, right?	
2.239	R2	Yes, yes.	PEUT
2.240	JM	If a new radiologist starts who has not had this introduction suddenly has to use it, would he have problems getting into it or would he learn it within a week?	
2.241	R2	No, you learn that within a week, faster I think.	PEUT
2.242	JM	Okay. In general, before lung cancer screening is now used in [city, anonymized] on a large scale then nationally, Germany-wide. So if lung cancer screening is now implemented throughout Germany and it is now really serious in [city, anonymized] and is no longer within the framework of a study and then probably more patients will come. Do you think there will be any problems again or should this actually just work well without any problems?	
2.243	R2	It's always a question of how much money is invested in personnel. If this is somehow well organized, it's not a problem at all. I don't think we're going to have any problems. That is then, oh yes what I haven't mentioned yet, that should I do that is that we have patients who have a stupid finding, and they have to come back again. We have to call them. And that, of course, is then really an additional expenditure of time, which also requires a bit of tact. You cannot simply call and say, something was strange, please come back in three months, but they need a bit of explanation. They are somehow worried and that also does [name,	

		anonymized] very much. So, when you talk to her, I would say something like that in addition. All the rest, which is a bit, I would say, a bit schematic, is no problem at all if you have the manpower. So, you sit down, do it and then you can also, well, depending on how many people you hire, you can also do it well. You just need, as I said, you always have to be clear, maybe we'll get up to six or seven. But I say, seven CTs are one hour, meaning one radiologist is busy for one hour. At the moment, anyway.	
2.244	JH	But, for example, these phone calls that you're talking about are something so delicate and sensitive that they should be made by a human being. It shouldn't, somehow, be an automated e-mail that goes out.	
2.245	R2	Yes, a person has to do that. The question is whether to leave it to the family doctor. That would be a possibility, of course. At the moment we are doing this, of course, this is also a study, the family doctors have nothing to do with it, etc. If this is something that is implemented in the screening, then perhaps in the insurance system or so that one then says, well the family doctor gets the information from the computer perhaps and then calls the patient again, somehow but the question is, exactly, there you need a very good patient guidance. It's not so easy when you've never seen the people before, and we don't see them at that moment. They come to a general practitioner and get their initial consultation, so to speak, and are then examined by an MTA in the truck, and the radiologist only sees the images but calls them afterward.	
2.246	JM	Jan, do you have something else?	
2.247	JH	I think we're done with our interview guide itself. Are you still missing any topics that we might not have touched on or that you might want to elaborate on a bit more?	
2.248	R2	The financing of such a system, but there I'm really not the right person.	
2.249	JM	Otherwise, we would stop the interview officially here and also the recording.	
2.250	JH	Do you have any other questions for us?	

2.251	R2	No.	

German

Row	Speaker	Transcribed text	Codes
2.1	JM	Können Sie uns in zwei bis drei Sätzen sagen, wer Sie sind und in welchen Zusammenhang Sie bis jetzt mit Lungencancer Screening zu tun hatten?	
2.2	R2	Also ich bin Radiologe. Ich bin schon lange im Job, Fachärztin für Radiologie, habe auch relativ viel Erfahrung mit Lungen Bildgebung oder mit Thorax Bildgebung, habe in verschiedenen Zentren gearbeitet, erst in [city, anonymized] und [city, anonymized] und jetzt hier in [city, anonymized]. Ja und seit August letzten Jahres betreue ich eine Studie mit, bei der wir mit einer AI, die ein Lung Screening durchführen, wo wir Raucher auf das Vorhandensein von Lungenkrebs untersuchen, einfach alle Leute, die einmal geraucht haben oder die noch rauchen im Alter von 55 bis ich glaube 79 Jahren, dürfen sich dann vorstellen und kommen vorbei. Und das wird ausgewertet in dieser Studie und daran nehme ich teil und dafür musste ich ein paar Zertifizierungskurse noch durchführen, von der ST, von der Thorax Society, das habe ich natürlich gemacht irgendwie, dass man dann so eine Art, also die einem so die Pitfalls,des Lungencancer Screenings, worauf man achten muss, das war einfach so ein Kurs mit mehreren Vorträgen von sehr renommierten Leuten und das war auch sehr interessant.	
2.3	JH	War das für den Prozess des Lungenkrebs Screenings, also für die Bildauswertung oder auch schon mit dem Umgang mit der AI? Gab es dazu auch Kurse?	
2.4	R2	Ja, doch, doch. Der, mit dem das ist, war so ein bisschen integriert. Das wurde so ein bisschen vorausgesetzt, dass das, worüber man sprach, dass das mit AI durchgeführt wurde. Das war jetzt nicht, das war kein spezieller Kurs über die AI, sondern das war einfach so, der eher so, der Arbeitskurs im täglichen Umgang mit dem Patienten, mit der AI, mit dem Wissen, dass man aus anderen größeren Studien hatte, von so Screenings.	
2.5	JM	Das lief im Vorfeld bevor das Ganze eingeführt wurde auch jetzt in [city, anonymized], bevor die Studie angefangen hat?	

2.6	R2	Ja, genau.	
2.7	JM	Okay. Und noch mal, damit ich das richtig verstanden habe. Dabei ging es allgemein um die Zusammenarbeit mit AI oder wie?	
2.8	R2	Nein, dabei ging es eigentlich um die Epidemiologie hinter diesen Studien mehr. Also das wurde dann gesagt, welche Studien wurden schon durchgeführt, die NELSON Studie, diese Studie, die große europäische Studie. Was war da die Ergebnisse? Was haben wir dann in dieser Studie? Worauf sollte man achten? Was ist noch zusätzlich zu untersuchen? So.	
2.9	JM	So, dann kommen wir mal zur ersten direkt etwas schwierigeren Frage. Wenn Sie sich jetzt überlegen müssten, wie Lungencancer Screening in Deutschland eingeführt wird, wo können Sie sich vorstellen? Oder was denken Sie, wo würde es eingeführt werden? Wo wird das angeboten und für wen fällt die zusätzliche Arbeit an? Würden Sie zum Beispiel sagen, das machen bestimmte Praxen, das machen alle Krankenhäuser, das machen Unikliniken? Oder wie können Sie sich vorstellen, wird das umgesetzt?	
2.10	R2	Das könnte ich mir ein bisschen wie das Mammographie Screening vorstellen. Es müssen ganz bestimmte Centers sein, es bräuchte eine Zertifizierung und dann würden die Leute extra dafür dahin kommen. Ich glaube, dass das dann so normalen Alltagsablauf, also ich weiß nicht, es könnte einfach komplett parallel laufen neben dem Arbeitsalltag der Röntgen Praxen sozusagen. Also wie es jetzt ja im Grunde auch so ein bisschen ist. Ich finde das eigentlich nicht so schlecht, dass sich Leute gezielt nur dafür hinsetzen und die Leute gezielt nur dafür kommen und dass das als eigenes Programm läuft, so als Untersuchungen, die neben dem normalen klinischen Arbeitsaufwand auch der Fall sind, es sind ja eigentlich keine Patienten in dem Sinne. Es sind ja gesunde Leute. Theoretisch gesunde Leute.	PT
2.11	JH	Das hört sich so an als ist das ja schon großer Mehraufwand wäre, denken Sie denn, die Kapazitäten würden aktuell reichen? Oder müsste man noch extra weitere Radiologen einstellen, die halt explizit dafür sozusagen beauftragt werden?	
2.12	R2	Ein bisschen schwierig zu sagen, also theoretisch, wenn man das in größerem Maßstab noch machen wird oder immer machen möchte. Ich meine im Grunde sind wir ja Radiologen, die extra dafür abgestellt sind. Es wird ein Teil meiner Arbeitszeit oder wird dafür vergütet irgendwie. Ich arbeite derzeit auch nur für die Studie. Ich mache das nicht irgendwie jetzt gucke ich noch mal das Knie an und dann jetzt gleich noch mal die Lunge, sondern	

		ich setze mich dann wirklich gezielt hin und mache ein paar Stunden nur das irgendwie, also das.	
2.13	JM	Ja, aber ich meine, du machst das dann vielleicht immer einen Tag in der Woche oder zwei Tage die Woche, oder?	
2.14	R2	Es kommt immer drauf an, der Truck ist ja nur alle sechs Wochen bei uns und dann ist er zwei Wochen da und wir sind drei Leute, die sich das teilen. Vier, eigentlich vier, drei, vier Leute und ähm, ja und jeder ist mal dran, sozusagen. Meistens bin ich da für zwei volle Tage eingeteilt. Das muss natürlich jetzt, was weiß ich, wenn ich Donnerstag, Freitag, dann kann ich Samstag und Sonntag da auch noch dran arbeiten, so ist das nicht. Das läuft ja nicht weg, aber es ist tatsächlich so, dass ich in der Zeit eigentlich nur das mache und das ist auch sinnvoll, weil wenn ich einmal die Oberfläche aufrufe und mir diese Arbeitsumgebung schaffe und so, dann ist es eigentlich sinnvoll, da erst mal bei zu bleiben und mit gewissen Schwung an Patienten wegzumachen.	
2.15	JM	Weil man sich da so ein bisschen einarbeitet oder wie?	
2.16	R2	Ja, das auch. Und vorallem man muss es erst mal öffnen mit 1000 Passwörtern, wie das immer so ist und dann, also das stelle ich mir so, wenn ich das also es wäre jedenfalls ein gewisser Aufwand, da hin und her zu switchen, wenn ich jetzt vorher kommen drei Leute mit dem Knie, da mache ich jetzt noch mal zwei Lungen aus einer ganz anderen Oberfläche und dann gehe ich noch mal zurück zum Knie.	PEUT PT
2.17	JM	Ja. Also, um das jetzt mal auf unser Thema, was ja auch um den Arbeitsablauf vom Radiologen geht, noch mal genauer irgendwie einzuordnen. Also du, du sagst es ist, sage ich mal viel Arbeit, sage ich mal jetzt da mit AI, die die Lunge zu diagnostizieren, wenn du kurz davor was ganz anderes gemacht hast, weil du diese ganzen Programme erst mal aufrufen musst. Und auch deshalb ergibt es Sinn, ganz viel davon hinterweg zu machen, damit du halt nicht immer diesen Switch hast, da zwischen den Systemen.	
2.18	R2	Genau. Genau. Also eigentlich ist das, was das ist. Das sind natürlich in Wirklichkeit nur so drei Minuten oder so, aber das ist natürlich, das läppert sich über den Tag und ich mach ja, ich mache ja fünf, 5 bis 6 pro Stunde schaffe ich von denen. Das heißt also so im Augenblick, das heißt, wenn ich. Genau. Also wenn ich das dann immer so oft hin und her switche, dann verliert man halt doch so 30 Minuten oder so, das ist jetzt nicht ganz so wenig.	PT
2.19	JM	Aber insgesamt so machst du das dann ja immer, wenn es wieder anfällt für zwei Tage weg. Aber das ist jetzt auch nicht super viel	

		insgesamt mehr Arbeit gerade für die Lungenklinik und du würdest sagen das allgemein, auch wenn es insgesamt implementiert wird in Deutschland, dass ihr das auch so halbwegs hinbekommen würde weiterhin, oder?	
2.20	R2	Ihr seid jetzt gerade Frozen. Ich hoffe mein Internet reicht hier.	
2.21	JM	Das liegt glaube ich leider bei uns.	
2.22	R2	Aber soweit ich euch noch höre, ist ja alles gut.	
2.23	JM	Genau. Ja.	
2.24	R2	Was wollte ich sagen? Was soll jetzt nochmal die Frage? Also doch, das ist ja eine ganz andere Arbeit eigentlich. Also in die [institution, anonymized] kommen die Leute ja mit einem klinischen Problem zu mir. Da habe ich ja einen ganz anderen Blick auf, diesen Screening Blick.	
2.25	JM	Ja	
2.26	R2	Das ist schon so, das ist Zeit, die ich nur für dieses Programm verwende, es wäre jetzt so, wenn jetzt ein Patient käme und hätte was ganz anderes und würde dann trotzdem noch dieses Lungencancer Screening haben, das ich dann so durchlaufen lassen würde, sage ich mal, ich hätte das in einem Programm drin. Ich öffne meine Oberfläche der hat ein Knie gekriegt und zum Lungen Screening kann man auch noch suchen. Vielleicht kann man sich das so vorstellen, dass ich dann erst das Knie angucke, so und dann switche ich zum Lungenscreening und dann gehe ich da einmal durch.	
2.27	JM	Ja.	
2.28	R2	Dann würde das mich wahrscheinlich zehn Minuten extra Zeit kosten.	PT
2.29	JM	Okay. Ja. Ja.	
2.30	R2	Das würde schon helfen, aber ist natürlich Zeit. Also es dauert wie gesagt, ich schaffe ungefähr fünf bis sechs pro Stunde. Das ist, also zehn Minuten pro Patient ist realistisch.	
2.31	JM	Ja, dann sind natürlich zehn extra Minuten noch Mal kritisch.	

2.32	R2	Aber das ist dann extra Zeit. Also das ist nicht so, dass das, nun ist es auch so, dass wir dann eine ganze Menge anderer, also wir gucken uns auch immer den Calcium Score an in dem Programm und wir gucken uns auch immer das Emphysem. Wenn ich jetzt rein auf Tumor gucken würde, würde es mich vielleicht nur, und die KI mir hilft, dann würde es mich vielleicht nur fünf Minuten kosten.	
2.33	JM	Ja.	
2.34	R2	Also, denke ich.	
2.35	JH	Wird das auch Teil, oder denken Sie, es wird auch Teil des offiziellen Lungenkrebs Screeningsprogramms sein, das man dann auch auf andere Sachen unabhängig vom Lungenkrebs guckt, weil man eh schon im Prozess dabei ist?	
2.36	R2	Das ist eine gute Frage. Wäre natürlich sinnvoll, weil alle Raucher haben auch diese Tendenz den Kalk an den Herzkranzgefäße irgendwie. Und wenn man das sowieso untersucht, dann kann man da auch leicht draufschauen. Das finde ich eine sinnvolle Sache. Und Emphysem haben sie sowieso alle. Also ich meine, das zu quantifizieren... Wir versuchen gerade in dieser Studie herauszufinden, wie sehr die Quantifizierung, also wenn man Ihnen einmal sagt, dass sie ein Lungenemphysem haben, ob das Ihr Verhalten ändert, das ist ja Teil der Studie, glaube ich irgendwie, dass einige, das wird ja randomisiert, einige kriegen es mitgeteilt, einige nicht.	
2.37	JH	Ja, meinen Sie das Rauchverhalten oder welches Verhalten?	
2.38	R2	Und ob sich das Rauchverhalten ändert. Genau, ob die Leute denen gesagt werden ui, ui, ui, da sind aber schon 17 % nicht mehr da, ob die dann eher sagen dann erst recht oder ob die dann sagen jetzt nicht mehr.	
2.39	JM	Können Sie dann noch mal genauer den gesamten Prozess der Untersuchung eines Patienten im Rahmen des Lungencancer Screening erläutern? Also Sie haben gesagt, es gibt drei verschiedene Funktionen und manchmal wird es dem Patienten gesagt, manchmal wird es dem Patienten nicht gesagt, können Sie da mal genau den Gesamtprozess irgendwie versuchen zu erläutern?	

2.40	R2	Also es ist komplexer, die kommen auch zu einer klinischen Untersuchung vorher und kriegen dann erst ihr CT. Da werden sie befragt und es gibt natürlich Einschluss- und Ausschlusskriterien auch. Die dürfen in den letzten zehn Jahren, glaube ich, keine Tumorerkrankungen gehabt haben. Sie müssen selbst nicht vom Tumor wissen und sonst werden sie zu Sachen gefragt. Also immer aufgeführt wird, ob sie Statine benutzen. Das ist aber glaube ich, im Rahmen dieses Calciumscores dann, ob sie das ist ein Medikament, das die Blutwerte senkt irgendwie. Und dann wird gefragt, ob sie rauchen natürlich oder nicht oder ob sie geraucht haben. Und dann, oder ich glaube noch, ob sie aktiv rauchen, ist auch die Frage. Und dann wird gefragt, ob Sie familiären Lungenkrebs haben.	
2.41	JM	Ja, okay.	
2.42	R2	Das sind die drei Sachen, die ich immer ansehen kann in der Studie, die, also die anderen Sachen könnte ich, also da gibt es ein extra Portal, das ist aber ein anderes Portal, wo die klinischen Daten gesammelt werden. Das schaue ich mir aber normalerweise nicht an, weil es mich ja eigentlich biased wenn ich mir das anschaue sozusagen. Und dann bekommen die Patienten ihr CT, und dann gehen sie wieder nach Hause und ihnen wird gesagt, sie bekommen in den nächsten Wochen dann das Resultat bekommt ihr Hausarzt zugeschickt.	
2.43	R2	Manchmal auch die Patienten selber. Ich glaube, das können die angeben, wenn sie das möchten.	
2.44	JM	Dann separat, wird zu einem späteren Zeitpunkt dieser CT-Scan ausgewertet?	
2.45	R2	Genau der wird zum späteren Zeitpunkt ausgewertet. Das wird dann tagsüber gefahren und wenn wenn es gut ist, dann sitze ich vielleicht auch online dran. Wenn das an, ich arbeite ja immer nur Donnerstag und Freitag dran, also meistens am Donnerstag, wenn ich am Donnerstag und am Freitag arbeite, werte ich mindestens die Resultate von Mittwoch, Donnerstag, Freitag aus, manchmal auch sogar die Resultate von Dienstag. So ist das vom Arbeitsaufwand. Es werden die Patienten werden zu viertelstündlich untersucht und an einem Tag werden so zwischen 15 bis 20 Patienten untersucht.	PT
2.46	JM	Ja. Ja.	

2.47	R2	Und ich arbeite dann immer so 3 bis 4 Tage weg, aber je nachdem, was gerade anfällt. Und manchmal sind meine Kollegen auch online. Dann habe ich nicht so viel zu tun und manchmal arbeite ich ein bisschen nach. Und das erste Mal muss man sagen, der erste Zyklus war furchtbar. Das funktionierte gar nichts. Das System brach ständig zusammen und wir hatten große Serverprobleme eigentlich. Das war blöd, weil dann konnten wir die Patienten gar nicht online benachrichtigen. Das war ein bisschen schwierig. Aber seither funktioniert das eigentlich gut und wir geben das manchmal wirklich fast tagesaktuell raus die Befunde. Das finde ich auch relativ wichtig.	RU SR
2.48	JM	Was waren das noch mal genau für Serverprobleme?	
2.49	R2	Das weiß ich gar nicht mehr, was da genau kaputt war, dass das hat die IT damals, es war ständig so, also ich öffnete das Programm und die Verbindung war weg.	SR
2.50	JM	Auf das AI Programm oder auf welches Programm oder das Programm um die Patient zu informieren oder das AI Programm. Sorry, das habe ich nicht ganz verstanden.	
2.51	R2	Ne, sondern das mein eigenes Programm, um die Software auszuwerten. Also das [AI software, anonymized] Programm ging dann, blieb dann schwarz. Man hat also keinen Zugriff auf die [AI software, anonymized] und dann war das und konnte man natürlich demzufolge auch nicht befunden, was ein bisschen schmerzhaft war.	SR
2.52	JM	[AI software, anonymized], also das ist doch die AI die unterstützend ist für die Befundung, oder?	
2.53	R2	Das ist die KI aus [country, anonymized], mit der wir arbeiten, genau.	
2.54	JM	Aber theoretisch kann man doch auch ohne die AI das befunden. Oder würde das zu lange dauern und deswegen wird es dann nicht gemacht?	
2.55	R2	Aber ich habe die Bilder dann ja gar nicht. Die Bilder waren zu dem Zeitpunkt also vom CT direkt auf diese, dieses, diesen [country, anonymized] Server gespielt.	
2.56	JM	In dachte die werden in ihrem PACS Reader wieder parallel sowieso noch mal ohne die AI, das es auch angezeigt wird.	
2.57	R2	Aber ne, das ist tatsächlich so, dass die erst, dass die theoretisch gar nicht ins allgemeine PACS überspielt wurden.	

2.58	JM	Okay, das heißt, die werden direkt mit der AI nur angezeigt werden.	
2.59	R2	Genau, sie werden direkt mit der [AI software, anonymized] angezeigt, sie werden nur ins PACS überspielt, wenn ich das händisch anordne, weil der Patient ein Problem hat und wiederkommen wird, dann kann ich das direkt von über DICOM ins PACS senden, sozusagen. Aber ich mache das, also theoretisch sind die im normalen PACS, sind die nicht vorhanden, die sind nur auf diesem einen Server und von da aus werden die Daten irgendwie jetzt inzwischen in [city anonymized], meine ich, gesammelt, ist jetzt das Vorgehen. Das ist aber am Anfang nicht so gewesen, das kam dann erst später.	
2.60	JM	Können Sie uns mal genau durchführen durch den Prozess, wie Sie das dann diagnostizieren? Also wenn Sie sich reinsetzen vor das Bild bzw. die Ausgangsituation ist ja anscheinend...	
2.61	R2	Genau kann ich gern. Ich öffne das Bild, da gibt es eine Liste mit Patienten.	
2.62	JM	Also Sie setzen sich dran und da ist bereits die AI drübergelaufen, richtig?	
2.63	R2	Ich lasse Sie in dem Augenblick drüber laufen. Also ich öffne den Patienten und dann sage ich und dann habe ich an der Seite, so ein Menü, was ich anklicke. Und dann ist es so Ich öffne den Patienten und dann sage ich, ich hätte jetzt gern die LAA, also die Emphysemauswertung. Und da klicke ich drauf und dann brauche der einen Augenblick um einmal durchzulaufen.	PT
2.64	JM	Okay, dann läuft die AI drüber, okay.	
2.65	R2	Erst mal muss ich es segmentieren. Das ist immer so, ich muss sie oder jedenfalls die Segmentierung überprüfen. Er zeigt mir dann eine Segmentierung an, die schaue ich dann an ob sie stimmt. Also Mittellappen, Oberlappen, Unterlappen, eben diese Lungensegmente und gebe die frei. Und dann kann ich von da aus das Emphysem bestimmen. Und dann gibt er mir auch immer an, in welchem Lappen wie viel Emphysem ist. Deswegen ist es wichtig, dass das stimmt. Und das läuft aber in dem Augenblick, jedenfalls muss ich immer ein bisschen warten, während das während das die Oberfläche, bis die Oberfläche sich öffnet. Ich glaube, das wird erst in dem Augenblick läuft es jedenfalls für mich durch. Irgendwie.	PT
2.66	JM	Wie lange dauert das dann ungefähr?	

2.67	R2	Eins, eine Minute so.	PT
2.68	JM	Oh, ist das eine verlorene Minute, wo Sie einfach nur dasitzen können, oder?	
2.69	R2	Ja, es geht relativ schnell, also die ersten gehen relativ schnell. Das Öffnen des Patienten braucht ein bisschen, dann öffnet sich die Oberfläche, dann die Lungen Segmentierung geht relativ schnell, dann die Emphysemauswertung braucht einen kleinen Augenblick. Dann wechsle ich auf den Calcium Score Auswertung. Das ist auch das wird alles angezeigt. CIC heißt es dann, dann klicke ich da drauf und dann öffnet sich ein anderes Bild. Also ist alles jetzt bisher im weichen Kernel, immer im weichen Kernel. Da muss ich auch am Anfang einmal wechseln, meine ich. Ich muss einmal. Es öffnet sich. Das Problem ist, dass ich einmal, wie ist denn das, ich öffnet das und sage Ich möchte diese Lungen Segmentation und dann will es immer schon das, dass ich den Patienten freigebe, weil es irgendwie versteht, ich bin schon fertig, aber irgendwie muss ich das dann wegklicken. Dann sage ich Nein, ich bin noch nicht fertig. Dann öffnet es mir diese Lungensegmentation. Dann gehe ich zum, dann bin ich im weichen Kernel und wir gucken das Lungenemphysem im weichen Kernel an, den Calciumscore im weichen Kernel. Dann gehe ich auf den den Lung Screen, also die Krebsauswertung. Dann öffnet er mir ein Bild wo ich sehe wo ich mir den Kernel aussuchen kann quasi. Und dann muss ich den anderen Kernel anklicken, den hochauflösenden, also den, Lungen Kernel, der ist hochauflösend.	PEUT
2.70	JM	Ja, ja.	
2.71	R2	Und hat ein anderes Signal zu Rausch und das dauert, das dauert dann richtig so also wenn sich dieses Lung Screen Programm dann öffnet, das dauert mindestens eine Minute, vielleicht sogar zwei, dass da sitzt man und wartet.	PT
2.72	JH	Okay.	
2.73	R2	Ja, genau. Und dann mache ich es eigentlich immer so, dass ich, ich könnte dann schon, ich kann dann diese CAD anklicken, ich gucke mir die Lunge aber immer einmal selber an. Einfach einmal durchscreenen um zu sehen, weil es gibt manchmal Sachen, die er nicht erkennt, ob mir irgendwelche Sachen auffällig sind, sozusagen dann gehe ich da einmal selber durch. Und dann mache ich einmal die CAD und gucke, was der mir gibt. Und manchmal gibt er auch Nonsens. Das klicke ich dann weg. Das kann ich dann auf der CAD rausnehmen und dann oder erkennt sehr große Herde, da hat er die Tendenz nicht zu erkennen, zum Beispiel. Also ich habe mal ein richtig, also zwei, zwei Mal	OQ

		hatten wir richtig große Herde, die er dann gar nicht als Lungenherd erkannt hat. Weil es war kein Rundherd, in dem Sinne war ja eine Rundherderkennung, sondern das war dann irgendwie so bizarr konfiguriert und war unschön.	
2.74	JH	Ich würde mir vorstellen, dass es bei den Großen dann insbesondere kritisch ist, wenn man nicht erkennt, oder?	
2.75	R2	Ja genau, natürlich. Aber andererseits, damit man die erkennt, da braucht man die CAD auch nicht zu nutzen.	OQ
2.76	JM	Aber trotzdem, na ja, fragt man sich dann auch, wie gut ist das System wirklich, was ich hier benutze, wenn so was Großes nicht erkennt, oder?.	
2.77	R2	Ja, aber ich glaube das ist so, das ist so groß, dass es das für ein Organ hält oder so ich glaube, das denkt, das ist eine Magen Hernie oder das ist ein, also das ist so, wenn das Ding, weiß nicht wie groß war das Teil, so sechs sieben Zentimeter hat, dann denkt er irrt sich wahrscheinlich. So ist das so eingegeben.	
2.78	JM	Was ich mich frage ist das System identifiziert und zeigt nur jetzt da auch die die Lungenknoten an oder? Oder zeigt an...	
2.79	R2	Nein, nein, nein, es zeigt auch ein Histogramm inzwischen, das ist neu. Es zeigt mir die Lunge, es zeigt mir dort dran, ob die Fett enthalten, es zeigt mir die Lungenknoten und es spezifiziert sie dann auch gleich. Also wenn ich auf die Einzelnen klicke, dann sagt er mir, das ist vom Volumen her und also da gibt es dieses Lung-RADS, das ist dann 1 bis 4 und dann gibt es noch 4A und B und X und genau. Und er sagt mir dann immer gleich, also wenn ich die einzelnen Knoten anklicke, sagt er mir immer was das für ein Knoten ist und ich guck halt, stimmt das sozusagen.	RU OQ
2.80	JM	Ach das klassifiziert das auch direkt nach diesem Lung-RADS 1.1, oder was gerade aktuell ist?	
2.81	R2	Es macht es, das macht es und das macht er dann also die die einzelnen Knoten klassifiziert er mir gleich, und ich gucke dann halt hin und sonst es gibt da so Fälle, wo man sagt, dass noch nie im Leben ein Knoten in Dreiergröße. Ist es einfach nicht. Und dann guckt man, was er da gemessen hat und da muss man es manchmal ein bisschen korrigieren, weil er da irgendwelche Zäsuren mit misst oder irgendwas Angrenzendes. Irgendwelche Gefäße, die da angrenzen oder so, und dann, also da muss man immer einmal drüber gucken, das lernt ja auch noch ein bisschen das System. Und dann, ähm und seit neuestem gibt es uns auch Histogramme an, das zeigt wie, wie dicht ist der Herd sozusagen, oder	L

		wieviel Weichgewebe, wieviel Kalk, wie viel also weg von dem visuellen Aspekt? Ist das ein dichter Herd? Ist da Kalk drin? Ist das also natürlich, wenn er richtig Kalk drin ist, dann zeigt er das an und dann wird automatisch ein 1-Herd, also als gutartig beurteilt. Aber wenn das so ein bisschen grell ist, also das kann man sehen, ob das eher ein nekrotischer Herd, ob das eher ein dichter Herd ist irgendwie, das zeigt das Histogramm an.	
2.82	JH	Also das Histogram hilft dann auch schon.	
2.83	R2	Das hilft auch schon, ja.	
2.84	JM	Ich kann mir vorstellen, dass insbesondere wenn man zum Beispiel jetzt mit der AI nicht zustimmt, über die die Klassifizierungen der Größe des Lungenknotens jetzt zum Beispiel, das ist dann eher mehr Zeit kostet, dann sich das mal genauer anzugucken, um irgendwie die Messung des AI-Systems zu ändern. Oder nimmt das trotzdem Zeit ab?	
2.85	R2	Nein, es ist einfach, man ändert das schon. Das kostet mehr Zeit. Es ist einfach klar, je mehr Knoten, desto mehr Zeit. Wenn ein Patient zwei Herde hat, dann ist das schnell gemacht. Wenn ein Patient irgendwie leider eine entzündliche Grunderkrankung hat und 20 Herde hat, dann muss man sich trotzdem jeden einzelnen angucken und dann und dann also gut, ich mach das, weil ich weiß, ja, der Patient kommt wieder und wenn ich das nicht mache, dann mache ich es beim nächsten Mal. Ich weiß nicht, wie ich das jetzt im täglichen Alltag hätte. Wirklich so, ob ich da nicht kurz drüber gucken würde und sagen würde, ach naja, der ist jetzt ein bisschen groß, der ist ein bisschen klein, aber so im Großen und Ganzen stimmts. Kann sein, dass ich das machen würde irgendwie. Wenn einer mit 20 Herden kommt und ich gucke so und sage, hmm ist das wirklich ein Zweier, hmm ist das wirklich ein Einer, sondern dass ich da einfach dann so drüber gucken würde und sage, naja Zwei ist schon okay.	
2.86	JH	Aber haben Sie denn auch früher schon mal so so was ähnliches wie eine Lungenkrebsvorsorge gemacht, bevor es diese AI-Tools gab? Also hätten sie da so einen Vergleich, ob das den Workflow geändert hat oder die Zeit auch. Oder waren die Systeme von Anfang an da?	
2.87	R2	Früher gab es ja diese gab es ja auch schon von [software, anonymized] Messsysteme, aber das war so umständlich, da musste man sich wirklich mit jedem noch mal auf einer extra Workstation hinsetzen und das machen. Das haben wir manchmal gemacht, aber meistens nicht. Muss ich ehrlich sagen.	PEUT

2.88	JH	Das hört sich so an, als würden Sie jetzt die Nutzerfreundlichkeit von dem aktuellen [AI software, anonymized] Tool schon deutlich besser finden.	PEUT
2.89	R2	Ja, ja, es ist besser. Eigentlich ist es ein gutes System, das hat, es hat seine Schwächen, irgendwie und die muss man irgendwie noch, deswegen versuchen wir das ja auch möglichst. So ein bisschen da in die Tiefe gehen, um möglichst genau zu machen, damit das System wirklich lernt und man schaut, wo sind die Fehlerquellen? Ich finde zum Beispiel die Emphysem Einschätzung kann auch deutlich verbessert werden. Nun ist es so, dass man ja auch oft das visuell falsch einschätzt wie die AI ist einem da ja manchmal überlegen irgendwie. Das darf man auch nicht vergessen.	PEUT L
2.90	JH	Hatten sie da schon ein paar Mal, zum Beispiel oh nein das ist doch absolut nicht so groß und haben Sie sich das genauer anguckt und dann waren Sie so, oh doch, das hat das komplett richtig eingeschätzt. Und da dachte ich erst das habe ich falsch eingeschätzt.	
2.91	R2	Nee, eigentlich nicht. Aber tatsächlich ist es so, dass ich ein paar Mal bei diesen Emphysemblasen denke irgendwie, okay, wenn man das, also das hätte ich vielleicht anders eingeschätzt, so wenn ich das jetzt im klinischen Kontext gesehen hätte, aber wenn man das genau ausmisst kann es sein, dass er recht hat. So habe ich das mal gedacht, dass man da eher noch mal so ein bisschen zurückging und sagte okay, kann schon sein irgendwie.	PU L
2.92	JM	Ja, da geht es beim Lungenemphysem ja glaube ich auch um die Volumenberechnung, richtig?	
2.93	R2	Da geht es auch um die Volumenberechnung. Das ist aber mehr beim Herd das Problem irgendwie.	
2.94	JM	Was ich mich noch gefragt habe, als Sie gerade über die, die Rundherde auch gesprochen haben, ist und wo sie meinten, dass ab und zu gehen Sie die Bilder durch den sehen ganz große Rundherde, die das System gar nicht erkannt hat. Wenn das wirklich gerade nur anzeigt, wo es denkt, wo Lungenknoten liegen und keine anderen Abnormalitäten anzeigt, wärs dann nicht hilfreich, wenn jetzt auch allgemein mal alle Konsolidierungen anzeigen würde, damit sie auch sicher sein können? Oh das hat es zum Beispiel gesehen, aber nicht als Lungenknoten identifiziert oder?	
2.95	R2	Das tut aber eigentlich das, eigentlich tut es das. Das ist dann wirklich so, dass das muss das irgendwie für irgendwas ganz anderes gehalten haben. Es sind schon nicht mehr Lungenparenchym oder so.	

2.96	JM	Achso das zeigt eigentlich allgemeine Abnormalitäten bei Konsolidierungen an?	
2.97	R2	Das ist ja das Problem. Sie sehen da 20 Herde und müssen da durchgehen und sagen, das ist schon mal gar nichts. Das ist irgendwie ein Anschnitt, dass zum Beispiel. Relativ viele ältere Leute haben ja mal eine Tuberkulose gehabt und da gibt es diese und die Tuberkulose macht so Vernarbungen der Lungenspitzen. Diese Vernarbungen sind aber nicht glatt, sondern die sind immer so ein bisschen mit so kleinen Ausläufern und das erkennt das System sehr, sehr gerne als Rundherd und das ist sage ich mal, bei jedem vierten Patienten muss man da durchgehen. Nun ist es so, dass sich da auch gerne mal Tumoren ansiedeln in dem Bereich und deswegen ist es ganz doof, wenn man das so einen Rundherd gesagt vorgesetzt kriegt und dann sagt nein, das ist aber nix. Da muss man schon sicher sein irgendwie. Und das ist dann genau. Und da muss man dann manchmal mit dem System so ein bisschen kämpfen.	OQ
2.98	JM	Ja.	
2.99	JH	Ich glaube, wir sollten mal mit unseren eigentlichen Fragen weitermachen bevor wir zu sehr abschweifen. Wir hatten noch eine Frage zum Lungenkrebs Screening allgemein, also funktioniert der Prozess da gut oder gibt es da noch irgendwelche Probleme?	
2.100	R2	Also ich finde eigentlich funktioniert er gut. Es gibt ja immer das Problem, was will man jetzt mit der Sensitivität. Also wie oft will man den Patienten beunruhigen und den anrufen? Sie haben vielleicht was und müssen wiederkommen. Und da muss man so ein gewisses Maß von von, nicht zu oft sagen, das könnte was Böses sein und die Patienten dann einbestellen und der dem eigenen Sicherheitsbedürfnis, sag ich mal, wenn man was sieht irgendwie, könnte das nicht doch was sein irgendwie? Oder wann will ich den Patienten wiedersehen? Reicht es, wenn er in einem Jahr wieder kommt? Muss er in sechs Monaten wiederkommen?	OQ
2.101	JH	Hmm.	
2.102	R2	Das aber sonst, also eigentlich funktioniert es gut. Und dieser, diese, diese Einschätzung, wir haben da jetzt ein ziemlich standardisiertes Protokoll und ich finde, das funktioniert ganz gut. In Einzelfällen diskutieren wir das dann im klinischen Team aus.	
2.103	JH	Dann würden wir mal weitergehen zu Fragen zu den...	

2.104	JM	Tatsächliche eine weitere Frage hätte ich dazu noch, weil auch unser letzter Interviewpartner hat uns auch darauf hingewiesen, wo AI noch weiter zusätzlich gut genutzt werden könnte. Zum Beispiel bei Vorschlägen dazu, wann ein Patient erneut wiederkommen sollte oder auch einfach bei der Risikoklassifizierung von Patienten. Oder? Würden Sie da zustimmen? Oder wie sehen Sie das?	
2.105	R2	Ja doch, das ist so, weil das ist wir machen das ja im Grunde so, also wir haben ja diese, diese Lung-RADS Einteilung. Und tatsächlich ist es ja so, ab Stadium drei müssen die in sechs Monaten wiederkommen und dann gibt es 4A und 4B und 4A kommt in drei Monaten wieder und ähm und 4B wird dann gleich klinisch einbestellt.	
2.106	JM	Ich mein jetzt zum Beispiel auch so wie das auch Alter oder wie viel der Patient geraucht hat, auch in der Entscheidung hat berücksichtigt werden. Wann sollte dieser Patient wiederkommen oder wie wahrscheinlich ist es wirklich, dass der ein Lungenkrebs entwickelt hat, dass das irgendwie integriert werden könnte in die bisherige Diagnose?	
2.107	R2	Das wäre schön. Ja. Ja.	
2.108	JH	Das ist wahrscheinlich die Frage, wie gut sich das dann mit den Rahmenbedingungen des Lungencancer Screening vereinbaren lässt.	
2.109	R2	Im Grunde weiß man ja das Alter des Patienten, auch wenn die randomisiert sind. Das Alter sehe ich irgendwie und im Grunde weiß ich Mann oder Frau. Und ich meine theoretisch, weiß das System, also es ist halt noch nicht so stratifiziert, wie viel hat er wirklich geraucht, wie lange? Das taucht da jetzt, glaube ich, nicht auf im Augenblick.	OQ
2.110	JM	Würden Sie denn sagen, für die drei Funktionen, die damit jetzt abgedeckt werden sollen, also die Koronarkalkbestimmung allgemein, die Lungenknotenidentifizierung und -segmentierung und auch Bestimmung des Lungenemphysem? Kriegen Sie da alle Informationen, die Sie für die Diagnose insgesamt brauchen, von diesem System? Oder würden Sie sagen, wenn es mir das noch geben würde, wäre das toll?	

2.111	R2	Nein, eigentlich finde ich das ganz gut. Also das ist irgendwie. Also wir haben ja diesen Agatston-Score für das Calcium. Das ist relativ gut. Ich finde den so ein bisschen vage, weil der ist ab, ich würde sagen, der ist ab 15 oder so positiv. Und tatsächlich habe ich dann die Werte undulieren zwischen 35 und 1000, sage ich mal, also wenn die positiv sind, da könnte man vielleicht auch so ein bisschen feiner segmentieren, denke ich irgendwie, aber das weiß ich nicht, weil ich eben kein Kardiologe bin. Damit kann ich mich jetzt nicht so gut aus. Ich mach das auch sonst nicht die Kalziumauswertung. Das Emphysem finde ich eigentlich ganz gut. Da könnte man auch noch mehr stratifizieren, weil eigentlich weiß das System ja, ist das ein Oberlappen, ist das ein Unterlappen und das sind verschiedene Krankheitsbilder eigentlich, oder das hat eine verschiedene Relevanz für den Patienten. Das könnte sicher auch noch ein bisschen verfeinert werden, aber ich sage mal für so ein basales Patientenrisiko einschätzen finde ich das jetzt schon ganz okay. Also das ist jetzt nicht so, das sind drei Werte, die wichtig sind, die wir auch haben.	OQ
2.112	JM	Ja, ja. Also alle Daten, die das System auch ausschmeißt, sind auch relevant für sie oder schmeißen die irrelevanten Daten raus, die sie gar nicht nutzen, wo Sie sagen. Das macht es dann nur noch undeutlicher oder unübersichtlicher.	
2.113	R2	Nö, nö, macht es nicht.	
2.114	JM	Auch sage ich mal auch Detailreichtum alles eigentlich ziemlich gut, oder würden Sie sagen irgendwo wäre, sage ich mal, generischer besser oder irgendwo ist noch nicht präzise genug, oder da könnte noch mehr, genauer irgendwas angeben?	
2.115	R2	Das weiß ich, wie gesagt bei diesem Calcium, ich mach sonst diese CP-Herzauswertung nicht, deswegen weiß ich das nicht genau. Aber das ist so ein Fach, das ich so ein bisschen vernachlässigt habe. Aber ob man da noch ein bisschen, ein bisschen genauer, welches Gefäß oder ich weiß, welches Gefäß betroffen ist, das kann ich alles sehen, aber ob man da noch bei der Auswertung das Risikoprofil noch ein bisschen feiner gehen könnte. Aber im Grunde finde ich das sehr detailreich. Wie gesagt, die Lunge ist fast manchmal ein bisschen zu detailreich. Ob man wirklich jeden winzigen Herd mit klassifizieren muss, weiß ich nicht.	OQ JR
2.116	JH	Finden Sie denn, dass der Output verständlich dargestellt ist von dem AI System?	

2.117	R2	Also unser Output ist ja sehr basal sag ich mal, den wir dem Patienten liefern, irgendwie. Das ist, vielen Dank, dass Sie teilgenommen haben. Sie haben gar kein Risiko, ein minimales Risiko und in einem Jahr reicht die Kontrolle. Sie haben einen wahrscheinlich gutartigen Befund und sollten sich aber in sechs Monate, sollte aber in sechs Monaten kontrolliert werden. Dann gibt es noch ein vermutlich gut, also auch wird auch beruhigt, aber Sie sollten sich in drei Monaten noch mal vorstellen und dann gibt es dieses Bitte kommen Sie, also das ist. Man muss dabei wissen, dass auch diese 4A-Befunde, wo wir die in drei Monaten wieder einbestellen oft nur ein wirkliches Erkrankungsmalignitätsrisiko von 2 bis 4 % haben. Also das ist nicht so, dass die dann auch wirklich Krebs haben.	
2.118	JM	Und bezogen auf den Output, den das System ihnen jetzt bei der Bildanalyse bietet. Würden Sie da sagen, dass da auch alles verständlich ist und dass Sie damit gut umgehen können und es auch gut verwendet werden kann?	
2.119	R2	Es ist ein ganz klein bisschen umständlich, die Messfunktionen sind umständlich, die muss ich immer anklicken und es gibt keine Shortcuts. Das könnte man natürlich wahrscheinlich einfach einrichten, irgendwie. Das gibt es jetzt nicht, weil ich, weil die das nicht eingerichtet haben. Ich schätze, das wäre eigentlich kein Problem, aber man muss halt immer, Messungen sind ein bisschen umständlich in dem System, sowohl Durchmessermessung als auch dichte Messungen. Also auch und ja, dann gibt es noch so Kleinigkeiten. Man man scrollt mit einer Taste, mit der man sonst nichts scrollt und so, aber das ist vielleicht auch normal in [country, anonymized]. Das weiß ich nicht genau.	PEUTE RU
2.120	JM	Ja. Wissen Sie, ob das nur Ihnen so geht? Oder haben Sie sich auch mal mit den Kollegen darüber unterhalten?	
2.121	R2	Ich weiß, dass es am Anfang alle ein bisschen gestört hat, dass man irgendwie, wenn man, wenn man scrollt, dass man da die andere Maustaste nimmt, irgendwie. Und es ist, ähm, ich sage mal, nicht für alle, ich bin nun so, ich klicke einfach auf alles einmal drauf und gucke was passiert. Aber ich sage mal, wenn Leute, gerade ältere Leute, vielleicht sieht man, was sag ich da, ältere Leute. Aber es gibt dann tatsächlich auch Leute, die noch weniger Erfahrung mit Computern haben. Für die ist das vielleicht nicht ganz so intuitiv. Bedienbar sag ich mal.	PEUT
2.122	JH	Also gibt es da irgendwie im Tool..	
2.123	R2	Also es gab eine kleine Einarbeitung, an die ich mich kaum erinnere und die dann irgendwie auch nicht nicht mehr erfolgte.	PEUT

2.124	JH	Aber gibt es eine Dokumentation, die man nachgucken kann. Oder im Tool Hilfe?	
2.125	R2	Ja, ja, die gibt es, das gibt es. Und man kann auch Fragen stellen. Wenn wir irgendwas nicht wussten, es gibt einen ITler, der für uns zuständig ist, dann kann man den auch einfach anschreiben und fragen. Also der ist sehr nett.	RU
2.126	JM	Und haben Sie das auch mal gemacht, funktioniert das schnell und hat auch gut funktioniert alles?	
2.127	R2	Ja. Ja.	
2.128	JM	Super. Ja.	
2.129	R2	Meistens, mehr, ich habe es mehr gemacht, wenn es Probleme gab, weil wieder irgendwas nicht funktioniert oder irgendwas sich nicht bewegte oder ich einen Patienten nicht schicken konnte. Oder manchmal werden mir auch, die werden ja vom CT ins System geschickt und kommen dann unverständlich an oder haben irgendwie einen Bug dadrin und das hat immer sehr gut funktioniert. Da gibt es einen Herren der sich drum kümmert.	RU
2.130	JM	Aber allgemein ist das System sehr buggy oder haben Sie da regelmäßig Probleme mit?	
2.131	R2	Nee, nee, nee.	
2.132	JM	Das sind dann Ausnahmefälle wo das System nicht funktioniert?	
2.133	R2	Ja, es war zum Beispiel einmal so, dass ein Patient einen Herzschrittmacher hatte, der sehr große Artefakte machte. Da ist es komplett zusammengebrochen und musste dann erst mal, also diese automatische Auswertung funktionierte gar nicht, weil diese ganzen, weil diese Auswertungsartefakte, da glaube ich alle das Auswertungssystem völlig überfordert haben.	SR
2.134	JM	Ja.	
2.135	R2	Das war ein Problem. Dann ist letztens mal einer dann irgendwie falsch ausgelesen worden, ich weiß nicht wo das war. Der lag irgendwie, ich glaube, dass das im CT die TA falsch eingestellt hatte. Dem wurde wo das Zentrum des Bildes war, irgendwie. Und da war dann hinten quasi der Rücken abgeschnitten, irgendwie und das musste dann noch mal geschickt werden. Aber sonst, nö. Also seit diese, dieses, diese Serververbindung oder dieses Verbindungsproblem behoben ist, hat es eigentlich, das war das	

		was am Anfang am schlimmsten war, dass es immer wieder das System neu gestartet werden musste. Und da war ein ITler ein ganzes Wochenende quasi On-Call und hat mich immer wieder, hat mir immer wieder das System hochgefahren, damit ich weiterarbeiten konnte.	
2.136	JM	Noch mal zu dem, was Sie davor gesagt hatten, weil fand ich ganz spannend. Sie meinten ja, es war ein bisschen störend, dass man mit der anderen Maustaste scrollen musste. Da habe ich mir gerade überlegt, das liegt ja nicht per se daran, dass das System irgendwie schlecht designt ist, aber das liegt daran, dass Sie normalerweise mit einem anderen System, die Sie benutzen, mit der anderen Maustaste scrollen, oder?	
2.137	R2	Ja genau. Die Funktionen sind einfach ein bisschen unterscheidlich belegt. Ich meine da hat man sich nach dem dritten Patienten auch wieder dran gewöhnt. Das ist ja nicht schwierig, aber es ist, aber es ist schon so, aber die andere Taste hat andere Funktionen. Also das ist dann irgendwie...	PEUT
2.138	JM	Das ist dann nochmal Umgewöhnung, ja.	
2.139	JH	Es sollte ja eigentlich auch standardmäßig belegte Tasten geben.	
2.140	R2	Genau, aber das ist, das ist lächerlich. Aber es ist natürlich dann in dem Augenblick so, die ersten zwei Male oder zum Beispiel wenn man die Nodules selber ausmisst. Man muss immer, man misst ein Nodule aus und dann muss man diese Funktion immer wieder wegklicken. Und das ist was, so blöd sich das anhört, was man oft vergisst, also es ist wirklich jedes Mal wieder und dann, dann, wie ist denn das. Und dann scrollt man mit derselben Maustaste nämlich weiter, glaube ich und wenn man das dann vergessen hat, zeichnet man plötzlich irgendwelche Nodules ins System, die man gar nicht zeichnen will. Das ist ein bisschen, das ist mir ein paarmal passiert, also irgendwie, das ist dann und dann muss man die erst mal wieder löschen. Das ist natürlich jetzt irgendwie, aber ich sage mal, dass es einfach, wenn man bedenkt, man möchte gerne von diesem Patienten 30 machen oder so. Hallo, jetzt sehe ich aber gar nichts mehr.	PEUT
2.141	JM	Ich, ich habe gerade überlegt, vielleicht wenn ich einmal das Video ausmachen und wieder anmache, ob dann das wieder flüssig ist. Ja es ist super. Weil ich dachte mir ist ja unschön, wenn Sie die ganze Zeit in dieses eingefrorene Bild gucken.	

2.142	JH	Dann noch eine Frage. Ähm, was erachten Sie so als die größten Probleme bei der Nutzung mit dem System? Sie haben jetzt ja schon ein paar aufgezählt, aber was stört Sie so am meisten noch an dem System?	
2.143	R2	Die Zeitverzögerung, gut da kann man wahrscheinlich nicht so viel machen. Das finde ich noch ein bisschen lästig. Ich, ne es sind eher so, es sind eher so bedienerunfreundliche Sachen, wie gesagt, wie zum Beispiel dieses, also das stört mich tatsächlich. Also, mich stört zum Einen, dass ich so viele Herde immer noch ausmessen muss, aber ich glaube, das ist so eine Kinderkrankheit. Das soll er ja irgendwann dann ganz alleine machen und nicht mehr korrigiert werden. Aber mich..	PEUT
2.144	JM	Da würden Sie auch sagen, das kann das System einfach noch nicht gut genug, oder?	
2.145	JH	Oder es tut das einfach noch nicht?	
2.146	R2	Ne, würde ich sagen, das kann es, das kann es noch nicht gut genug irgendwie. Das ist so das, das, also sagen wir immer so, oder es erkennen zu viele Herde. Man muss ja immer dran denken, die Leute kommen also in einem Jahr wieder und sie haben dann 20 Herde, die zum Teil winzig klein sind, die dann im Verlauf, die das System aber speichert und dann im Verlauf kontrolliert und von diesen winzig kleinen Herden, wenn so ein Herd zwei Millimeter groß ist, da muss der Patient nur einmal ein bisschen anders schnaufen. Plötzlich ist er drei Millimeter groß und hat ne Verdopplungs... und wird einem als maligner Herd angezeigt. Also, ich sehe da noch, wir haben jetzt die ersten Drei-Monatskontrollen gehabt. Das war zum Teil ein bisschen schwierig irgendwie, also mit dem mit der Vergleichbarkeit. Und da muss man wirklich aufpassen, dass man nicht zu kleine Herde dann nimmt irgendwie, weil dann ist das mit diesem, wo sich vielleicht einfach nur so ein bisschen Lungenparenchym weiter verklebt oder so, aber das ist wirklich eine sehr empfindliche Größezunahme, in der Größenzunahme ist ist das wirklich sehr sehr empfindlich und das ist irgendwie, das ist natürlich auch so gedacht, dass es besser sieht als das menschliche Auge, aber das ist dann in dem Augenblick, muss man da eben sehr aufpassen, dass man nicht übertreibt mit den ganz kleinen Herden, das ist glaube ich schwierig.	OQ JR PU
2.147	JM	Ja okay. Wie würden Sie sagen, verändert die Nutzung dieses Systems Ihre Ihre finale Diagnose allgemein oder wie verändert das quasi oder wie, wenn Sie sich überlegen würden, wenn Sie das alles ohne das AI System machen würden? Wie sehr würde jetzt quasi das Endresultat der Diagnose, inwiefern sieht das anders aus als quasi das, wenn Sie es ohne AI machen würden?	

2.148	R2	Also er sieht mehr Herde als ich, weil ich einfach ganz viele Herde in meinem Unterbewusstsein schon mal wegklicke irgendwie und das ist so, ich scrolle da durch und wenn da so eine Fissur, so eine kleine Verdickung hat oder so, das das, da gucke ich keine Minute länger drauf irgendwie und das holt er mir alles raus und macht mir aber schon noch mal bewusst und ganz das passiert selten, aber manchmal sieht er tatsächlich kleine Herde, die aber vielleicht potenzial haben, die ich übersehe. Gerade so, die so eben gerade so zwischen zwei und drei sind, sag ich mal so zwischen Herd aber nicht so wichtig und könnte was sein irgendwie und da ist er denke ich besser als ich. Größere Herde sehe ich auch und er sieht eben mehr unwichtige Herde als ich, muss man sagen.	OQ PU
2.149	JM	Okay, würden Sie da ich sagen dadurch erhöht sich, naja, die Anzahl der falsch positiven oder wie denken Sie? Also kann man vielleicht bis jetzt schlecht sagen, aber wie denken Sie beeinflusst das insgesamt Output?	
2.150	R2	Ohne Kontrolle würde es sicher die Anzahl der falsch positiven erhöhen, bin ich ganz sicher. Aber da wir das ja alles nochmal durcharbeiten und dann irgendwie zum Teil dann auch, wenn man so einen Herd dann ändert, dann gibt man den immer noch mal jemand anderen draufgucken eigentlich, dass man da ja auch so sicher ist, dass das wirklich nichts ist irgendwie, dass man tatsächlich, also ich glaube es würde schon mehr falsch positive geben, wenn wir nicht draufgucken würden. Es würde nicht, nicht wesentlich, also ich weiß gar nicht, ob ich bisher einen einzigen wirklich bösen Befund übersehen hätte.	PU
2.151	JM	Okay, ja.	
2.152	R2	Und das andere ist und da muss man abwarten, die Kontrolle in einem halben Jahr oder im Jahr, ob die kleinen Herde, die so ein bisschen als verdächtig eingestuft wurden...	PU
2.153	JM	Gibt es Patienten die schon zum Zweiten Mal kamen, wo sie eine Einschätzung haben, ob insgesamt die Performance allgemein sehr gut ist, zusammen mit dem System oder können Sie dazu noch nichts sagen.	

2.154	R2	Also jetzt kommen ja gerade erst, das hat ja im August angefangen, jetzt kommen ja gerade erst die sechs Monatsverläufe rein, die kommen jetzt gerade und da habe ich jetzt nicht so viel gesehen leider, weil wir hatten uns das so ein bisschen aufgeteilt, dass das sind natürlich deutlich weniger als die Gesamtmasse der Patienten und eine Kollegin, die nicht ganz so viel Zeit hat, die macht dann lieber die, diese Kontrollen sozusagen, weil es eben auch, das ist, die hat sehr viel Erfahrung. Das ist eben [name, anonymized], mit der Sie auch noch sprechen werden, die hat sehr viel Erfahrung und gerade diese etwas kritischen Fälle, das ist immer so, glaube ich, sehr gut, wenn sie das macht. Aber natürlich lernen wir anderen da auch dran. Es ist nicht so, dass wir das komplett teilen, also komplett weggeben. Aber es ist, es ist dann doch so, dass ich mich dann zum Beispiel mal einen Tag hinsetze und nur das mache und sie eigentlich mehr so in ihrem Arbeitsfluss, dann am Nachmittag, wenn noch Zeit ist, sich die schwierigen Fälle anguckt.	
2.155	JM	Ja, ja und ähm, würden Sie aber sagen, dass Ihnen mit dem AI System insgesamt Zeit abgenommen wird? Oder was denken Sie, wie viel schneller arbeiten Sie zusammen mit dem AI System als ohne das AI System?	
2.156	R2	Es wird aber auf jeden Fall, wenn ich das jetzt nur machen würde und also jedenfalls für die Lungenherde, wenn ich das jetzt nur machen würde, das Ding anklicke und, und das gibt mir die Herde und ich kontrolliere dann wirklich nur die Herde. Das kommt auch so ein bisschen darauf an, wie gesagt, wenn der Patient 20 Herde sind, von denen ich vielleicht zehn weggeklickt hätte oder gar nicht, gar nicht beachtet hätte, wäre es natürlich mehr Arbeit, aber ich glaube, so im Großen und Ganzen macht es, erleichtert es mir die Arbeit schon, weil ich dann nur noch das durchschaue, was er mir zeigt. Und da ich natürlich sowieso drüber über einmal über die Lunge drüber gehe, wurden mir so Sachen, die er nicht sieht, wahrscheinlich auch auffallen. Also das.	OQ RD PU
2.157	JH	Gibt es denn eigentlich einen vorgeschriebenen Prozess, wie Sie als Radiologin, die die AI benutzen müssen? Ist es vorgeschrieben, dass sie sie überhaupt benutzen? Oder benutzten sie sie, weil es ihnen die Arbeit erleichtert.	
2.158	R2	Ich mache das, weil es mir die Arbeit, ich muss das gar nicht, also so im normalen Arbeitsfluss, also abgesehen von dem, aber auch die Mammographeure, ne müssen tun wir es gar nicht, also zum Beispiel für die Emphysemauswertung, läuft jetzt gerade in [city, anonymized], das weiß [name, anonymized] sicher besser, weil sie das mehr macht. Dann wünschen sich die Pneumologen immer eine maschinelle Auswertung sozusagen. Also weil sie das	PU JR SV

		besser quantifiziert, aber ich glaube das sonst, also wenn ich so einen Patienten krieg und die Frage ist, da wurde ein Herd gesehen und der hustet zu viel oder so und schauen Sie mal darauf, ob der was hat, dann benutze ich keine AI.	
2.159	JM	Ja. Auch mal bezogen auf andere Radiologen wissen sie auch, ob die da auch alle genauso wie sie die AI benutzen ähnliche und die auch gerne benutzen? Oder sehen die das kritischer teilweise, oder?	
2.160	R2	Also in [city, anonymized] benutzt das glaube ich keiner ehrlich gesagt. Ich weiß, es kommt auch immer darauf an, wo das Gerät gerade rumsteht und zur Verfügung steht. Also das ist, wie das ist und das ist oft auch in Absprache mit dem Überweiser. Es gibt Überweiser die da Wert drauf legen. Ich glaube in [hospital, anonymized] haben wir so eine syngo.via stehen, wo dann einige Sachen auch dann tatsächlich auf syngo.via immer noch mal extra angeschaut werden. Dann aber wie gesagt, es ist immer ein extra Arbeitsschritt dann. Dann setzen Sie sich da einmal hin, dann rufen Sie den Patienten auf, dann wird der meist noch irgendwie ausgelesen und dann müssen Sie es noch einmal durchgucken. Also das ist tatsächlich mehr Arbeit und das macht man meistens auch nicht einfach so.	PU JR SV PT
2.161	JH	Und das ist dann wahrscheinlich auch der Grund, wieso Andere das nicht so gerne benutzen, weil es einfach mehr Zeitaufwand bedeutet.	
2.162	R2	Genau.	PT
2.163	JM	Aber jetzt beim Lung Cancer Screening, ist das ja deutlich besser integriert, oder? Und da wo Sie alles weg arbeiten, benutzen das schon wahrscheinlich die meisten, oder?	
2.164	R2	Das ist aber bei diesem Lung Cancer Screening gehts es ja im Augenblick nur darum, das ist ja im Grunde eine, schwer zu sagen. Aber wenn ich jetzt so, wenn jetzt Leute kommen würden, nur dafür, dann würde ich das denke ich schon machen, weil das die Arbeit ja auch wirklich erleichtert. Und in 80 % der Fälle das würde ich so sagen.	SV PU
2.165	JH	Sehen Sie irgendwie noch Potential wie das irgendwie noch besser, also dieses AI Tool, noch besser in den Arbeitsprozess integriert werden kann?	

2.166	R2	Man müsste es einfach, also wenn Sie Ihre normale Oberfläche haben, im Grunde müsste es hinter der Oberfläche ablaufen. Also das ist dann so, dass Sie dann dem Patienten aufrufen und dann gucken Sie sich irgendwie irgendwas anderes an und dann in dem Augenblick, wo sie dann die die KI brauchen. Also das ist immer schwierig. Ich guck ja nicht nur Lungen an, es ist irgendwie. Es müsste sich im Grunde und auf der gleichen Benutzeroberfläche aufrufen lassen, und zwar schnell mit einem Klick und jetzt irgendwie klicke ich hierhin und dann öffnet sich die AI, und dann mache ich das schnell so. Und dann mache ich in meinem normalen Programm weiter, auf demselben Bildschirm, auf derselben Benutzeroberfläche, ohne dass ich irgendwie irgendnoch was und hin und herschalten muss oder irgendwie etwas wegklicken muss.	PU
2.167	JM	Auch nochmal, weil ich nicht sicher bin, ob ich die die Antwort von Ihnen zu der Frage davor ganz verstanden habe. Also jetzt, wenn auch die anderen Radiologen, die die CT Scans im Rahmen vom Lung Cancer Screening beurteilen, da benutzen ja alle jetzt [AI software, anonymized] richtig und gucken sich das daüber an.	
2.168	R2	Aber weil wir mit diesem System, das ist die Studie. Also da gibt es keine andere, kein anderes System, das da ja integriert wird. Wir nehmen immer das. Also das ist quasi vorgegeben.	SV
2.169	JM	Aber ganz kurz. Sie meinten ja auch, weil das ja auch direkt so integriert ist damit beziehungsweise, das heißt sie müssen das eigentlich im Rahmen der Studie benutzen?	
2.170	R2	Ja, es geht gar nicht anders. Die Studie ist, dass in den, in den Anträgen und bei allem ist diese Software vorgeschrieben.	SV
2.171	JM	Ja, aber auch wenn, wenn das irgendwann mal eingeführt wird und die AI nicht Pflicht ist, würden Sie dann sagen, dass es trotzdem genutzt werden würde, weil es ja irgendwie Zeit einspart? Oder denken Sie eher nicht?	
2.172	R2	Also ich glaube für dieses Screening Programm ist es ganz wichtig, dass man was standardisiertes benutzt. Irgendwie, also wenn, wenn dann sollte man bei jedem Patienten, der kommt, dann auch dasselbe Programm benutzen. Und insofern finde ich das schon. Und wenn die zum Screening kommen, dann sollte man auch Screening machen. Also das ist dann tatsächlich irgendwie, dann reicht es halt nicht, wenn der Radiologe da eben mal kurz drüber schaut, auch mal immer eingedenkt, dass ja nicht immer der selbe Radiologe beim nächsten Mal, wenn der Patient kommt, zur Verfügung steht und das ist ja meist das ist das soll ja, so ein Lungen Screening ist ja nur sinnvoll, wenn das alle paar Jahre wiederholt wird irgendwie.	JR SV

2.173	JH	Wenn Sie meinen, dass dieses [AI software, anonymized] Programm durch die Studie vorgeschrieben ist, heißt dann vorgeschrieben, dass jede Praxis, die teilnimmt, halt dieses Programm anschafft? Aber es ist für die Radiologen an sich nicht vorgeschrieben, dass sie es benutzen. Aber wenn sie es nutzen sollten sie das benutzen?	
2.174	R2	Es ist so, für unsere Studie, die wir gerade machen, ist es vorgeschrieben für das Lungencancer Screening selber ist es glaube ich eigentlich egal, was sie benutzen. Es sollte nur dann, es sollte dann nur sein, dass wenn ein Patient zum Screening kommt, sollten sie die KI benutzen und wenn er das nächste Mal kommt, sollte dieselbe KI angewendet werden.	SV
2.175	JM	Ja.	
2.176	R2	Das ist wichtig, das und damit es so ein bisschen wird, sagen wir mal so, der Patient ist jetzt irgendwie Mitte 50 und kommt dann alle paar Jahre, der wird nun glückliche 85 oder so irgendwie und dann ist halt dann halt irgendwann ist das dann nicht mehr derselbe Radiologe, sitze dann nicht mehr, weil ich dann auch schon meine Füße hochlege.	
2.177	JM	Sehen Sie allgemein der Nutzung von AI im Rahmen von Lungen Cancer Screening entgegen? Also finden Sie es ist schön mit der AI zusammenzuarbeiten, so, weil es das effizienter macht? Oder sagen Sie so, früher war alles besser, als alles händisch noch gemacht wurde. Das ist die wahre Radiologenarbeit.	
2.178	R2	Ne, gar nicht, ich finde so was ja immer spannend. Ich bin da nicht so ganz, also nicht so ganz fest davon überzeugt. Am schönsten wäre es natürlich man lässt die Leute kommen. Da läuft das Programm drüber und spuckt am Ende dann was, ein paar Zahlen aus und man kann sich da hundertprozentig drauf verlassen. So ist es halt leider nicht irgendwie und Ich weiß auch nicht, ob wir da jemals hinkommen, dass wenn das das Ziel ist, ist das natürlich hoch sinnvoll, weil das Zeit und Manpower und alles Mögliche spart irgendwie. Und dann finde ich das super. Ich meine, ich rationiere mich da gerade selber so ein bisschen weg, aber das ist ja, dann kann man Zeit auf andere Dinge verwenden. Ich meine, ich sage jetzt mal, Lungenknoten zählen ist jetzt auch eine Arbeit, die so ein bisschen.	SV
2.179	JH	Also Sie haben keine Angst, dass das allgemein Ihren Job ersetzt, sondern nur die lästigen Arbeiten eher, und dann können Sie sich auch auf sinnvollerere Sachen fokussieren?.	

2.180	R2	Es wäre schön, wenn nicht, wenn man die Sachen, wo man so ein bisschen, die einem Spaß machen oder wo man ein bisschen nachdenken muss oder so, wenn man die behält und Sachen, wo man Punkte zählt, dann der AI übernimmt.	SV
2.181	JM	Solange das dann in soeiner Unterstützungsfunktion gerade ist, ist das doch super, oder?	
2.182	R2	Ja, also finde ich.	SV
2.183	JM	Aber hätten Sie dann Angst, dass es irgendwann das dann tatsächlich kann? Also dass es die Diagnose selber stellt? Macht Ihnen das Angst?	
2.184	R2	Ach bis dahin arbeite ich nicht mehr.	
2.185	JM	Wie gehts Ihren Kollegen denn mit dem Thema?	
2.186	R2	Ja, genau. Tatsächlich ist es so jetzt. Ich habe ja auch Kinder, die studieren wollen, die wollen ja auch irgendwann studieren und meine Tochter möchte wahrscheinlich auch Medizin studieren. Und dann ist es natürlich so, dass man schon drüber nachdenkt, kann ich ihr empfehlen Radiologin zu werden? Oder gehe ich davon aus, dass das in 30 Jahren schon nicht mehr so als eigene Speciality existiert oder ob nicht auf jeden Fall im sehr, im ganz anderen Maßstab existiert, ob es da viel weniger von gibt und die ein ganz viel interdisziplinäres Aufgabenfeld haben.	
2.187	JH	Wahrscheinlich muss man nur darauf achten, dass man auch irgendwie noch technische Fächer dazu hat, mit mit man auch diese AI verstehen und bedienen kann. Ist dann wahrscheinlichlich eine wichtige Kompetenz in dem Bereich.	
2.188	R2	Ich finde das eigentlich grundsätzlich gut, weil kein Mensch wertet gerne Scores aus oder zählt gerne Lungenknoten oder irgendwie, oder misst gerne Sachen aus. Um mal ganz ehrlich zu sein, das ist ja alles so ein bisschen Deppenarbeit, um das mal freundlich zu sagen, irgendwie, das ist, das kann der Computer besser und soll er auch tun. Also das, genau und sonst weiß ich nicht genau, wo das hingeht irgendwie. Das bleibt ja auch spannend.	
2.189	JH	Ja. Ist es denn eigentlich im Programm so, dass wenn die AI irgendwas anzeigt, dass Sie das nachvollziehen können? Also gibt es irgendeine Erklärung? Oder wenn Sie das sehen, dass es angezeigt wird, dann sind Sie auch so ah ja okay, das ergibt Sinn, dass das irgendwie als bedenklich markiert wurde?	
2.190	JM	Zum Beispiel bei der Berechnung des Kornarkalkscores oder so.	

2.191	R2	Ja. Nö, eigentlich hat es, also das macht schon Sinn. Das ist irgendwie, das Problem, das Problem im Augenblick ist eigentlich, dass es die ganzen Stents immer mitmisst, wo ich noch nicht genau weiß, wieso das nicht implementiert ist, dass das, also manche nimmt er raus, manche nicht und ich korrigiere eigentlich die ganze Zeit nur die Stents raus, oder lösche da die ganze Zeit nur die Stents raus dem Programm. Aber das ist natürlich ist das, natürlich ist das super sinnvoll und dass man dann, wenn man so ein CT gleich fährt, dass dann gesagt wird, Sie haben ein Risiko einen Herzinfarkt zu kriegen, nehmen Sie schon Medikamente? Wenn nicht, sollten Sie es tun, ist das vielleicht eine sinnvolle Sache.	OQ L RU
2.192	JM	Aber zum Beispiel auch wenn es sagt, so der Lungenherd ist so und so groß, dann sehen Sie da quasi auch die Umrisse, die er da genau identifiziert hat, oder?	
2.193	R2	Ja. Ja. Ja. Kann ich mir in drei Ebenen anschauen. Es ist tatsächlich, ich kann mir das wahrscheinlich auch dreidimensional anschauen, wenn ich auf drei Knöpfe klicke. Das brauche ich halt nicht wirklich, tatsächlich arbeitet man ja auch... Das Programm ist, glaube ich, noch viel breiter angelegt, als ich es wirklich benutze irgendwie. Also das ist jetzt nicht so, dass ich mit allen, mit allen Features spiele, die es mir oben zeigt. Ich mache ja auch nicht jede Messung und so oder so, ich benutze da natürlich die Sachen, die Sinn machen.	L PEUT
2.194	JH	Bringt das dann auch so ein bisschen Spaß damit so rumzuspielen mit dem Tool?	SV PEUT
2.195	R2	Ja, ja, klar.	SV PEUT
2.196	R2	Was ich mich halt immer noch frage ist, wenn Sie zum Beispiel einer Messung jetzt nicht nicht zustimmen oder sagen so ne, so groß ist der Lungenherd nicht. Was genau würden Sie dann machen? So gehen Sie dann in das System rein und verziehen Sie dann die, die die Umrisse im System oder messen sie das selber noch mal aus?	

2.197	R2	Genau, ne ich messe das, ich messe das tatsächlich selber noch mal aus, das ist so, ich gucke mir erst mal an und denk mir, woran liegt es, dass der Herd falsch erkannt wird irgendwie, weil der für mich deutlich, also deutlich kein risikoreicher Herd ist und dann gucke ich so,warum schätzt der den falsch ein? Ist es wirklich die Größe? Irre ich mich vielleicht und dann messe ich es meistens noch mal nach und wenn es dann und wenn es dann, also wenn ich dann sehe, okay es liegt wirklich über der Stelle, dann gucke ich noch mal, sind irgendwelche Gefäße mit drin irgendwie. Warum denke ich, dass es kein, manchmal irre ich mich auch und dann muss ich das zugeben und manch mal, aber ganz misst er halt irgendwelche Nachbarstrukturen mit. Und das muss ich mir in drei Ebenen angucken, dann sehe ich gut, der geht da mit, der hat ein Ausläufer auf die Fissur oder er misst dieses Lungengefäß, das da direkt vorbeiläuft mit oder so und man das halt versuchen rauszuschneiden, das ist...	L
2.198	JM	Aber dann nutzen, also dann nutzen Sie nicht das System für die Messung, sondern dann eher händisch gucken Sie..	
2.199	R2	Es gibt eine Nodulemessung. Man kann @Nodule machen und dann kann man den ausmessen und man kann es auch ändern. Ich kann immer die Shape ändern und die Size ändern.	
2.200	JM	Also nutzen Sie schon das System dafür?	
2.201	R2	Ja ich nutze dafür das System.	
2.202	JH	Also manuell dann sozusagen?	
2.203	R2	Genau. Ich kann ganz rausgehen und es neu messen. Das mache ich manchmal, wenn ich ganz verzweifelt bin. Aber ich kann auch die Messung nehmen, die er mir gibt und dann kann in Maßen das ändern. Es gibt mir Plus/Minus und das kann ich in die eine Richtung zweimal in Richtung dreimal machen oder so, ich kann dann immer so herspielen und dann gibt er mir unterschiedliche Formen, zum Beispiel, weil er sagt gut dann liegt das [inaudible] und dann gibt er mir einen Stern oder eine Ellipse oder so und dann sage ich kommt hin.	JR
2.204	JM	Ja.	
2.205	R2	So ungefähr.	

2.206	JH	Hilft das Programm dann eigentlich nur bei der Erkennung oder generiert es am Ende auch einen Report, sozusagen etwas Schriftliches, was Sie sonst händisch selber machen müssten?	
2.207	R2	Ne genau, es gibt einen Report, der ist sehr kurz gehalten. Dann gibt es eben wirklich nur, Sehr geehrter Herr Soundso, wir haben bei der Auswertung das und das rausgefunden und dann eben randomisiert kommt dazu und dann wird immer ein Bild ausgedruckt und einen Satz über das über das Emphysem und dann wird gesagt, ein signifikantes Emphysem liegt ab einem Schwellenwert von 7% vor und dann wird der Emphysemwert ausgespuckt oder es wird eben dazu geschrieben, es liegt eben kein Lungenemphysem vor und dann bekommt es einen Kalziumwert und dann wird eben gesagt ein signifikantes Herzinfarktrisiko liegt dann ab einem Wert von so und so vor. Oder es liegt kein, und das ist randomisiert und bedeutet, dass auch die negativen Befunde mitgeteilt kriegen oder nicht.	
2.208	JH	Das sind ja schon einige Informationen. Würden Sie das, wenn Sie das Tool nicht hätten, müssten Sie das alles selber schreiben? Also das hört sich ja auch so an, als würde das einiges an Zeit einsparen.	
2.209	R2	Also tatsächlich ist es so, dass die Software, es gibt so Sachen, es gibt so Grunderkrankungen, wo die Software das Emphysem nicht richtig erkennt und wenn ich finde und ich sehe das und ich finde, das Emphysem ist deutlich falsch eingeschätzt, dann schreibe ich und der hat wirklich ein ausgeprägtes Emphysem, dann schreibe ich es händisch drunter und dann schreibe ich das drunter, dass aus den und den Gründen vermutlich von der Software nicht erkannt liegt doch ein relevantes Emphysem vor. Das schreibe ich dann so hin. Also ich kann so Additional Findings hinzufügen am Schluss, das kann ich machen. Oder wenn ich sehe, der Patient hat eine riesige Schilddrüse zum Beispiel. Dann kommt das auch noch unten mit rein, das macht das System, sowas schreibt, das erkennt das System natürlich nicht und schreibt es nicht selber rein. Und das ist dann wieder so ein bisschen die Klinikerarbeit daran, aber also über die Lungen, die die die Lungenrundherdinformation ist immer drin in dem Befund. Die ist immer drin, das andere ist randomisiert.	
2.210	JM	Ich glaube, Sie meinten das dauert so ungefähr zehn Minuten pro Patient, richtig?	
2.211	R2	Ja, ich kann es in fünf Minuten schaffen, wenn da wirklich, wenn der wirklich nichts hat und dann kommt der nächste Patient, der 20 Herde hat, das dauert fünfzehn Minuten. Also im Schnitts sind es zehn.	

2.212	JM	Darauf wollte ich eingehen. Wofür werden diese ganzen zehn Minuten oder fünf Minuten oder 20 Minuten denn verwendet? Also noch mal einfach alle Bilder durchgucken, alle Lungenknötchen überprüfen? Was genau ist da noch am zeitintensivsten?	
2.213	R2	Na ich gucke, ich gucke, einmal die Emphysem, also das geht eigentlich ziemlich schnell, die Segmente angucken, wenn da nichts ist, dauert das wirklich nur ein paar Sekunden. Dann gucke ich das Lungenemphysem durch und das kommt immer drauf an, wenn da nicht viel ist, geht das auch ziemlich schnell und dann warte ich bis es umspringt. Dann gucke ich mir den Calcium Score an und das ist ein bisschen zeitaufwendig, weil dab ich auch immer alle, alle Organe mit angucke. Das dauert vielleicht so drei vier Minuten würde ich sagen und dann springe ich auf diese...	
2.214	JM	Was genau müssen Sie sich da angucken, sorry.	
2.215	R2	Die Leber, ob da irgendwie Wasser in der Lunge ist. Dann schaue ich mir an ob das die Oberbauchorgane die mit abgebildet sind, ob die irgendwelche Pathologien haben. Dann schaue ich mir die Schilddrüse an und schaue ich mir die Knochen einmal an.	
2.216	JM	Das heißt Sie gehen alle Bilder auf Unregelmäßigkeiten durch, richtig?	
2.217	R2	Ja genau. Also die Wirbelkörperfrakturen und so was, das macht das System halt nicht. Aber das ist natürlich eine Information, die ich habe und die kann ich ja nicht einfach übersehen irgendwie das geht ja nicht.	OQ
2.218	JM	Also das machen Sie, das machen sie, aber das macht das System noch nicht.	
2.219	R2	Das macht das System noch nicht. Das System misst in diesem Bild nur den Calcium Score.	OQ
2.220	JM	Das heißt das wäre eigentlich eine Funktion die Sie sich noch wünschen würden, könnte man so sagen, oder?	
2.221	R2	Ich weiß gar nicht, ob das System das kann ehrlich gesagt. Das ist schon eine ganz komplexe Sache. Es muss ja mehrere Organsysteme angucken. Also ich guck da ja tatsächlich, ich gucke die Knochen, die Weichteile, also das ist, und da sind Information drin, zum Beispiel schaue ich, ob der Patient einen Gallenstau hat, wenn ich das sehe, oder ich schaue, ob der Patient, ob der Magen, ob der einen Gleitmagen hat. Das ist auch was relativ häufig ist da, was die auch oft nicht wissen, die Leute. Was man dann wirklich nur in so einer Untersuchung sieht und das wäre ja	OQ

		dumm, das dann nicht hinzuschreiben. Solche Dinge schaue ich an. Das dauert so drei, vier Minuten, dann schalte ich wieder um und das dauert jetzt richtig eins, zwei Minuten und dann gucke ich mir die Lunge einmal an und wenn die in Ordnung ist, dann geht das super fit.	
2.222	JM	Mit dem Umschalten, was meinen Sie noch mit dem Umschalten?	
2.223	R2	Ich muss ja immer von dem weichen Kernel auf den harten Kernel schalten, um mir die Lunge anzuschauen. Die Lunge muss ich immer in so einem hochauflösenden Modus anschauen und das dieses Umspringen dauert bei dem System aus irgendeinem Grund.	
2.224	JM	Und diese Zeit ist wirklich einfach nur Warten, richtig? Wo Sie nichts anderes machen können?	
2.225	R2	Genau und dann, wenn da alles gut ist, brauche ich höchstens noch zwei Minuten, sage ich mal und wenn das dann, wenn da was ist und ich dann irgendwie was ändern muss, also dann ist es davon abhängig. Aber wenn ich das einfach nur einmal drauf gucke, einmal runter gucke, andere Seite und die CAD anklicke und die findet nichts, dann bin ich auch durch und dann muss ich noch anschauen, stimmt nicht es gibt noch was und ich muss das noch bestätigen, die Messwerte ist. Er sagt mir immer Lungenemphysem und Calcium Score muss ich noch mal EXIST oder nicht. Das muss ich immer noch mal bestätigen. Dann sieht es CONFIRMED CONFIRMED, sonst meckert er und dann kann ich den Patienten [inaudible], dann sage ich bitte einen Report erstellen. Und tatsächlich ist es in unserem Zentrum so, diesen Report, den drucken wir nicht direkt aus, sondern ich speichere ihn erst mal zwischen und die werden dann in einem großen Batch ausgedruckt. Das ist dann noch mal, ich mache dann so einen Download einfach in einen Folder auf dem Computer und sammle die da unter dem Datum, meistens unter dem ich das, unter dem sie untersucht worden sind. Und dann wird das alles von einer Study Nurse ausgedruckt, hinterher. Das ist nicht mehr unser Job.	
2.226	JM	Sie meinten, jetzt bei einem komplexen Fall 20 Minuten dauern. Das dauert einfach nur, weil Sie, da fällt die Zeit mehr auf die Lungen...	

2.227	R2	Weil ich da diese ganzen Lungenherde angucken muss, weil ich vielleicht die ganzen Stents im Calcium Score rausnehmen muss, weil ich die Lungensegmente nicht stimmen, weil der irgendwelche Lappen nicht erkennt und ich das händisch einzeichnen muss, was auch ein großer Spaß ist, irgendwie. Also das müsste ich wahrscheinlich gar nicht unbedingt machen, aber ich meine, man will das ja so anwenden, wie das angewendet wird, also das da alles richtig ist.	PEUT
2.228	JH	Wie oft kommt es so vor, dass Sie das Gefühl haben, dass aufgrund von Fehlentscheidungen der AI es länger dauert, als es eigentlich müsste?	
2.229	R2	Bei jedem zehnten Patienten vielleicht. Also das kommt immer ein bisschen drauf an. Also die Lungenherde, das ist einfach, die soll die AI ja erkennen, sozusagen. Das ist nur für mich mehr Arbeit, aber diese Lungensegmentation, dass das mal falsch läuft, das ist relativ viel Arbeit. Aber so bei jedem zehnten Patienten, und das mit den Bypässen oder mit den Stents, die dann falsch erkannt werden, das kommt dann auch noch mal so bei jedem, das ist vielleicht so, ja, also ich schätze mal, im Schnitt bei jedem zehnten Patienten habe ich da mehr Arbeit.	PU PT
2.230	JM	Was ich mich gerade auch noch einmal frage. Sie meinten ja, Sie hatten damals vor der ganzen Studie einen Vorbereitungskurs, aber der war allgemein dazu, was es schon für Studien gab.	
2.231	R2	Ja was es für Studien gab. Auch so ein bisschen die Pitfalls, auch so ein bisschen darauf, was muss man beachten, wenn man zum Beispiel eine bestimmte Art von Herd sieht, was ist da ein Zeichen, dass es bösartig sein könnte? Was, also es gibt ja auch diese Klassifikation 4X. Das heißt, dass Sie ein Herd haben, der ähm, der an und für sich von der Größe her Sie schon auch beunruhigt, aber der vor allen Dingen klare Malignitätskriterien aufweist, die werden dann noch mal erklärt. Das waren also ganz unterschiedliche Vorträge. Es ging auch in einem um diese, dass man die Patienten nicht übermäßig beunruhigen darf irgendwie. Das wird dann auch sehr oft mit Daten hinterlegt, irgendwie, wo ich dann manchmal so ein bisschen abschalte.	
2.232	JM	Wurden Sie auch irgendwie auf die Nutzung des Systems vorbereitet oder wurden Sie da einfach so reingeschmissen, dann eines Tages?	

2.233	R2	Nee, nee, die Benutzung des Systems. Da kam da ne nette Dame mit irgendwie. Also es kamen dann auch, aus [country, anonymized] waren einige Leute da. Also das war dann so eine Art kleiner Empfang, denn dann wurde das System geöffnet, man bekam ein Passwort und dann wurden so die Basic Features erklärt sag ich mal. Das ist natürlich dann einmal irgendwie vorher passiert und dann nie wieder und dann diese Frau [name, anonymized] die uns dann da eingearbeitet hat, die kam dann noch ein paar Mal irgendwie.	RU
2.234	JM	Vor Ort dann zu Ihnen hin?	
2.235	R2	Zu irgendwelchen, wenn das dann noch mal neu, wenn es neue Features gab zum Beispiel, das kam am Anfang noch ein paar neue Features dazu, dass sie das noch mal erklärt hat irgendwie	RU
2.236	JH	Also wenn, wenn es Updates gibt, werden Sie auch darüber unterrichtet.	
2.237	R2	Dann wurden wir noch mal, es gibt auch so ein Bedienungsmanuell, aber ich muss ehrlich sagen, ich habe mir das am Anfang einmal haben wir uns das alle selber uns hingesetzt. Alles, was wir davon noch erinnert haben, haben wir da mal so reingeschmissen und so ein bisschen hin und her probiert und im Grunde ja, es ist halt wichtig, dass so was sich intuitiv bedienen lässt, sage ich mal, man kann nicht jedes Mal.	PEUT
2.238	JM	Aber Sie würden schon sagen, allgemein ist das intuitiv, oder?	
2.239	R2	Ja ja.	
2.240	JM	Wenn jetzt ein neuer Radiologe anfängt der diese Einführung nicht hatte und es auch plötzlich benutzen soll, hätte der Probleme damit reinzukommen oder würde der es innerhalb einer Woche lernen?	
2.241	R2	Nein, das lernt man innerhalb von einer Woche, schneller glaube ich.	PEUT
2.242	JM	Okay. Allgemein, bevor jetzt Lungenkrebs Screening, sage ich mal bei Ihnen in [city, anonymized] eingesetzt werden im großen Stil dann national, deutschlandweit. Also wenn Lungenkrebs Screening jetzt deutschlandweit umgesetzt wird und es auch jetzt wirklich ernst wird in [city, anonymized] und nicht mehr im Rahmen einer Studie ist und dann wahrscheinlich auch noch mal mehr Patienten kommen. Denken Sie da kommen nochmal irgendwelche Probleme auf Sie zu oder sollte das eigentlich einfach ohne Probleme gut funktionieren können?	

2.243	R2	Das ist ja immer die Frage, wie viel Geld da als Personal investiert wird. Wenn das irgendwie gut organisiert ist, ist das gar kein Problem. Ich glaube nicht, dass da Problem auf uns zukommen. Das ist dann, ach ja was ich noch nicht erwähnt habe, dass sollte ich das tun ist, dass wir Patienten, die einen blöden Befund haben und die noch mal wieder kommen müssen. Die müssen wir anrufen. Und das ist dann natürlich wirklich auch ein zeitlicher Mehraufwand, der auch so ein bisschen Fingerspitzengefühl erfordert. Da können Sie ja nicht einfach anrufen und sagen, bei Ihnen war was komisch, kommen Sie doch bitte in drei Monaten wieder, sondern die haben dann ein bisschen Erklärungsbedarf. Sind beunruhigt irgendwie und das macht auch [name, anonymized] sehr viel. Also wenn Sie mit der sprechen und ja ich sag mal so was noch zusätzlich einzuflechten. Der ganze Rest, der so ein bisschen, so ein bisschen, sag ich mal, so ein bisschen schematisch abläuft, ist ja, wenn man die Manpower hat, überhaupt kein Problem. Also da setzt man sich dann hin, macht das und dann kann man da auch, also nachdem, wie viele Leute man da einstellt, kann man das auch gut durchführen. Man braucht halt, wie gesagt, man muss sich immer klarmachen, vielleicht kommen wir dann auf sechs oder sieben hoch. Aber ich sag mal, sieben CTs sind eine Stunde, sind eine Stunde ein Radiologe beschäftigt. Im Augenblick jedenfalls.	
2.244	JH	Aber zum Beispiel diese Telefonate, die Sie meinen, ist so etwas feinfühliges und sensibles, das sollte schon ein Mensch machen. Da sollte nicht, irgendwie, eine automatisierte E-Mail rausgehen.	
2.245	R2	Ja, das muss ein Mensch machen. Die Frage ist halt, ob man das dann dem Hausarzt überlässt. Das wäre natürlich eine Möglichkeit. Im Augenblick machen wir das, das ist natürlich jetzt auch eine Studie, da haben die Hausärtzte nichts mit zu tun etc. Wenn das eine Sache ist, die im Screening, dann vielleicht ins Kassensystem irgendwie implementiert wird oder so, dass man dann sagt, gut der Hausarzt bekommt die Information vom Computer vielleicht und ruft dann den Patienten noch mal an, irgendwie aber die Frage ist, genau da braucht man schon eine ganz gute Patientenführung. Das ist auch nicht so einfach, wenn man die Leute noch nie gesehen hat und wir sehen die Leute ja gar nicht in dem Augenblick. Die kommen ja zu einem Allgemeinarzt und bekommen da ihr Erstgespräch sozusagen, werden dann von einer MTA im Truck untersucht und der Radiologe sieht nur die Bilder, aber ruft da hinterher an.	
2.246	JM	Jan hast du noch mehr?	

2.247	JH	Ich glaube, mit unserem Interviewguide an sich sind wir durch. Fehlen Ihnen noch irgendwelche Themen, die wir vielleicht noch nicht angesprochen haben oder die Sie vielleicht noch ein biss-chen ausführen würden?	
2.248	R2	Die Finanzierung von solchen System, aber da bin ich wirklich nicht der richtige Ansprechpartner.	
2.249	JM	Ansonsten würden wir das Interview offiziell hier mal mit dem Recording auch stoppen.	
2.250	JH	Haben Sie sonst irgendwelche Fragen an uns?	
2.251	R2	Nein.	

Appendix F - Interview 3

Duration: 37 minutes

Date: May 3, 2022

Interviewee: R3

English

Row	Speaker	Transcribed text	Codes
3.1	JM	And then we'll go straight to the first question, namely, could you tell us briefly, in two or three sentences, the extent to which you have been in contact with Lung Cancer Screening so far?	
3.2	R3	Yes, a large study is currently underway in Germany, the HANSE study in three centers in northern Germany, in Hanover, Großhansdorf near Hamburg, and in Lübeck. And here we are one of the study radiologies involved in this project. This means that we will jointly examine a total of 5000 patients over the next two years. And here in [city, anonymized] so far, I think, we've included a good 1000 patients in this study.	
3.3	JM	How do you think lung cancer screening will actually be introduced in Germany at some point? Where will it happen then? At your institute too, for example? Or is this generally done via university clinics or special practices? Or how do you think that will be handled?	
3.4	R3	So I won't be able to decide that, but politicians will certainly give us very precise guidelines on how this should be organised. What I would like for lung cancer screening to remain linked to lung cancer centers, that you also have the expertise for the entire process chain, not only take the pictures, but also interpret the images, discuss them in an interdisciplinary manner and also draw the right conclusions from them. However, this will not achieve nationwide supply in Germany. This means that it will presumably amount to such a system, where there will be satellites to every site that has a lung cancer center, where the examinations will also be carried out, and then the findings that are being discussed will then have to be viewed centrally. It will probably be such a solution at the end of the day.	
3.5	JH	Do you think this process will work well, or are you already seeing any issues that might occur?	

3.6	R3	Oh, no. I think it will. This will work well because we have simply learned in recent years how to handle images, how to send pictures, how to communicate with colleagues via video conferences. I believe that is now the case, the technical requirements are in place and we have become accustomed to working in this way.	
3.7	JH	Yes. Can you possibly lose another sentence or two about yourself, what is your day-to-day work and what contact or role do you have in lung cancer screening?	
3.8	R3	So my ex-boss always said he did desk medicine and that's ultimately my job. I have to say that I am hardly involved in the process itself of dealing with the cases. I was involved, of course, in the conception of the study, the development of the idea, and the implementation. To create the conditions here in [city, anonymized] for this study to be carried out and what I do from time to time if my calendar allows it. I'm going into this appointment, which is called the Nodule Conference, where the unusual findings are looked at and discussed in an interdisciplinary way, how high we assess the risk for the patient, and what the further procedure should be. But otherwise, I'm hardly involved in practical work.	
3.9	JM	Exactly where you just mentioned the nodules. I think several things are done as part of lung cancer screening. Can you explain again in detail what a patient goes through or what is being watched when they come to you for lung cancer screening.	
3.10	R3	Yes, I think that we need to differentiate a little bit between what we are now doing in the context of this study. Yes, which may go beyond what will also take place in lung cancer screening afterward. So if you look at the lung cancer screening now and limit it to a minimal program, then the patient would have to receive a CT scan. That means he lies down on the examination couch once. Once for three seconds, the thorax is scanned, the images are viewed, evaluated and then a decision is made as to whether there will be a routine follow-up or whether something is now required. Of course, the data that is acquired offers significantly more information, i.e. what we are doing, for example, as part of the HANSE study. Many people who smoke not only get carcinoma, but they also damage their lungs as a whole. This means that they develop pulmonary emphysema, i.e. the pulmonary alveoli are destroyed. This can also be recorded very well in CT and quantified. That's what we do. And of course, all people know that a major cause of this development of lung cancer and also pulmonary emphysema is smoking. That's why we only include high-risk patients with pack years. And what people all develop is, of course, vascular diseases, which is why, as part of the	

		HANSE study, we are now looking at the calcium on the vessels, for example, in order to then, for example, the risk of a heart attack.	
3.11	JM	Yes. That could all make sense so far, but you don't think it will all be looked at in the lung cancer screening when it comes?	
3.12	R3	Yes, it will be looked at. You have to look at what is then evaluated in a structured manner. Um, it's always about that, things like that don't happen by themselves, require a bit of work. And, of course, it must then also be financed and, of course, that is always the case with these processes, you always have to look carefully at what evidence is there for? Because nothing is introduced for which there is no evidence. And there is now evidence for lung cancer screening. I don't know whether it makes sense to quantify emphysema. Whether there is really evidence that it benefits patients at the end of the day. It is, therefore, necessary to look at what is then demanded and actually carried out afterward.	
3.13	JH	Can you explain the process of radiologists again? So if they've done a CT scan now, how do they evaluate these images. Is there a radiologist involved in this? Do several do it? And how do they then work together with a computer system?	
3.14	R3	So the pictures are taken and then sent to the corresponding evaluation workstation. And an AI algorithm immediately runs over these images, which detects exactly the steps that we have just described, the round foci, and also carries out the emphysema quantification and coronary calcification virtually automatically. But of course, that's not enough. The algorithms are pretty good, but they're not one hundred percent perfect. This means that the images are always reviewed by a radiologist and compared with what the computer has found. And in the case of the pathological findings, of course, also in the second diagnosis and then it goes into this, into this tumor board. So I can also offer you, if you are interested, to drop by here for half a day when the truck is back	OQ JR

		on site, then you can simply look over the shoulders of the colleagues who are evaluating the studies here.	
3.15	JM	When would that be in theory?	
3.16	R3	I don't even know now, so it's always two weeks in Hanover, two weeks in Groß-Hansdorf, and two weeks in Lübeck. So he has to be back here in six weeks at the latest. I'm not entirely sure whether that is, but somehow in two or three weeks, this would have to be done again here on-site if there had to be another investigation here on site.	
3.17	JM	Pretty much that's where we have to hand in our master's thesis. So thank you for the offer, but we are currently on a very narrow path.	
3.18	R3	It seems that way.	
3.19	JM	Are the CT scans always evaluated directly as soon as the patient has passed through or are they only evaluated later at another point in time, or...?	
3.20	R3	No, we are already trying to evaluate things on a daily basis, but as a huge mountain of investigations is simply piling up and that is quite some work. As part of the HANSE study, we now perform about 25 such CTs per day for the patients that the truck is here on site.	
3.21	JM	Yes, yes.	
3.22	R3	And that costs a good half day of work, at least.	
3.23	JM	Yes, and then a radiologist is simply assigned, who then gets a picture of it from time to time and does that?	
3.24	R3	Exactly.	
3.25	JM	Does the radiologist then only do that or does something else in parallel?	

3.26	R3	So, because additional tasks were added in the context of the study, they are actually quite busy with the HANSE study. In other words, in the two weeks that the truck is on-site because the patients must then also be called, they will be informed of their examination results. There is also a certain amount of organizational tasks involved. Well, that's fine, they're moderately busy then.	
3.27	JM	Yes, if lung cancer screening is introduced at some point, it sounds as if it would really take radiologists the whole day, and there are really a lot of smokers in Germany who might come along. Do you think that would just work? Or are there not enough radiologists? Or do you think it could work somehow?	
3.28	R3	So the thing is, I've already said that. The time that is now invested in the study is certainly higher than what you will do afterward in the clinical routine. There's no question about that. And then you have to take a look at how you really optimize the process for throughput, which is not the case now in the study. Um, then I would think that you can certainly do 50 such CTs a day. Um, if everything goes perfectly and the radiologists are really trained, maybe even more so. You might even get 100 CTs that you can do a day. But then, you really have square eyes. If you've watched 100 CTs then, um, I don't think much more is possible.	
3.29	JH	Yes. Let's get back to these AI systems. What exactly are their functions? You already said that they can recognize the nodules and perhaps also say whether they are benign or malignant. Do they have any other functions that support the radiologists?	
3.30	R3	No, after all, when you look at a thorax CT like that, when you look at a single layer, there are always dots and lines. Yes, the strokes are vessels hit lengthwise. There are two options for the points that you see in the picture. Either it's a sphere and then it can be a tumor, or it's an orthogonally struck vessel. And that's what you have to differentiate between. And if you want to do that on 500 or 800 layers, where there are 100 such structures in each case, that is complicated and a computer can do that pretty perfectly and [inaudible].	
3.31	JH	This means that without this AI, these lung cancer screening processes would actually not be as feasible in terms of effort?	
3.32	R3	Yes, I think that then you wouldn't have to think about whether you can somehow get towards 100 CTs. So that's only possible with IT support.	OQ PU PT

3.33	JM	You meant at the moment first the CT scan and then immediately the AI also runs over, right? And then together with the AI is that already directly evaluated by the radiologists or do you also look at each again without the AI evaluation?	
3.34	R3	So we're already watching it at the moment without the support of Artificial Intelligence, that's how we've all grown up. And the fact that the technology has turned out good, but we do not yet trust that alone one hundred percent. Radiologist interaction is still required.	
3.35	JM	Yes, how can I imagine that? So on one screen without AI, on another screen with AI, right?	
3.36	R3	So I don't know how people really like to do it now when I look at such cases in clinical routine, uh, I usually look over without artificial intelligence to make an impression for myself, to see where do I end up? And then I look at what artificial intelligence has found. Does it match with what I found? And in most cases, I throw out findings that artificial intelligence has marked for me as things that I should look at because I think they are irrelevant.	SV OQ JR PU
3.37	JH	Is there actually...	
3.38	R3	So, much more sensitive, but not super specific yet.	OQ JR PU
3.39	JH	Yes. Is there actually a prescribed process as to whether you should use this AI or if you use it, how to use it? Or is that, as it were, left to radiologists themselves?	
3.40	R3	So far, that is up to us. Whether it will then be determined in the context of Lung Cancer Screening, I am also curious to see whether a clear regulation will be made. But that would be something that is really new, that it is then actually demanded and is also demanded in a certain format.	SV
3.41	JM	Yes, do you know, as is the case with radiologists on your premises, do they actually use AI, are they happy with it, or are there any who say that I am a hard-established person? I see that better than AI.	
3.42	R3	No, I think everyone has now realized this, especially this topic of pulmonary artery detection. That is simply good. As I said, you have to sort out the false positives. But it is a real support and shortens the reporting process. There is absolutely no question about that. And other things, i.e. lung emphysema quantification, to do that manually, no one will even get the idea. Because that will take an hour to paint its contours somehow. This is only	PU RD JR PT

		possible with automatic segmentation, which now always works with AI support.	
3.43	JM	Yes. You said the system is very good, but not perfect. What would you say? Are these the points of contact with AI at the moment where there are problems or where it is not perfect?	
3.44	R3	The false positives that you have to sort out.	
3.45	JM	That is the only problem? Otherwise, cooperation with radiologists also always works super well, doesn't it?	
3.46	R3	Yes, so that's this interaction. Of course, there are still things where you might be able to develop more modern user interfaces, but basically, that works well. It's fast, it's available everywhere. So we're on a super good way there.	PEUT
3.47	JM	But what do you mean about the user interface? Was it said somehow that it wasn't so...	
3.48	R3	Yes, so how soon can I accept whether the nodule is benign or malignant? Do I need it there..? Do I have different levels there that I can set? Can I generate a structured report from this right away? And questions like that. There is certainly still room for improvement. But the software itself works very well.	PEUT
3.49	JM	Okay, yes.	
3.50	R3	You have to say that quite honestly, too, and this also applies to many manufacturers. There are a whole lot of systems that do something like this. We have opted for a system from a company from [country, anonymized] for the HANSE study.	
3.51	JM	[AI software, anonymized], yes? We've already read into that, too.	
3.52	R3	[AI software, anonymized], that's right. Because, on the one hand, they have experience from the program in [country, anonymized] and they were also the most flexible to now respond to our needs in the context of the study.	RU

3.53	JH	Yes, great. Then we would have a few more questions about the characteristics of the system or the quality. Do the systems provide you with all the information you would like to have, or would you like to have other information?	
3.54	R3	Nope. In fact, they can now deliver anything that interests us.	OQ JR
3.55	JM	With regard to if a diagnosis has to be made now, then there are certainly various things being looked at. Does the system support the radiologist, I say, at every point of reference? Or is there anything where you could say AI, if it still could, that would be cool too?	
3.56	R3	Yes, I think they simply have more to learn now, to detect the relevant findings, because I mean in the context of this study, um, in about yes, I don't know, just under 10% of the patients or the subjects we see, we have some findings. And then, of course, we also learn from it, and the software can of course learn from that too. What are relevant findings, what were carcinomas in the end, and what turned out to be benign? And with even more help with differentiation and with treatment recommendations or control recommendations, the systems can certainly learn more and get even better.	OQ
3.57	JM	Yes. You said that the user interface may not be perfect, you would say, but in general, the output of the system is understandable, or is it sometimes a bit less understandable?	L PEUT
3.58	R3	No, that, so we understand each other.	
3.59	JM	Okay, do you actually have to be trained to use the system? Or do you say you sit in front of it and that's intuitive like a book or something?	
3.60	R3	No, that is definitely not the case. You would like that for many of these software systems, but that is, the software is not written for me, but it is written for worldwide use. And there are so many colleagues who have different opinions, have different requirements, and want the process to be different. That is why it is usually the case that such software offers an incredible number of possibilities and you can also achieve a result in completely different ways. In order to be able to simply fulfill all the wishes that are placed on the software, if necessary. Yes, yes, I actually always use the same and for me personally, it would be enough if the software could do this one way, and then it would probably be much more intuitive and easier to use. But because it offers so many options, most IT solutions have too many buttons, too many options, and that's why you simply need an introduction.	PEUT E

3.61	JH	So was there some kind of workshop or is there a manual where you can always look things up? Or is there also some kind of IT support for the radiologists if you get stuck somehow?	
3.62	R3	There was. There was a training on the software where everything was shown what it can do, how it works, and also how you can set up your own workflow there.	PEUT RU
3.63	JM	Have there actually been any technical problems with the system or did it always run cleanly? Has never crashed?	
3.64	R3	No, we've already had a few technical problems, that because the data is also stored centrally and then data transfers, which can get stuck in many places, especially the connection from the truck, where there only has to be a data socket out there, so to speak, in the courtyard. And yes, there were a few such a few hick-ups, but on the whole, it worked.	SR
3.65	JM	So just small stumbles or something worse or something or something?	
3.66	R3	No, simply that data could not be sent. So no.	
3.67	JM	However, this could always be solved quickly on site. There was also enough support there?	
3.68	R3	Yes.	SR RU
3.69	JM	Okay, yes good. How would you rate the changes in the diagnostic output that happens as a result of AI support? So I'd say it's definitely going to be faster, right?	
3.70	R3	Yes, definitely.	RD
3.71	JM	Would you also say the quality is improving too, right? Of the diagnosis.	
3.72	R3	Yes, I think so. It will be less overlooked. That is certain. But the decisive thing is, what do I do now with the findings? And is there still a lot of human intelligence in interdisciplinary conferences?	PU
3.73	JM	Yes, but yes. Let's say, even if it would take longer. A finding that a person would only look at alone and then process, against the current process, where a radiologist might need another ten minutes with the support of an AI system. How would the two findings differ? Would you say that the person with AI recognizes more, but also recognizes more where nothing is, right?	

3.74	R3	He recognizes more where nothing is. So the false-positive rate is the big challenge with all these systems. Yes, and it is, they are support systems and they won't replace treatment.	PU
3.75	JH	Yes. Does it mean additional work for you? You will then have these false-positive results. What happens then? Do the patients then have to come back earlier for a second CT scan and then there is, so to speak, additional work for you?	
3.76	R3	No, so we would sort out the false-positives when checking the findings of AI. In other words, this does not actually result in any disadvantage for patients.	
3.77	JH	Yes, but does it happen that some of them would actually take longer due to AI than without it because they simply have these false positives that they first have to sort out?	
3.78	R3	Yes, that can happen when the AI finds an infinite number of findings. Yes, that is rare in this population that we have here. But if you imagine that you have other diseases that cause nodules in the lungs and that may also cause diffuse nodules, then the AI will of course find an insane number of findings and then you can almost forget about using it. Because if it shows me 100 or 200 findings, then it becomes really tedious and then I have to look manually, of course. But these are really absolute exceptions.	PU
3.79	JM	Yes. What would you say are the biggest challenges or problems in general when working with this AI system? Or the AI systems that support that.	
3.80	R3	Yes, I think what you would still wish for is that you, that is, we still live very much with these textual findings. That is the way we communicate. And, um, that's usually free text. It is a great challenge to immediately make a structured text finding out of these AI findings, but that will help significantly further.	OQ JR PU
3.81	JM	AI does not actually support it already by providing such a structure for the text, i.e. preformulating the text?	
3.82	R3	Yes, such systems exist, but there is always this compromise between, on the one hand, I want to be complete, and on the other hand, it should be quick. This is contradictory and flexibility is lost a bit.	OQ JR PU
3.83	JM	Yes. Yes.	

3.84	R3	And also the weighting when you have to cling too much to a structure. A finding must not only be correct, but it must also be clinically exploitable. My ex-boss has written partly, partly one-word findings, normal findings. It is a very difficult finding. You have to be 100% sure that nothing is really relevant at all. But this is of course in terms of speed and clinical usability, the result is great because you can read it in no time. And every doctor who gets hold of the findings knows why I don't have to worry anymore. It's all there. Yes, but that just doesn't work, not in all cases and sometimes it just turns into very complex constructs such findings.	
3.85	JM	Yes	
3.86	R3	Because you also have to weigh up...	
3.87	JM	Does it otherwise cost so much time to write down the texts? So there's definitely potential where AI could support that even better, right? When formulating these texts.	
3.88	R3	Definitely.	
3.89	JM	What else costs a lot of time in the process of diagnosis, the reports?	
3.90	R3	Now in general in radiology?	
3.91	JM	No, specific for lung cancer screening.	
3.92	R3	No, everything is primarily looking through the pictures. That is already the most important factor there.	
3.93	R3	Okay, that takes the most time, but AI supports that as well as it can with highlighting. Okay, okay. And otherwise just the, well, the text that has to be written in the end takes a lot of time. Okay.	
3.94	R3	Exactly.	
3.95	JH	Do you still see any problems with integrating these AI applications into the radiologists' workflow, that they might not be optimized yet, that they could work together even better?	
3.96	R3	It works really, really great. In other words, the data goes to the AI system, is processed, and is also available for review of the results a few minutes later. All of this can be perfectly organized.	JR

3.97	JH	Awesome. Are these systems actually so fast that the radiologist is always busy? Or does it happen that you just sit in front of the computer for a minute because the system simply loads?	
3.98	R3	No. Depending on what, what they are processing and how many images are in it, it can take five minutes for the test to go through all of this. But the fact that we always have a bit of delay with the image acquisition and really not when it is acquired in the scanner means that there are no waiting times.	
3.99	JM	Exactly. You said the AI system runs immediately. I'll say this will be watched 10 minutes later, but five minutes before that, it's just gone through, right?	
3.100	R3	Yes, exactly.	
3.101	JM	I believe that most of them are looking forward... Your colleagues and radiologists in general, who are already looking forward to using AI in the context of lung cancer screening, so they are all looking forward to it as well. No one is afraid that someone will lose their job due to AI, right?	SV
3.102	R3	No.	SV
3.103	JM	Okay, you also said that you think that these are clear support systems. Don't you see the chance at some point that AI could really take over the entire job or actually create a report from zero to a hundred and be pretty good at it?	
3.104	R3	So, um, for lung cancer screening, I can imagine that.	
3.105	JM	That AI can completely take over, possibly?	
3.106	R3	Yeah, well, I'm still going to have to pick out a few false-positives. I am quite sure of that, but basically, you can certainly push this even further, because if you now, I see the chances of using such systems, when it is actually a simple decision, when I have to decide yes/no for cancer. Such a system can certainly learn that at some point. But if you now have as another extreme example, a patient who has already been operated on his stomach five times and where nothing is the same as it was originally and you then put such an examination into a system, a correct result will never be able to come out. It simply lacks the additional information that we use, what has been done. How has it been done, how is the patient and so on. I don't see that we'll get there at some point, that a system all these factors that we acquire along the way, um, just from the pictures. Because that's just the way it is! The information is not only in the pictures, at least looking at the entire	

		area of radiology. Now with such a question as lung cancer screening, the information is already in the pictures, we know everything, it is a high-risk collective, they all smoked and so on.	
3.107	JM	Nevertheless, we have already heard that it would be really cool to have a system that still says a bit more about how much the patient has smoked and how old he is so that you really have a risk profile like that for the patient. Could that help in any way? With the findings, too.	
3.108	R3	So, uh, I don't think so, since it doesn't really matter if he has 30 or 50 pack-years, who only has 30 pack-years, if he's had bad luck, he has worse emphysema or a higher risk than the one with the 50 pack-years. So that's not entirely linear now somehow that you could make that dependent on each other. I think I know risk factors, yes, that's important, but I don't think it makes any sense or doesn't really help you.	OQ
3.109	JM	Do you actually anticipate problems in some way before lung cancer screening is implemented in Germany and that is actually carried out at your clinic in this context, let me say when the number of patients increases significantly again? Do you think that you will have any problems again or do you think you are actually prepared for this now?	
3.110	R3	So technically I think you can do that. There is absolutely no question about that. Politically and also when it comes to the question of who pays for it and at what price? There will be even longer discussions about this. I'm quite sure of that.	
3.111	JH	Okay. yeah. yeah.	
3.112	R3	And also the question of who does it, who assigns it, who checks the risk factors, and so on. Things like this, which payers also demand that it be combined with a smoking cessation program, which of course also makes perfect sense because we don't want to screen people so that they can continue smoking, but rather also do something for patient health and combinations like that, that's all in the overall package have to be in there.	

3.113	JH	Yes, thank you very much. I think we're finished with our questionnaire by now. Do you have another question, Joachim?	
3.114	JM	Me? No.	
3.115	JH	Do you have any other questions on the subject that you would like to have addressed? What has come up short?	
3.116	R3	Nope, I'm curious what you write in your master's thesis.	
3.117	JM	And that's exactly what we wanted to ask if they would be willing to do member checking with us, that is, as soon as we have roughly an initial analysis of this conversation or what our findings from the conversation are, we'll send you a quick email and you say yes, that's exactly what I said, that's exactly what I meant. Or you write like this, no, you completely misunderstood me, that meant otherwise.	
3.118	R3	Yes, you are welcome to do that.	
3.119	JH	Great, thank you very much. Thanks again for taking the time, it really helped a lot, and we gained lots of cool insights.	
3.120	JM	And again sorry that we're going over a bit here during the conversation. I hope you still have a full morning and can enjoy the day.	
3.121	R3	I wish you the same. Have a nice day and good luck with your writing. Hurry up.	
3.122	JH	Thank you. Bye.	

German

Row	Speaker	Transcribed text	Codes
3.1	JM	Und dann gehen wir auch mal direkt zur ersten Frage über, nämlich genau. Können Sie uns einmal kurz in zwei, drei Sätzen sagen, inwiefern Sie bis jetzt mit Lung Cancer Screening im Kontakt waren?	

3.2	R3	Ja, es läuft ja in Deutschland im Moment eine große Studie, die HANSE Studie in drei Zentren in Norddeutschland, in Hannover, Großhansdorf in der nähe von Hamburg und in Lübeck. Und da sind wir eine der Studienradiologien, die an diesem Projekt beteiligt sind. Das heißt, wir werden gemeinsam in den nächsten zwei Jahren insgesamt 5000 Patienten untersuchen. Und hier in [city, anonymized] haben wir bisher, ich glaube, gute 1000 Patienten in diese Studie eingeschlossen.	
3.3	JM	Wie denken Sie, wenn dann irgendwann Lung Cancer Screening auch wirklich in Deutschland eingeführt wird? Wo wird es dann passieren? Auch bei Ihrem Institut zum Beispiel? Oder wird das allgemein über Unikliniken gemacht oder spezielle Praxen? Oder wie denken Sie, wird das gehandelt werden?	
3.4	R3	Also das werde ich nicht entscheiden können, aber die Politik wird uns da sicherlich sehr genaue Vorgaben machen, wie das organisiert werden sollte. Was ich mir wünschen würde, dass das Lungenkrebs Screening angebunden bleibt an Lungenkrebs Zentren, dass man auch die Expertise hat für die gesamte Prozesskette nicht nur die Bilder machen, sondern halt eben auch die Bilder interpretieren, interdisziplinär besprechen und auch die richtigen Konsequenzen daraus ziehen. Damit wird man aber keine flächendeckende Versorgung in Deutschland hinbekommen. Das heißt, es wird vermutlich auf so ein System hinauslaufen, wo es zu jedem Standort, der ein Lungenkrebs Zentrum hat, Satelliten geben wird, wo die Untersuchungen auch durchgeführt werden und dann aber die Befunde, die diskutiert werden, müssen dann zentral angeschaut werden. So eine Lösung wird das vermutlich am Ende des Tages werden.	
3.5	JH	Denken Sie, dieser Prozess wird so gut funktionieren oder sehen Sie da schon irgendwelche Probleme, die auftreten könnten?	
3.6	R3	Oh nein. Ich glaube, das wird. Das wird gut funktionieren, weil wir einfach in den letzten Jahren gelernt haben, wie man mit Bildern umgehen kann, wie man Bilder versenden kann, wie man über Videokonferenzen auch mit Kollegen kommunizieren kann. Ich glaube, das ist inzwischen, die technischen Voraussetzungen sind da und wir haben uns daran gewöhnt, so zu arbeiten.	
3.7	JH	Ja. Können Sie vieleicht noch einmal eins, zwei Sätze zu Ihrer Person verlieren, was Ihr Arbeitsalltag ist und welchen Kontakt oder welche Rolle Sie genau beim Lungenkrebs Screening haben?	

3.8	R3	Also mein Ex Chef hat immer gesagt er macht Schreibtischmedizin und das ist auch letztendlich mein Job. Ich bin in die in den Prozess selber in das Abarbeiten der Fälle kaum eingebunden, muss ich sagen. Ich habe natürlich, war ich beteiligt an der Konzeption der Studie, an der Entwicklung der Idee, an der Umsetzung. Hier in [city, anonymized] die Voraussetzungen zu schaffen, dass diese Studie durchgeführt werden kann und was ich ab und zu mache, wenn es mein Kalender zulässt. Ich gehe in diesen Termin, der bei uns Knötchen Konferenz heißt, da, wo die auffälligen Befunde angeschaut werden und interdisziplinär diskutiert werden, wie hoch wir das Risiko für den Patienten einschätzen und wie das weitere Vorgehen sein sollte. Aber sonst, in die praktische Arbeit bin ich kaum eingebunden.	
3.9	JM	Genau, wo Sie gerade auch die Knötchen angesprochen haben. Es wird ja im Rahmen des Lungen Cancer Screenings glaube ich auch mehrere Sachen gemacht. Können Sie es noch mal ganz genau erklären, was ein Patient da durchläuft oder was bei dem alles angeguckt wird, wenn der für das Lung Cancer Screening zu Ihnen kommt.	
3.10	R3	Ja, ich glaube, da muss man so ein bisschen differenzieren zwischen dem, was wir jetzt im Rahmen dieser Studie machen. Ja, was vielleicht über das hinausgeht, was nachher auch im Lungenkrebs Screening stattfinden wird. Also wenn man jetzt das eine Lungenkrebs Screening sich anschaut und es auf ein Minimalprogramm begrenzen würde, dann müsste der Patient die CT Untersuchung erhalten. Das heißt er legt sich einmal auf die Untersuchungsliege. Einmal drei Sekunden wird der Thorax gescannt, die Bilder werden angeschaut, ausgewertet und dann wird eine Entscheidung getroffen, ob es nun eine Routine Follow-Up geben wird oder ob jetzt irgendwelche Sachen erforderlich sind. Natürlich bieten die Daten, die dabei akquiriert werden, noch deutlich mehr Informationen, also was wir zum Beispiel im Rahmen der HANSE Studie auch machen. Viele Leute, die rauchen, kriegen nicht nur ein Karzinom, sondern sie schädigen auch ihre Lunge insgesamt. Das heißt sie entwickeln ein Lungenemphysem, also die Lungenbläschen werden zerstört. Das kann man auch im CT sehr gut erfassen und auch quantifizieren. Das tun wir. Und natürlich wissen alle Leute, dass eine Hauptursache für diese Entwicklung von Lungenkarzinomen und auch dem Lungenemphysem ist das Rauchen. Darum schließen wir ja auch nur Risikopatienten mit pack-years ein. Und was die Leute natürlich auch alle entwickeln ist, sind Gefäßerkrankungen und darum im Rahmen der HANSE Studie jetzt zum Beispiel schauen wir uns auch den Kalk an den Gefäßen an, um dann zum Beispiel das Risiko für einen Herzinfarkt.	

3.11	JM	Ja. Das könnte ja auch soweit alles sinnvoll, aber Sie denken nicht, es auch alles im Lung Cancer Screening schlussendlich dann mit angeguckt wird, wenn es kommt?	
3.12	R3	Ja angekuckt wird es werden. Was davon dann strukturiert ausgewertet wird, muss man schauen. Ähm, da geht es natürlich auch immer darum, solche Sachen passieren nicht von alleine, erfordern ein bisschen Arbeitsaufwand. Und natürlich muss es dann auch finanziert sein und man muss natürlich, das ist auch immer bei diesen Prozessen so, es wird immer genau geschaut, wofür gibt es Evidenz? Weil es wird nichts eingeführt, wofür es keine Evidenz gibt. Und für das Lungenkrebs Screening ist jetzt die Evidenz belegt. Ob es Sinn macht, das Emphyse zu quantifizieren, weiß ich nicht. Ob es dafür wirklich Evidenz gibt, dass es Patienten am Ende des Tages nutzt. Darum muss man schauen, was dann nachher gefordert und wirklich durchgeführt wird.	
3.13	JH	Können Sie da noch mal den Prozess der Radiologen an sich erklären? Also wenn die jetzt ein CT durchgeführt haben, wie sie diese Bilder auswerten. Ist da ein Radiologe daran beteiligt? Machen es mehrere? Und wie arbeiten die dann mit einem, mit einem Computersystem zusammen?	
3.14	R3	Also die Bilder werden gemacht und werden dann an die entsprechende Auswertung Workstation geschickt. Und da läuft sofort über diese Bilder ein KI Algorithmus, der genau diese Schritte, die wir eben beschrieben haben, die Rundherde erkennen und auch die Emphysemquantifizierung und auch die Koronarkalkbestimmung quasi automatisch durchführt. Aber natürlich reicht das nicht. Die Algorithmen sind ziemlich gut geworden, aber hundertprozentig perfekt sind sie nicht. Das heißt, es erfolgt immer dann die Durchsicht der Bilder auch von einem Radiologen und der Abgleich mit dem, was der Computer gefunden hat. Und bei den pathologischen Befunden natürlich auch in die Zweitbefundung und dann geht es in dieses, in dieses Tumor Board. Also ich kann Ihnen auch anbieten, wenn Sie Interesse haben, mal hier einen halben Tag vorbeizuschauen, wenn der Truck wieder hier vor Ort ist, dann können Sie den Kollegen, die die Studien hier auswerten, einfach mal über die Schulter schauen.	OQ JR
3.15	JM	Wann wäre das denn theoretisch?	
3.16	R3	Ich weiß es jetzt gar nicht, also der ist ja immer zwei Wochen in Hannover, zwei Wochen in Groß-Hansdorf, zwei Wochen in Lübeck. Also spätestens in sechs Wochen muss er wieder hier sein. Ich bin nicht ganz sicher, ob das, aber jetzt irgendwie in zwei oder drei Wochen müsste das hier vor Ort wieder, müsste es hier vor Ort wieder Untersuchung geben.	

3.17	JM	Ziemlich genau da müssen wir auch unsere Masterarbeit schon abgeben. Also vielen Dank für das Angebot, aber wir sind da zeitlich auf einem sehr schmalen Pfad unterwegs zurzeit.	
3.18	R3	Es scheint so.	
3.19	JM	Werden die CT-Scans denn immer direkt ausgewertet, sobald der Patient durchgelaufen ist oder werden die erst später zu einem anderen Zeitpunkt ausgewertet, oder..?	
3.20	R3	Nein, wir versuchen die Sachen schon tagesaktuell auszuwerten, aber da sich einfach einen riesen Berg an Untersuchungen anhäuft und das ist aber eine ganz ordentliche Arbeit. Wir machen hier im Rahmen der HANSE Studie jetzt bei den Patienten ungefähr 25 solcher CTs pro Tag, den der Truck hier vor Ort ist.	
3.21	JM	Ja, ja.	
3.22	R3	Und das kostet dann schon mal einen guten halben Tag Arbeit, mindestens.	
3.23	JM	Ja, und dann wird einfach ein Radiologe abgestellt, der dann ab und zu mal so ein Bild davon rein bekommt und das macht?	
3.24	R3	Genau.	
3.25	JM	Macht der Radiologe dann nur das oder macht auch was anderes parallel noch?	
3.26	R3	Also, die sind insgesamt, weil ja auch noch Aufgaben im Rahmen der Studie dazugekommen sind, die mit der HANSE Studie eigentlich ganz gut beschäftigt. Also in den zwei Wochen, wo der Truck hier vor Ort ist, weil die Patienten müssen dann ja auch angerufen werden, über ihre Untersuchungsergebnisse informiert werden. Es ist ein gewisser Teil auch an organisatorischen Aufgaben da dran. Also das ist schon, die sind dann gut beschäftigt.	
3.27	JM	Ja. Wenn irgendwann mal Lung Cancer Screening eingeführt wird, das klingt ja so, als würde es schon Radiologen auch wirklich den ganzen Tag beanspruchen, und es gibt ja wirklich sehr viele Raucher in Deutschland, die ja eventuell dann vorbeikommen würden. Denken Sie, das klappt einfach so? Oder fehlen dafür die Radiologen? Oder denken Sie, das könnte irgendwie funktionieren?	

3.28	R3	Also die Sache ist, das habe ich ja eben schon gesagt. Die Zeit, die man jetzt im Rahmen der Studie investiert, ist sicherlich höher als das, was man nachher in der klinischen Routine tun wird. Das ist gar keine Frage. Und man muss dann mal gucken, wenn man den Prozess wirklich auf Durchsatz optimiert, was jetzt im Rahmen der Studie nicht der Fall ist. Ähm, dann würde ich schon denken, dass man 50 solcher CTs ganz sicher am Tag schaffen kann. Ähm, wenn das alles perfekt läuft und auch die Radiologen wirklich trainiert sind, vielleicht sogar noch mehr. Vielleicht kommt man sogar auf 100 CTs, die man am Tag schaffen kann. Aber dann, dann hat man wirklich eckige Augen. Wenn man 100 CTs angeguckt hat dann, ähm, viel mehr geht glaube ich nicht.	
3.29	JH	Ja, ja. Noch mal zurück zu diesem AI Systemen. Was haben die denn genau für Funktionen? Sie meinten schon, die können die Rundherde erkennen und vielleicht auch schon sagen gutartig, bösartig. Haben die noch weitere Funktionen, die die Radiologen unterstützen?	
3.30	R3	Nein, letztendlich geht es ja darum, wenn man sich so ein Thorax CT anguckt, dann gibt es dort, wenn man eine einzelne Schicht sich anschaut, immer Punkte und Striche. Ja, die Striche sind der Länge nach getroffene Gefäße. Bei den Punkten, die man auf dem Bild sieht, gibt es zwei Möglichkeiten. Entweder es ist eine Kugel und dann kann es ein Tumor sein, oder es ist ein orthogonal getroffenes Gefäß. Und das muss man letztendlich unterscheiden. Und wenn man das auf 500 oder 800 Schichten machen will, wo jeweils 100 solche Strukturen sind, das ist aufwendig und das kann ein Computer ziemlich perfekt und [inaudible].	
3.31	JH	Das heißt, ohne diese AI wären diese Lungenkrebs Screening Prozesse auch eigentlich gar nicht so vom Aufwand machbar?	
3.32	R3	Ja, ich glaube, dann bräuchte man nicht drüber nachdenken, ob man irgendwie auch in Richtung 100 CTs kommen kann. Also das das geht nur mit einer IT Unterstützung.	OQ PU PT
3.33	JM	Sie meinten momentan erst der CT Scan und dann läuft direkt die AI auch rüber, oder? Und dann zusammen mit der AI wird das schon direkt von den Radiologen beurteilt oder gucken Sie sich auch jedes noch mal an ohne ohne die AI Auswertung?	
3.34	R3	Also wir gucken es uns schon im Moment auch einmal ohne die Unterstützung der Artificial Intelligence an, das ja, so sind wir noch alle groß geworden. Und, dass die, dass die Technik gut geworden ist, aber so ganz hundertprozentig dem allein trauen tun wir noch nicht. Die Interaktion des Radiologen ist immer noch erforderlich.	

3.35	JM	Ja, wie kann ich mir das vorstellen? So, also auf einem Screen ohne AI, auf einem anderen Screen mit AI, oder?	
3.36	R3	Also ich weiß nicht, wie die Leute es jetzt wirklich am liebsten tun, wenn ich mir in der klinischen Routine solche Fälle anschaue, äh, ich gucke meistens erst einmal ohne Artificial Intelligence rüber, um mir selber einen Eindruck zu machen, um zu gucken wo lande ich? Und dann schaue ich mir an, was die künstliche Intelligenz gefunden hat. Gleiche das mit dem ab, was ich gefunden habe? Und in den meisten Fällen ist es dann so, dass ich Befunde, die die künstliche Intelligenz mir markiert hat, als Sachen, die ich mir angucken soll, rausschmeiße, weil ich der Meinung bin, die sind irrelevant.	SV OQ JR PU
3.37	JH	Gibt es denn eigentlich...	
3.38	R3	Also, deutlich sensitiver, aber noch nicht super spezifisch.	OQ JR PU
3.39	JH	Ja. Gibt es denn eigentlich einen vorgeschriebenen Prozess, ob man diese AI benutzen soll oder wenn man sie benutzt, wie man sie benutzen soll? Oder ist das sozusagen Radiologen selber überlassen?	
3.40	R3	Also das ist uns selber überlassen bisher. Ob es dann im Rahmen der der Lungenkrebs Früherkennung, irgendwie festgelegt werden wird, da bin ich auch mal, auch mal gespannt, ob man da eine klare Regelung macht. Das wäre aber etwas, das wirklich neu ist, dass das dann wirklich eingefordert wird und auch in einem gewissen Format eingefordert wird.	SV
3.41	JM	Ja, wissen Sie es, wie es bei den Radiologen, bei Ihnen vor Ort ist, nutzen die eigentlich die AI, sind die damit glücklich oder gibt es auch welche, die sagen so, ich bin hart alteingesessen? Ich sehe das besser als die AI.	
3.42	R3	Nein, ich glaube, inzwischen ist es bei allen angekommen, dass diese, gerade dieses Thema Lungenrundherddetektion. Das ist einfach gut. Wie gesagt, man muss die Falsch-Positiven aussortieren. Aber es ist eine echte Unterstützung und verkürzt den Befundungssprozess. Das ist überhaupt gar keine Frage. Und andere Dinge, also Lungenemphysemquantifizierung, das händisch zu machen, da wird niemand auch nur auf die Idee kommen. Weil das wird ja eine Stunde dauern, bis man da mal irgendwie seine Konturen gemalt hat. Das geht nur mit einer automatischen Segmentierung, die inzwischen immer mit AI Unterstützung funktioniert.	PU RD JR PT

3.43	JM	Ja. Sie meinten ja schon, das System ist sehr gut, aber nicht perfekt. Was würden Sie denn sagen? Sind momentan noch so die die Berührungspunkte mit der AI, wo es Probleme gibt oder wo es nicht perfekt ist?	
3.44	R3	Die Falsch-Positiven, die man raussortieren muss.	
3.45	JM	Das ist das einzige Problem. Ansonsten auch die Zusammenarbeit mit den Radiologen, funktioniert die auch immer super gut, oder?	
3.46	R3	Ja, also das ist diese Interaktion. Klar gibt es immer noch mal Dinge, wo man vielleicht auch modernere User Interfaces entwickeln kann, aber grundsätzlich funktioniert das gut. Das ist schnell, das ist überall verfügbar. Also da sind wir auf einem super guten Weg.	PEUT
3.47	JM	Aber was meinen Sie zum User Interface? Wurde denn da mal irgendwie geäußert, dass das nicht so...	
3.48	R3	Ja, also wie schnell kann ich akzeptieren, ob der Herd gutartig oder bösartig ist? Brauche ich dort..? Habe ich dort unterschiedliche Abstufungen, die ich festlegen kann? Kann ich daraus gleich einen strukturierten Befund generieren? Und solche Fragen. Da ist sicherlich noch Luft nach oben. Aber die Software an sich, die funktioniert ganz hervorragend.	PEUT
3.49	JM	Okay, ja.	
3.50	R3	Das auch und das auch über viele Hersteller, muss man ganz ehrlich sagen. Es gibt ja ganz viele Systeme, die so etwas machen. Wir haben uns hier für die HANSE Studie für ein System einer Firma aus [country, anonymized] entschieden.	
3.51	JM	[AI software, anonymized], ja? Wir haben uns da auch schon auch eingelesen.	
3.52	R3	[AI software, anonymized], genau. Weil die haben einerseits Erfahrungen aus dem Programm in [country, anonymized] und die waren auch am flexibelsten, um jetzt auf unsere Bedürfnisse im Rahmen der Studie einzugehen.	RU
3.53	JH	Ja, super. Dann hätten wir noch einige Fragen zu den Eigenschaften des Systems oder zur Qualität. Stellen Ihnen die Systeme denn alle Informationen zur Verfügung, die Sie gerne hätten oder würden Sie sich da noch andere Informationen wünschen?	
3.54	R3	Nö. Eigentlich können die inzwischen alles liefern, was uns interessiert.	OQ JR

3.55	JM	In Bezug auf wenn jetzt eine Diagnose gestellt werden muss, dann werden sicher verschiedene Sachen angeguckt. Unterstützt das System den Radiologen in, ich sag mal an jedem Anhaltspunkt? Oder gibt es irgendetwas, wo irgendwo Sie sagen könnten AI, wenn es das auch noch könnte, wäre das auch noch cool.	
3.56	R3	Ja, ich glaube, die haben jetzt einfach noch mehr zu lernen, die relevanten Befunde zu detektieren, weil ich meine im Rahmen dieser Studie, ähm, wir haben ja bei ungefähr ja weiß gar nicht, knapp 10% der Patienten oder der Probanden, die wir sehen, haben wir irgendwelche Findings. Und dann lernen wir natürlich auch daraus und daraus kann natürlich auch die Software lernen. Was sind relevante Befunde, was waren am Ende Karzinome und was hat sich als gutartig herausgestellt? Und da noch mehr Hilfestellungen bei der Differenzierung und bei der Therapieempfehlung oder der Kontrollempfehlung, da können die Systeme sicherlich noch noch dazulernen und noch besser werden.	OQ
3.57	JM	Ja Sie meinten ja schon, dass das User Interface eventuell nicht perfekt ist, würden sie sagen, aber allgemein der Output des Systems ist verständlich oder ist es manchmal auch ein bisschen nicht so verständlich?	L PEUT
3.58	R3	Nein, das, also wir verstehen uns.	
3.59	JM	Okay, muss man da auch erst mal angelernt werden eigentlich, um das System zu benutzen? Oder sagen Sie man setzt sich davor und das ist intuitiv wie ein Buch oder so?	
3.60	R3	Nein, das ist definitiv nicht der Fall. Sas würde man sich wünschen bei vielen dieser Software Systeme, aber das ist, die Software wird ja nicht für mich geschrieben, sondern die wird geschrieben für eine weltweite Nutzung. Und es gibt so viele Kollegen, die andere Meinungen haben, andere Anforderungen haben, den Prozess anders haben wollen. Darum ist es in der Regel so, dass so eine Software unheimlich viele Möglichkeiten bietet und man auch auf ganz verschiedenen Wegen zu einem Ergebnis kommen kann. Um einfach alle Wünsche, die an die Software gestellt werden, gegebenenfalls erfüllen zu können. Ja, ja, ich nutze eigentlich immer denselben und mir persönlich würde es reichen, wenn die Software diesen einen Weg können würde und dann würde sie vermutlich auch viel intuitiver und einfacher zu bedienen sein. Aber dadurch, dass sie so viele Möglichkeiten bietet, haben die die meisten IT Lösungen zu viele Knöpfe, zu viele Optionen und darum braucht man einfach eine Einführung.	PEUT E

3.61	JH	Also gab es dann so eine Art Workshop oder gibt es ein Handbuch, wo man immer nachgucken kann? Oder gibt es da auch so einen IT Support für die Radiologen, falls man mal irgendwie nicht weiter kommt?	
3.62	R3	Es gab eine. Es gab eine Schulung in die Software, wo alles gezeigt wurde, was sie kann, wie sie arbeitet und auch wie man sich seinen eigenen Workflow dort einrichten kann.	
3.63	JM	Gab es eigentlich mal technische Probleme mit dem System oder lief das immer sauber? Ist nie abgestürzt?	
3.64	R3	Ne, wir haben schon ein paar technische Probleme gehabt, dass die, weil die Daten ja auch zentral gespeichert werden und dann Datenübertragungen, was an ganz vielen Stellen haken kann, gerade auch so die Anbindung vom Truck, wo dann draußen quasi auf dem Hof nur Datendose sein muss. Und ja, es gab schon ein paar so ein paar Hänger, aber im Großen und Ganzen lief es.	SR
3.65	JM	So nur kleine Strauchler oder auch mal was irgendwie Schlimmeres oder so oder oder was?	
3.66	R3	Nein, einfach, dass mal Daten nicht verschickt werden konnten. Also nein.	
3.67	JM	Das könnte man aber auch immer schnell vor Ort lösen. Da war auch genügend Support da?	
3.68	R3	Ja.	SR RU
3.69	JM	Okay, ja gut. Wie würden Sie denn die Veränderungen des Diagnose Outputs bewerten, den diese, den durch die Unterstützung AI passiert? Also ich würd mal sagen, es wird auf jeden Fall schneller, richtig?	
3.70	R3	Ja, auf jeden Fall.	RD
3.71	JM	Würden Sie auch sagen, die Qualität verbessert sich auch, oder? Der Befunde.	
3.72	R3	Ja, ich glaube. Es wird weniger übersehen werden. Dass ist sicher. Aber die, das entscheidende ist ja, was mache ich jetzt aus dem Befund? Und da ist viel noch humane Intelligenz in den interdisziplinären Konferenzen?	PU

3.73	JM	Ja, aber ja. Sagen wir mal, auch wenn es länger dauern würde. Ein Befund, den ein Mensch nur alleine sich angucken würde und dann bearbeiten würde, gegen den derzeitigen Prozess, wo vielleicht ein Radiologe noch zehn Minuten braucht mit der Unterstützung eines AI Systems. Wie würden sich die beiden Befunde unterscheiden? Würden Sie sagen so der mit der AI erkennt mehr, erkennt aber auch mehr, wo gar nichts ist, oder?	
3.74	R3	Der erkennt mehr wo nichts ist. Also die Falsch-Positiven-Rate ist bei all diesen Systemen die große Herausforderung. Ja, und es sind, es sind Unterstützungssysteme und sie werden die Behandlung nicht ersetzen.	PU
3.75	JH	Ja. Bedeutet es denn Mehraufwand für Sie? Sie haben dann ja diese Falsch-Positiven Ergebnisse. Was passiert dann? Müssen die Patienten dann eventuell schon früher wieder herkommen, zu einem zweiten CT und dann entsteht da sozusagen Mehraufwand für Sie?	
3.76	R3	Nein, also wir würden die ja bei der Kontrolle der Befunde der AI aussortieren, die Falsch-Positiven. Also für die Patienten entsteht daraus eigentlich kein Nachteil.	
3.77	JH	Ja, aber kommt es dann vor, dass sie teilweise durch die AI eigentlich länger bräuchten als ohne, weil sie halt diese Falsch-Positiven haben, die sie erst mal aussortieren müssen?	
3.78	R3	Ja. Also, das kommt, das kann mal sein, wenn die KI unendlich viele Befunde findet. Ja, das ist jetzt bei diesem Population, die wir hier haben, kommt das selten vor. Aber wenn man sich jetzt vorstellt, dass man ja andere Erkrankungen hat, die Knötchen in der Lunge machen dann und die das auch vielleicht diffus machen, dann findet die AI natürlich wahnsinnig viele Befunde und dann kann man es sich eigentlich auch fast schenken, sie zu nutzen. Weil wenn die mir 100 oder 200 Befunde anzeigt, dann wird es wirklich mühsam und dann muss ich natürlich manuell gucken. Das sind aber, das sind echt absolute Ausnahmen.	PU
3.79	JM	Ja. Was würden Sie denn sagen Sie noch allgemein so die größten Herausforderungen oder Probleme allgemein in der Zusammenarbeit mit diesem AI System? Oder denen AI Systemen, die da unterstützen.	
3.80	R3	Ja, ich glaube, was man sich noch wünschen würde, ist dass man, also wir leben ja immer noch sehr mit diesem Textbefunden. Das ist ja unsere Art und Weise der Kommunikation. Und ähm und das ist in der Regel Freitext. Dann aus diesen Findings der AI sofort einen strukturierten Text Befund zu machen, das ist eine große Herausforderung, aber das wird noch mal deutlich weiter helfen.	OQ JR PU

3.81	JM	Unterstützt, die AI nicht eigentlich schon, indem sie so eine Struktur vorgibt für den Text, also den Text vorformuliert?	
3.82	R3	Ja, solche Systeme gibt es, aber da ist dann immer so dieser kompromiss zwischen, einerseits will ich vollständig sein, andererseits soll es schnell gehen. Das widerspricht sich und es geht so ein bisschen die Flexibilität verloren.	OQ JR PU
3.83	JM	Ja. Ja.	
3.84	R3	Und auch die Gewichtung, wenn man sich zu sehr an eine Struktur klammern muss. Ein Befund muss ja nicht nur richtig sein, sondern er muss auch klinisch verwertbar sein. Mein Ex-Chef hat teilweise, hat teilweise Ein-Wort-Befunde geschrieben, Normalbefund. Es ist ein, ist ein sehr schwieriger Befund. Da muss man sich hundertprozentig sicher sein, dass wirklich überhaupt nichts Relevantes ist. Aber der ist natürlich von der, von der Geschwindigkeit und von der klinischen Verwertbarkeit her ist der Befund super, weil den kann man in Nullkommanix lesen. Und jeder Arzt, der den Befund in die Hand kriegt, weiß warum muss ich mir keine Gedanken mehr machen. Da ist alles takko. Ja, aber das geht halt eben nicht, nicht in allen Fällen und manchmal werden es dann einfach auch sehr komplexe Konstrukte solche Befunde.	
3.85	JM	Ja.	
3.86	R3	Weil man auch abwägen muss...	
3.87	JM	Kostet das sonst so viel Zeit die Texte runter zu schreiben? Also da ist dann auf jeden Fall Potenzial, wo AI das dann noch besser unterstützen könnte, richtig? Beim Formulieren dieser Texte.	
3.88	R3	Auf jeden Fall.	
3.89	JM	Was was kostet denn ansonsten noch viel Zeit allgemein in dem Prozess der Diagnose, des Befundes?	
3.90	R3	Jetzt bei allgemein in der Radiologie?	
3.91	JM	Ne jetzt gerade auch bei Lung Cancer Screening spezifisch.	
3.92	R3	Nein, da ist alles in erster Linie das Durchschauen der Bilder. Das ist schon der wichtigste Faktor dort.	

3.93	JM	Okay, das braucht am meisten Zeit, aber unterstützt das schon die AI so gut wie es kann, mit der Hervorhebung. Okay, okay. Und ansonsten nur noch der, na ja, der Text, der geschrieben werden muss am Ende kostet viel Zeit. Okay.	
3.94	R3	Genau.	
3.95	JH	Sehen Sie denn noch irgendwelche Probleme bei der Integration dieser AI Anwendungen in den Workflow der Radiologen, dass das vielleicht noch nicht optimiert ist, dass das, dass die noch besser zusammenarbeiten könnten?	
3.96	R3	Das funktioniert wirklich, wirklich super. Also die Daten gehen an das AI System, werden prozessiert und stehen ein paar Minuten später auch zur Durchsicht der Ergebnisse zur Verfügung. Das kann man alles perfekt organisieren.	JR
3.97	JH	Super. Sind diese Systeme eigentlich so schnell, dass der Radiologe immer beschäftigt ist? Oder kommt das mal vor, dass man irgendwie eine Minute einfach vor dem Computer sitzt, weil das System einfach lädt?	
3.98	R3	Nein. Die, also die je nachdem was, was die nun prozessieren und wie viel Bilder drin sind, kann das schon mal fünf Minuten dauern, bis der Test das komplett alles durchgelaufen ist. Aber dadurch, dass wir immer ein bisschen Verzug zu dem, zu der Bild Akquisition haben und das ja wirklich nicht, wenn es im Scanner akquiriert wird, entstehen da keine Wartezeiten.	
3.99	JM	Genau. Sie meinten ja das AI System läuft ja direkt natürlich. Das wird halt, sage ich Mal, 10 Minuten später angeguckt, aber fünf Minuten vorher is es halt durchgelaufen, richtig?	
3.100	R3	Ja genau.	
3.101	JM	Ich glaube, die meisten schauen sogar bei Ihnen allgemein die Kollegen und die Radiologen, die freuen sich schon auf die Nutzung mit der AI im Rahmen vom Lung Cancer Screening, also die sehen das auch alle positiv entgegen. Keiner hat Angst, dass irgendjemand seinen Job verliert durch AI, oder?	SV
3.102	R3	Nein.	SV
3.103	JM	Okay, Sie meinten ja auch, dass Sie meinen, genau das sind klare unterstützende Systeme. Sehen Sie nicht die Chance irgendwann, dass AI wirklich den kompletten Job übernehmen könnte oder auch wirklich einen Befund von null auf hundert selbst erstellt und ziemlich gut da drin ist?	

3.104	R3	Also ähm, für Lungenkrebs Screening kann ich mir das vorstellen.	
3.105	JM	Dass AI das komplett übernehmen kann, eventuell?	
3.106	R3	Ja, also man wird mir immer noch ein paar Falsch-Positive raussuchen müssen. Da bin ich ganz sicher, aber grundsätzlich kann man das sicherlich noch weiter, weiter vorantreiben, weil wenn man jetzt, also ich sehe die Chancen, solche Systeme einzusetzen, wenn es eigentlich um eine einfache Entscheidung geht, wenn ich Krebs Ja/Nein entscheiden muss. Das kann so ein System irgendwann sicherlich lernen. Wenn man aber jetzt als anderes Extrembeispiel hat, einen Patienten, der schon fünfmal am Bauch operiert worden ist und wo nichts mehr so ist, wie es ursprünglich mal war, und man so eine Untersuchung dann in ein System hinein gibt, da wird niemals ein korrekter Befund rauskommen können. Da fehlt einfach die, die zusätzliche Information, die wir nutzen, was ist gemacht worden. Wie ist es gemacht worden, wie geht es dem Patienten und und und. Das sehe ich nicht, dass wir da irgendwann hinkommen, dass ein System all diese Faktoren, die wir ja nebenbei mit akquirieren, ähm, nur aus den Bildern. Weil das ist einfach so! Die Information steckt nicht nur in den Bildern, zumindest über das Gesamtgebiet der Radiologie anschaut. Jetzt bei so einer Frage Lungenkrebs Screening, da steckt die Information schon in den Bildern, wir wissen alles, es ist ein Hochrisikokollektiv, die haben alle geraucht und und und.	
3.107	JM	Da haben wir trotzdem schon mal gehört, es wäre noch mal ganz cool, auch irgendwie ein System zu haben, was noch mal trotzdem ein bisschen mehr sagt, wie viel der Patient geraucht hat und wie alt er ist, damit man auch wirklich noch mal so ein Risikoprofil hat zu dem Patienten. Könnte das irgendwie vielleicht helfen? Bei dem Befund auch.	
3.108	R3	Also, äh, ich glaube, ich glaube nicht, da es ist dann letztendlich egal ob der 30 oder 50 Pack-Years hat, der nur 30 Pack-Years hat, wenn er Pech gehabt hat, hat er ein schlimmeres Emphysem oder ein höheres Risiko als der mit den 50 Pack-Years. Also das ist ja nicht ganz linear jetzt irgendwie, dass man das voneinander abhängig machen könnte. Ich glaube Risikofaktoren zu kennen, ja, das ist wichtig, aber das jetzt noch weiter zu erfassen, glaube ich macht keinen Sinn oder hilft einem nicht wirklich weiter.	OQ
3.109	JM	Antizipieren Sie eigentlich noch irgendwie Probleme bevor Lung Cancer Screening in Deutschland umgesetzt wird und das an Ihrer Klinik in diesem Zusammenhang dann auch wirklich durchgeführt wird, ich sag mal, wenn die Patientenmenge auch noch Mal deutlich ansteigt? Denken Sie da kommen noch mal irgendwelche	

		Probleme auf Sie zu oder denken Sie sind jetzt eigentlich dafür vorbereitet?	
3.110	R3	Also technisch glaube ich, kann man das wuppen. Das ist überhaupt gar keine Frage. Politisch und auch wenn es um die Frage geht, wer bezahlt das und zu welchem Preis? Das wird noch längere Diskussionen geben. Da bin ich ganz sicher.	
3.111	JH	Okay. Ja. Ja.	
3.112	R3	Und auch die Frage, wer macht es, wer weist zu, wer prüft die Risikofaktoren und und und. Solche Sachen, was ja auch die Kostenträger verlangen, dass man es mit Raucherentwöhnung Programm kombiniert, das natürlich auch absolut sinnvoll ist, weil wir wollen ja nicht die Leute screenen, damit sie weiter quarzen können, sondern eher auch was für die Patientengesundheit tun und solche Kombinationen, dass das wird ja alles in dem Gesamtpaket mit drin sein müssen.	
3.113	JH	Ja, vielen Dank. Ich glaube, wir wären mit unserem Fragenkatalog an sich jetzt durch. Hast du noch eine Frage Joachim?	
3.114	JM	Ich? Ne.	
3.115	JH	Haben Sie sonst noch irgendeine Frage zum Thema, was Sie gerne noch angesprochen hätten? Was zu kurz gekommen ist?	
3.116	R3	Nö, ich bin mal gespannt, was Sie in ihrer Masterarbeit schreiben.	
3.117	JM	Genau das wollten wir auch noch fragen, ob sie bereit wären, mit uns, Member Checking durchzuführen, das heißt, sobald wir ungefähr eine erste Analyse von diesem Gespräch haben oder so, was unsere Findings aus dem Gespräch sind, dass wir Ihnen eine kurze Mail schicken und Sie sagen so Ja, genau so habe ich das gesagt, genau so habe ich das gemeint. Oder Sie schreiben so, nein, da haben Sie mich komplett missverstanden, das war anders gemeint.	
3.118	R3	Ja das können Sie gerne tun.	
3.119	JH	Super, vielen Dank. Noch mal vielen Dank, dass Sie sich die Zeit genommen haben, dass hat wirklich sehr geholfen, wir haben viele coole Insights gewonnen.	

3.120	JM	Und nochmal sorry, dass wir hier etwas überziehen hier beim Gespräch. Ich hoffe Sie haben trotzdem noch einen erfüllten Morgen und können den Tag genießen.	
3.121	R3	Das wünsche ich Ihnen auch. Einen schönen Tag und viel Erfolg beim Schreiben. Geben Sie Gas.	
3.122	JH	Dankeschön. Tschüss.	

Appendix G - Interview 4

Duration: 18 minutes

Date: May 3, 2022

Interviewee: R4

English

Row	Speaker	Transcribed text	Codes
4.1	JM	Yes, exactly. Can you tell us very briefly in two or three sentences what kind of contact you've had with lung cancer screening so far and what kind of role you play there?	
4.2	R4	I have been involved in lung cancer screening for many years and, I don't know, seven or eight years ago I was involved in the S3 guidelines for lung cancer screening in Germany, I was a representative of the [organization, anonymized] and was instrumental in bringing an optional decision for lung cancer screening. I was there [action, anonymized] and then also pleaded for a can-do decision, because at that time there were only the data from the US Lung Cancer Screening Trial, which had such a high false-positive rate and also had the complications with the follow-up so that then that and the Nelson data were not yet out. And then there were a number of European data that were not significant and there was no plausible meta-analysis at that time. And that's when I said this is a can-do recommendation. Many wanted a must or should or what recommendation, but I said we are not a political body here, we are a scientific body. And then my motion was [description, anonymized]. And with that, all the other votes sort of lapsed and everybody is actually quite happy with that decision. What I also did, is that it should only be done in a structured program, that you connected this first and second, you seem to know the S3 guidelines, the first and second question, that it is a can-do recommendation, but only in a structured program. And then in the second question, how such a structured, who all should participate in this. The radiology, surgery, and so on that it's national and what risk group and so on down. Yes, that was the first and then is yes IQWiG and BFS the votes out there. That is a parallel process in Germany for the introduction of lung cancer screening in Germany. And now we are waiting for the BMU, so the BMU on the	

		regulation of the BMU and it is a question now only how and when it comes but no longer whether it comes in Germany and um yes and then we said yes, we were approached in the German Center for Lung Research by the Lung Ambition Alliance, which is a public-private partnership, which also, where AstraZeneca also participates and there we had a workshop, how we can bring the lung cancer screening on the road, in an interdisciplinary group from northern Germany. And then it came out that we do an implementation trial and there [action, anonymized]. In addition, the BFS vote was informally available to the professional societies from May of last year, I think it was, and it was published in December, and so I was able to [action, anonymized] before the first participant on July 21, I remember, because [personal information, anonymized]. So, we adjusted the findings letter a little bit more. But everything was known for the most part, but I then shaped it in that respect, and also with the Reader First and Second Read, as they want to have that, we also adjusted that. I have the threshold, but I think I am also in the majority when I talk to my colleagues. We have now decided on Lung-RADS, there are different classifications to do this radiological reporting. We have decided on the Lung-RADS classification, but we have modified it slightly. The paper is in the press right now, that will probably come out soon now as well.	
4.3	JM	May I interrupt you for a moment? Exactly, because for us it would be a bit more relevant if we actually talked more about the processes and also about the actual processes and the work of the radiologists. Namely, you have already said, exactly within the framework of the study, that patients are also already being examined. We would be interested to know how exactly the process works. How do you examine patients in the context of lung cancer screening? What process do they go through on-site with you?	
4.4	R4	Do you want to know the inclusion process or what do you want? How to win the subjects or?	
4.5	JM	No, the process of how the, when they're on-site with you, how they're examined, what all is done there and then we would also go into it afterward…	

4.6	R4	The process before is more important. The process beforehand is much more important. Because you have to, you can have the best screening program, in the U.S., if you've done your research, there is a screening program and hardly anyone went, it was in the single-digit percentage range, they only went there and that was because it was poorly structured and probably also poorly paid, and you didn't really reach the target group. And we're looking at exactly that in the HANSE study. How do we get the target group to the truck or into the program? We have different things there. That's media, that's through internet presences, YouTube, and whatever. That works very well. But also, direct personal writing. We know that we got everyone between 55 and 79 from the residents' registration offices, we don't know whether they smoke or whether they are at risk, but we write to them and say, if you don't agree, you can still distribute it among your acquaintances, so to speak. And it works very well there. So, we will be able to include over half of the test subjects, we have 5,000 high-risk test subjects, in one year, which no one has ever done in a study like this before, and we have really managed to do it because we have organized the front end correctly. There are various channels, television appearances, and so on. But essentially also these register-based letters and so on and so forth. And then, then and then web portal is also important, a low-threshold. You probably looked at the website. "I'm participating." Have you tried signing up? It takes five minutes. And that's how we find out whether you have the threshold of a risk subject or not? Of course, no doctor checked you, but we all asked the questions very simply. And then when they get in the truck, a doctor checks to make sure that all of that is correct. We ask the same questions again on the spot and see if it's true and if they cross the threshold. But that is the case with most of them, I can tell you that already, without going into more detail. And then you are included in the study by pulmonologists, risks and side effects are discussed and then you get a holistic approach with us. You get a lung function, a pulse measurement, a blood pressure measurement, and then questionnaires and so on.	
4.7	JM	That's what we would actually be more interested in as well. Precisely for the examination and how AI can support it. As soon as we know that there is also [AI software, anonymized], I think, which is also used to support lung nodule detection, for example.	
4.8	R4	Yes.	

4.9	JM	That is used also with you, yes exactly. Do you know how the radiologists stand on that, whether that supports them really well or how they stand on that in general?	
4.10	R4	So, it always depends on which software. Yes, but that comes afterward. At first, they do get into the truck. We also have tablets and that is then triggered automatically and so on and so forth. Then you go into the truck, get the exam, and then you come, the image comes back and that is analyzed by an AI computer for 20 minutes. And we do the lung calcium, the emphysema, and the nodules, so tumor, so we do three in one, hole in one, and when the radiologist comes, he sees the analysis of the computer and he just has to check if that's right I would say. Of course, we have a slightly different workflow in that he has to find the nodules himself and is then shown the AI result. That is very well accepted, because it is the right workflow because the radiologist is still the one who has to take responsibility for that and then also nodules, if he says, oh that is not recorded correctly or whatever and that is not a solid nodule. That wouldn't bother at all, the category of the nodule is wrong. Then he can rather change that manually and then it's calculated for him and then it automatically outputs which Lung-RADS category it is, and the report letter is also, we programmed it that way in the back, automatically spit out. So, if it's a normal finding, you just click on it two or three times, scroll through it and press print and then it's done. Also, with this reader system, I can do that with so many. We have 120 per week. I can't do that at all, they all get mixed up. And these on this findings system, that it is then ready for ready, then First Read, the assistants read that in the universities. We have two universities and one, a lung clinic, where there are no assistants, but with us, the assistants read it, then it's the First Read, but then the Official Read is then looked through by the senior physician, and if it's then above a threshold, in Lung-RADS, Lung-RADS 3 we make the threshold, we have set it ourselves. The BFS has a slightly different approach, but I think we are in good company when I look at the X-ray congress. I think others see it similarly but then set the threshold for the second read. Then we have a folder that automatically goes into Second Read and then my colleague looks at it, who is blinded to my results. And then the computer, yes, can see my result and that of her and then it goes right out the letter. Or if we are discrepant, then we still have to find the discrepancies, so the workflow is also already forewarned and that is very well accepted by the radiologists.	OQ PU JR

4.11	JH	Do you know how well the system is performing, i.e., is it, are there a lot of false-positives or is it already working quite well?	
4.12	R4	It works great. When they sold it to me, it had an RC curve of 0.92 in lung nodule detection. We are highly satisfied. It saves work and by now I think that's over and done with that there is a danger that you have to remove all the nodules because they are not nodules. 92% of the nodules are correct.	RD
4.13	JM	Okay. In general, how would you say now AI overall changes the end result of the diagnosis? I guess it's faster, right? So, does it make it more time-efficient?	
4.14	R4	It also makes it better because radiologists, the radiologist traditionally, we have also now done a survey, only 20% of radiologists do volume-based evaluations and I have to say quite honestly, we in the clinic also do not do volume-based evaluations, when it comes to the, when it comes to tumor follow-up, for example, we only do two, three parameters, that's it. But since we know scientifically that volume is much more accurate, because differences are better recognized than with the diameter and so it improves the quality. And just this Lung-RADS version that we have, you can give diameter once and volume. But we have the volume first and I tell you, the software manages to calculate the volume adequately in almost 100%, 99.8% maybe if I could estimate. The squiggles are shown there, so you can set that. But what is important is that we make sure that everyone is scanned with a CT, that the CT protocol is the same and that the software is the same. That will be the big story, say in Germany we will manage to make a national tender and agree on one or two software or we won't manage it, because if you then go from A to B, you can't have growth all at once just because you get a different CT protocol or a different nodule and that can happen quickly with the so medium-sized, if it then goes so between six and ten millimeters, there is a lot happening in the early detection. You can adjust the threshold, and you'll notice that it's also very interactive. You can set the threshold a bit higher, manually, if the radiologist doesn't like it. And therefore, the radiologist has full control. There's quite a lot of other software, I screened the market back then, was that the absolute frontrunner back then. It's a volatile market, of course, but I think they are. There's another EU project, independent of us called 4-IN-THE-LUNG-RUN, you've probably heard of it, and independent of them I chose the same software as them. But I did not know it. But it's not supposed to be a [AI software, anonymized] or something...Leave that out of it, that's actually, I don't care if they're called [AI software, anonymized] or [AI	RD JR PU PEUT

		software, anonymized] or [AI software, anonymized] or something, they all offer something there, it's the content that's important.	
4.15	JM	It is merely important and yes also exactly that all also have the same, that one has then also really comparable results. Yes. Would you say in general that the AI software is also easy to use? So, it's also clear, that you can easily familiarize with it?	
4.16	R4	Yes, and we have now trained a new assistant there. She came back completely flashed from the two weeks and says wow super and also the workflows are different. And actually the radiologists, that's super, the the the and of course then there will be a few in the background the dollar signs, I say if that then after BFS should first start from me in Germany. Yes, I am not financially driven, I am rather content-driven and I think that it will also run well, that this is anchored in the young centers, institutionalized, and that there such locusts can not pull the money from this process.	PEUT
4.17	JM	So, the radiologists, they all like to use the system too, or is there anyone who is afraid that it might take away their job?	
4.18	R4	It loads a bit slow when loading a trial. It's web-based, it's web-based, it's not an app yet. Don't know how technical you are on it, haven't been told now, as long as you run it in the browser, is always a tad slower and I stopped it once. One needs about a minute of loading time in total between upload, series load and then save again. That is half a minute or 40 minutes upload, 40 seconds upload and 20 seconds save. If you really...	PT
4.19	JM	How exactly does that work, the CT scan goes through, and later when it's evaluated, then the AI system goes over it first when the radiologists are looking at it or what?	
4.20	R4	No, before that. So, we don't have any emergencies. The CT runs from the scanner and then it runs on the server here with us, it is security compliant and there it is then analyzed simultaneously and then analyzed and then we have the result is then fully automatic, is then displayed when uploading.	
4.21	JM	But if you access it in the browser, you still have the loading times.	

4.22	R4	Yes, then you have the loading time and 20 minutes it's post-processing, but we don't care because you then, because we don't have… What we also have is a cloud, where we have connected the three locations, but on top, there is pseudo-sympathetic data.	
4.23	JH	Are you missing any features in AI or does it already do everything you want it to?	
4.24	R4	Right now, it's doing what we need it to do.	OQ PU
4.25	JM	But can you think of any potential for improvement? I think you also have a very good overview of all the sites. Does that work great everywhere, or was it once said that the radiologists aren't quite ideally supported here?	
4.26	R4	Yes, it's so natural with the connection, with the IT, and with the technology. There has been a bit of jerking and hailing and so and then a new port or whatever. We have an IT commission, we have the radiology groups, we have pneumology groups, and everything together. The important thing is that in the back, the backbone is organized. We have an [inaudible] in the back that also did the website and taps the ICF right in the back. All the data that we collect is also entered into this ICF, i.e., all these volumes of all the nodules and into the categorization, which are tons of parameters, including scoliosis, I almost said that now, but skewness, the heterogeneity of the nodules, all this is automatically recorded, much more than the radiologist even wants to know and we push all this into our database. It's automatically deducted and then we have a gigantic database, and we do research afterward. But that's not so important for the, for the implementation and implementation is already important that you also have a research register nationally to develop further. Yes.	SR
4.27	JM	Perfect. Otherwise, I think we've asked all the important questions so that you can get to the next appointment on time. Do you have any other questions you would like to ask?	
4.28	JM	It's a master's thesis you're doing now, right?	
4.29	JH	Exactly, yes.	

4.30	R4	And you do that as a team? In Germany, you always have to do it alone.	
4.31	JM	That's a very special thing in Sweden because the master's only lasts for one year. Exactly, by the way, we wanted to ask you if we could do member checking with you. That is, by having a first analysis of this conversation that we can tell you briefly, we have taken the following, would you agree that you have said so, or have we misunderstood something? I wonder if we could send that to you once?	
4.32	R4	You would send that to me then, right?	
4.33	JH	That's right, a short summary, then you nod it through, so to speak. Not that we have completely misunderstood you.	
4.34	R4	Yes. Great. Yes, good luck then. I know that such interviews are always very important. Good luck.	
4.35	JM	Thank you very much, bye.	
4.36	JH	Thank you, thank you again for your time.	

German

Row	Speaker	Transcribed text	Codes
4.1	JM	Ja, genau. Können Sie uns ganz kurz in zwei, drei Sätzen sagen, in was für einem Kontakt Sie denn bis jetzt mit Lung Cancer Screening zu tun hatten und in was für einer Rolle Sie da tätig sind.	
4.2	R4	Also ich habe seit vielen Jahren mit Lungenkrebs Früherkennung zu tun und ich habe, weiß ich nicht, vor sieben, acht Jahren mitgearbeitet bei den S3 Leitlinien Lungenkarzinom Screening für Deutschland war ich ein Vertreter der [organization, anonymized] und habe da maßgeblich eine Kann-Entscheidung für das Lungenkrebs Screening herbeigeführt. Ich habe da [action, anonymized] und dann auch plädiert für eine Kann-Entscheidung, weil damals waren es nur die Daten aus den USA vom US Lung Cancer Screening Trial da, die ja so eine hohe Falsch-Positive-Rate hatten und	

		auch die Komplikationen mit dem Follow-Up hatten, so dass dann das und die Nelson Daten waren noch nicht raus. Und dann gab es eine Reihe europäischer Daten, die nicht signifikant waren und es gab keine plausible Metaanalyse damals. Und da habe ich gesagt, das ist eine Kann-Empfehlung. Viele wollten da eine Muss- oder Soll- oder was Empfehlung, aber ich habe gesagt wir sind hier nicht ein politisches Gremium, sondern ein wissenschaftliches Gremium. Und dann war mein mein Antrag [description, anonymized]. Und damit waren alle anderen Abstimmungen sozusagen hinfällig und alle sind eigentlich recht zufrieden mit dieser Entscheidung. Was ich auch noch gemacht habe, dass es nur in einem strukturierten Programm durchgeführt werden soll, dass man diese erste und zweite, Sie scheinen ja die S3 Leitlinien zu kennen, die erste und zweite Frage miteinander verbunden hat, dass es eine Kann-Empfehlung ist, aber nur in einem strukturierten Programm. Und dann wird dann in der zweiten Frage wie so ein strukturiert, wer da alles mitmachen soll. Die Radiologie, Chirurgie und so weiter, dass es national ist und welche Risikogruppe und so weiter runter. Ja, das war das erste und dann ist ja IQWiG und BFS die Voten da rausgegangen. Das ist ja ein paralleler Prozess in Deutschland für die Einführung des Lungenkarzinom Screenings in Deutschland. Und jetzt sind wir dabei warten auf die BMU, also das BMU auf die Verordnung des BMU und es ist eine Frage jetzt nur noch wie und wann es kommt aber nicht mehr ob es kommt in Deutschland und ähm ja und dann haben wir gesagt ja, da wurden wir angesprochen im Deutschen Zentrum für Lungenforschung von der Lung Ambition Alliance, das ist eine Public Private Partnership, die auch, wo AstraZeneca auch mitmacht und da hat man nen Workshop, wie wir den Lung Karzinom Screening auf die Straße bringen können, in interdisziplinären Gruppe aus Norddeutschland. Und dann kam raus, dass wir doch ein Implementierungs-Trial machen und da [action, anonymized]. Dazu kam noch, dass das BFS Votum informell den Fachgesellschaften zur Verfügung stand ab Mai letzten Jahres war das glaube ich und das wurde im Dezember veröffentlicht und so konnte ich [action, anonymized] bevor der erste Proband am 21. Juli, das weiß ich noch, weil [personal information, anonymized]. So haben wir den Befundbrief ein bisschen angepasst noch. Aber es war ja weitestgehend alles bekannt, aber ich habe das dahingehend dann noch geshaped und auch mit dem Reader First und Second Read, wie die das haben wollen, auch das haben wir angepasst. Ich habe die Schwelle, aber ich glaube, da bin ich auch in der Mehrheit, wenn ich mal so mit meinen Kollegen spreche. Wir haben jetzt Lung-RADS entschieden, es gibt verschiedene Klassifikationen, diese radiologische Befundung durchzuführen. Wir haben die Lung-RADS Klassifikation entschieden, aber wir haben sie leicht modifiziert. Das Paper ist gerade im Press, das wird wahrscheinlich jetzt auch bald rauskommen.	

4.3	JM	Darf ich Sie kurz unterbrechen? Genau, weil für uns wäre es noch ein bisschen relevanter, wenn wir tatsächlich mehr über die Prozesse und auch auf die, auf die tatsächlich Prozesse und auf die Arbeit auch der Radiologen zu sprechen kommt. Nämlich Sie haben ja schon gesagt, genau im Rahmen der Studie wird ja auch bei Ihnen, werden auch schon Patienten untersucht. Wir würden uns dafür interessieren, wie genau der Prozess da abläuft. Wie werden bei Ihnen gerade Patienten im Rahmen von Lung Cancer Screening untersucht. Welchen Prozess durchlaufen die da vor Ort bei Ihnen?	
4.4	R4	Sie wollen den Einschlussprozess wissen oder was wollen Sie? Wie man die Probanden gewinnt oder?	
4.5	JM	Ne den Prozess, wie die, wenn die vor Ort bei Ihnen sind, wie die untersucht werden, was da alles gemacht wird und dann würden wir auch anschließend darauf eingehen...	
4.6	R4	Der Prozess vorher ist ja wichtiger. Der Prozess vorher ist ja viel wichtiger. Weil Sie müssen ja, Sie können ja das beste Screening Programm haben, in den USA, wenn Sie sich kundig gemacht haben, gibt es ja ein Screening Programm und kaum einer ist hingegangen, es waren im einstelligen Prozentbereich, sind die da nur hingegangen und das lag daran, dass es schlecht strukturiert war und wahrscheinlich auch schlecht bezahlt war und man die Zielgruppe nicht richtig erreicht hat. Und wir untersuchen genau das in der HANSE Studie. Wie bekommen wir die Zielgruppe zum Truck oder ins Programm rein? Da haben wir verschiedene Sachen. Das ist Medien, das ist durch Internetauftritte, YouTube und was auch immer. Das funktioniert sehr gut. Aber auch direktes persönliches Anschreiben. Wir wissen, wir haben von den Einwohnermeldeämter alle gekriegt zwischen 55 und 79, wir wissen nicht, ob die rauchen oder ob die Risikoprobanden sind, aber wir schreiben die an und wir sagen, wenn es für Sie nicht zustimmt, dann können Sie es ja noch in Ihrem Bekanntenkreis verteilen, sozusagen. Und da taugt es sehr gut. Also über die Hälfte der Probanden, wir haben 5000 Hochrisikoprobanden, werden wir schaffen in einem Jahr einzuschließen, hat niemand bisher gemacht in so einer Studie und wir haben es wirklich geschafft, weil wir halt das Frontend richtig organisiert haben. Es sind verschiedenste Kanäle, Fernsehauftritte und und und. Aber im Wesentlichen halt auch diese registerbasierte Anschreiben und so weiter und so fort. Und dann werden die dann und dann Webportal auch wichtig, niedrigschwellig. Sie haben wahrscheinlich auf die Webseite geguckt. "Ich mache mit". Haben Sie mal versucht sich anzumelden? Das dauert fünf Minuten. Und da kriegen wir raus, ob sie nun die Schwelle eines Risikoprobanden haben oder nicht? Natürlich hat Sie das kein Arzt kontrolliert, aber die Fragen haben wir alle sehr einfach gestellt. Und wenn sie dann in den Truck kommen,	

		kontrolliert ein Arzt, ob das alles auch stimmt. Wir stellen die gleichen Fragen vor Ort wieder und gucken, ob das stimmt und ob sie die Schwelle überschreiten. Das ist aber bei den meisten der Fall, kann ich Ihnen schon mal so sagen, ohne da mich näher zu äußern. Und dann werden Sie in die Studie eingeschlossen von Pneumologen Risiken und Nebenwirkungen diskutiert und dann kriegen Sie bei uns noch einen ganzheitlichen Ansatz. Sie kriegen halt noch eine Lungenfunktion, eine Pulsmessung und Blutdruckmessung und dann Fragebogen und so weiter.	
4.7	JM	Genau dafür würden wir uns tatsächlich auch genauer interessieren. Genau für die Untersuchung und wie AI da unterstützen kann. Sobald wir nämlich auch wissen gibt es da ja glaube ich auch [AI software, anonymized], wird ja auch eingesetzt, um da zum Beispiel Lungenknotenerkennung zu unterstützen zum Beispiel.	
4.8	R4	Ja.	
4.9	JM	Das wird genau auch bei Ihnen benutzt, ja genau.Wissen Sie, wie die Radiologen dazu stehen, ob das die wirklich gut unterstützt oder wie sie allgemein dazu stehen?	
4.10	R4	Also kommt immer drauf an welche Software. Ja gut, aber das kommt hinterher. Erst dann kommen sie in den Truck. Auch wir haben so Tablets und das wird dann automatisch ausgelöst und so weiter und so fort. Dann gehen Sie in den Truck, kriegen die Untersuchung und dann kommen Sie, kommt das Bild zurück und das wird durch einen AI Computer 20 Minuten lang analysiert. Und wir machen den Lungenkalcium, das Emphysem und die Knötchen, also Tumor, also wir machen drei in One, hole in one, und wenn der Radiologe kommt, dann sieht er die Analyse des Computers und muss nur noch checken, ob das stimmt sage ich mal so. Wir haben den Workflow natürlich ein bisschen anders, dass er die Knötchen dann selber auch finden soll und dann das von dem AI Ergebnis gezeigt kriegt. Das wird sehr gut angenommen, weil es der richtige Workflow ist, weil der Radiologe immer noch derjenige ist, der das zu verantworten hat und dann auch Knötchen, wenn er sagt, oh das ist nicht richtig erfasst oder was auch immer und das ist kein solider Knoten. Das würde gar nicht stören, Kategorie des Knötchen ist falsch. Dann kann eher das manuell ändern und dann wird ihm das so berechnet und dann wird automatisch ausgegeben, welche Lung-RADS Kategorie das ist und der Befundbrief wird auch, hinten haben wir extra so programmiert, automatisch ausgespuckt. Also wenn es ein Normalbefund ist, klicken Sie da nur zwei drei mal rum, scrollen dadurch und drücken auf Befund ausdrucken und dann ist fertig. Auch dieses Reader System, ich kann das ja bei so vielen. Wir haben 120 pro Woche. Da kann ich gar nicht, die kommen ja alle durcheinander. Und diese auf dieses Befundsystem, dass es dann ready ist für fertig,	OQ PU JR

		dann First Read, die Assistenten lesen das in den Unis. Wir haben zwei Unis und einen, eine Lungenklinik, da wo keine Assistenten sind, aber bei uns lesen es die Assistenten, dann ist es der First Read, dann aber der Official Read ist dann beim Oberarzt durchgeguckt, und wenn es dann über einer Schwelle ist, in Lung-RADS, Lung-RADS 3 machen wir die Schwelle, die haben wir selber so festgelegt. Das BFS hat da ein bisschen anders und aber ich glaube, wir sind da, wenn ich so auf dem Röntgenkongress gucke in guter Gesellschaft. Ich glaube, andere sehen das ähnlich, aber dann die Schwelle gesetzt für den Second Read. Dann haben wir so einen Folder, da kommt das automatisch in Second Read und dann guckt meine Kollegin das an, die geblendet ist zu meinen Ergebnissen. Und dann macht der, kann der Computer, ja mein Ergebnis und das von Ihr sehen und dann geht es gleich raus, der Brief. Oder wenn wir diskrepant sind, dann müssen wir es noch die Diskrepanzen befunden, also der Workflow ist da auch alles schon vorgewarnt und das wird sehr gut angenommen von den Radiologen.	
4.11	JH	Wissen Sie denn wie gut das System performt, also wird es, gibt es da viele falsch-positive Befunde oder arbeitet das schon ganz gut?	
4.12	R4	Es arbeitet super. Es hatte eine RC-Curve, als sie es mir verkauft haben von 0,92 in der Lungenknötchendetektion. Wir sind hochzufrieden. Es erspart Arbeit und mittlerweile glaube ich ist der Drops gelutscht, dass das da dass da die Gefahr ist, dass man so die Knötchen alle wegmachen muss, weil es keine Knötchen sind. 92 % der Knötchen sind richtig.	RD
4.13	JM	Okay. Wie würden Sie allgemein sagen, verändert jetzt die AI insgesamt das Endresultat der Diagnose? Ich schätze mal, es geht schneller, richtig? Also es macht es zeiteffizenter?	
4.14	R4	Es macht es auch besser, weil Radiologen, der Radiologe traditionell, wir haben auch jetzt eine Umfrage gemacht, nur 20 % der Radiologen machen volumenbasierte Auswertungen und ich muss ganz ehrlich sagen, wir in der Klinik machen auch keine volumenbasierten Auswertungen, wenn es um die, wenn es um die Tumornachsorge zum Beispiel geht, machen wir nur zwei, drei Parameter, das war's. Aber da man weiß wissenschaftlich, dass Volumen viel genauer ist, weil da Unterschiede besser erkannt werden als beim Diameter und so verbessert es die Qualität. Und eben diese Lung-RADS Version, die wir haben, die kann man einmal Diameter angeben und Volumen. Aber wir haben das Volumen zuerst und ich sage Ihnen, die Software schafft es in nahezu 100 %, 99,8% vielleicht, wenn ich schätzen dürfte, das Volumen adäquat auszurechnen. Die Kringel werden ja gezeigt da, da kann man das dann auch setzen. Was aber ist, das wir darauf Wert legen, dass	RD JR PU

		alle mit einem CT gescannt werden, dass das CT Protokoll gleich ist und dass die Software gleichwertig ist. Das wird die große Geschichte sagen in Deutschland schaffen wir es, eine nationale Ausschreibung zu machen und uns auf ein oder zwei Softwaren zu einigen oder schaffen wir es nicht, weil wenn man dann von A nach B geht, kann man kein Wachstum dann auf einmal haben nur weil man ein anderes CT Protokoll bekommt oder ein anderes knötchen und das kann schnell passieren bei der so mittelgroßen, wenn es dann so zwischen sechs und zehn Millimeter geht, da spielt sich ja viel ab in der Früherkennung. Da kann man ja, wenn man den Threshold und das merken Sie, das ist auch sehr interaktiv. Man kann den Treshold ein bisschen höher setzen, manuell, wenn einem dem Radiologen das nicht gefällt. Und deshalb der Radiologe hat volle Kontrolle. Es gibt noch ganz viele andere Software, ich hab den Markt damals gescreent, war das der absolute Frontrunner damals. Das ist natürlich ein volatiler Markt, aber ich glaube, die sind. Es gibt noch ein EU Projekt, unabhängig von uns 4-IN-THE-LUNG-RUN heißt das, haben Sie wahrscheinlich auch gehört und unabhängig von denen habe ich mich für das gleiche Programm wie die entschieden. Wusste es aber nicht. Aber es soll jetzt keinen [AI software, anonymized] oder so...Lassen Sie das mal raus, das ist ja eigentlich, mir ist es egal, ob die [AI software, anonymized] heißen oder [AI software, anonymized] oder oder [AI software, anonymized] oder so, alle bieten da was an, es ist der Inhalt, der wichtig ist.	
4.15	JM	Es bloß wichtig und ja auch genau, dass alle auch das Gleiche haben, das man dann auch wirklich vergleichbare Ergebnisse hat. Ja. Würden Sie denn sagen allgemein, dass die AI Software auch gut zu benutzen ist? Also auch klar, dass man sich da gut reinarbeiten kann?	
4.16	R4	Ja, und wir haben jetzt wieder eine neue Assistentin da eingearbeitet. Die kam ganz geflasht von den zwei Wochen wieder und sagt wow super und auch der Workflows ist anders. Und eigentlich die Radiologen, das ist super, die die die und natürlich wird dann da paar im Hintergrund die Dollarzeichen, sage ich mal, wenn das dann nach BFS erst mal von mir anfangen sollte in Deutschland. Ja, ich bin jetzt nicht finanziell getrieben, ich bin eher inhaltlich getrieben und ich denke, dass es auch gut ablaufen wird, dass das in den jungen Zentren verankert wird, institutionalisiert, und dass da solche Heuschrecken sich nicht an diesem Prozess das Geld abziehen können.	PEUT
4.17	JM	Also die Radiologen, die nutzen das System auch alle gerne oder hat da auch mal irgendjemand Angst, dass es vielleicht einem den Job wegnehmen könnte?	

4.18	R4	Es lädt ein bisschen langsam, wenn sie eine Studie laden. Es ist web-basiert, ist web-basiert, ist noch keine App. Weiß nicht, wie technisch Sie drauf sind, hat man mir jetzt noch nicht erzählt, solange man es im Browser läuft, ist immer ein Tick langsamer und ich habe es mal gestoppt. Man braucht ungefähr eine Minute Ladezeit insgesamt zwischen hochladen, Serien laden und dann wieder speichern. Ist eine halbe Minute oder 40 Minuten hochladen, 40 Sekunden hochladen und 20 Sekunden speichern. Wenn man wirklich...	PT
4.19	JM	Wie genau funktioniert das denn, der CT Scan läuft durch und später wenn es ausgewertet wird, dann läuft erst das AI System drüber, wenn die Radiologen sich das angucken oder wie?.	
4.20	R4	Ne, vorher. Also wir haben ja keine Notfälle. Der CT läuft vom Scanner und dann läuft es auf dem Server hier bei uns, sicherheitskonform ist es und da wird es dann analysiert und gleichzeitig und dann analysiert und dann haben wir das Ergebnis ist dann vollautomatisch, wird dann angezeigt beim Hochladen.	
4.21	JM	Aber wenn man das in dem Browser dann darauf zugreift, da hat man trotzdem die Ladezeiten.	
4.22	R4	Ja, da hat man dann die Ladezeit und 20 Minuten ist es Post Processing, aber das ist uns egal, weil man dann, weil wir ja keine... Was wir auch noch haben, ist eine Cloud, wo wir die drei Standorte miteinander verbunden haben, aber oben sind dann pseudonymisierte Daten.	
4.23	JH	Fehlen Ihnen noch irgendwelche Funktionen in der AI oder macht das schon alles was Sie haben wollen?	
4.24	R4	Im Moment macht es das, was wir brauchen.	OQ PU
4.25	JM	Kann sich da aber noch irgendwo Verbesserungspotential vorstellen? Ich glaube, Sie haben ja auch von allen Standorten da einen ganz guten Überblick. Funktioniert das überall super, oder wurde mal gesagt hier unterstützt das die Radiologen doch nicht ganz ideal?	
4.26	R4	Ja, es ist so natürlich mit der Anbindung, mit der IT und mit der Technik. Da hat es ein bisschen ruckelt und gehagelt und so und dann mal einen neuen Port oder was auch immer. Wir haben eine IT-Kommission, wir haben die Radiologengrüppchen, wir haben Pneumologengrüppchen und alles zusammen. Wichtig ist, dass hinten das Backbone organisiert ist. Wir haben eine [inaudible] im Hintergrund, die auch die Webseite gemacht haben und hinten das ICF gleich abgreift. Alle Daten, die wir erheben, werden auch in dieses ICF, also diese ganzen Volumina von allen Knötchen und	SR

		in die Kategorisierung, das sind ja Tonnen von Parametern, auch die Skoliose, hätte ich jetzt fast gesagt, aber Skewness, die Heterogenität der Knötchen, das alles wird automatisch erfasst, viel mehr als der Radiologe überhaupt wissen will und das schieben wir alles in unsere Datenbank. Wird automatisch abgezogen und dann haben wir eine gigantische Datenbank und machen da hinterher Forschung. Aber das nicht so wichtig für die, für die Implementierung und Implementierung ist schon wichtig, dass man national auch ein Forschungsregister hat, um sich weiterzuentwickeln. Ja.	
4.27	JM	Super. Ansonsten, ich glaube, wir hätten damit alle wichtigen Fragen gestellt, damit Sie auch pünktlich in den nächsten Termin dann können. Haben Sie noch irgendwelche Fragen die Sie stellen wollen?	
4.28	JM	Es ist einen Masterarbeit die Sie machen jetzt, oder?	
4.29	JH	Genau, ja.	
4.30	R4	Und die macht man im Team? In Deutschland muss man immer alleine machen.	
4.31	JM	Das ist in Schweden genau eine ganz besondere Sache, weil der Master auch nur ein Jahr lang geht. Genau, wir wollten Sie übrigens noch fragen, ob wir noch Member Checking mit Ihnen machen dürfen. Das heißt, indem wir eine erste Analyse dieses Gesprächs haben und dass wir Ihnen kurz sagen können, wir haben das Folgende mitgenommen, würden Sie so zustimmen, dass Sie das so gesagt haben oder haben wir da was falsch verstanden? Ob wir Ihnen das einmal zuschicken dürfen?	
4.32	R4	Das würden Sie mir dann zuschicken, ja?	
4.33	JH	Ja genau, eine kurze Zusammenfassung, dann nicken Sie das einmal ab sozusagen. Nicht, dass wir Sie ganz falsch verstanden haben.	
4.34	R4	Ja. Super. Ja, dann wünsche ich viel Glück. Ich weiß dass solche Interviews immer sehr wichtig sind. Viel Erfolg.	
4.35	JM	Vielen Dank, tschüss.	
4.36	JH	Dankeschön, vielen Dank noch Mal für Ihre Zeit.	

Appendix H - Interview 5

Duration: 39 minutes

Date: May 5, 2022

Interviewee: R5

English

Row	Speaker	Transcribed text	Codes
5.1	JH	Okay, could you say again in two to three sentences who you are and in what context have you had contact with the lung cancer screening up to now?	
5.2	R5	Yes, exactly. So I am the medical director of radiology, at the [institution, anonymized] of the [institution, anonymized] and I have been working with another employer for 15 years and have gained expertise in lung radiology and I am an internist and radiologist. And yes, and now the HANSE study has been running since last summer. That's, um, but I've already participated in screening studies of the BGs before, since this one study was already running through the professional associations, I had already participated, and now the HANSE study from last summer. But that's now the first study, when, let's say, AI is then used on a larger scale.	E
5.3	JM	Yes, if lung cancer screening also comes across Germany, how do you think that will happen? What kind of clinics? Will it be introduced in university hospitals or in large practices?	
5.4	R5	So it is bound to a low-dose process. So you need a low-dose CT for this, which of course must also be of good quality. And I think this equipment must be necessary. But I think that this will be common in ten years, also in practices and hospitals. This will not be limited to university hospitals.	
5.5	JM	Exactly. But now I say for the introduction, I mean, this is actually already supposed to come at best in the next two years or so and then first of all university clinics or where do you see the introduction there?	
5.6	R5	Not necessarily, actually.	

5.7	JM	Or where would you say? I mean, I mean, exactly, a lot of sites just don't have the hardware for the CT scans.	
5.8	R5	So you, the question is, let's say, do you continue to link this to AI as well, as the study is now structured? That is not absolutely necessary. And if that is linked to AI, then it would of course be the university sector first. But you can also simply look through a scan like this and see if there is a round focus there. This is our radiology business and I believe that it will spread relatively quickly on a larger scale.	PU
5.9	JM	Okay. Can you explain what the examinations of a patient as part of lung cancer screening in the HANSE study currently look like? What is being done to the patient or what exactly is being looked at?	
5.10	R5	Oh, there is a big clinical part, I can't say that much about it, they do, they also record functional parameters. Surveys about pre-existing conditions will of course... The inclusion and exclusion criteria must be met. So there's a big part that I don't have anything to do with as a radiologist.	
5.11	JM	Then, right then, we are more interested in the process anyway, where you as a radiologist also have contact points, i.e. what is actually examined on the patient, in the end.	
5.12	R5	Yes, that's right. These are actually three, three areas. First of all, it is this lung cancer screening, which means looking for nodules through the machine and then there is also a check for pulmonary emphysema and which also quantifies lobes, segments them. That means, the lungs are divided, so to speak, into their own lobes, and then emphysema is calculated for the individual lobes. And the third point is that coronary calcifications are determined. These are, so to speak, the three pillars of software as well. There are separate sub-programs of the software for this and, um, and the decisive or central point is, of course, the search for, the detection of the nodules, and also the characterization. The software does this too.	OQ
5.13	JH	You have already told us what functions the software has or what it is used for. Can you say again how we can imagine that? How exactly does the software then display this? So this indicates, for example, okay, here is a nodule, whether it is benign, malignant, or what exactly is displayed to you?	
5.14	R5	No, so that's that information. So the size and the volume and the density, the density are determined. Um and size and volume is the key point, so to speak. And then are the other criteria, that is, the morphological criteria, so to speak. And that would be density. And there is also a distinction between ground-glass, which	JR

		means flatter or solid parts, and then also the edge contour. So this is information that you also enter, but which I also enter as a diagnostician. One part is recognized by the machine itself and I also make supplementary, um, additions. So I can also enter, for example, whether it contains calcium. Very important point.	
5.15	JH	If I can intervene very briefly, do you say that there are things that AI doesn't fully recognize and that's why it's not complete, or that's just what the AI does not do yet and that's why it's not complete?	
5.16	R5	Um. Yes, both. So the program is focused on something specific for now and the program does a lot of things itself. But it is provided from the outset that, for example, the data for limitation, that I enter them. I think it's probably very difficult to program that too.	OQ
5.17	JM	What exact data do you enter again, for example? Can you give a few more specific examples?	
5.18	R5	So, for example, for the contour, i.e. whether there are additional malignancy criteria. I'll enter if it's directly on a fissure, the nodule. Then it becomes, then it often has benign behavior and then calcium yes/no, but the machine can also read that largely by itself. But I can also enter that manually.	OQ
5.19	JM	If you, when you type something like that, well, it is close to a fissure, is that information directly used by the system, or is it just for you?	
5.20	R5	Right then it jumps immediately to, then immediately corrects its endpoint. So the endpoint is that it does a nodule classification according to the probability of malignancy. That's that these Lung-RADS that you have, but you're familiar with it, right?	OQ
5.21	JH	Yes.	
5.22	R5	And it immediately jumps over. So if I enter additional information and what I can also enter is, for example, atelectasis in the dependent parts or I can also enter that, then it immediately jumps to another value.	
5.23	JM	So is this a lot of manual work that you still have to enter in addition or are these just short adjustments in the beginning?	
5.24	R5	No, these are important clinical parameters that complement the picture of a malignancy.	

5.25	JM	Exactly. Yes, but is it a time-consuming process to re-enter everything manually into the system, or do you say, I do this for seconds in the beginning and then the diagnosis is pretty good?	
5.26	R5	No that's quick, that's fast.	PEUT
5.27	JH	Then let's go back to the process of how this system is used. So the images, after they come out of CT, are automatically analyzed by AI, or do you have to initiate it yourself, so to speak, that the individual...	
5.28	R5	Exactly. I'm initiating that, right. I'm initiating this. I get the data, the image data loaded onto a workstation and then I open the test person and then I initiate that too. These three pillars, which I explained to you, I have to click on all of them.	
5.29	JM	It has to go through all three functions individually, or how can I imagine that?	
5.30	R5	That's exactly what I have to initiate, so to speak.	
5.31	JH	Does that take time for now? So you have to wait a few minutes, or when you call it up, does it show up immediately?	
5.32	R5	Exactly. No, no. The loading times, that is, that matters. That costs time. That's loading times, that's right.	PEUT
5.33	JM	About how long?	
5.34	R5	Oh god, there I have the...	
5.35	JM	Like a minute or five minutes?	
5.36	R5	Yes, I would say about a minute. Exactly.	
5.37	JM	Per function or all together?	
5.38	R5	No per, per function.	
5.39	JM	Can you load at least one function before, then call it, and let the other two run in the background, or do you have to, Is it unused waiting time that you then have there?	
5.40	R5	Yes that is, that is unused waiting time and I also switch between two different images in the lung window. So there is always a	PU PT

		way to do that with a harder edge display. This is how you like to look at lungs or with a softer presentation. You can also switch to soft tissues and that takes time. When I switch between them, it completely reloads again. So then only one function, but that also takes time. So one minute to one and a half minutes.	
5.41	JM	Is this in proportion to the amount of time saved by automatic evaluation, or do you lose time overall because it all takes so long?	
5.42	R5	Yes, I would say that it is still in proportion. You just wish that this would be faster in further development.	PT
5.43	JM	We have also been told that when you open the AI system, you must first verify or authenticate themselves with lots of passwords. And that it costs time again?	
5.44	R5	No. I have to authenticate myself, of course. I have a password to get in there. But this is the case with every workstation. This is a very common procedure. It's no more complicated than usual.	PEUT
5.45	JM	Is the workstation, is it, is it your workplace in general or how can I understand it? Or is that something now specialized for lung cancer screening?	
5.46	R5	No, no. This is a separate workstation, only for this software for the [AI software, anonymized], and only people who have a password have access to it. That's just the three of us doctors where I told you the [name, anonymized], no MTA either. That's a separate workstation with its own, yes, it's not really, it's like a mini PACS. These are these archiving systems. But it is not a usual PACS. It doesn't have all PACS functions. But that's the kind of thing you have to imagine. It's basically a workstation with a piece of software. But one password.	
5.47	JM	If we say now that you just had to diagnose a foot again or must immediately thereafter and then immediately go back to lung cancer screening, then you always have to log in again, right? Does that cost time again?	
5.48	R5	Exactly. Yes, that is actually a small problem. I go in and out again and yes.	
5.49	JM	Does that also cost a lot of time? How much does that cost or is that also critical?	
5.50	R5	So we are working with two people in the department at the moment and then one is in the study for a longer period of time and then does the routine.	PU PT

5.51	JM	Exactly, that's exactly why you make sure that a person is only sitting at the lung cancer screening during the course. Do you think that would also make sense for the introduction of a person who only focuses on lung cancer screening for one day? Or do you think that you can somehow integrate this into everyday clinical practice for general purposes?	
5.52	R5	Yes, I think so. You could also organize it in such a way that you have it open, for example, in a working room, then on a computer there, and that different colleagues operate it, so I could imagine that. So that can certainly be integrated well into everyday work.	
5.53	JM	But it would probably be more efficient if only one person did that, you probably get used to the field a little bit, right?	
5.54	R5	Yes, but you can also train several people there. I think so, yes. After all, it is nodule detection. So that's the core area, nodule detection, and that you actually have to control. And you have to look, so to speak, that all nodules have been detected first and the fact is that the software doesn't either, so it sometimes doesn't capture the larger, especially the larger nodules. That's an interesting phenomenon. And then you also have to watch what I told you. The software also recognizes characteristics and that they are correctly recorded. You also have to check and readjust that. And then, of course, it's about the...	OQ
5.55	JM	But then it is also evident for you how it calculates these characteristics, right? Is that clear? Or by characteristics, you mean, for example, the volume of pulmonary emphysema, right?	
5.56	R5	For example, the proportion of the soft tissue, and the proportion of the solid part.	
5.57	JM	Okay, and you can see exactly, let me say, what contours the AI system draws, how it calculated that? That you can understand it?	
5.58	R5	Exactly, and if that's correct. There are two circles, there are two circles. Or in the case of a purely solid nodule, there is only one circle and then I can see whether the volume is completely captured or not, for example, whether it has accidentally included a vessel, which then makes it a too high value. That is a classic mistake. So I have to correct the machine and then it is the case that I always see what I see as an image, that I have to classify it clinically, of course. So there are pneumonia that are only like nodules and they can look like a tumor, and of course, I have to give a rough assessment based on what the machine delivers to me. And in the other two fields, where this is the coronary calcium calculation, the machine only provides me with figures, I	OQ L

		accept them and how good the software is, is very difficult for me to assess. I don't know that, of course, but I'm basically dependent on the machine.	
5.59	JM	Okay. That is, there, you simply have to rely on that, or is there any way to verify that?	
5.60	R5	Exactly, that's exactly what I can do with emphysema when calculating emphysema, because we had the impression that emphysema is somewhat underestimated. That there is a systematic mistake in it, but proving that now is not easy at all. And of course, it is always the question of how big and how big this error is then. Is that one, two, or three percent?	L
5.61	JM	Do you think that is acceptable or do you say that it is really the case, that should definitely be improved again?	
5.62	R5	I think that's acceptable. So the software, let's say, the image data too, it's really excellent, and the image processing and the structure of the system, that's really excellent, but I think that there's still, for the future you wish maybe even sharper or possibilities of verification or control.	OQ JR L
5.63	JM	Yes. What we've also heard is, or what I'm wondering, is how good is the AI system and how well is its output actually connected to other systems that you have internally. So can you actually send the report that you generate well to patients in the end or is the AI system not yet somehow connected? You have to print this out somehow and then fax it, or how can I imagine that?	
5.64	R5	Exactly. Well, it's like I'm printing this or one of the study nurses. And at the moment, there is only one printer that can do this because it is connected to the [institution, anonymized]'s system, so to speak. So logistics actually have to be established, but I think that is more so inherent in the clinic, or that has nothing to do with this actual software.	
5.65	JM	Yes, so is that generally, I say, the internal PACS system, or something?	
5.66	R5	Yes, internal logistics.	
5.67	JM	Okay, but isn't that the AI system's fault now that it's not properly integrated yet?	
5.68	R5	No exactly, that's two different pairs of shoes. So basically that's exactly what I do, too, I just press a button, and then I have the report and then I have to transfer it back to my computer system. Exactly.	

5.69	JM	Exactly. Exactly that, the AI, exactly, actually sets a structure for the report. It's almost writing this report, right?	
5.70	R5	Exactly. That is the structure of the report, that is all programmed. Exactly.	OQ JR RD PU
5.71	JM	Does that also save a lot of time because the report is then preformulated in this way?	
5.72	R5	Yes. Yes. Yes. Totally. Yes, that is really great and I think that is very good. So this is another interesting point that we have actually still been working on and improving since the introduction of this software. We kept reporting to [country, anonymized] that we would like something else, and can you change that again? And that the software has actually been continuously improved and edited. I think that's something, actually something so special, that during a study, so to speak, it was still being worked on.	OQ JR RD PU RU
5.73	JH	So that was also feedback from you as a radiologist? What was that, for example?	
5.74	R5	They also want to have data, then later for scientific evaluation, of course, it is always about storing this data, for example, or even things that you change something again in the report or even in the software, that you, yes, I remember... So there was a problem, we did, we then also make comparisons. This means that if we have people whom we check after three or six months, then you have to make a comparison. And the plan is to repeat the test after one year anyway and to compare the result. This means that you also carry out a comparative study and for this, you have to be sure that the machine correlates the two nodules, that they are also compared, that they are matched, and that worked very well at first, and all of the sudden not. Then you set it manually and measured it separately. That wasn't good, of course. But now they've been working on it. Now there is a way to set this cleanly by hand so that exactly the nodules are matched. So that's something that was really developed.	JR PU RU
5.75	JH	Is the patient data automatically connected to the system so that it then accesses the data from the previous examination? Or how can I imagine that?	
5.76	R5	It accesses them, it has saved them.	
5.77	JM	What is currently taking in the process when you go through the CT scans with AI, takes the longest, would you say? Or how long	

		does it generally take per patient? How much do you need approximately? And which of them takes the longest?	
5.78	R5	Yes, that's difficult now, because I don't make these first reports where you routinely look through people, but I do these so-called Second Reads, which means that there are always things where there are findings and where it can be a bit more difficult to classify. I need at least ten minutes to do that.	
5.79	JM	Okay, by Second Read, do you mean when AI is the First Reader or even if one of your colleagues is the First Reader?	
5.80	R5	That is exactly what is included in the system, is it intended that if you have a higher Lung-RADS stage, that means, you have a nodule and if that is at least 3, that goes up to 4B, and if that is at least 3, then a Second Reader comes, a Second Read.	
5.81	JM	Okay, okay.	
5.82	R5	And these are the findings I'm doing. That is 3, 4A, 4B, and 4X. And of course, there may be something in between, where I say well, that's here now, that's pneumonia, that's pneumonia. Let's take this out completely, we won't let it run under Lung-RADS now. This means that the medical interpretation is already in order and is also very important, and there are also extra pneumonic findings. We also had a patient the other day who suddenly has large lesions in his chest wall and then, I called him, and it turned out in conversation, that he had malignant lymphoma years ago and that was obviously a relapse. This had nothing to do with the lungs at all, but it is an essential finding that needs to be worked up and this patient urgently needs medical treatment. And these are these things that I then practically work on.	
5.83	JM	Okay, and in the cases where the cases are usually less critical, then is there only one person looking at that with AI?	
5.84	R5	Exactly, that's the machine and one person and that actually takes 5 to 10 minutes.	
5.85	JM	Okay, or do you know what's taking the longest time in the process? So have you talked to your colleagues and they said to themselves that AI works great, but it still takes three minutes per finding to do that and that?	

5.86	R5	So everyone is a bit annoyed about the loading times. Otherwise, I think that there is now the machine for certain operations, otherwise, that's just the loading times. That they are otherwise there for certain processes somehow, it is natural if you have to correct when you start correcting, then it automatically takes longer. So if the machine now, the software, it can happen that it does not correctly detect a nodule. Then it automatically takes longer.	PU
5.87	JM	Yes, we've also heard that the system is now scrolled with the right mouse button, that this is also unusual.	
5.88	R5	Yes, exactly, I think so too.	PU PEUT
5.89	JM	Does that annoy? Is that really something that annoys you actively or is that something like that comes where you think could be different, right?	
5.90	R5	Yes, that's unfortunate, because it's the other way around, as with the other devices, and then of course you quickly make a mistake.	PEUT
5.91	JM	Wasn't that even somehow communicated to the manufacturers? Why hasn't that been changed?	
5.92	R5	Yes, no, did we... It was a comparatively small problem and in that respect, yes, you threw yourself into the decisive things after all.	RU
5.93	JM	Yes, okay, yes.	
5.94	R5	Well, so that the matching was not working was a relatively big thing, of course.	
5.95	JH	Then back to the list of questions, because we don't have that much time anymore. We would still be interested to know whether the AI systems provide you with all the functions or whether you are missing something that could really help you with the analysis.	
5.96	R5	No, the software provides everything necessary, that's practically the content of the study, so that's what it covers. The content of the study is covered.	PU
5.97	JH	So are they also relevant? Because we have also heard that sometimes there are just so many functions or information that it is actually almost overwhelming, because some of them are not relevant at all, the information and that then it tends to distract you.	

5.98	R5	I don't think so. But well, that's my business. I think I think that's very, very clear. Yes, and they have, of course, they also draw more data because they want to work with it scientifically, and now, for example, these emphysema determinations, I am simply familiar with them and so that's not too much for me either. Not if I have image data or a list of where the individual lobes are named with the percentage of emphysema values. Well, I don't think so now. I find it, i.e. the [AI software, anonymized] clearly arranged.	PEUT JR
5.99	JH	Okay, yes, then within this study, is there actually a prescribed process that you have to use the AI, or if you use it, how to use it?	
5.100	R5	So within this study? The HANSE study runs with this, with this software. Exactly, that is coupled. All three centers also use this software.	SV
5.101	JM	Do you and your colleagues actually like to use the software, or is it so, okay, we have to use it now as part of the study, or is everyone saying in general, that helps us tremendously, that we all really like to use it? What is your impression?	
5.102	R5	Yes, but that's actually fun to use, yes. But that is actually, is actually quite popular. Of course, if you sort of get the 50th case, then the satisfaction curve drops a bit, but it's basically fun to work with.	PEUT
5.103	JH	Was there actually training or an introduction to the software, or were you simply put in front of it and you had to teach yourself that?	
5.104	R5	Yes, there was. No, there was a training session, that's right.	PEUT
5.105	JH	Okay. Would you say in general that the software is intuitive? So you could theoretically do that too, is it like with an iPhone that you simply understand it right away, or is it a bit more complicated?	
5.106	R5	So it's those details, details that you actually have to know, how to document something again, or, in that respect, and you then have to know that, you can probably acquire the basics relatively quickly, but these details, that you do it correctly. So I mean, such a small detail is that we have agreed that patients will be called when they have 4, Lung-RADS 4 and that practically has to be entered, the telephone call and the date of the call are documented and how to enter that, for example, that also appears in the report. And these are things that you simply have to know how to do that.	PEUT

5.107	JM	If you say in general that there are details you should know, then there has probably been some confusion about something in the system. Is there any way or do you have a contact point where you can get help using the system?	
5.108	R5	There is, exactly. Well, once it is...	
5.109	JM	Does that work well, would you say?	
5.110	R5	All right, there is a point of contact for technical problems. I would say that works very well and in the application, these are questions of the application, I would say it works medium. I would say medium.	RU
5.111	JM	Okay, can you elaborate on that again briefly, please?	
5.112	R5	Yes, I didn't want that now, okay, that it's not that easy to communicate the problem and that it's not really understood, and then you've usually found that yourself or so that it doesn't really help that much to call and yes.	RU
5.113	JM	And on the technical issues, you say that it works well. But are there then many technical problems with the system? So does it often crash or what is there?	
5.114	R5	Yes, or there is a patient with no status. We always have a processing status that, for example, dropped out or there was also that the data was loaded over and then it didn't appear on the workstation. So there have been technical problems before.	SR
5.115	JM	Were these individual cases and were these critical individual cases?	
5.116	R5	Yes, no. I would say that these are individual cases, but nevertheless, every patient, the subject, must be worked on and must not fall behind.	
5.117	JM	They could then be solved quickly, the problems and then everything worked again?	
5.118	R5	Yes, within two days.	PT SR
5.119	JM	Oh, okay. Yes, but that is not time-critical for lung cancer screening.	
5.120	R5	It is a study after all. Yes.	

5.121	JM	Yes. Going back to that, how would you say, is the use of AI generally changing the final diagnosis and the whole process? You said there were time savings here, and there, but again, it takes a bit longer over time overall, but it evens out, right? Or are there no time savings overall? Or what would you say?	
5.122	R5	Yes, I mean, considering how much information this software provides, it's definitely a time saver. And I think the machine is doing work for us but you get around this corrective moment of the doctor, you can't avoid it at all. Yes, this classification and, above all, the interpretation of the findings and you always have to be careful that the machine doesn't do any nonsense after all.	RD PU
5.123	JM	Yes, we've also heard that there are sometimes problems with false positives. Would you say that the quality of your diagnosis is actually improved by the system, or?	
5.124	R5	So this is now about the detection of nodules. So what is a false-positive? It may be that they have a nodule compaction. But it's not a tumor, it's pneumonia. That is my interpretation and so to that extent this is called false-positive, false-negative, that is not relevant for this software now.	JR PU
5.125	JM	It can, for example, mistakenly recognize any vessel or something, or not?	
5.126	R5	Oh, yes, that's right. Yes, that is correct. Yes, it can be. That is my interpretation again. Yes, well I mean, the vessel, it is actually more about that the vessel is also measured. And then it would be a wrongly too high value. But purely theoretically, it could of course also be a small vascular malformation that is measured as a tumor. But then again, the machine recognizes the nodule but the nodule interpretation is my job.	JR PU
5.127	JM	I think by that we have finished the largest list of questions and I think we have now made good use of the half an hour. One more question at the end. Before lung cancer screening is now introduced throughout Germany and is then treated so full-time at your clinic, do you think again that there will be any major problems that will occur for you? Any hurdles? Or do you think that zack, you've already practiced all that here, you can do that right away?	
5.128	R5	So actually I don't see any hurdle. So I always think you have to differentiate whether you are doing lung cancer screening now with software support. You can also do it by looking through the scan yourself. We usually do that too. There are questions for us, is there a tumor, a long-term smoker or what do I know? Or has breast cancer, metastasis search. It's all the same and we're trained on that. That is our daily bread. In other words, the	PU

		screening will soon be carried out, which is different, whether you use software for this and whether there are still software problems, so I actually think that this software is already very well developed.	
5.129	JH	Yes.	
5.130	R5	And yet, of course, improvements are always desirable here and there, but that it is actually already very well developed.	OQ
5.131	JH	So you think there would be enough radiologists in theory to be able to work through even such a big rush for lung cancer screening? Don't you necessarily need AI software that does the work for you?	
5.132	R5	Oh, that's a tough question. I can't actually answer that. I don't know that. Well, that's where it's going to start, won't it? It won't start abruptly, and even now it's the case that many doctors simply provide this indication. They have a long-term smoker and then simply send him to a low-dose CT. Take a look at him whether, it is already being practiced, that is not something that is now being kicked off in one stroke. Um, then the other thing, of course, if you institutionalize this correctly, is that people get messaged like with breast screening, there's the question of how you're going to institutionalize that. And that's exactly what you have to prepare for, that's important. Yes.	
5.133	JH	I mean, what was that again? I believe that 5,000 people will be treated for the first time as part of the HANSE study. And I believe that there are 2 million people in Germany who would generally be targeted by the target group. Well, that's exactly what remains to be foreseen, and I also believe in the context of the HANSE study who is actually aware of this.	
5.134	R5	But they're not all willing to do this. They are not all willing to do so, and it is not entirely clear at what intervals should be screened, as those are the last questions. Yes.	
5.135	JH	Yes. Okay, thank you very much. Do you have any other questions for us? Or do you have another topic that has perhaps not been addressed far enough? But what do you still find important?	
5.136	R5	Yes, I also think that I actually got the decisive point across, so a few things, yes, this further development while working and even minor technical deficits, even though the software is already very highly developed. So that certainly. So also these image printouts, what it delivers, also in very high quality and yet, as I said, also with the matching, I also said.	OQ

5.137	JH	Exactly.	
5.138	R5	The decisive factor again is that it converges with the doctors in the interpretation, that we then, even today we have 14 people who were at Second Read, I worked on it today and then discussed several of them again with my colleague, how do you proceed with this in detail now? But that is also this Lung-RADS categorization, i.e. classification. How do you apply them correctly now? This is not always easy and there are possibilities for interpretation. I actually still find that. It is, of course, a hurdle, these Lung-RADS too, and it also makes it easier to communicate the findings, and yet they are not able to record every finding well with it.	OQ
5.139	JM	The Lung-RADS was also integrated, right? Into the software?	
5.140	R5	This is exactly what is integrated into the software. So he spits out a Lung-RADS stadium at the end.	
5.141	JM	Yes, exactly. But that means even if Lung-RADS 1.2 is introduced, for example, it has to be adjusted or how?	
5.142	R5	Exactly, that would have to be adjusted, correct, and exactly, and it even does volume calculations, which is not a matter of course with Lung-RADS. So it even does volume calculations.	OQ
5.143	JM	Yes, great. Thank you so much for the interview. It was very, very insightful. So thank you for taking the time out in the evening and yes, I think it would also be great if you could get us in touch with your colleague. That would also be immensely valuable once again and yes then...	
5.144	R5	Exactly, then I wish you all the best. When do you have to hand in this master's thesis?	
5.145	JH	End of May.	
5.146	R5	End of May. Good. Okay. All right. yeah. Well then, all the best.	
5.147	JH	Thank you.	
5.148	R5	And I don't know, does one get it somehow or would you also provide it? The master's thesis.	
5.149	JH	We'll be happy to send it to you at the end when it's ready.	

5.150	R5	After all, we are a team. That's enough for me once. [name, anonymized], [name, anonymized] and I are a team. I think we would be interested in that.	
5.151	JM	I think that after we have analyzed your interview or our interview and I'll say that we've almost written down the core findings, we would like to write you a short email so that you can confirm, that's exactly what I meant, or no, there you have misunderstood something if it is okay and you just approve that we understood it correctly.	
5.152	R5	So I'm on vacation for 14 days now.	
5.153	JM	Oh yes.	
5.154	R5	I told you that.	
5.155	JM	Yes. Exactly. That's not that bad. I also think that we're not going to get something completely wrong now. But if you can then simply confirm again briefly at the end of May. So yes, that's true.	
5.156	JH	This is also just a brief summary.	
5.157	R5	That there is no gross blunder in there, so to speak.	
5.158	JM	Yes, that's right.	
5.159	R5	Yes.	
5.160	JM	Awesome. Perfect.	
5.161	R5	Yes, have a good evening.	
5.162	JH	Have a nice evening. Bye.	

German

Row	Speaker	Transcribed text	Codes
5.1	JH	Okay, könnten Sie dann noch einmal in zwei bis drei Sätzen sagen, wer Sie sind und in welchem Zusammenhang Sie bis jetzt mit dem Lungenkrebs Screening Kontakt hatten?	
5.2	R5	Ja genau. Also ich bin ärztliche Leitung, Radiologie, Standort [institution, anonymized] von der [institution, anonymized] und die Tätigkeit übe ich aber schon mit anderem Arbeitgeber schon seit 15 Jahren aus und habe da eine Expertise in Lungenradiologie gewonnen und ich bin von Haus aus bin ich Internistin und Radiologen. Und ja und jetzt die HANSE Studie läuft ja seit letztem Sommer. Das ist, ähm, Ich habe aber vorher schon an Screening Studien der BGs teilgenommen, da lief schon über die Berufsgenossenschaften diese eine Studie, da hatte ich schon teilgenommen und jetzt die HANSE Studie ab letzten Sommer. Genau das ist aber mit KI jetzt die erste Studie, wo doch, sagen wir, in größerem Stile KI dann genutzt wird.	
5.3	JM	Ja, wenn Lungenkrebs Screening auch deutschlandweit kommt, wie denken Sie wird das kommen? Was für Kliniken, wird es in Universitätskrankenhäusern kommen oder in Großpraxen?	
5.4	R5	Also es ist ja gebunden an ein Low-Dose Verfahren. Also man braucht ja ein Niedrigdosis CT dafür, was natürlich auch qualitativ gut sein muss. Und ich denke, diese apparativen Vorrichtungen müssen notwendig sein. Aber ich denke, dass das in zehn Jahren verbreitet sein wird, auch in den Praxen und in Krankenhäusern. Das wird nicht auf Universitätskliniken beschränkt bleiben.	
5.5	JM	Genau. Aber jetzt sage ich mal für die Einführung, ich meine, das soll ja eigentlich schon bestenfalls in den nächsten zwei Jahren oder so kommen dann erst mal Unikliniken oder wo sehen Sie da die Einführung?	
5.6	R5	Eigentlich nicht unbedingt.	
5.7	JM	Oder wo würden Sie sagen? Ich meine, ich meine, genau, viele Standorte haben ja einfach nicht die, die Hardware mit den CT Scans.	

5.8	R5	Also Sie, es ist ja die Frage auch, sagen wir mal, koppelt man das weiterhin auch an KI, so wie die Studie jetzt aufgebaut ist? Das ist ja nicht unbedingt notwendig. Und wenn das an KI gekoppelt ist, dann wäre es natürlich der universitäre Bereich zunächst. Aber man kann ja auch in so einen Scan einfach durchgucken und sehen, ob ein Rundherd da ist. Das ist unser radiologisches Geschäft und ich glaube, dass das auch relativ zügig in größerem Stile Verbreitung haben wird.	PU
5.9	JM	Okay. Können Sie mal erklären, wie denn momentan die Untersuchungen eines Patienten im Rahmen des Lung Cancer Screening in der HANSE Studie aussieht? Was da alles so an den Patienten gemacht wird oder was genau sich anguckt wird?	
5.10	R5	Oh, das ist da ist ja ein großer klinischer Teil, zu dem kann ich gar nicht so viel sagen, die machen ja auch, die erheben ja auch Funktionsparameter. Da werden Befragungen natürlich zu Vorerkrankungen, es müssen ja die Ein- und Ausschlusskriterien gegeben sein. Also da ist ein großer Teil, mit dem ich als Radiologin nichts zu tun habe.	
5.11	JM	Dann, genau dann, uns interessiert sowieso mehr der Prozess, wo Sie als Radiologe dann auch Kontaktpunkte haben, also was wirklich an den Patienten dann untersucht wird, schlussendlich.	
5.12	R5	Ja, genau. Das sind eigentlich drei, drei Bereiche. Es ist einmal dieses Lung Cancer Screening, das heißt das Aufsuchen von Knoten durch die Maschine und dann wird aber auch nach einem Lungenemphysem geschaut und das quantifiziert auch Lappen, segmentiert. Das heißt, es wird die Lunge sozusagen aufgeteilt in ihre eigene, in ihre Lappen und dann wird das Emphysem berechnet für die einzelnen Lappen. Und der dritte Punkt ist, dass der Koronarkalk bestimmt wird. Das sind sozusagen die drei Säulen auch der Software. Dafür gibt es eigene Unterprogramme der Software und äh, und das entscheidende oder der zentrale Punkt ist natürlich das Aufsuchen, die Detektion der Knoten und auch die Charakterisierung. Also das macht die Software auch.	OQ
5.13	JH	Sie haben uns ja schon aufgezählt, welche Funktionen die Software hat oder wofür die eingesetzt wird. Können Sie noch mal sagen, wie wir uns das vorstellen können? Wie genau zeigt die Software das dann an? Also die zeigt zum Beispiel an okay, hier ist ein Rundherd, ob der auch gutartig ist, bösartig oder was genau wird Ihnen angezeigt?	
5.14	R5	Nee, das ist also diese Information. Also die Größe und das Volumen und die Dichte, die Dichte werden bestimmt. Ähm und Größe und Volumen ist sozusagen der entscheidende Punkt. Und dann sind die anderen Kriterien, das sind sozusagen die morphologischen Kriterien. Und das wäre einmal die die Dichte. Und da	JR

		unterscheidet man ja auch zwischen Milchglas, das sind so flauere oder solide Anteile und dann auch die Randkontur. Also das sind Informationen, die man auch, die ich aber auch als Befunder dann auch zusätzlich eingebe. Ein Teil erkennt die Maschine von sich aus und ich gebe auch, mach aber auch ergänzende ähm, also Ergänzungen. Also ich kann ergänzend eingeben auch zum Beispiel ob der Kalk enthält. Ganz wichtiger Punkt.	
5.15	JH	Wenn ich da mal ganz kurz eingreifen kann, sagen Sie einfach, es gibt Sachen, die da die AI nicht komplett erkennt und deswegen ist es nicht vollständig oder das macht die AI einfach noch nicht und deswegen ist nicht vollständig?	
5.16	R5	Ähm. Ja beides. Also das Programm ist ja auf was Bestimmtes erst mal fokussiert und vieles macht das Programm selber. Aber es ist eben von vornherein vorgesehen, dass zum Beispiel die Daten zur Begrenzung, dass ich die eingebe. Ich denke, das ist sehr schwer wahrscheinlich, das dann auch mit zu programmieren.	OQ
5.17	JM	Was für Daten geben Sie zum Beispiel noch mal genau ein? Können Sie da noch ein paar konkrete Beispiele nennen?	
5.18	R5	Also zum Beispiel gebe ich also zur Kontur, also ob zusätzliche Malignitätskriterien bestehen. Ich gebe ein, ob der direkt an einer Fissur liegt, der Knoten. Dann wird er, dann hat er häufig ein benignes Verhalten und dann auch Kalk Ja/Nein, aber das kann die Maschine auch großenteils selber lesen. Aber das kann ich auch händisch eingeben.	OQ
5.19	JM	Wenn Sie, wenn Sie sowas eingeben, wie zum Beispiel, dass das na ja, nah an einer Fissur dran liegt, wird es dann direkt diese Information vom System verwertet oder ist es dann nur für Sie?	
5.20	R5	Genau dann springt er sofort auf, dann korrigiert er sofort seinen Endpunkt. Also der Endpunkt ist ja, dass er macht ja eine eine Knotenklassifizierung nach der Wahrscheinlichkeit der Malignität. Das ist dieses Lung-RADS, das haben, aber das Ihnen geläufig, ne?	OQ
5.21	JH	Ja.	
5.22	R5	Und der springt dann sofort um. Also wenn ich eine zusätzliche Information eingebe und ich, was ich auch eingeben kann, sind zum Beispiel eine Atelektase in den abhängigen Partien oder das kann ich auch eingeben, dann springt er nämlich auch sofort auf einen anderen Wert um.	

5.23	JM	Also ist das viel händische Arbeit, die Sie da immer noch zusätzlich eingeben müssen oder sind das nur kurze Anpassungen anfangs?	
5.24	R5	Nein, das sind wichtige klinische Parameter, die das Bild eines Malignoms ergänzen.	
5.25	JM	Genau. Ja, aber ist es ein zeitintensiver Prozess da händisch das nochmal alles einzugeben in das System oder sagen Sie das mache ich zehn Sekunden am Anfang und danach ist die Diagnose ziemlich gut?	
5.26	R5	Ne das geht schnell, das geht schnell.	PEUT
5.27	JH	Dann vieleicht nochmal zum Prozess, wie dieses System eingesetzt wird. Also werden die Bilder, nachdem die aus dem CT kommen, werden die automatisch von der AI analysiert oder müssen Sie es sozusagen selber anleiten, dass die einzelnen...	
5.28	R5	Genau. Ich initiiere das, genau. Ich initiiere das. Ich bekomme die Daten, die Bilddaten auf eine Workstation geladen und dann rufe ich den Probanden auf und dann initiiere ich das auch. Diese drei Säulen, was ich Ihnen erklärt hatte, das muss ich alles anklicken.	
5.29	JM	Es müssen alle drei Funktionen einzeln durchlaufen, oder wie kann ich mir das vorstellen?	
5.30	R5	Genau die muss ich sozusagen initiieren.	
5.31	JH	Dauert das dann erstmal? Also müssen Sie ein paar Minuten warten oder wenn Sie es aufrufen, wird es eigentlich auch angezeigt?	
5.32	R5	Genau. Nein, nein. Die Ladezeiten, das ist, das schlägt zu Buche. Das kostet Zeit. Das sind Ladezeiten, genau.	
5.33	JM	Wie lange ungefähr?	
5.34	R5	Oh gott, da habe ich die...	
5.35	JM	So eine Minute oder fünf Minuten?	
5.36	R5	Ja, ich würde sagen so eine Minute. Genau.	
5.37	JM	Pro Funktion oder alles zusammen?	

5.38	R5	Ne pro pro pro Funktion.	
5.39	JM	Können Sie wenigstens eine Funktion schon mal laden, die dann aufrufen, die anderen beiden im Hintergrund laufen lassen oder müssen Sie explizit auch wirklich? Ist es ungenutzte Wartezeit, die Sie dann da haben?	
5.40	R5	Ja das ist , das ist ungenutzte Wartezeit und ich wechsel auch zwischen zwei verschiedenen Bilddarstellungen im Lungenfenster. Also da gibt es immer eine Möglichkeit, das mit einer härteren Kantendarstellung. So guckt man sich gerne Lunge an oder mit einer weicheren Darstellung. Da kann man auch auf Weichteile umstellen und das dauert auch. Wenn ich dazwischen wechsle, dann lädt er wieder ganz komplett neu. Also dann nur die eine Funktion, aber das dauert eben auch. Also eine Minute bis eineinhalb Minuten.	PU PT
5.41	JM	Steht das im Verhältnis mit der Zeit, die man sich einspart, durch die automatische Auswertung oder verlieren Sie dadurch insgesamt Zeit, dass das alles so lange geht?	
5.42	R5	Doch, das würde ich sagen. Das steht noch im Verhältnis. Man wünschte sich nur in der Weiterentwicklung, dass das schneller geht.	PT
5.43	JM	Uns wurde auch schon zu Ohren getragen, dass Sie sich, wenn Sie das AI System öffnen, erst mal verifizieren müssen oder authentifizieren müssen mit ganz vielen Passwörtern. Und das das auch noch Mal Zeit kostet?	
5.44	R5	Nein. Ich muss mich natürlich authentifizieren. Ich habe ein Passwort, um dort hineinzugehen. Das ist aber bei jeder Workstation so. Das ist ein ganz übliches Verfahren. Das ist nicht komplizierter, als als üblich ist.	
5.45	JM	Ist die Workstation, ist die, ist die allgemein Ihr Arbeitsplatz oder wie kann ich es verstehen? Oder das spezialisiert etwas jetzt für das Lung Cancer Screening?	
5.46	R5	Nein, nein. Das ist eine eigene Workstation, nur für diese Software für die [AI software, anonymized] und da kommen kommen auch nur Leute, die dann ein Passwort haben ran. Das sind nur wir drei Ärzte wo ich Ihnen sagte der [name, anonymized], auch keine MTA. Das ist eine eigene Workstation mit einem eigenen, Ja, das ist eigentlich kein, es ist wie so ein Mini PACS. Das sind diese Archivierung Systeme. Aber es ist nicht ein übliches PACS. Es hat nicht alle PACS Funktionen. Aber so ein bisschen muss man sich das vorstellen. Es ist im Grunde eine Workstation mit einer Software. Aber ein Passwort.	

5.47	JM	Wenn wir jetzt mal sagen, Sie mussten gerade oder müssen danach direkt wieder einen Fuß diagnostizieren und danach aber direkt wieder Lung Cancer Screening, dann müssen Sie sich ja auch immer umloggen, oder? Das kostet auch noch mal Zeit?	
5.48	R5	Genau. Ja, genau das ist tatsächlich ein kleines Problem. Ich geh rein und wieder raus und ja.	
5.49	JM	Kostet das auch viel Zeit? Wie viel kostet das so oder ist das auch kritisch?	
5.50	R5	Also wir arbeiten im Moment mit zwei Leuten da in der Abteilung und dann ist eben einer auch längere Zeit dann in der Studie und macht dann die Routine.	PU PT
5.51	JM	Genau, genau deswegen machen Sie dann auch, dass eine Person das durchgänging nur an dem Lung Cancer Screening sitzt. Denken Sie, das würde auch für die Einführung dann Sinn ergeben, dass man eine Person dann auch abstellt, die sich für einen Tag immer nur auf Lung Cancer Screening fokussiert? Oder denken Sie, man kann das auch irgendwie für allgemein in den klinischen Alltag integrieren?	
5.52	R5	Ja, das glaube ich schon. Man könnte das ja auch so organisieren, dass man zum Beispiel in einem Arbeitsraum, dann an einem Computer dort die offen hat und dass unterschiedliche Kollegen das bedienen, also das könnte ich mir so vorstellen. Also das kann man sicherlich im Arbeitsalltag gut integrieren.	
5.53	JM	Aber effizienter wäre wahrscheinlich wenn nur eine Person das macht, man guckt sich da wahrscheinlich auch so ein bisschen ein auf das Feld, oder?	
5.54	R5	Ja, aber da kann man auch mehrere Leute einarbeiten. Das denke ich schon, ja. Es ist ja eine Knotendetektion. Also das ist ja der Kernbereich, die Knotendetektion und die Sie eigentlich kontrollieren müssen. Und Sie müssen gucken, sozusagen, dass erst mal alle Knoten erfasst sind und da ist es tatsächlich so, dass die Software auch nicht, also die größeren, gerade die größeren Knoten erfasst sie manchmal nicht. Das ist ein interessantes Phänomen. Und dann muss man aber auch gucken, was ich Ihnen sagte. Die Software erkennt ja auch Charakteristika und dass die auch richtig erfasst sind. Das muss man auch mal nachkontrollieren, nachstellen. Und dann geht es natürlich auch um die...	OQ
5.55	JM	Dann ist aber auch für Sie ersichtlich, wie es diese Charakteristika berechnet, oder? Das ist klar? Oder mit Charakteristika meinen Sie jetzt zum Beispiel Volum des Lungenemphysem, oder?	

5.56	R5	Zum Beispiel Anteil Weichteile, Anteil solider Teil.	
5.57	JM	Okay und da sehen Sie aber genau, sage ich mal, welche Konturen das AI System zieht, wie es das berechnet hat? Dass Sie es nachvollziehen können?	
5.58	R5	Genau und ob das korrekt ist. Da sind zwei Kreise, da sind zwei Kreise. Oder eben bei einem rein soliden Herd ist nur ein Kreis und dann kann ich ja sehen, ob das Volumen komplett erfasst ist oder nicht, oder ob er zum Beispiel versehentlich ein Gefäß mit einbezogen hat, was dann ja einen zu hohen Wert macht. Das ist ein klassischer Fehler. Also da muss ich die Maschine korrigieren und dann ist es so, dass ich immer das, was ich als Bild sehe, dass ich das natürlich klinisch einordnen muss. Also es gibt Lungenentzündungen, die nur so herdförmig sind und die können aussehen wie ein Tumor und da muss ich natürlich schon anhand dessen, was die Maschine mir liefert, eine grobe Einschätzung geben. Und bei den anderen beiden Feldern, also wo das die Koronarkalk Berechnung ist, da liefern liefert die Maschine mir ja nur Zahlen, die nehme ich hin und wie gut da die Software ist, das ist schon für mich sehr schwer zu ermessen. Da habe ich natürlich nicht Ahnung, aber da bin ich im Grunde von der Maschine abhängig.	OQ L
5.59	JM	Okay. Das heißt, da, da müssen Sie sich einfach drauf verlassen, oder gibt es da irgendeine Möglichkeit das zu überprüfen?	
5.60	R5	Genau das kann ich natürlich bei dem Emphysem, bei der Emphysemberechnung, da hatten wir den Eindruck, dass das etwas unterschätzt wird das Emphysem. Das da ein systematischer Fehler drin ist, aber das jetzt zu belegen, ist gar nicht einfach. Und es ist natürlich auch immer die Frage, wie groß, wie groß dann dieser Fehler ist. Sind das ein, zwei drei Prozent?	
5.61	JM	Denken Sie das im Rahmen oder sagen Sie, das ist schon wriklich so, das sollte auf jeden Fall noch mal verbessert werden?	
5.62	R5	Ich denke, das ist im Rahmen. Also die Software, sagen wir mal so, auch die Bilddaten, die sind wirklich hervorragend und die Bildaufarbeitung und die Struktur des Systems, das ist wirklich hervorragend, aber ich glaube, dass da noch so, da wünschte man sich noch in die Zukunft hinein, vielleicht so noch schärfere oder Möglichkeiten der Verifizierung oder der Kontrolle.	OQ JR L

5.63	JM	Ja. Was wir auch schon gehört haben ist, oder was ich mich frage ist, wie gut ist das AI System und sag ich mal, der, der Output eigentlich angeschlossen an andere Systeme, die Sie intern haben. Also können Sie eigentlich den Report, den Sie da mit generieren, schlussendlich auch gut an die Patienten schicken, oder? Oder ist das AI System noch gar nicht irgendwie angebunden? Sie müssen irgendwie das ausdrucken und dann faxen, oder wie kann ich mir das vorstellen?	
5.64	R5	Genau. Also, es ist so, dass ich das ausdrucke oder eine der Study Nurses. Und da gibt es auch im Moment nur einen Drucker, über den laufen kann, weil es sozusagen an das System der [institution, anonymized] angeschlossen ist. Also da muss tatsächlich eine Logistik hergestellt werden, aber ich denke, das ist mehr so Klinik immanent, oder das hat nichts mit diesem eigentlichen Softwareprogramm zu tun.	
5.65	JM	Ja, also ist das allgemein, sage ich mal, das interne PACS System, oder so?	
5.66	R5	Ja, interne Logistik.	
5.67	JM	Okay, aber das ist jetzt nicht die Schuld des AI Systems, dass das noch nicht richtig integriert ist?	
5.68	R5	Nein genau, das sind zwei verschiedene Paar Schuhe. Also im Grunde ist das, genau das ist auch ich druck einfach auf einen Knopf und dann habe ich den Report und dann muss ich den aber noch mal in mein Computersystem übertragen. Genau.	
5.69	JM	Genau. Also genau das. die AI genau, die gibt eigentlich schon eine Struktur für den Report vor. Die schreibt schon fast diesen Report, richtig?	
5.70	R5	Genau. Das ist die Struktur des Reports, das ist alles programmiert. Genau.	OQ JR RD PU
5.71	JM	Nimmt das auch viel Zeit ab, dadurch, dass der Report dann so vorformuliert ist?	
5.72	R5	Ja. Ja. Ja. Total. Ja, das ist wirklich sehr und da ist das finde ich schon sehr gut. Also noch ein interessanter Punkt, dass wir auch seit der Etablierung jetzt dieser Software eigentlich immer noch wieder dran gearbeitet haben und verbessert haben. Wir haben immer wieder nach [country, anonymized] gemeldet, da hätten wir gerne noch was anderes und können Sie das noch mal umstellen? Und dass eigentlich die Software auch laufend noch	OQ JR RD PU RU

		verbessert wurde und bearbeitet wurde. Ich glaube, das ist etwas, eigentlich sowas besonderes sogar, das während einer Studie sozusagen, da auch noch dran gearbeitet wurde.	
5.73	JH	Also das war dann auch Feedback von Ihnen als als Radiologe? Was war das zum Beispiel?	
5.74	R5	Die wollen ja auch Daten haben, dann später für die wissenschaftliche Auswertung, da geht es natürlich auch immer um die Speicherung dieser Daten, da zum Beispiel, oder auch Sachen, die, dass man im Report noch mal was ändert oder auch sogar in der Software, dass man ja, das fällt mir... Also es war ein Problem, wir haben, wir machen dann ja auch Vergleich. Das heißt, wenn wir Leute haben, die wir nach drei oder sechs Monaten kontrollieren, dann muss man ja einen Vergleich machen. Und geplant ist ja, dass eh nach einem Jahr noch mal gefahren wird und verglichen wird. Das heißt, man macht auch eine vergleichende Untersuchung und dazu müssen Sie ja sicher sein, dass die Maschine die beiden Herde korreliert, die auch verglichen, die gematcht werden und das hat erst sehr gut funktioniert und auf einmal nicht. Dann hat man das händisch eingestellt, hat das getrennt gemessen. Das war natürlich nicht gut. Jetzt haben die aber dran gearbeitet. Jetzt gibt es eine Möglichkeit, das aber händisch sauber einzustellen, dass exakt die Noduli gematcht sind. Also das ist so was, wo wirklich entwickelt wurde.	JR PU RU
5.75	JH	Sind die Patientendaten automatisch ans System angeschlossen, dass es sich dann die Daten von der vorigen Untersuchung greift? Oder wie kann ich mir das vorstellen?	
5.76	R5	Die greift er sich, die hat er gespeichert.	
5.77	JM	Was dauert momentan im Prozess, wenn Sie die CT Scans dann mit der AI durchgehen, was dauert da am längsten, würden Sie sagen? Oder wie lange dauert es allgemein pro Patient? Wieviel brauchen Sie da ungefähr? Und was davon am längsten?	
5.78	R5	Ja, das ist jetzt schwierig, weil ich mache diese ersten Befundungen mache ich praktisch nicht, wo man routinemäßig die Leute so durchguckt, sondern ich mache diese sogenannten second Reads, das heißt, es sind ja immer Sachen, wo Befunde sind und wo es auch ein bisschen schwieriger in der Einordnung sein kann. Also ich brauche dafür dann mindestens zehn Minuten.	
5.79	JM	Okay, meinen Sie mit Second Read jetzt, wenn die AI der First Reader ist oder auch wenn einer Ihrer Kollegen der First Reader ist?	

5.80	R5	Genau das ist ja praktisch im System, ist das so vorgesehen, dass Sie, wenn Sie ein höheres Lung-RADS Stadium haben, das heißt, Sie haben einen Knoten und wenn der mindestens 3 ist, das geht ja bis 4B und wenn der mindestens 3 ist, dann kommt ein zweiter Leser ein Second Read.	
5.81	JM	Okay, okay.	
5.82	R5	Und das sind die Befunde, die ich mache. Also 3, 4A, 4B, 4X. Und da kann natürlich auch mal was dazwischen sein, wo ich sage na, das ist doch hier jetzt, das ist doch eine Pneumonie, das ist eine Lungenentzündung. Dies nehmen wir mal komplett raus, den lassen wir jetzt nicht unter Lung-RADS laufen. Das heißt, da ist schon die ärztliche Interpretation angesagt und auch ganz wichtig und es gibt auch extra pneumolane Befunde. Wir hatten neulich auch einen Patienten, der hat mit ein mal in der Brustwand große Raumforderungen und dann hat sich, habe ich den angerufen, hat sich im Gespräch herausgestellt. Er hatte vor Jahren ein malignes Lymphom und das war offensichtlich dann ein Rezidiv. Das hat überhaupt nichts mit der Lunge zu tun gehabt, ist aber ja ein wesentlicher Befund, der aufgearbeitet werden muss und der muss ja auch dringlich dieser Patient in ärztliche Behandlung. Und das sind diese Dinge, die praktisch ich dann bearbeite.	
5.83	JM	Okay, und in den Fällen, wo die die Fälle weniger kritisch sind normalerweise, dann gibt es nur eine Person, die sich das anguckt mit AI?	
5.84	R5	Genau, das ist die Maschine und eine Person und das dauert eigentlich so 5 bis 10 Minuten.	
5.85	JM	Okay, oder wissen Sie da, was da so am längsten Zeit noch beansprucht wird in dem Prozess? Also haben Sie da mal mit den Kollegen geredet und die haben sich gesagt so AI funktioniert super, aber es dauert noch drei Minuten pro Befund noch das und das zu machen?	
5.86	R5	Also es sind alle etwas genervt über die Ladezeiten. Ansonsten denke ich, dass da jetzt also die Maschine für bestimmte Vorgänge, Jetzt sonst, das sind nur die Ladezeiten. Dass sie sonst da für bestimmte Vorgänge irgendwie, es ist natürlich, wenn man korrigieren muss, wenn man anfängt zu korrigieren, dann dauert das automatisch länger. Also wenn die Maschiene jetzt, die Software, das kann ja passieren, dass die mal einen Herd nicht korrekt erfasst. Dann dauert es automatisch gleich länger.	PU

5.87	JM	Ja, wir haben auch mal gehört, dass man bei dem System jetzt mit der rechten Maustaste scrollt, dass das auch ungewohnt ist.	
5.88	R5	Ja, genau, das finde ich auch.	PU
5.89	JM	Nervt das? Ist das wirklich etwas was Sie aktiv nervt oder ist das so etwas kommt, wo sie denken, könnte auch anders sein, oder?	
5.90	R5	Ja, das ist unglücklich, weil es anders herum ist, wie an den anderen Apparaturen und dann macht man natürlich schnell mal einen Fehler.	
5.91	JM	Wurde das nicht mal irgendwie vielleicht an die Hersteller kommuniziert? Wieso wurde das nicht geändert?	
5.92	R5	Ja, ne, haben wir... Es war ein vergleichsweise kleines Problem und insofern ja, man hat sich dann doch auf die entscheidenden Sachen geworfen.	RU
5.93	JM	Ja, okay, ja.	
5.94	R5	Also, ich glaube, dass das Matching nicht funktionierte war natürlich ein relativ großes Ding.	
5.95	JH	Dann noch mal zurück zu dem Fragenkatalog, weil wir ja nicht mehr so viel Zeit haben. Uns würde noch interessieren, ob die AI Systeme Ihnen alle Funktionen zur Verfügung stellen oder ob Sie dann noch irgendwas vermissen, was Ihnen noch wirklich weiterhelfen könnte bei der Analyse.	
5.96	R5	Nö da stellt die Software alles Notwendige, das ist ja praktisch der Studieninhalt, also insofern das wird abgedeckt. Der Studieninhalt wird abgedeckt.	PU
5.97	JH	Also sind sie auch relevant? Weil wir haben auch gehört, dass es teilweise einfach so viele Funktionen oder Informationen gibt, dass das eigentlich auch fast überfordernd ist, weil die teilweise gar nicht relevant sind, die Information und dass dann einen eher ablenkt.	

5.98	R5	Das finde ich nicht. Aber gut, das ist so mein Geschäft. Ich finde, ich finde, das ist sehr, sehr übersichtlich ist. Ja und die haben natürlich, die ziehen auch weitere Daten, weil sie damit wissenschaftlich arbeiten wollen und nun sind zum Beispiel diese Emphysembestimmungen, ich bin einfach damit vertraut und insofern ist das auch für mich keine Überforderung. Nicht, wenn ich da so Bilddaten habe oder eine Auflistung, wo die einzelnen Lappen genannt sind mit den prozentualen Emphysemwerten. Also das finde ich jetzt nicht. Ich finde es, also die [AI software, anonymized] finde ich schon übersichtlich.	PEUT JR
5.99	JH	Okay, ja, gibt es dann im Rahmen der Studie eigentlich einen vorgeschriebenen Prozess, dass man die AI benutzen muss oder wenn man sie benutzt, wie man sie benutzt?	
5.100	R5	Also innerhalb dieser Studie? Die HANSE Studie, die läuft mit dieser, mit dieser Software. Genau, das ist gekoppelt. Auch alle drei Zentren benutzen diese Software.	SV
5.101	JM	Benutzen Sie und Ihre Kollegen die Software denn eigentlich gerne oder ist es so, okay, wir müssen sie jetzt im Rahmen der Studie benutzen oder sagen alle allgemein, das hilft uns ungemein, wir benutzen sie alle super gerne? Was ist da so Ihr Eindruck?	
5.102	R5	Ja doch das macht eigentlich schon Spaß, sie zu benutzen, ja. Doch das ist eigentlich, ist eigentlich ganz beliebt. Natürlich, wenn man jetzt irgendwie den 50. Fall macht, dann sinkt so ein bisschen die Zufriedenheitskurve, aber im Grunde macht es Spaß, damit zu arbeiten.	PEUT
5.103	JH	Gab es dann eigentlich auch ein Training oder eine Einführung in die Software oder wurden Sie einfach davor gesetzt und Sie mussten sich das selber beibringen?	
5.104	R5	Ja gab es. Ne, es gab ein Training, genau.	PEUT
5.105	JH	Okay. Würden Sie denn allgemeinen sagen, dass die Software intuitiv ist? Also könnte man sich das theoretisch auch, ist es so wie bei einem iPhone, dass man das einfach auf Anhieb versteht oder ist das schon ein bisschen komplizierter?	
5.106	R5	Also es sind so Details, Details, die man eigentlich dann wissen muss, wie man noch mal was dokumentiert, oder, also insofern, und das muss man dann wissen, also die Grundzüge kann man wahrscheinlich relativ zügig sich aneignen, aber diese Details, dass Sie das so korrekt dann machen. Also ich meine, so ein kleines Detail ist, wir haben vereinbart, dass ja die Patienten dann angerufen werden, wenn sie 4 haben, Lung-RADS 4 und da muss	PEUT

		praktisch eingegeben werden, dokumentiert werden das Telefonat und das Datum des Telefonats und wie man das zum Beispiel eingibt, das erscheint dann auch im Report. Und das sind Sachen, die man einfach wissen muss, so wie das geht.	
5.107	JM	Wenn Sie allgemein sagen, es gibt Details man wissen sollte, dann dann gab es wahrscheinlich schon mal Unklarheiten zu irgendwas im System. Gibt es da irgendwie oder haben Sie da eine Anlaufstelle, wo Sie sich Hilfe holen können für die Benutzung des Systems?	
5.108	R5	Gibt es genau. Also, es ist einmal..	
5.109	JM	Funktioniert das auch gut, würden Sie sagen?	
5.110	R5	Also gut, es gibt eine Anlaufstelle für technische Probleme. Das würde ich sagen läuft sehr gut und in der Applikation, das sind ja Fragen der Applikation, ich sage mal so mittel. Da würde ich mal so sagen mittel.	RU
5.111	JM	Okay, können Sie es noch mal näher ausführen kurz, bitte?	
5.112	R5	Ja, ich wollte da jetzt nicht so, also gut, dass es nicht so einfach ist, das Problem rüberzubringen und dass es dann auch nicht richtig verstanden wird und dann hat man meistens selber irgendwie schon auch das selber gefunden oder so, dass das gar nicht so viel bringt, dann noch groß anzurufen und ja.	RU
5.113	JM	Und zu den technischen Problemen sagen Sie, das funktioniert gut. Aber gibt es dann viele technische Probleme mit dem System? Also stürzt oft ab oder was gibt es da?	
5.114	R5	Ja oder es ist mal ein Patient ohne Status da. Wir haben ja immer einen Bearbeitungsstatus, dass der da zum Beispiel rausgefallen ist oder es gab auch, dass die Daten rüber geladen waren und dann sind die aber auf der Workstation nicht erschienen. Also es gab schon schon mal technische Probleme.	SR
5.115	JM	Waren das jetzt so Einzelfälle und waren das kritische Einzelfälle?	
5.116	R5	Ja, nein. Ich würde sagen, das sind Einzelfälle, aber trotzdem muss ja jeder Patient, der Proband bearbeitet werden und darf nicht hinten runtfallen.	

5.117	JM	Die könnten dann auch schnell gelöst werden, die Probleme und dann hat auch alles wieder funktioniert?	
5.118	R5	Ja, also innerhalb von ein zwei Tagen.	PT SR
5.119	JM	Oh, okay. Ja, aber das genau ist ja nicht zeitkritisch bei einem Lung Cancer Screening.	
5.120	R5	Genau es ist eben letztlich eine Studie. Ja.	
5.121	JM	Ja. noch mal darauf zurück, wie würden Sie sagen, verändert allgemein die Nutzung der AI, die schlussendliche Diagnose und den ganzen Prozess. Sie haben gesagt, Zeiteinsparungen gibt es hier, da, da dauert es aber wieder ein bisschen länger mit der Zeit insgesamt, aber gleich es sich ungefähr aus, richtig? Oder es gibt insgesamt keine Zeiteinsparungen? Oder was würden Sie sagen?	
5.122	R5	Ja, ich meine, wenn man bedenkt, wie viel Informationen diese Software dann ja liefert, ist es auf jeden Fall eine Zeiteinsparung. Und ich denke, die Maschine arbeitet zu. Aber man kommt um dieses korrigierende Moment des Arztes, kommt man überhaupt nicht herum. Ja, diese Einordnung und und vor allen Dingen die Interpretation des Befundes und man muss auch immer aufpassen, dass die Maschine nicht doch, nicht doch irgendwelchen Unsinn macht.	RD PU
5.123	JM	Ja, wir haben auch schon gehört, dass da teilweise auch Probleme mit falsch-positiven gibt. Würden Sie sagen, dass eigentlich die die Qualität Ihrer Diagnose verbessert wird durch das System, oder?	
5.124	R5	Also hier geht es jetzt ja um die Detektion von Knoten. Also was ist da falsch-positiv? Es kann sein, dass die eine Herdverdichtung haben. Die ist aber kein Tumor, sondern es ist eine Lungenentzündung. Das ist meine Interpretation und also insofern ist dieses mit falsch-positiv, falsch-negativ, das ist jetzt nicht bei dieser Software relevant.	JR PU
5.125	JM	Es kann ja auch zum Beispiel irgendein Gefäß oder so fälschlicherweise erkennen oder nicht?	
5.126	R5	Ach so, ja, richtig. Ja, das ist richtig. Ja, das kann sein. Das ist wieder meine Interpretation. Ja, das kann, also gut ich meine, das Gefäß, da ist es eigentlich mehr, dass das Gefäß mit gemessen wird. Und dann wäre es ein falsch zu hohe Wert. Aber rein theoretisch könnte es natürlich auch sein oder eine kleine Gefäßmissbildung, dass die gemessen wird als Tumor. Das ist aber dann	JR PU

		wiederum die Maschine erkennt den Herd. Aber die Herdinterpretation ist mein Ding.	
5.127	JM	Ich denke, damit haben wir so den größten Fragenkatalog und haben jetzt ja auch die halbe Stunde, glaube ich auch gut ausgenutzt. Noch mal eine Frage zum Ende. Denken Sie, bevor Lung Cancer Screening jetzt deutschlandweit eingeführt wird und dann auch bei Ihrer Klinik so fulltime behandelt wird, noch mal, dass dann noch mal irgendwelche großen Probleme gibt, die auf Sie zukommen? Irgendwelche Hürden? Oder denken Sie, dass Zack, das haben Sie schon alles hier geübt, das können Sie direkt?	
5.128	R5	Also eigentlich sehe ich keine Hürde. Also ich denke immer, man muss unterscheiden, ob Sie das Lung Cancer Screening jetzt mit einer Software Unterstützung machen. Sie können es ja auch machen, indem Sie selber den Scan durchgucken. Das machen wir ja sonst auch. Es gibt ja Fragen an uns, besteht da ein Tumor, langjähriger Raucher oder oder was weiß ich? Oder oder hat eine Mammakarzinom, Metastasensuche. Das ist ja alles das gleiche und darauf sind wir ja geschult. Das ist unser tägliches Brot. Das heißt, in dem Sinne wird das Screening ja auch dann bald durchgeführt werden, was anderes ist, ob man eine Software nutzt dafür und ob da noch Softwareprobleme, also ich finde eigentlich, dass diese Software schon sehr weit entwickelt ist.	PU
5.129	JH	Ja.	
5.130	R5	Und trotzdem natürlich immer hier und da auch noch Verbesserungen wünschenswert aber, dass die eigentlich schon doch sehr weit entwickelt ist.	
5.131	JH	Also denken Sie, da wären theoretisch genug Radiologen vorhanden, um auch so einen großen Ansturm auf das Lung Cancer Screening dann abarbeiten zu können? Da braucht man nicht zwingend eine AI Software, die einem die Arbeit abnimmt?	
5.132	R5	Oh, das ist eine schwierige Frage. Das kann ich eigentlich nicht beantworten. Das weiß ich nicht. Da wird ja auch erst mal anlaufen, nicht? Es wird ja nicht schlagartig losgehen und auch jetzt ist es ja schon so, dass viele Ärzte auch diese Indikation einfach stellen. Die haben einen langjährigen Raucher und schicken den dann einfach mal zum Niedrigdosis CT. Guckt mal drauf, ob der, das heißt, das wird ja schon praktiziert, das ist ja nichts was jetzt mit einem Schlag dann losgetreten wird. Ähm, dann die andere Sache ist natürlich, wenn man das richtig institutionalisiert, dass so wie bei Mammascreening die Leute angeschrieben werden, da ist es die Frage, in welcher Form man das institutionalisieren	

		wird. Und darauf, genau darauf müsste man sich dann vorbereiten, das ist wichtig. Ja.	
5.133	JH	Ich meine, wie war das noch mal? Ich glaube, im Rahmen der HANSE Studie werden 5000 Leute erst mal behandelt. Und ich glaube, deutschlandweit gibt es 2 Millionen Personen, auf die die Zielgruppe allgemein zutreffen würde. Also, genau das bleibt ja auch abzusehen, ich glaube auch im Rahmen der HANSE Studie, wer überhaupt das wahrnimmt.	
5.134	R5	Die sind aber nicht alle bereit. Die sind auch nicht alle bereit dazu und es ist ja auch nicht ganz klar in welchen Abständen da gescreent werden sollte,das sind ja noch letzte Fragen. Ja.	
5.135	JH	Ja. Okay, vielen Dank. Haben Sie sonst noch irgendwelche Fragen an uns? Oder haben Sie noch ein Thema, was vielleicht noch nicht weit genug angesprochen wurde? Was Sie aber trotzdem als wichtig empfinden.	
5.136	R5	Ja ne ich glaube auch, ich habe eigentlich glaube ich so das Entscheidende eigentlich so rübergebracht, also so ein paar Sachen ja auch diese Weiterentwicklung noch während des Arbeitens und schon auch mal kleinere technische Defizite, obwohl die Software schon sehr hoch entwickelt ist. Also das mit Sicherheit. Also auch diese Bildausdrucke, was sie liefert, auch in sehr hoher Qualität und trotzdem, wie ich erzählte, auch mit dem Matching, habe ich auch gesagt.	
5.137	JH	Genau.	
5.138	R5	Das Entscheidende eigentlich auch immer wieder, dass es in der Interpretation bei den Ärzten zusammenläuft, dass wir dann, auch heute haben wir 14 Leute, die im Second Read waren, habe ich heute bearbeitet und da dann doch mehrere auch mit meinen Kollegen noch mal besprochen, wie geht man bei dem im Einzelnen jetzt vor? Aber das ist auch diese Lung-RADS Kategorisierung, also Klassifikation. Wie wendet man die jetzt auch korrekt an? Das ist nicht immer ganz einfach und da sind Interpretationsmöglichkeiten. Das finde ich tatsächlich noch. Es ist natürlich ein Krückstock, auch diese Lung-RADS und erleichtert auch die Verständigung über die Befunde und trotzdem können sie auch nicht jeden Befund damit gut erfassen.	OQ
5.139	JM	Das Lung-RADS das war auch integriert, richtig? In die Software?	
5.140	R5	Genau das ist integriert in die Software. Also der spuckt am Ende ein Lung-RADS Stadium aus.	

5.141	JM	Ja genau. Das heißt aber auch wenn Lung-RADS 1.2 kommt zum Beispiel muss es aber auch wahrscheinlich angepasst werden oder wie?	
5.142	R5	Genau, da müsste das müsste das angepasst werden, völlig richtig und genau und er macht sogar Volumenberechnung, was ja nicht auch nicht bei Lung-RADS selbstverständlich ist. Also der macht sogar Volumenberechnungen.	OQ
5.143	JM	Ja super. Vielen, vielen Dank für das Interview. Es war sehr, sehr insightful. Also vielen Dank dafür, dass Sie sich am Abend noch die Zeit genommen haben und ja, ich denke, es wäre auch großartig, wenn Sie uns noch den Kontakt zu Ihrem Kollegen herstellen könnten. Das wäre auch ungemein wertvoll noch mal und ja dann..	
5.144	R5	Genau, dann wünsche ich Ihnen alles Gute. Wann müssen Sie denn diese Masterarbeit abgeben?	
5.145	JH	Ende Mai.	
5.146	R5	Ende Mai. Gut. Okay. Alles klar. Ja. Also dann, alles, alles Gute.	
5.147	JH	Dankeschön.	
5.148	R5	Und ich weiß nicht, bekommt man dann irgendwie oder würden Sie die auch zur Verfügung stellen? Die Masterarbeit.	
5.149	JH	Die können wir Ihnen am Ende sehr gerne schicken, wenn sie fertig ist.	
5.150	R5	Wir sind ja ein Team. Das reicht ja auch einmal an mich. [name, anonymized], [name, anonymized] und ich sind ein Team. Also das würde uns glaube ich interessieren.	
5.151	JM	Ich glaube, wir würden sowieso ganz gerne auch noch mal, nachdem wir Ihr Interview oder unser Interview jetzt hier analysiert haben und da sage ich mal, die Core Findings quasi niedergeschrieben haben, auch gerne, dazu noch mal Ihnen eine Mail kurz schreiben, damit Sie bestätigen können, so genau so, habe ich das gemeint, oder nein, da haben Sie was falsch verstanden, wenn es okay wäre und einfach nur absegnen, dass wir das richtig verstanden haben.	
5.152	R5	Also ich bin jetzt 14 Tage im Urlaub.	

5.153	JM	Oh ja.	
5.154	R5	Das hatte ich Ihnen gesagt.	
5.155	JM	Ja. genau. Das ist auch ansonsten auch kein Ding, wie gesagt. Ich denke auch mal dass wir jetzt nicht großartig jetzt was falsch verstehen werden. Aber wenn Sie dann Ende Mai einfach nur noch mal kurz bestätigen können. So Ja, das stimmt so.	
5.156	JH	Das ist auch nur eine kurze Zuammenfassung.	
5.157	R5	Das da kein grober Schnitzer drin ist sozusagen.	
5.158	JM	Ja, genau.	
5.159	R5	Ja.	
5.160	JM	Super. Klasse.	
5.161	R5	Ja, dann schönen Abend noch.	
5.162	JH	Schönen Abend. Tschüss.	

References

4-IN-THE-LUNG-RUN. (2021). Clinical Trials Essential Information, Available Online: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ce87cc5f&appId=PPGMS [Accessed April 21 2022]

Aberle, D.R., Abtin, F., & Brown, K. (2013). Computed tomography screening for lung cancer: has it finally arrived? Implications of the national lung screening trial, *Journal of Clinical Oncology*, vol. 31, no. 8, pp.1002-1008.

Abouzahra, M. (2011). Causes of failure in Healthcare IT projects, 3rd International Conference on Advanced Management Science, vol. 19, pp.46-50, Singapore: IACSIT Press.

Agarwal, R., Gao, G., DesRoches, C., & Jha, A K. (2010). Research commentary—The digital transformation of healthcare: Current status and the road ahead. *Information systems research*, vol. 21, no. 4, pp.796-809.

AIS. (2019). Member Code of Conduct: AIS Code of Ethics and Professional Conduct, Available online: https://aisnet.org/page/MemberCodeOfConduct [Accessed 24 April 2022].

Al Mohammad, B., Hillis, S. L., Reed, W., Alakhras, M., & Brennan, P.C. (2019). Radiologist performance in the detection of lung cancer using CT, *Clinical radiology*, vol. 74, no. 1, pp.6775.

Aldosari, B. (2012). User acceptance of a picture archiving and communication system (PACS) in a Saudi Arabian hospital radiology department, *BMC medical informatics and decision making*, vol. 12, no. 1, pp.1-10.

ALLEA. (2017). The European code of conduct for research integrity, revised edn, Berlin: ALLEA.

Angen, M.J. (2000). Evaluating interpretive inquiry: Reviewing the validity debate and opening the dialogue, *Qualitative health research*, vol. 10, no. 3, pp.378-395.

Ardila, D., Kiraly, A.P., Bharadwaj, S., Choi, B., Reicher, J.J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G. & Naidich, D.P. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography, *Nature medicine*, vol. 25, no. 6, pp.954-961.

Awai, K., Murao, K., Ozawa, A., Komi, M., Hayakawa, H., Hori, S., & Nishimura, Y. (2004). Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance, *Radiology*, vol. 230, no. 2, pp.347-352.

Bandyopadhyay, T., & Zadeh, B. (2014). Acceptance of Mobile Health Technology in the Value Chain, *AMCIS Proceedings*, no. 15, Available Online: https://aisel.aisnet.org/amcis2014/HealthIS/GeneralPresentations/15 [Accessed 26 April 2022]

Bærøe, K., Jansen, M., & Kerasidou, A. (2020). Machine Learning in Healthcare: Exceptional Technologies Require Exceptional Ethics, *The American Journal of Bioethics*, vol. 20, no. 11, pp.48-51.

Becker, N., Motsch, E., Gross, M.L., Eigentopf, A., Heussel, C.P., Dienemann, H., Schnabel, P.A., Pilz, L., Eichinger, M., Optazaite, D.-E., Puderbach, M., Tremper, J., & Delorme, S. (2012). Randomized study on early detection of lung cancer with MSCT in Germany: study design and results of the first screening round, Journal of cancer research and clinical oncology, vol. 13, no. 9, pp.1475-1486.

Becker, N., Motsch, E., Trotter, A., Heussel, C.P., Dienemann, H., Schnabel, P.A., Kauczor, H., Maldonado, S.G., Miller, A.B., Kaaks, R., & Delorme, S. (2020). Lung

cancer mortality reduction by LDCT screening—results from the randomized German LUSI trial, International journal of cancer, vol. 146, no. 6, pp.1503-1513.

BenMessaoud, C., Kharrazi, H., & MacDorman, K.F. (2011). Facilitators and barriers to adopting robotic-assisted surgery: contextualizing the unified theory of acceptance and use of technology, *PloS one*, vol. 6, no. 1, p.e16395.

Black, N. (1996). Why we need observational studies to evaluate the effectiveness of health care, *British Medical Journal*, vol. 312, no. 7040, pp.1215-1218.

Bonavita, I., Rafael-Palou, X., Ceresa, M., Piella, G., Ribas, V., & Ballester, M.A.G. (2020). Integration of convolutional neural networks for pulmonary nodule malignancy assessment in a lung cancer classification pipeline, *Computer methods and programs in biomedicine*, vol. 185, no. 105172.

Bogoni, L., Ko, J.P., Alpert, J., Anand, V., Fantauzzi, J., Florin, C.H., Koo, C.W., Mason, D., Rom, W., Shiau, M., Salganicoff, M., & Naidich, D.P. (2012). Impact of a computer-aided detection (CAD) system integrated into a picture archiving and communication system (PACS) on reader sensitivity and efficiency for the detection of lung nodules in thoracic CT exams, *Journal of digital imaging*, vol. 25, no. 6, pp.771-781.

Boughzala, I. (2014). How Generation Y Perceives Social Networking Applications in Corporate Environments, *Integrating Social Media into Business Practice, Applications, Management, and Models*, pp.162-179, Hershey: IGI Global.

BREATH. (2021). North German Lung Cancer Centers start the HANSE Lung Check, Available Online: https://www.breath-hannover.de/en/news-media/news/news-details.html?tx_news_pi1%5Bnews%5D=560&cHash=c7ab4d02ce93a371f10b50863aa5eeb4 [Accessed 1 April 2022]

Brocklehurst, P., Field, D., Greene, K., Juszczak, E., Keith, R., Kenyon, S., Linsell, L., Mabey, C., Newburn, M., Plachcinski, R., Quigley, M., Schroeder, E., & Steer, P. (2017). Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial, *The Lancet*, vol. 389, no. 10080, pp.1719-1729.

Brown, M., Browning, P., Wahi-Anwar, M.W., Murphy, M., Delgado, J., Greenspan, H., Abtin, F., Ghahremani, S., Yaghmai, N., da Costa, I., & Becker, M. (2019). Integration of chest CT CAD into the clinical workflow and impact on radiologist efficiency, *Academic radiology*, vol. 26, no. 5, pp.626-631.

Buchanan, B.G. (2005). A (very) brief history of artificial intelligence, *Ai Magazine*, vol. 26, no. 4, pp.53-60.

Buck, C., Hennrich, J. & Kauffmann, A.L. (2021). Artificial Intelligence in Radiology–A Qualitative Study on Imaging Specialists' Perspectives, *ICIS 2021 Proceedings*, vol. 20.

Bundesamt für Strahlenschutz. (2021). Lungenkrebsfrüherkennung mittels Niedrigdosis-Computertomographie - Wissenschaftliche Bewertung des Bundesamtes für Strahlenschutz gemäß § 84 Absatz 3 Strahlenschutzgesetz, Available online: http://doris.bfs.de/jspui/handle/urn:nbn:de:0221-2021082028027 [Acessed 3 April 2022].

Chen, Y., Stavropoulou, C., Narasinkan, R., Baker, A., & Scarbrough, H. (2021). Professionals' responses to the introduction of AI innovations in radiology and their implications for future adoption: a qualitative study, *BMC Health Services Research*, vol. 21, no. 1, pp.1-9.

Creswell, J.W. (2013). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 3rd edn, California: Sage.

Creswell, J.W. (2007). Qualitative Inquiry & Research Design: Choosing Among Five Approaches, 2nd edn, California: Sage.

Cruickshank, A., Stieler, G., & Ameer, F. (2019). Evaluation of the solitary pulmonary nodule, *Internal Medicine Journal*, vol. 49, no. 3, pp.306-315.

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare, *Future healthcare journal*, vol. 6, no. 2, pp.94-98.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology, *MIS quarterly*, vol. 13, no. 3, pp.319-340.

de Koning, H.J., van der Aalst, C.M., de Jong, P.A., Scholten, E.T., Nackaerts, K., Heuvelmans, M.A., Lammers, J.W.J., Weenink, C., Yousaf-Khan, U., Horeweg, N. and van't Westeinde, S., (2020). Reduced lung-cancer mortality with volume CT screening in a randomized trial, *New England journal of medicine*, vol 382, no. 6, pp.503-513.

Deo, R. C. (2015). Machine learning in medicine, *Circulation*, vol. 132, no. 20, pp.1920-1930.

Deutscher Ethikrat (2017). Big Data und Gesundheit–Datensouveränität als informationelle Freiheitsgestaltung, Available Online: https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-big-data-und-gesundheit.pdf [Accessed 11 May 2022]

Deutsches Krebsforschungszentrum. (n.d.). Hintergrund und Ziel der Studie, Available Online: https://www.dkfz.de/de/epidemiologie-krebserkrankungen/arbeitsgr/Lungenkrebsscreening/4-IN-THE-LUNG-Run-Ziele.html?m=1637228158& [Accessed 16 April 2022].

Dishaw, M.T., & Strong, D.M. (1999). Extending the technology acceptance model with task–technology fit constructs, *Information & management*, vol. 36, no. 1, pp.9-21.

Dou, Q., Chen, H., Jin, Y., Lin, H., Qin, J., & Heng, P. A. (2017). Automated pulmonary nodule detection via 3d convnets with online sample filtering and hybrid-loss residual learning, in M. Descoteaux, L. Maier-Hein, A., Franz, P., Jannin, D.L. Collins, S. Duchesne (eds), *Medical Image Computing and Computer Assisted Intervention - MICCAI 2017*, Quebec City: Springer, pp.630-638.

Duffy, S.W. & Field, J.K. (2020). Mortality reduction with low-dose CT screening for lung cancer. *New England Journal of Medicine*, vol. 382, no. 6, pp.572-573.

Dünnebeil, S., Sunyaev, A., Blohm, I., Leimeister, J.M., & Krcmar, H. (2012). Determinants of physicians' technology acceptance for e-health in ambulatory care. International journal of medical informatics, vol. 81, no. 11, pp.746-760.

Eitle, V., & Buxmann, P. (2020). Cultural Differences in Machine Learning Adoption: An International Comparison between Germany and the United States, *Proceedings of the 28th European Conference on Information Systems (ECIS)*, Available Online: https://aisel.aisnet.org/ecis2020_rp/138 [Accessed 17 May]

El-Gayar, O.F., Deokar, A.V., & Wills, M.J. (2010). Evaluating task-technology fit and user performance for an electronic health record system, *International Journal of Healthcare Technology and Management*, vol. 11, no.1, pp. 50-65.

Espinoza, J.L., & Dong, L.T. (2020). Artificial intelligence tools for refining lung cancer screening, *Journal of clinical medicine*, [e-journal] vol. 9, no. 12, Available online: https://www.mdpi.com/2077-0383/9/12/3860 [Accessed 3 April 2022].

European Commission. (2021). 4-IN THE LUNG RUN: towards INdividually tailored INvitations, screening INtervals, and INtegrated co-morbidity reducing strategies in

lung cancer screening, Available Online: https://cordis.europa.eu/project/id/848294 [Accessed 21 April 2022]

European Lung Foundation. (2021). Lung cancer screening, Available Online: https://europeanlung.org/en/information-hub/factsheets/lung-cancer-screening/ [Accessed 04 April 2022]

Fan, W., Liu, J., Zhu, S., & Pardalos, P.M. (2020). Investigating the impacting factors for the healthcare professionals to adopt artificial intelligence-based medical diagnosis support system (AIMDSS), *Annals of Operations Research*, vol. 294, no. 1, pp.567-592.

Faqih, K.M, & Jaradat, M.I.R.M. (2015). Assessing the moderating effect of gender differences and individualism-collectivism at individual-level on the adoption of mobile commerce technology: TAM3 perspective, *Journal of Retailing and Consumer Services*, vol. 22, pp.37-52.

Fischer, A.M., Yacoub, B., Savage, R.H., Martinez, J.D., Wichmann, J.L., Sahbaee, P., Grbic, S., Varga-Szemes, A., & Schoepf, U.J. (2020). Machine learning/deep neuronal network: routine application in chest computed tomography and workflow considerations, *Journal of Thoracic Imaging*, vol. 35, pp.21-27.

Fishbein, M., Ajzen, I. (1975). Belief, Attitude, Intention and Belavior: An Introduiction to Theory and Research, MA: Addison-Wesley.

Gagnon, M.P., Orruño, E., Asua, J., Abdeljelil, A.B., & Emparanza, J. (2012). Using a modified technology acceptance model to evaluate healthcare professionals' adoption of a new telemonitoring system, *Telemedicine and e-Health*, vol. 18, no. 1, pp.54-59.

Garbuio, M., Lin, N. (2019). Artificial Intelligence as a Growth Engine for Health Care Startups: Emerging Business Models, California Management Review, vol. 61, no. 2, pp.59-83.

Gartner. (2020). Gartner Identifies the Top Strategic Technology Trends for 2021, Available Online: https://www.gartner.com/en/newsroom/press-releases/2020-10-19-gartner-identifies-the-top-strategic-technology-trends-for-2021 [Accessed 20 May 2022].

Gaube, S., Suresh, H., Raue, M., Merritt, A., Berkowitz, S.J., Lermer, E., Coughlin, J.F., Guttag, J.V., Colak, E., & Ghassemi, M. (2021). Do as AI say: susceptibility in deployment of clinical decision-aids, *NPJ digital medicine*, vol. 4, no. 1, pp.1-8.

Goldfarb, A., & Teodoridi, F. (2022). Why is AI adoption in health care lagging?, Brookings, Available Online: https://www.brookings.edu/research/why-is-ai-adoption-in-health-care-lagging/ [Accessed 3 April 2022].

Goncalves, S., Fong, P.C., & Blokhina, M. (2022). Artificial intelligence for early diagnosis of lung cancer through incidental nodule detection in low-and middle-income countries-acceleration during the COVID-19 pandemic but here to stay, *American Journal of Cancer Research*, vol. 12, no. 1, p.1-16.

Goodhue, D.L, & Thompson, R.L. (1995). Task-technology fit and individual performance, *MIS quarterly*, vol. 19, no. 2, pp.213-236.

Gregor, S. (2006). The nature of theory in information systems, *MIS Quarterly*, vol. 30, no. 3, pp.611-642.

Guo, J., Wang, C., Xu, X., Shao, J., Yang, L., Gan, Y., Zhang, Y., & Li, W. (2020). DeepLN: an artificial intelligence-based automated system for lung cancer screening, *Annals of Translational Medicine*, [e-journal] vol. 8, no. 18, Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7576052/ [Accessed 3 April 2022].

Hamet, P, & Tremblay, J. (2017). Artificial intelligence in medicine, *Metabolism*, vol. 69, pp.36-40.

Harst, L., Lantzsch, H., & Scheibe, M. (2019). Theories predicting end-user acceptance of telemedicine use: systematic review, *Journal of medical Internet research*, vol. 21, no. 5.

Hassan, N.R., Mingers, J., & Stahl, B. (2018). Philosophy and information systems: where are we and where should we go?, *European Journal of Information Systems*, vol. 27, no. 3, pp.263-277.

Hecht, H.S., Henschke, C., Yankelevitz, D., Fuster, V., & Narula, J. (2014). Combined detection of coronary artery disease and lung cancer, *European heart journal*, vol. 35, no. 40, pp.2792-2796.

Heinsch, M., Wyllie, J., Carlson, J., Wells, H., Tickner, C., & Kay-Lambkin, F. (2021). Theories informing eHealth implementation: systematic review and typology classification, *Journal of Medical Internet Research*, vol. 23, no. 5, p.e18500.

Hercheui, M. & Mech, G. (2021). Factors Affecting The Adoption Of Artificial Intelligence In Healthcare, *Global Journal of Business Research*, vol. 15, no. 1, pp.77-88.

Holden, R.J., & Karsh, B.T. (2010). The technology acceptance model: its past and its future in health care, *Journal of biomedical informatics*, vol. 43, no. 1, pp.159-172.

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L.H., & Aerts, H.J. (2018). Artificial intelligence in radiology, *Nature Reviews Cancer*, vol. 18, no. 8, pp.500-510.

Huang, P., Lin, C. T., Li, Y., Tammemagi, M.C., Brock, M.V., Atkar-Khattra, S., Xu, Y., Hu, P., Mayo, J.R., Schmidt, H., Gringras, M., Pasian, S., Stewart, L., Tsai, S., Seely, J.M., Manos, D., Burrowes, P., Bhatia, R., Tsao, M.-S., & Lam, S. (2019). Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method, *The Lancet Digital Health*, vol. 1, no. 7, pp.e353-e362.

Hufnagl, P. (2020). Digitalisierung in der Medizin, E-Health, Künstliche Intelligenz–das Forschungsprojekt EMPAIA–Ecosystem for Pathology Diagnostics with AI Assistance, *Gesundheitswesen aktuell*, Available Online: https://www.bifg.de/media/dl/Gesundheitswesen%20aktuell/2020/GWA%202020-Kapitel%20Hufnagl.pdf [Accessed 11 May 2022]

Huisman, M., Ranschaert, E., Parker, W., Mastrodicasa, D., Koci, M., Pinto de Santos, D., Coppola, F., Morozov, S., Zins, M., Bohyn, C., & Koç, U. (2021). An international survey on AI in radiology in 1041 radiologists and radiology residents part 2: expectations, hurdles to implementation, and education, *European Radiology*, vol. 31, no. 11, pp.8797-8806.

Hwang, E.J., & Park, C.M. (2020). Clinical implementation of deep learning in thoracic radiology: potential applications and challenges, *Korean Journal of Radiology*, vol. 21, no. 5, p.511-525.

Institute for Quality and Efficiency in Health Care. (2020). Lung cancer screening with low-dose computed tomography, Available Online: https://www.iqwig.de/download/s19-02_lung-cancer-screening-with-low-dose-ct_extract-of-final-report_v1-0.pdf [Accessed 31 March 2022]

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future, *Stroke and vascular neurology*, vol. 2, no. 4, pp.230-243.

Jorritsma, W., Cnossen, F., & van Ooijen, P.M. (2015). Improving the radiologist–CAD interaction: designing for appropriate trust, *Clinical radiology*, vol. 70, no. 2, pp.115-122.

Kalis, B., Collier, M., & Fu, R. (2018). 10 promising AI applications in health care, Harvard business review, Available Online: https://www.investkl.gov.my/clients/asset_28B5D799-69B3- 4BCB-B61B-D284619547A3/uploads/10-Promising-AI-Applications-in-HealthCare.PDF [Accessed 3 April 2022].

Kaplan, B., & Maxwell, J.A. (2005). Qualitative research methods for evaluating computer information systems, in J.G. Anderson & C.E. Aydin (eds), *Evaluating the organizational impact of healthcare information systems*, New York: Springer, pp.30-55.

Karantanas, A.H., & Efremidis, S. (2022). The concept of the invisible radiologist in the era of artificial intelligence, *European Journal of Radiology*, Available Online: https://doi.org/10.1016/j.ejrad.2021.110147 [Accessed 20 May 2022].

Karsh, B.T., Escoto, K.H., Beasley, J.W., & Holden, R.J. (2006). Toward a theoretical approach to medical error reporting system research and design, *Applied ergonomics*, vol. 37, no. 3, pp.283-295.

Kassania, S.H., Kassania, P.H., Wesolowskic, M.J., Schneidera, K.A., & Detersa, R. (2021). Automatic Detection of Coronavirus Disease (COVID-19) in X-ray and CT Images: A Machine Learning Based Approach, *Biocybernetics and Biomedical Engineering*, vol. 41, no. 3, pp.867-879.

Kauffmann, A.L., Hennrich, J., Buck, C., & Eymann, T. (2022). Künstliche Intelligenz in der Radiologie und Strahlentherapie aus der Perspektive von Ärzten und Medizinphysikexperten–Eine Interviewstudie, *Künstliche Intelligenz im Gesundheitswesen*, pp.617-638, Wiesbaden: Springer Gabler.

Khanijahani, A., Iezadi, S., Dudley, S., Goettler, M., Kroetsch, P., & Wise, J. (2022). Organizational, professional, and patient characteristics associated with artificial intelligence adoption in healthcare: A systematic review, *Health Policy and Technology*, vol 11, no. 1.

Kitzmiller, R., Hunt, E, & Sproat, S.B. (2006). Adopting best practices: "Agility" moves from software development to healthcare project management, *CIN: Computers, Informatics, Nursing*, vol. 24, no. 2, pp.75-82.

Klein, F. (2021). Screening auf Lungenkrebs: Start noch in diesem Jahr?, *DGIM-Kongress*, Available Online: https://www.springermedizin.de/dgim-2021/praevention-und-rehabilitation-in-der-pneumologie/screening-auf-lungenkrebs--start-noch-in-diesem-jahr-/19085104 [Accessed 31 March 2022]

Klopping, I.M, & McKinney, E. (2004). Extending the technology acceptance model and the task-technology fit model to consumer e-commerce, *Information Technology, Learning & Performance Journal*, vol. 22, no. 1, pp.35-48.

Krumm, S., & Dwertmann, A. (2019). Perspektiven der KI in der Medizin, *Künstliche Intelligenz*, pp.161-175, Berlin: Springer.

Lai, P.C. (2017). The literature review of technology adoption models and theories for the novelty technology, JISTEM-Journal of Information Systems and Technology Management, vol. 14, pp.21-38.

Laï, M.C., Brian, M., & Mamzer, M.F. (2020). Perceptions of artificial intelligence in healthcare: findings from a qualitative survey study among actors in France, *Journal of translational medicine*, vol. 18, no. 1, pp.1-13.

Lancaster, H.L., Zheng, S., Aleshina, O.O., Yu, D., Chernina, V.Y., Heuvelmans, M.A., de Bock, G.H., Dorrius, M.D., Gratama, J.W., Morozov, D.P., Gombolevskiy, V.A., Silva, M., Yi, J. & Oudkerk, M. (2022). Outstanding negative prediction performance of solid pulmonary nodule volume AI for ultra-LDCT baseline lung cancer screening risk stratification, *Lung Cancer*, vol. 165, pp.133-140.

Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). Is AI ground truth really "true"? The dangers of training and evaluating AI tools based on experts' know-what, *MIS Quarterly*, vol. 45, no. 3, pp.1501-1525.

Lee, D., & Yoon, S.N. (2021). Application of artificial intelligence-based technologies in the healthcare industry: Opportunities and challenges, International Journal of Environmental Research and Public Health, vol. 18, no. 1.

Lee, J. G., Jun, S., Cho, Y. W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: general overview, *Korean journal of radiology*, vol. 18, no. 4, pp.570-584.

Lee, R. S. (2020). Artificial intelligence in daily life, Singapore: Springer.

Lee, Y., Kozar, K.A., & Larsen, K.R. (2003). The technology acceptance model: Past, present, and future, *Communications of the Association for information systems*, vol. 12, no. 50, pp.752-780.

Li, D., Pehrson, L.M., Lauridsen, C.A., Tøttrup, L., Fraccaro, M., Elliott, D., Zając, H.D., Darkner, S., Carlsen, J.F., & Nielsen, M.B. (2021). The added effect of artificial intelligence on physicians' performance in detecting thoracic pathologies on CT and chest X-ray: A systematic review, *Diagnostics*, vol. 11, no. 12, Available online: https://www.mdpi.com/2075-4418/11/12/2206 [Accessed 11 April 2022].

Li, X., Hess, T.J., & Valacich, J.S. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems, *The Journal of Strategic Information Systems*, vol. 17, no. 1, pp.39-71.

Liao, C., Palvia, P., & Chen, J.L. (2009). Information technology adoption behavior life cycle: Toward a Technology Continuance Theory (TCT), *International Journal of Information Management*, vol. 29, no. 4, pp.309-320.

Liu, K., Li, Q., Ma, J., Zhou, Z., Sun, M., Deng, Y., Tu, W., Wang, Y., Fan, L., Xia, C. and Xiao, Y. (2019a). Evaluating a fully automated pulmonary nodule detection approach and its impact on radiologist performance, *Radiology: Artificial Intelligence*, vol. 1, no. 3.

Liu, X., Faes, L., Kale, A.U., Wagner, S.K., Fu, D.J., Bruynseels, A., Mahendiran, T., Moraes, G., Shamdas, M., Kern, C., & Ledsam, J.R. (2019b). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis, *The lancet digital health*, vol. 1, no. 6, pp.271-297.

Liu, S., Setio, A.A.A., Ghesu, F.C., Gibson, E., Grbic, S., Georgescu, B., & Comaniciu, D. (2021). No surprises: Training robust lung nodule detection for low-dose CT scans by augmenting with adversarial attacks, *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp.335-345.

Lockey, S., Gillespie, N., Holm, D., & Someh, I.A. (2021). A review of trust in artificial intelligence: Challenges, vulnerabilities and future directions, *Hawaii International Conference on System Sciences*, vol. 54, pp.5463-5472.

Lucas, N. (2021). What Does Software As A Service Mean For Healthcare?, Available Online: https://www.forbes.com/sites/forbestechcouncil/2021/01/04/what-does-software-as-a-service-mean-for-healthcare/ [Accessed 8 May 2022]

Lund University. (2021). Guidelines for the processing of matters relating to suspected deviation from good research practice at Lund University, Available from: https://www.medarbetarwebben.lu.se/forska-och-utbilda/stod-till-forskning/forskningsetik-och-djurforsoksetik/avvikelser-fran-god-forskningssed [Accessed 24 April 2022].

Ma, J., Song, Y., Tian, X., Hua, Y., Zhang, R., & Wu, J. (2020). Survey on deep learning for pulmonary medical imaging, *Frontiers of medicine*, vol. 14, no. 4, pp.450-469.

Maier, S.B., Jussupow, E., & Heinzl, A. (2019). Good, bad, or both? Measurement of physician's ambivalent attitudes towards AI, *Proceedings of the 27th European Conference on Information Systems*, Available Online: https://aisel.aisnet.org/ecis2019_rp/115 [Accessed 7 April 2022].

Marangunić, N. and Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013, *Universal access in the information society*, vol. 14, no. 1, pp.81-95.

Mathew, C. J., David, A. M., & Mathew, C. M. J. (2020). Artificial Intelligence and its future potential in lung cancer screening, *EXCLI journal*, [e-journal] vol. 19, pp.1552-1562, Available online: https://www.excli.de/vol19/excli2020-3095.pdf [Accessed 3 April 2022].

McKinsey & Company, & EIT Health. (2020). Transforming healthcare with AI: The impact on the workforce and organizations, Available Online: https://eithealth.eu/wpcontent/uploads/2020/03/EIT-Health-and-McKinsey_Transforming-Healthcare-with-AI.pdf [Accessed 13 January 2022].

Medizinische Hochschule Hannover. (2021). HANSE Studie, Available online: https://www.hanse-lungencheck.de/ [Accessed 25 April 2022].

Moore, G.C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation, *Information Systems Research*, vol. 2, no. 3, pp.192–222.

Morabia, A. & Zhang, F.F. (2004). History of medical screening: from concepts to action, *Postgraduate medical journal*, vol. 80, no. 946, pp.463-469.

Morley, J., Machado, C.C., Burr, C., Cowls, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review, *Social Science & Medicine*, vol. 260.

Mumm, J.N., Rodler, S., Mumm, M.L., Bauer, R.M., & Stief, C.G. (2021). Digitale Innovation in der Medizin–die COVID-19-Pandemie als Akzelerator von „digital health", *Journal für Urologie und Urogynäkologie/Österreich*, vol. 28, no. 1, pp.1-5.

Murphy, K.P. (2012). Machine learning: a probabilistic perspective. London: MIT Press.

Myers, M.D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft, *Information and organization*, vol. 17, no. 1, pp.2-26.

Nasrullah, N., Sang, J., Alam, M. S., Mateen, M., Cai, B., & Hu, H. (2019). Automated lung nodule detection and classification using deep learning combined with multiple strategies. Sensors, [e-journal] vol. 19, no. 17, Available online: https://www.mdpi.com/1424-8220/19/17/3722 [Accessed 3 April 2022].

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*, vol. 15, no. 1, pp.49-52.

Nunes, I., Ayres-de-Campos, D., Ugwumadu, A., Amin, P., Banfield, P., Nicoll, A., Cunningham, S., Sousa, P.B., Costa-Santos, C., & Bernardes, J. (2017). Central fetal monitoring with and without computer analysis, *Obstetrics & Gynecology*, vol. 129, no. 1, pp.83-90.

Pagani, M. (2006). Determinants of adoption of high speed data services in the business market: evidence for a combined technology acceptance model with task technology fit model, *Information & Management*, vol. 43, no. 7, pp.847-860.

Panch, T., Mattie, H., & Celi, L.A. (2019). The “inconvenient truth” about AI in healthcare, *NPJ digital medicine*, vol. 2, no. 1, pp.1-3.

Parikh, R. B., Teeple, S., & Navathe, A. S. (2019). Addressing Bias in Artificial Intelligence in Health Care, *JAMA*, vol. 322, no. 24, pp.2377–2378.

Park, C.W., Seo, S.W., Kang, N., Ko, B., Choi, B.W., Park, C.M., Chang, D.K., Kim, H., Kim, H., Lee, H., & Jang, J. (2020). Artificial intelligence in health care: Current applications and issues, *Journal of Korean medical science*, vol. 35, no. 42.

Patton, M.Q. (2015). Qualitative research & evaluation methods, 4th edn, California: Sage publications.

Pekka, A.P., Bauer, W., Bergmann, U., Bieliková, M., Bonefeld-Dahl, C., Bonnet, Y., Bouarfa, L. (2018). The European Commission’s high-level expert group on artificial intelligence: Ethics guidelines for trustworthy AI, *Working Document for stakeholders’ consultation, Brussels,* pp.1–37.

Pinsky, P.F., & Berg, C.D. (2012). Applying the National Lung Screening Trial eligibility criteria to the US population: what percent of the population and of incident lung cancers would be covered?, *Journal of medical screening*, vol. 19, no. 3, pp.154-156.

Pinsky, P.F., Gierada, D.S., Nath, P.H., Kazerooni, E., & Amorosa, J. (2013). National lung screening trial: variability in nodule detection rates in chest CT studies, *Radiology*, vol. 268, no. 3, pp.865-873. Pumplun, L., Fecho, M., Wahl-Islam, N., & Buxmann, P. (2021). Machine learning systems in clinics–how mature is the adoption process in medical diagnostics?, *Proceedings of the Hawaii International Conference on System Sciences*, vol. 54, pp.6317-6326.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential, Health information science and systems, vol. 2, no. 1, pp.1-10.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine, *New England Journal of Medicine*, vol. 380, no. 14, pp.1347-1358.

Recker, J. (2013). Scientific Research in Information Systems: A Beginner's Guide, Berlin: Springer.

Reis, L., Maier, C., Mattke, J., Creutzenberg, M., & Weitzel, T. (2020). Addressing User Resistance Would Have Prevented a Healthcare AI Project Failure, *MIS Quarterly Executive*, vol. 19, no. 4, pp.279-296.

Rubin, D.L. (2019). Artificial intelligence in imaging: the radiologist’s role, *Journal of the American College of Radiology*, vol. 16, no. 9, pp.1309-1317.

Russell, S., & Norvig, P. (2020). Artificial intelligence: a modern approach, 4th edn, New Jersey: Prentice-Hall.

Saunders, M., Lewis, P. & Thornhill, A. (2007). Research methods for Business Students, 4th edn, London: Pearson.

Schultze, U., & Avital, M. (2011). Designing interviews to generate rich data for information systems research, *Information and organization*, vol. 21, no. 1, pp.1-16.

Seneviratne, M.G., Shah, N.H., & Chu, L. (2020). Bridging the implementation gap of machine learning in healthcare, *BMJ Innovations*, vol. 6, no. 2, pp.45-47.

Setio, A.A.A., Traverso, A., de Bel, T., Berens, M.S.N., Bogaard, C.van, Cerello, P., Chen, H., Dou, Q., Fantacci, M.E., Geurts, B., Gugten, R.van, Heng, P.A., Jansen, B., de Kaste, M.M.J., Kotov, V., Lin, J.Y.-H., Manders, J.T.M.C., Sóñora-Mengana, A., García-Naranjo, J.C., Papavasileiou, E., Prokop, M., Saletta, M., Schaefer-Prokop, C.M., Scholten, E.T., Scholten, L., Snoeren, M.M., Torres, E.L., Vandemeulebroucke, J., Walasek, N., Zuidhof, G.C.A., van Ginneken, B. & Jacobs, C. (2017). Validation, Comparison, and Combination of Algorithms for

Automatic Detection of Pulmonary Nodules in Computed Tomography Images: The luna16 Challenge, *Medical Image Analysis*, vol. 42, pp.1–13

Shibl, R., Lawley, M., & Debuse, J. (2013). Factors influencing decision support system acceptance, *Decision Support Systems*, vol. 54, no. 2, pp.953-961.

Siegmund-Schultze, N. (2019). Onkologie: Wendepunkt beim Lungenkrebs-Screening für Risikogruppen, Deutsches Ärzteblatt, vol. 116, no. 14, pp.686-688.

Stang, A., Schuler, M., Kowall, B., Darwiche, K., Kühl, H., & Jöckel, K.H. (2015). Lung Cancer Screening Using Low Dose CT Scanning in Germany: Extrapolation of Results From the National Lung Screening Trial, *Deutsches Ärzteblatt International*, vol. 112, no. 38, p.637-644.

Strauss, A., & Corbin, J. (1998). Basics of qualitative research techniques, London: SAGE Publications Ltd.

Strohm, L., Hehakaya, C., Ranschaert, E.R., Boon, W.P., & Moors, E.H. (2020). Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors, *European radiology*, vol. 30, no. 10, pp.5525-5532.

Sturm, T., & Peters, F. (2020). The impact of artificial intelligence on individual performance: Exploring the fit between task, data, and technology, *Proceedings of the 28th European Conference on Information Systems (ECIS)*, no. 200, Available Online: https://aisel.aisnet.org/ecis2020_rp/200/ [Accessed 26 April 2022].

Svoboda, E. (2020). Artificial intelligence is improving the detection of lung cancer, *Nature Outlook*, vol. 587, pp.20-22.

Taherdoost, H. (2018). A review of technology acceptance and adoption models and theories, *Procedia manufacturing*, vol. 22, pp.960-967.

Tang, A., Tam, R., Cadrin-Chênevert, A., Guest, W., Chong, J., Barfett, J., Chepelev, L., Cairns, R., Mitchell, J.R., Cicero, M.D, & Poudrette, M.G. (2018). Canadian Association of Radiologists white paper on artificial intelligence in radiology, *Canadian Association of Radiologists Journal*, vol. 69, no. 2, pp.120-135.

Tanoue, L.T., Tanner, N.T., Gould, M.K., & Silvestri, G.A. (2015). Lung cancer screening, *American journal of respiratory and critical care medicine*, vol. 191, no. 1, pp.19-33.

Thrall, J.H., Li, X., Li, Q., Cruz, C., Do, S., Dreyer, K., & Brink, J. (2018). Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success, *Journal of the American College of Radiology*, vol 15, no. 3, pp.504–508.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence, *Nature medicine*, vol. 25, no. 1, pp.44-56.

Tubaishat, A. (2018). Perceived usefulness and perceived ease of use of electronic health records among nurses: application of technology acceptance model, *Informatics for Health and Social Care*, vol. 43 no. 4, pp.379-389.

US Preventive Services Task Force. (2021). Screening for lung cancer: US Preventive Services Task Force recommendation statement, *Jama*, vol. 325, no. 10, pp.962-970.

van der Aalst, C.M., Ten Haaf, K., & de Koning, H.J. (2021). Implementation of lung cancer screening: what are the main issues?, *Translational Lung Cancer Research*, vol. 10, no. 2, p.1050-1063.

Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating perceived behavioral control, computer anxiety and enjoyment into the technology acceptance model, *Information Systems Research*, vol. 11, pp.342–365.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions, *Decision sciences*, vol. 39, no. 2, pp.273-315.

Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems, *MIS Quarterly*, vol. 37, no. 1, pp. 21-54.

Venkatesh, V., Davis, F.D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies, *Managing Science*, col. 46, no. 2, pp.186-204.

Venkatesh, V., Morris, M.G., Davis, G.B, & Davis, F.D. (2003). User acceptance of information technology: Toward a unified view, *MIS quarterly*, pp.425-478.

Vogelsang, K., Steinhüser, M., & Hoppe, U. (2013). A qualitative approach to examine technology acceptance, *AIS Electronic Library*, Available Online: https://core.ac.uk/download/pdf/301361231.pdf [Accessed 12 May 2022]

Walsham, G. (1995). Interpretive case studies in IS research: nature and method, *European Journal of information systems*, vol. 4, no. 2, pp.74-81.

Webster, J., & Watson, R.T. (2002). Analyzing the past to prepare for the future: Writing a literature review, *MIS quarterly*, vol. 26, no. 2, pp.xiii-xxiii.

Weller, D.P., Patnick, J., McIntosh, H.M., & Dietrich, A.J. (2009). Uptake in cancer screening programmes, *The lancet oncology*, vol. 10, no. 7, pp.693-699.

Wender, R., Fontham, E.T., Barrera Jr, E., Colditz, G.A., Church, T.R., Ettinger, D.S., Etzioni R, Flowers, C.R., Gazelle, G.S., Kelsey, D.K., LaMonte, S.J., Michaelson, J.S., Oeffinger, K.C., Shih, Y.C., Sullivan, D.C., Travis, W., Walter, L., Wolf, A.M., Brawley, O.W., & Smith, R.A. (2013). American Cancer Society lung cancer screening guidelines, *cancer journal for clinicians*, vol. 63, no. 2, pp.106-117.

World Health Organization. (2020a). Cervical cancer screening - Response by country, Available Online: https://apps.who.int/gho/data/view.main.UHCCERVICAL-CANCERv [Accessed 03 April 2022]

World Health Organization. (2020b). Existence of national screening program for breast cancer, Available Online: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/existence-of-national-screening-program-for-breast-cancer [Accessed 03 April 2022]

World Health Organization. (2021). Ethics and governance of artificial intelligence for health: WHO guidance, Available Online: https://www.who.int/publications/i/item/9789240029200 [Accessed 24 April 2022].

Wolff, J., Pauling, J., Keck, A., & Baumbach, J. (2021). Success Factors of Artificial Intelligence Implementation in Healthcare, *Frontiers in digital health*, vol. 3, no. 51.

Wu, P.F. (2011). A Mixed Methods Approach to Technology Acceptance Research, *Journal of the AIS*, Available Online: https://ssrn.com/abstract=1937656 [Accessed 12 May 2022]

Xiao, M., Meredith, R., & Gao, S. (2017). An exploratory study investigating how and why managers use tablets to support managerial decision-making, *Australasian journal of information systems*, vol. 21.

Xie, Y., Zhang, J., & Xia, Y. (2019). Semi-supervised adversarial model for benign–malignant lung nodule classification on chest CT. Medical image analysis, vol. 57, pp.237-248.

Ye, T., Xue, J., He, M., Gu, J., Lin, H., Xu, B., & Cheng, Y. (2019). Psychosocial factors affecting artificial intelligence adoption in health care in China: Cross-sectional study, *Journal of medical Internet research*, vol. 21 no. 10, p.e14316.

Yin, J., Ngiam, K.Y., & Teo, H.H. (2020). Work Design in Healthcare Artificial Intelligence Applications: The Role of Advice Provision Timing, *ICIS 2020 Proceedings*, Available online: https://aisel.aisnet.org/icis2020/is_health/is_health/10/ [Accessed 11 April 2022].

Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare, *Nature biomedical engineering*, vol. 2, no. 10, pp.719-731.

Zaltman, G., Duncan, R., & Holbek, J. (1973). Innovations and organizations, New York: Wiley.

Zoom Video Communications. (2022). About Zoom, Available online: https://explore.zoom.us/en/about/ [Accessed 02 May 2022].