Challenges of AI adoption in healthcare organisations

A case study of ML adoption for medical diagnosis of rare diseases in the UK

Master Thesis submitted to University of Bern

Supervisors:
Prof. Dr. Rudolf Blankart
Sandra Gillner, M.Sc.

Center of Competence for Public Management
Schanzeneckstrasse 1
CH-3001 Bern

by:
Nicole Widmer
from Hildisrieden (LU)
16-323-479

Bern, 23.09.2022

Abstract

Artificial intelligence (AI) has the potential to significantly change the healthcare industry by allowing for more effective and efficient diagnoses of different types of diseases. As people living with rare diseases usually endure long diagnostic processes, they would particularly benefit from machine learning (ML) technologies, as these technologies could shorten the time to diagnosis and treatment. However, despite the high potential of ML for medical diagnostics, ML adoption in healthcare organisations is only increasing slowly. To learn why the widespread adoption of ML for rare disease diagnosis is slow, this paper conducted a case study as part of the Screen4Care research project. This case study explored which factors of the context-appropriate Technology-Organisation-Environment (TOE) framework are hindering the adoption of a particular ML product for rare disease diagnosis in the United Kingdom (UK). The analysis showed that the ML product faces similar market access challenges as other AI technologies in healthcare. These are mainly challenges related to data. However, because the ML product is explicitly needed for the diagnosis of rare diseases, thus, only for a minority of the population, market access for the ML product is particularly difficult and requires much more time and resources.

Table of contents

List of figures

List of tables

List of abbreviations

AI	Artificial intelligence
AMC	Academic medical centre
EFPIA	European Federation of Pharmaceutical Industries and Associations
EHR	Electronic health record
EU	European Union
FDA	Food and Drug Administration
GDPR	General Data Protection Regulation
GMSA	Genomic Medicine Service Alliance
GP	General practitioner
HIT	Health information technology
HR	Human resources
ICD	International Classification of Diseases
IT	Information technology
IVDR	In Vitro Diagnostic Medical Device Regulation
MDD	Medical Device Directive
MDR	Medical Device Regulation
MHRA	Medicines and Healthcare products Regulatory Agency
ML	Machine learning
NHS	National Health Service
TOE	Technology-Organisation-Environment
UK	United Kingdom

1. Introduction

A major challenge in the treatment of rare diseases is receiving the right diagnosis (Schaefer et al., 2020, p. 2). Patients with rare diseases report many years of diagnostic odyssey; they are often diagnosed too late or not at all (Evans & Rafi, 2016, p. 550). The problem with rare diseases lies in the large number of different diseases, each of which is only associated with a small number of cases. Due to the low prevalence and high clinical complexity of rare diseases, doctors often lack knowledge and experience which leads to delayed diagnosis or misdiagnosis (Zhang et al., 2022, p. 2). However, early rare disease diagnosis would allow to treat the patients with appropriate measures so that symptoms could be stopped or at least slowed down. This could lead to reduced healthcare costs and sometimes even save lives (Decherchi et al., 2021, p. 1).

Given the diagnostic and treatment challenges, patients with rare diseases could particularly benefit from AI and ML technologies. Indeed, while it is virtually impossible for a doctor to memorise information about numerous rare diseases, modern computers can easily recall large amounts of digital information (Schaefer et al., 2020, p. 2). If computers were able to extract this information and make sense of it, for example by classifying patients into disease groups or predicting treatment outcomes, this would have great potential for improving both diagnosis and treatment (Schaefer et al., 2020, p. 2).

However, despite the praise and media attention AI and its potential have received in recent years, few medical practices have yet adopted ML (Lebcir et al., 2021; Sun & Medaglia, 2019). There are currently several AI pilot projects underway, but only a few of them succeed and go into production (Benbya et al., 2020). Thus, the adoption of AI technologies in healthcare organisations is increasing, but only slowly. Since ML systems seem to have a particularly great potential for the diagnosis of patients with rare diseases, the following question arises:

Why is the adoption of ML-based diagnostics for rare diseases increasing slowly in healthcare organisations?

This paper attempts to answer the research question based on findings of a case study embedded in the Screen4Care research project, a project that offers an innovative research approach to accelerate rare disease diagnosis. The case study examines the factors that hinder the adoption of a particular ML product for medical diagnosis of rare diseases in primary care in the UK. Only when the factors that hinder the adoption of ML-based diagnostics for rare diseases are

known, appropriate measures can be taken to overcome these obstacles and thus accelerate adoption.

This paper is structured as follows, the next chapter, chapter two, provides background information on ML technologies for medical diagnosis and, in this context, presents the vision of the Screen4Care research project and of the particular area of focus on which this paper concentrates. Chapter three is dedicated to the theoretical foundation, namely the Technology-Organisation-Environment (TOE) framework, contextualised with findings of previous studies on the adoption of AI in the healthcare sector. This is followed by chapter four, the research methodology. Chapter five provides the empirical findings of the case study. These findings will be discussed in chapter six. Finally, chapter seven concludes by highlighting the principal findings and limitations of the paper as well as opportunities for future research.

2. Background information

This chapter first provides a basic understanding of ML technologies for medical diagnosis. After a general description, the second subchapter will provide a more detailed explanation of how the Screen4Care research project, which this study is embedded in, aims to accelerate the diagnosis of rare diseases with the help of AI algorithms.

2.1 Machine learning for medical diagnosis

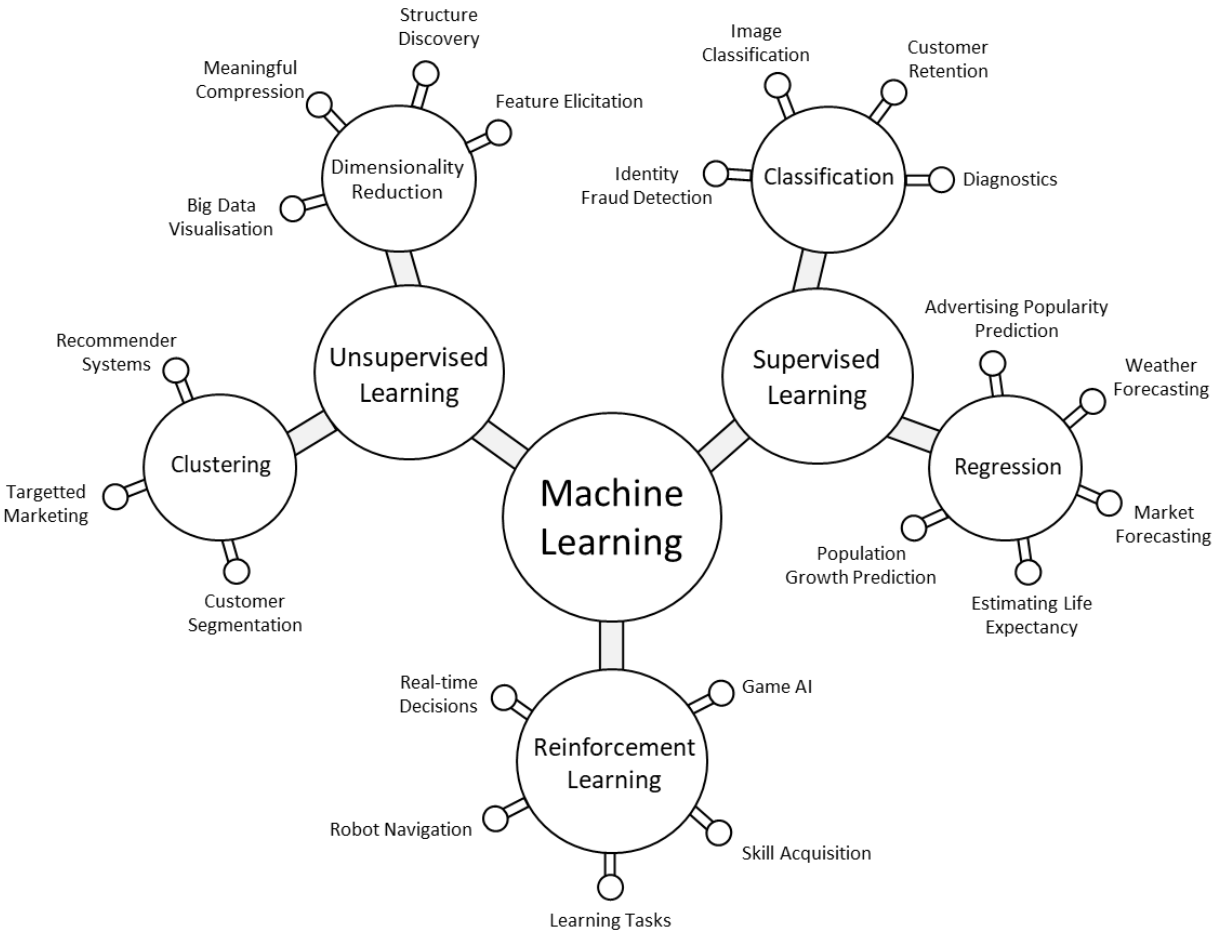
ML is defined as a discipline of AI that enables machines to automatically learn from data and past experience to recognise patterns and make predictions with minimal human intervention (Bini, 2018, p. 2359).

ML algorithms have existed for a very long time. However, the possibility of automatically applying complex mathematical calculations to enormous amounts of data, repeatedly and at an increasing pace is new (Döbel et al., 2018, p. 14-15). Further, it owes its success to new and, above all, more powerful computer technology (Döbel et al., 2018, p. 14-15).

In general, ML has the potential to be used in almost all industries. However, ML has particularly great potential in the health sector, where it can help to detect diseases earlier, provide people with better care, and reduce healthcare expenditure (Verma & Verma, 2021, p. 2144). Examples of ML applications in healthcare include medical imaging diagnosis, improved radiotherapy, personalised treatment, crowdsourced data gathering, smart health records, ML-based behavioural modification, and ML-based medical diagnostics (Verma & Verma, 2021, p. 2144).

As illustrated by Figure 1 below, ML can be roughly divided into three different types of learning, namely reinforcement learning, unsupervised learning, and supervised learning (Stuart, 2010). In reinforcement learning, the goal of the algorithm is to find the most appropriate action to maximise a reward, which in turn depends on the action (Brasil et al., 2019, p. 2). In unsupervised learning, there is input data but no targets. It is then the algorithm's task to reveal an underlying structure in the data (Brasil et al., 2019, p. 2). In supervised learning, the data is labelled, which means that the algorithm receives the input data together with the corresponding target data (Brasil et al., 2019, p. 2).

Figure 1: *Different types of machine learning*



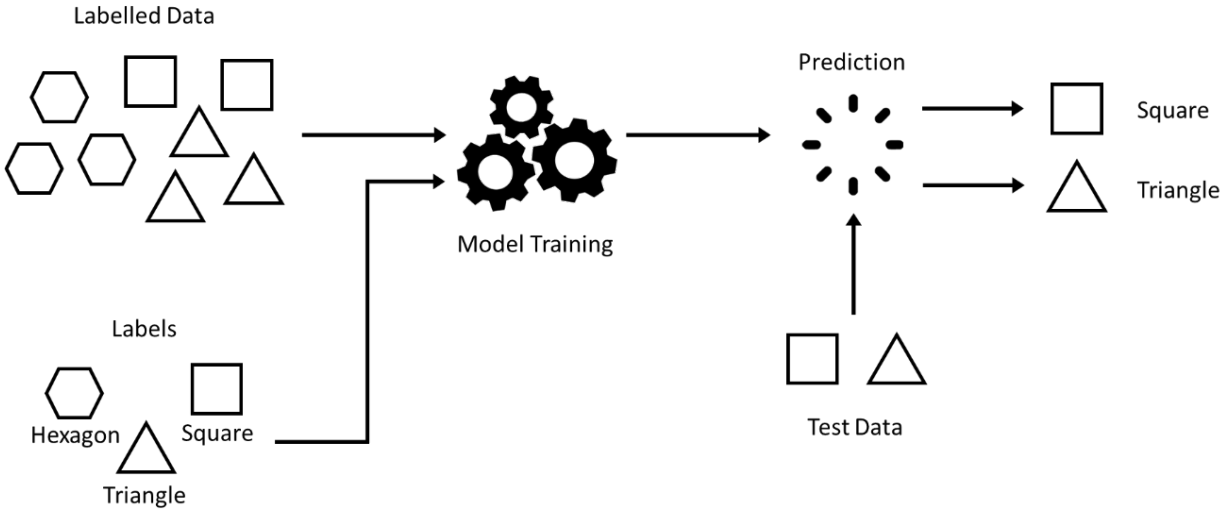
Source: Own representation based on Dhruv, 2021

The choice of the learning algorithm depends on the particular problem that needs to be solved. In medical research, reinforcement learning is used for maximizing favourable results, for example to optimise the overall waiting time of patients in the emergency room. Unsupervised learning is used for phenotyping a disease (Alanazi, 2022, p. 2-3). Supervised learning, however, is commonly used for prognoses and diagnoses (Wiens & Shenoy, 2018, p. 150). As the use of

ML systems for medical diagnosis of rare diseases is of particular interest in this paper, supervised machine learning will be further looked at.

In supervised learning, two types can be distinguished, namely classification and regression (Stuart, 2010; Murphy, 2012; Marsland, 2011). As shown in Figure 1, medical diagnostics are categorised as classifications. For the software to learn and find solutions on its own, prior action by humans is necessary. Thus, if you want to teach a computer how to recognise a rare disease such as, for example, cystic fibrosis, the programmer first needs to find out the features, i.e. the diagnostic criteria, of cystic fibrosis (Nasteski, 2017, p. 4). To teach the computer to recognise as many rare diseases as possible, the programmer must know the characteristics of as many rare diseases as possible and must record them as classes in a labelled data set, as shown in Figure 2 below (Nasteski, 2017, p. 4).

Figure 2: *Supervised machine learning*



Source: Own representation based on Supervised Machine learning - Javatpoint, n.d.

This labelled dataset is called the "training dataset" or "ground truth" (Bini, 2018, p. 2359). The model then tries to predict the result on its own, using test data and can compare its guess with the stored correct results. This process is called model training (see Figure 2). If the output of the trained model deviates from the desired result, the learning algorithm adjusts the model (Stuart, 2010; Murphy, 2012; Marsland, 2011). After a successfully completed learning process of the training dataset, the software has learned, for example, which combination of features is associated with which type of rare diseases. The programmer then receives a model as output that can be used to evaluate new and unknown patient data from electronic health records (EHRs) (Stuart, 2010; Murphy, 2012; Marsland, 2011).

Therefore, as demonstrated, ML algorithms can, similar to doctors, learn to see patterns. One major difference, however, is that the algorithms must learn from a large number of concrete examples (L'heureux et al., 2017, p. 7777). Thus, big data is of great relevance. The more training data a learning algorithm receives, the more likely it is to improve its model and reduce the error rate (Döbel et al., 2018, p. 11). Moreover, the data must be neatly digitised, as machines are unable to read between the lines in textbooks (L'heureux et al., 2017, p. 7777). Thus, ML is particularly helpful where the diagnostic information examined by the doctor has already been digitised. Further, it is essential to keep the model general enough for it to work well on new data that was not included in the training phase (Döbel et al., 2018, p. 11). In addition, the models should be robust, which means they should show similar reactions to similar inputs. The quality of a model also depends on the quality of the training data. If the algorithm is shown too many wrong examples, it cannot learn the correct answers (Döbel et al., 2018, p. 47). If the examples are not representative, the outputs are also subject to greater uncertainty for novel inputs. Only some models, together with their output, can provide an estimate of how sound the output is (Döbel et al., 2018, p. 47). Moreover, ML systems are sometimes associated with 'black boxes', because they are becoming increasingly complex and the mechanisms with which ML systems make their predictions are often opaque to humans (Holzinger et al., 2018, p. 296). For instance, ML systems based on deep neural networks make predictions with millions of parameters and humans are unable to understand every single calculation (Bini, 2018, p. 2359). The general comprehensibility of the models and their results in individual cases is important. Decision trees can be interpreted particularly well, deep neural networks, on the contrary, are often understood poorly (Döbel et al., 2018, p. 22).

ML-based technologies for medical diagnostics raise new legal issues from different perspectives. These include liability in the events of damage and defects, responsibility of content and copyright issues, transparency of decisions, data and consumer protection or the question of the extent to which the decisions of such machines must be followed by medical professionals (Döbel et al., 2018, p. 11). Thus, the central challenge is to design ML systems in a way that they are compatible with our concepts of society, law, and values (Bitkom, 2017, cited in Döbel et al., 2018, p. 11).

2.2 Vision of the Screen4Care research project

In Europe, a disease is considered rare when it affects less than one in 2'000 people (Sernadela et al., 2017, p. 2). More than 7'000 rare diseases exist worldwide (Groft et al., 2021, p. 2711). Although individually rare, in total rare diseases affect over 30 million people in Europe and an

estimated 350 million people worldwide (Ronicke et al., 2019, p. 2). Rare diseases are in over 70% of cases genetic and affect predominately, in 75% of the cases, children (Sernadela et al., 2017, p. 2). Moreover, rare diseases are often serious, to the extent of being life-threatening, especially if they are not diagnosed and treated (Melnikova, 2012, p. 267). However, the low prevalence of individual cases of rare diseases leads to a lack of research and expertise in the field (Brasil et al., 2019, p. 2). It is due to this neglect of rare diseases in medical research, that they are also called health orphans (Schieppati et al., 2008, p. 2040). Unsurprisingly, the diagnostic odyssey and uncertainty associated with the onset of symptoms place a great burden on the affected people as well as their families, carers, doctors, and society as a whole (Groft et al., 2021, p. 2711). On average, people with rare diseases endure eight years of inconclusive consultations and possible misdiagnosis, which can lead to ineffective treatments and inefficient use of health resources (*IMI Innovative Medicines Initiative | Screen4Care,* 2021).

The Screen4Care research project aims to put an end to this diagnostic odyssey (*Sitem Insel*, n.d.). The project is carried out by an international public-private consortium of 35 partners and has a duration of five years, i.e. it will run until September 30, 2026 (*What Is Screen4Care? | Screen4Care*, n.d.). Moreover, the project has a total budget of €25 million, which is provided by the Innovative Medicines Initiative, a joint venture of the European Union (EU) and the European Federation of Pharmaceutical Industries and Associations (EFPIA) (*What Is Screen4Care? | Screen4Care*, n.d.). In order to accelerate the diagnosis of patients with rare diseases, the research project relies on genetic newborn screening and advanced analysis methods such as ML (*What Is Screen4Care? | Screen4Care*, n.d.). The Screen4Care research project is organised into six interconnected areas of focus, also known as work packages, involving all relevant stakeholders, such as researchers, clinicians, patients, data scientists, pharmaceutical companies, patient organisations, academics, small and medium-sized enterprises, public health policy makers, regulators and health technology assessment experts (*Areas of Focus | Screen4Care*, n.d.).

This paper is a contribution to the first area of focus of the project. This work package one seeks to understand how diagnostic algorithms are made available to healthcare providers and aims to identify the current and future role of public funding and policy for ML technologies for rare diseases in Europe (*Areas of Focus | Screen4Care*, n.d.). For this purpose, research staff located in Bern, Switzerland, are conducting two-way interviews on the expectations and experiences of developers of AI-based screening and diagnostic tools as well as of clinical sites that have adopted such algorithms. Regardless of their product's current development status, respondents

are ideally familiar with the challenges and opportunities which their organisation has experienced in adopting AI-based diagnostic algorithms.

3. Theoretical foundation

The acquisition of ML-systems for medical diagnosis of rare diseases represents essentially an organisational innovation. Therefore, this paper should not draw on a theory of individual perceptions, but on a theory of organisational adoption of innovations. The following chapter first discusses the most commonly used theory of organisational information technology (IT) adoption, namely the Technology-Organisation-Environment (TOE) framework. To suit the context of this paper, the framework is adapted based on the existing literature on AI adoption in healthcare. The second sub-chapter is dedicated to the summary of the context-adapted TOE framework and, in this context, also provides an overview of the literature that is used in this paper.

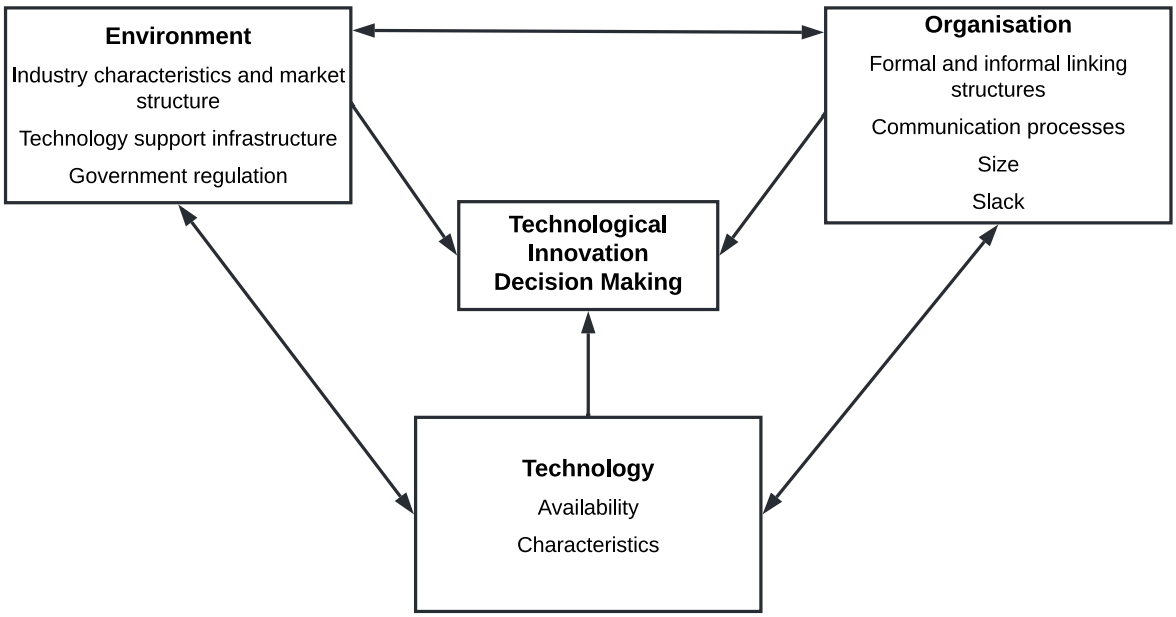
3.1 Technology-Organisation-Environment (TOE) framework

The TOE framework was originally developed by Louis G. Tornatzky and Mitchell Fleischer in 1990 (Ahmadi et al., 2015, p. 53) to identify the dominant factors that influence the technology adoption process and its implementation in firms (Sukardi et al., 2021). Today, however, the framework is applied to different types of organisations. The framework focuses on higher level attributes, i.e. macro-level analysis, rather than detailed behaviours of individuals, i.e. micro-level analysis, in an organisation.

Tornatzky and Fleischer (1990) assume that technology adoption at the organisational level can be influenced by factors related to the technological, organisational, and environmental context. As shown in Figure 3 on the next page, factors in the technological context include *availability* and *characteristics of the technology*. In the organisational context, they include *formal and informal linking structures*, *communication processes* as well as organisational *size* and *slack* (see Figure 3). Finally, factors in the environmental context include *industry characteristics and market structure*, *technology support infrastructure* and *government regulation* (see Figure 3). According to Tornatzky and Fleischer (1990), these factors influence the adoption process of the respective technology either positively or negatively. It is of significance to mention that factors are not only interlinked in the same context, but also to factors from other contexts. This is illustrated by the arrows in Figure 3. For example, government regulation may depend on the

characteristics of a technology; a technology that carries many potential risks usually requires more regulation than vice versa.

Figure 3: *Technology-Organisation-Environment (TOE) framework*



Source: Own representation based on Tornatzky & Fleischer, 1990, p. 153

While the framework provides a useful lens through which to view technology adoption in organisations, the framework is not flawless. This is because different types of innovations may have different factors influencing their adoption, also depending on their environmental setting (Baker 2012, p. 236). Therefore, relevant factors for any specific research question must be defined with the help of previous studies and theoretical implications or even with further theories (Baker 2012, p. 236).

Indeed, as outlined in chapter two, ML systems for medical diagnosis have several specific characteristics that cannot be compared with those of other health information technologies (HITs). Thus, the given factors of the TOE framework, shown in Figure 3, are not sufficient to capture the characteristics of ML systems for medical diagnosis of rare diseases; adjustments and extensions to the framework are needed. The literature research shows that a few studies investigating the adoption process of AI technologies in healthcare have already been conducted. Therefore, the potentially relevant factors for the adoption of ML systems for medical diagnosis of rare diseases will be defined using existing studies on the adoption of AI in healthcare. In the following sections, the respective factor of the original TOE framework is always initially explained, then followed by corresponding findings of studies on the adoption of AI in healthcare.

3.1.1 The technological context

The technological context includes micro-level assessments made by actors of an organisation when deciding whether to adopt a technology. According to Tornatzky and Fleischer (1990), this context includes both the *availability* and the *characteristics of the technology*.

Availability

The first component of the technological context is the availability of the technology, which is described as the extent of an organisation's technical opportunities in the external marketplace (Tornatzky & Fleischer, 1990, p. 163). The number, quality, and applicability of innovations available to a particular organisation may vary. Research on human behaviour has examined how difficult it is for individuals to make decisions under complex market conditions, for example when there is an imbalance of demand and supply. Theoretically, the origin of the impact of market complexity on human behaviour can be traced back to the argument of bounded rationality. If an organisation perceives a technology provider's market to be complex, thus, difficult to obtain the suitable technology on normal conditions, it is more likely to not adopt the innovation (Tornatzky & Fleischer, 1990, p. 163-164). So there is a positive relationship between availability of the technology under normal conditions and innovation adoption (Tornatzky & Fleischer, 1990, p. 163-164).

No study has addressed the factor *availability* in relation to AI adoption in healthcare. As AI technologies are still in their infancy, many AI pilot projects are currently being carried out. However, only a small number of these projects succeed and go into production. In pilots, potential participants are selected, contacted, and invited to participate and thus to adopt the technology in question. Hence, in pilot projects, adopting organisations do not have to deal with complex market conditions to obtain the technology. For this reason, the factor *availability* of the technology is not considered relevant in the adoption of AI in healthcare.

Characteristics

Regarding the characteristics of technology, the TOE framework was found to be consistent with the Diffusion of Innovations (DOI) theory of Rogers (1995). In this context, Rogers (1995) proposes five variables including *relative advantage, compatibility, complexity*, and *observability*.

- *Relative advantage*

Relative advantage refers to the extent to which the innovation is perceived to have improved compared to its predecessor, because it is more productive, more efficient, less costly, or advances in some other way on existing practices (Rogers, 1995, p. 213). According to Rogers (1995, p. 217), there is a positive relationship between relative advantage and innovation adoption.

Interview findings of several studies on AI adoption in healthcare have demonstrated that the perceived relative advantages of AI technologies positively influence their adoption in healthcare organisations. One of the influencing opportunities of AI technologies, which is often mentioned, is its increases in effectiveness (Al Badi et al., 2022, p. 203; Fan et al., 2020, p. 584; Hofmann et al., 2019, p. 12; Pumplun et al., 2021a, p. 10; Pumplun et al., 2021b, p. 6323; Hercheui et al., 2021, p. 80). ML technologies are prone to bias from the input data, but they are robust to human bias and cognitive shortcomings such as fatigue and thus seem to be effective (Hofmann et al., 2019, p. 6; Hemmer et al., 2022, p. 4). In the study by Hercheui et al. (2021, p. 81), a clear majority of the clinicians, 75% of respondents, is convinced that AI solutions can improve the accuracy of diagnosis. Similarly, the interviewed physicians in the study by Hemmer et al. (2022, p. 4) argued that AI can in some cases provide a second medical opinion on ambiguous clinical findings, for example by recognising patterns in high-dimensional data. Further, there is the possibility of mutual improvement through bilateral feedback loops; medical experts can contribute new insights and training data to increase the accuracy of the AI model, and conversely, the originally flawed expert can also learn from the AI-initiated feedback (Hemmer et al., 2022, p. 5). According to experts, another opportunity of ML systems that have an even greater impact on adoption is the increase in efficiency through time savings and lower costs, and patients can benefit from faster decision-making (Fan et al., 2020, p. 584; Hemmer et al., 2022, p. 4; Hofmann et al., 2019, p. 7; Pumplun et al., 2021b, p. 6323; Pumplun et al., 2021a, p. 80; Hercheui et al., 2021, p. 84; Weinert et al., 2022, p. 1). Clinicians in the interview analysis by Hercheui et al. (2021, p. 80) even expect AI algorithms to be able to diagnose people remotely, and consequently, they no longer have to physically go to the doctor's office. So far, however, generating income for the organisation using AI technologies is not among the relative advantages of AI technologies, according to hospital managers and doctors interviewed in Sun and Medaglia (2019, p. 374). The opinions of the chief information officers questioned in Weinert et al. (2022, p. 6) differed in this regard. Indeed, half of respondents argued that AI does not lead to financial savings in the hospital, while the other half claimed the opposite. Nevertheless, as argued by many interviewees, the error rates of ML systems are

still too high, which can have devastating consequences in the healthcare sector, where human lives are at stake. It can thus be assumed that the relative advantage of future ML systems will lie primarily in their use as intellectual decision-making aids and not in the complete automation of medical diagnosis (Pumplun et al., 2021b, p. 6323; Morrison, 2021, p. 651; Watson et al., 2020, p. 170; Hercheui et al., 2021, p. 84; Hemmer et al., 2022, p. 5).

- *Compatibility*

Compatibility refers to the extent to which the innovation is perceived compatible with the current values, past experiences, and needs of potential users (Rogers, 1995, p. 223). According to Rogers (1995, p. 226), there is a positive relationship between compatibility and innovation adoption.

The majority of chief information officers questioned in Weinert et al. (2022, p. 5) agreed that lack of compatibility of AI models is a barrier to their adoption, because the technologies either do not fit the setting or are clinically unusable. Interview participants in Morrison (2021, p. 650) emphasised that the adoption of AI technologies in primary care faces different challenges than in secondary care. One expert of the regulatory body argued that the cooperation between technology suppliers and general practitioner (GP) practices is not easy, because of the high number of GP practices (Morrison, 2021, p. 650). There are relatively few hospitals in secondary care compared to the number of GP practices in primary care. Accordingly, collaboration between technology providers and secondary care seems to be easier. Moreover, one doctor interviewed in Morrison (2019, p. 650) explained that GPs often deal with mental health issues, chronic illnesses, disabilities and other medical conditions which are non-digitizable healthcare problems. Following this argument, AI better suits secondary healthcare, as secondary healthcare includes different medical specialties. Further, digitalisation is more advanced in secondary healthcare than in primary healthcare, which is another reason why AI can be adopted more easily in secondary healthcare. One doctor in the interview analysis by Morrison (2021, p. 650) argued that those specialties, in which doctors work with patterns, such as dermatology, radiology, pathology and ophthalmology, are particularly well suited to AI technologies. The survey analysis by Weinert et al. (2022, p. 7) revealed that only 7% of the questioned chief information officers felt that the supply of applicable AI solutions on the technology market was sufficient for their needs. Another 58% responded that they were unsure (Weinert et al., 2022, p. 7). AI technology offers on the market that do not meet the requirements of potential customers will not be adopted (Weinert et al., 2022, p. 7-8). The leaders from academic medical centres (AMCs) interviewed in Watson et al. (2020, p. 169) pointed out that an AI model is only compatible

with the values of a healthcare organisation, if it is clinically useful. Respondents argued that the involvement of clinicians and other stakeholders throughout the model development cycle is essential to create a successful clinically implemented model (Watson et al., 2020, p. 169). Another factor affecting compatibility is the difficulty of configuring alerts in AI models. It is challenging to find the right balance between over- and under-alerting when action is required, as over-alerting can unnecessarily distract healthcare professionals from their work (Watson et al., 2020, p. 169). According to the chief information officers questioned in Weinert et al. (2022, p. 169), it is significant that an AI model is compatible with the workflows of medical experts, that it is integrated into their work and does not distract them from it. Lack of time is one of the most important stressors in physicians' daily work according to medical professionals interviewed in Hemmer et al. (2022, p. 5). Indeed, the interview analysis with physicians by Hemmer et al. (2022, p. 5) revealed that the time which is consumed by collaboration with AI needs to be kept to a minimum. Further, the collaboration should be integrated into existing workflows and not disturb medical professionals, otherwise it may also hinder the adoption of such technologies (Hemmer et al., 2022, p. 5).

- *Complexity*

Complexity refers to the perceived degree of difficulty in understanding and using the innovation (Rogers, 1995, p. 230). According to Rogers (1995, p. 231), there is a negative relationship between complexity and innovation adoption.

A survey conducted with healthcare professionals by Fan et al. (2020, p. 584) indicates that the ease of use and understanding of AI technology do not have an impact on its adoption. However, half of the interviewed clinicians in Hercheui et al. (2021, p. 81) argued that the fact that AI is considered to be complex and difficult to understand is a barrier to its adoption, as this is a main reason why medical professionals do not implement it in their daily use. Experts in the study by Pumplun et al. (2021a, p. 8) argued that ML systems based on neural networks can consist of several processing layers and up to billions of numerical weights, which makes ML systems difficult for humans to understand. Thus, the term 'black box' is often used in the context of ML technologies. Experts claimed that this characteristic could hinder the adoption of such technologies in healthcare organisations (Morrison, 2021, p. 651; Pumplun et al., 2021a, p. 8; Pumplun et al., 2021b, p. 6320; Watson et al., 2020, p. 169). Similarly, government policymakers and IT firm managers interviewed by Sun and Medaglia (2019, p. 373) identified a lack of knowledge in the general public on the values and advantages of AI due to its complexity. They view this as a barrier to the adoption of AI. However, the interviewed hospital managers

did not raise any concerns in this regard (Sun & Medaglia, 2019, p. 373). Contrasting opinions were also expressed in the study by Morrison (2021, p. 651). Some interview participants believe that the 'black-box' nature of AI technologies poses a difficulty for the adoption in healthcare, while others denied this concern and believe that 'black-box' AI is theoretical, and AI will never be completely unexplainable. Further, half of the clinicians interviewed in Hercheui et al. (2021, p. 81) have a pragmatic approach to dealing with the complexity of AI and believe that there is no need to understand the AI algorithm to see its advantages and usability. However, most respondents claimed the "lack of transparency" of ML systems due to their complexity is a major barrier to their adoption in healthcare organisations (Al Badi et al., 2022, p. 203; Pumplun et al., 2021a, p. 8; Pumplun et al., 2021b, p. 6320; Hemmer et al., 2022, p. 5; Watson et al., 2020, p. 169). In the study by Sun and Medaglia (2019, p. 376), government policymakers, hospital managers and doctors also mentioned the challenge of lack of transparency, while IT company managers argued that there is no reason to have a different attitude towards AI technologies than other popular technologies whose inner workings are also unknown to users. Nevertheless, many experts wonder how a doctor can explain a diagnosis or treatment recommendation to their patients using AI solutions, if they do not understand the algorithm (Pumplun et al., 2021a, p. 8; Pumplun et al., 2021b, p. 6320; Hemmer et al., 2022, p. 5; Watson et al., 2020, p. 169). Moreover, patient groups may not accept AI diagnoses, if they do not trust doctors to understand the AI algorithms.

- *Trialability*

Trialability refers to the scope and possibility of testing the innovation on a limited scale (Rogers, 1995, p. 231). Innovations are easier to adopt if they can be tried out in part, can be used on a temporary basis, or are easily dispensable after being tried out. According to Rogers (1995, p. 232), there is a positive relationship between trialability and innovation adoption.

No study has addressed the factor *trialability* in the context of AI adoption in healthcare. As AI pilots are about testing AI innovations in a less critical context before applying them on a larger scale, the factor *trialability* is not decisive in determining whether a healthcare organisation participates in an AI pilot, and thus adopts an AI technology or not. This factor can therefore be excluded in the adoption process of AI technologies in healthcare.

- *Observability*

Observability refers to the degree to which the results of an innovation are visible to others (Rogers, 1995, p. 232). Observability can improve the diffusion effect, a crucial component of technology transfer. According to Rogers (1995, p. 232), there is a positive relationship between observability and innovation adoption.

Some experts claimed that real-world examples of AI used in the healthcare sector would contribute to better understanding of such technologies and dispel some fears, as people would be able see the benefits of the technologies, thus making adoption more likely (Morrison, 2021, p. 651; Watson et al., 2020, p. 169). For this reason, in the study by Watson et al. (2020, p. 169) leaders from AMCs highlighted that some institutions started to develop evaluation procedures of their models to measure clinical impact. According to the interviewed experts, there is a need for education and communication with the public to explain the benefits of AI technologies and the opportunities they can offer healthcare professionals in their work (Hercheui et al., 2021, p. 82; Morrison, 2021, p. 651; Watson et al., 2020, p. 169). This would help to resolve the distorted clinicians' expectations towards AI technologies, which are caused by the hype around AI and promote adoption (Hercheui et al., 2021, p. 82; Morrison, 2021, p. 651; Watson et al., 2020, p. 169).

- *Application*

Another significant factor that is omitted in the DOI theory of Rogers (1995), but needs to be added in the context of AI adoption in healthcare, is the *application* of AI technologies to different tasks and contexts. One IT expert questioned in Hofmann et al. (2019, p. 9) argued that the fact that ML models for radiology usually deal with a single case and are thus not transferable to other tasks is an obstacle to their adoption. To transfer the ML models to other tasks, the integration and testing process would have to be repeated several times, which is not economically justifiable (Hofmann et al., 2019, p. 8). This transferability of an AI tool from one environment to another was also highlighted as a challenge by several interview participants in the study by Morrison (2021, p. 651). One expert of the regulatory body mentioned the example of a ML provider in radiology in North London, where the sensitivity and specificity was very high. When the ML provider was used in South London with a different ethnic group, and therefore with different data, scanner and radiology positioning, the provider did not work well (Morrison, 2021, p. 651). In addition, one government policymaker interviewed in Sun and Medaglia (2019, p. 373) explained that because of racial differences between China and

Western countries, the causes of diseases may be different. Therefore, in Western countries, less data is available on disease patterns that are more common in China, but less common in Western countries. Thus, since the Watson system, a specialised AI system, was originally trained with data in Western countries, its adoption in China is challenging (Sun & Medaglia, 2019, p. 373). ML systems are indeed limited in their ability to diagnose specific diseases in different contexts, as they can adapt their functions when trained with new data (Pumplun et al., 2021a, p. 8; Pumplun et al., 2021b, p. 6323). This characteristic of limited applicability is not only relevant when a ML system is transferred to another clinic, but also when it needs to be re-trained after some time, e.g. when new medical research results are obtained or the demographic structure of the patients changes (Pumplun et al., 2021a, p. 8). Physicians in the study by Pumplun et al. (2021a, p. 8) argued that to adopt ML systems, clinics need a clear strategy on how to deal with the opacity and adaptability of self-learning ML systems. Further, the experts interviewed in Hemmer et al. (2022, p. 5) mentioned the adaptability of the technology depending on the user as another important factor for adoption. According to these experts, AI-based clinical decision support systems should take into account different needs by offering different levels of information density depending on the level of knowledge and background of the collaborators, i.e. whether a nurse or a physician is using the system (Hemmer et al., 2022, p. 5). Another challenge for the widespread application of AI technologies is the different ways of working of the individual care units, such as hospitals, GP practices, etc. It is challenging to implement AI technologies in different contexts, as clinicians in Hercheui et al. (2021, p. 82) elaborated. Only more sophisticated AI systems can combine multiple data sources, and according to these experts, this lack of data interoperability hinders the application or adoption of AI solutions on a large scale at national level. Thus, AI is more likely to be adopted according to the usual case-by-case approach (Hercheui et al., 2021, p. 82).

3.1.2 The organisational context

The organisational context encompasses macro-level assessments made by actors of an organisation when they decide whether to adopt a certain technology, including the following four components: *formal and informal linking structures*, *communication processes*, *size* and *slack* (Tornatzky & Fleischer, 1990, p. 153).

Formal and informal linking structures

The first component of the organisational context is formal and informal linking structures. These structures are processes outside an organisation that allow the organisation to scan the

environment for information about the needs and opportunities for technological change. These structures ultimately enable the organisation to process and relay this information to support decisions about technology adoption (Tornatzky & Fleischer, 1990, p. 157). An example of such a linking structure is the direct contact and information flow between managers of different organisations (Galbraith, 1973, cited in Tornatzky & Fleischer, 1990, p. 157). There is a positive relationship between formal and informal linking structures and innovation adoption (Tornatzky & Fleischer, 1990, p. 155).

In the context of AI adoption in healthcare, formal and informal linking structures can be contextualized as formal relationships that facilitate the flow of patient data, i.e., *data sharing*, from one medical facility to others. Several interviewees in the study by Morrison (2021, p. 651) talked about the "fragmented data pool" within the National Health Service (NHS) in the UK. As these interviewees explained, the fragmented data pool poses a barrier to data collection and thus to AI adoption, as there is not just a single point of contact where you can get access to every patient record (Morrison, 2021, p. 651). However, to facilitate data collection, the respective organisations must first be willing to share their data. Indeed, lack of willingness for data sharing is identified by several studies as a major obstacle to AI adoption (Morrison, 2021, p. 651; Sun & Medaglia, 2019, p. 651; Pumplun et al., 2021b, p. 6324). In the study by Sun and Medaglia (2019, p. 375), government policymakers and IT firm managers argued that there seems to be tension between the need for data sharing and the interest of individual organisations. In China, a patient's data is owned by the healthcare organisation where the patient was treated. The healthcare organisation views data as a valuable resource. Therefore, healthcare providers are not willing to share their patient data with other providers (Sun & Medaglia, 2019, p. 375). In Europe, in contrast, the General Data Protection Regulation (GDPR) only permits the processing of health data if this is explicitly accepted by the patient or if the clinic can provide specific reasons for the use of the data (Pumplun et al., 2021a, p. 10). However, anonymised personal data is not covered by the GDPR. The interview analysis by Pumplun et al. (2021b, p. 6324) revealed that doctors fear data misuse, if they share their patient data with other healthcare organisations. Regarding international sharing of patient data, hospital managers and doctors as well as government policymakers interviewed in the study by Sun and Medaglia (2019, p. 375) even expressed concerns about possible national security threats to China from foreign AI companies when they manage sensitive Chinese patient data.

Communication processes

The second component of the organisational context is the communication processes. While the formal and informal linking structures are external to the organisation, the communication processes are internal to the organisation and relate to how information about an innovative technology is disseminated within the organisation (Tornatzky & Fleischer, 1990, p. 159-160). In particular, top management can foster innovation by creating an organisational environment that welcomes change and supports innovation that promotes the organisation's core mission and vision (Tornatzky & Fleischer, 1990, p. 160). In general, there seems to be a positive relationship between innovation-promoting communication processes in an organisation and innovation adoption (Tornatzky & Fleischer, 1990, p. 160).

In the context of AI adoption in healthcare, communication processes can be contextualised as the *support* given by a healthcare organisation to AI technologies. An innovation must be socially accepted to be adopted and some innovations require much time and discussion before they are accepted by society. Experts argue that acceptance of AI technologies is sometimes not given because doctors perceive these kinds of technologies as a threat to their professional identity, i.e. doctors fear that such technologies will replace them (Morrison, 2021, p. 651; Pumplun et al., 2021b, p. 6323; Sun & Medaglia, 2019, p. 375). However, according to the findings by Fan et al. (2020, p. 684), the substitution crisis does not seem to influence the adoption of AI. Many experts argued that the lack of transparency or misunderstandings about AI technologies in particular leads to mistrust as well as fears and thus to a lack of acceptance of such technologies (Fan et al., 2020, p. 584; Hercheui et al., 2021, p. 84; Hofmann et al., 2019, p. 10; Morrison, 2021, p. 651; Pumplun et al., 2021b, p. 6323; Sun & Medaglia, 2019, p. 375). In addition, many respondents elaborated that ethical concerns further lead to a lack of acceptance of AI technologies and hence constitute a barrier to AI adoption (Al Badi et al., 2022, p. 203; Hercheui et al., 2021, p. 84; Hofmann et al., 2019, p. 11; Pumplun et al., 2021a, p. 10; Pumplun et al., 2021b, p. 6322; Sun & Medaglia, 2019, p. 374). Underrepresentation of ethnic or economic minorities in the training of the algorithm may lead to unethical results (Hofmann et al., 2019, p. 11; Hercheui et al., 2021, p. 82). Unethical results can also arise if the algorithm incorporates other data besides health information (Hofmann et al., 2019, p.11). For example, if cost is considered a relevant factor for treatment recommendations, patients with basic insurance could be recommended a less effective but cheaper treatment option for the same condition than people with premium insurance (Kohli et al., 2017, cited in Hofmann et al., 2019, p. 11). Further, experts interviewed in Pumplun et al. (2021b, p. 6322) claimed that ML systems fed with patient data could determine whether a patient is susceptible to a disease and that this kind

of medical application would contradict the patient's "right not to know". In the context of genetic testing, the "right not to know" refers to the idea that adults should be able to decide for themselves whether to receive genetic information, in particular information about the risk of future diseases, and that their wish not to know certain information should be respected. Finally, experts also expressed ethical concerns about data protection, as hackers could gain access to the AI systems (Al Badi et al., 2022, p. 203; Hercheui et al., 2021, p. 82). In summary, various reasons can hinder the acceptance of AI technologies and their adoption. Therefore, in the study by Pumplun et al. (2021b, p. 6321), a doctor interviewed emphasised the importance of an innovative culture within an organisation and, in particular ML support from medical directors, as they are responsible for paving the way for clinics to be ready to adopt ML systems. Thus, according to these findings, the trust and acceptance of medical directors towards ML technologies seem to have a positive influence on the adoption of ML (Pumplun et al., 2021b, p. 6321).

Size

The third component of the organisational context is the size of an organisation. Scholars disagree on the best method for assessing an organisation's size. However, the TOE framework argues that size refers to the amount of work done in an organisation (Tornatzky & Fleischer, p. 1990, p. 169). Large organisations have been found to be more innovative due to their increased access to resources and ability to employ expert staff (Tornatzky & Fleischer, 1990, p. 169). Thus, there is a positive relationship between large organisational size and innovation adoption (Tornatzky & Fleischer, 1990, p. 169).

Experts in the study by Pumplun et al. (2021b, p. 6321) argued that a larger clinic size facilitates the adoption of ML technologies. Large clinics usually have more resources and care for more patients than small clinics, and thus also have access to more patient data (Pumplun et al., 2021, p. 6321). Leaders from AMCs interviewed in Watson et al. (2020, p. 170) mentioned that small organisational size represents a barrier to AI adoption, as smaller institutions indeed expressed more concerns about financial resources than larger ones.

Slack

The fourth component in the organisation is slack. In the TOE framework, both financial and human resources (HR) are referred to as slack (Tornatzky & Fleischer, 1990, p. 161). The presence of financial slack, the financial resources in an organisation, are a significant driver of technology adoption as it gives an organisation the opportunity to adopt an innovation and to

integrate it (Tornatzky & Fleischer, 1990, p. 161). Besides financial slack, HR slack is also considered to be an important driver of technology adoption. Within the context of the TOE framework, HR slack refers to specialized and skilled personnel resources that are rare and absorbed because "the resources are tied up in the organisation's current operations" (Tornatzky & Fleischer, 1990, p. 161). Thus, there is a positive relationship between slack and innovation adoption (Tornatzky & Fleischer, 1990, p. 161).

In the context of AI adoption in healthcare, slack can be contextualized as the availability of *resources,* such as a financial resources, specialized and skilled personal resources, as well as digital patient data and powerful technical infrastructure in a healthcare organisation (Pumplun et al., 2021b, p. 6321; Weinert et al., 2022, p. 7). A clear majority, 80%, of the chief information officers questioned in Weinert et al. (2022, p. 7) agreed that their hospital lacked financial resources and 90% of them argued that this represented a barrier for AI adoption. Hospital managers and doctors in the study by Sun and Medaglia (2019, p. 374) also elaborated that the adoption of AI tools is resource-intensive for the hospital management with limited financial benefits that do not match the costs. In further studies, physicians argued that the current funding structure of clinics leads to strict budget constraints, and thus represents a barrier to ML adoption for medical diagnosis (Pumplun et al., 2021, p. 9; Pumplun et al., 2021b, p. 6321). In this context, one physician explained that one part of the clinic's budget is used for daily costs such as medication and the other part is used to buy large medical equipment, such as X-ray systems; however, the development and establishment of ML systems is not covered by either part, and a specific ML budget cannot be requested (Pumplun et al., 2021a, p. 9; Pumplun et al., 2021b, p. 6321). Concerning specialized and skilled personnel resources, experts claimed that an insufficient number of personnel with expertise in the field of medicine, but also on the technical side, for instance in data science, hinders the adoption of AI technologies (Pumplun et al., 2021a, p. 9; Pumplun et al., 2021b, p. 6321; Sun & Medaglia, 2019, p. 376; Weinert et al., 2022, p. 5). Moreover, the chief information officers in the study by Weinert et al. (2022, p. 5) added that this lack of skilled personnel leads to a fear of potentially high costs for training and learning phase of AI and this may further hinder AI adoption. Another obstacle mentioned in connection to AI adoption was the insufficient digitalization of healthcare organisations, i.e. a lack of digitized patient data, as well as the low quality of the IT infrastructure in healthcare organisations (Morrison, 2021, p. 650; Weinert et al., 2022, p. 5). A lack of infrastructure, or a lack of the kind of advanced infrastructure in healthcare organisations that is required for optimal AI adoption, is considered a challenge by experts (Morrison, 2021, p. 650; Weinert et al., 2022, p. 5).

3.1.3 The environmental context

The environmental context includes macro-level assessments made by actors of an organisation when deciding whether to adopt a technology and consists of three components: *industry characteristics and market structure*, *technology support infrastructure* and *government regulation* (Tornatzky & Fleischer, 1990).

Industry characteristics and market structure

The first component of the environmental context is industry characteristics and market structure. This component refers to differences in competitive and market conditions (Tornatzky & Fleischer, 1990, p. 167). It is assumed that these differences manifest themselves in the form of mimetic pressure. Mimetic pressure exists when an organisation imitates the actions of similarly structured organisations operating in the same economic network and industry. Imitative behaviour enables an organisation to reduce search costs and is typically more pronounced when issues of relative advantage are at stake (Wolverton & Lanier, 2019, p. 409). It is argued that organisations in fast-growing industries tend to innovate faster; in mature or declining industries, innovation practices are not clear-cut (Tornatzky and Fleischer 1990, p. 168). Hence, there is a positive relationship between competitive pressure and innovation adoption (Tornatzky & Fleischer, 1990, p. 169).

In the context of AI adoption in healthcare, competitive pressure can be contextualized as *public pressure*. No study has yet addressed the factor of competitive pressure in relation to AI adoption in healthcare. The use of AI technologies in healthcare is increasing, but only slowly. It can therefore be assumed that there is no competitive pressure for the adoption of AI technologies in the healthcare sector. However, public pressure either for or against an AI technology can immensely impact the innovation adoption. Compared to other industries, the customer's or, in this case, patient's opinion of the innovation is of enormous importance to organisations in the healthcare sector (Pumplun et al., 2021b, p. 6322). For example, the EU GDPR only permits the processing of personal health data if the patient explicitly accepts or if the clinic can provide particular reasons for the use of the data (Pumplun et al., 2021a, p. 10). Moreover, if healthcare professionals feel that patients are not supportive of such technologies, this may negatively impact healthcare professionals' support for AI technologies. Thus, previous studies found that patient's reluctance towards AI technologies clearly hinders AI adoption (Pumplun et al., 2021a, p. 8; Pumplun et al., 2021b, p. 6323; Hemmer et al., 2022, p. 5; Watson et al., 2020, p. 169). The same applies to policy support. Most studies conclude that public policies

play an indirect, but decisive role in AI regulation and public funding (Pumplun et al., 2021a, p. 9; Pumplun et al., 2021b, p. 6321, Sun & Medaglia, 2019, p. 375; Watson et al., 2020, p. 170). Policymakers can thus encourage the adoption of AI through large public funding and favourable regulations, or hinder the process through insufficient public funding and strict regulations.

Technology support infrastructure

The second component of the environmental context is the technology support infrastructure, which refers to the constraints or opportunities that an organisation must consider when developing its technology sourcing strategy. This strategy depends on labour costs, the skills of the available labour force and access to providers of technology-related services (Tornatzky & Fleischer, 1990, p. 171). Success in acquiring a technology and satisfaction with the technological product are likely when an organisation's employees believe that the purchasing environment is normal or favourable (Tornatzky & Fleischer, 1990, p. 171-173). There is a positive relationship between existing technology support infrastructure and innovation adoption (Tornatzky & Fleischer, 1990, p. 169).

Unlike in software development, ML models cannot be "frozen" at a certain level. Data changes and therefore the models must also be continuously adapted. The maintenance and care of AI technologies must not be neglected under any circumstances. Models that are not updated may do more harm than good. Many experts interviewed in Morrison (2021, p. 650) argued that in some cases the adoption of AI fails only because the technology is funded in isolation, but not the subsequent maintenance and care of the innovation. According to respondents in the study by Morrison (2021, p. 650), it is of significance for healthcare organisations to be able to count on the support of IT and engineering project managers for technical issues. A lack of this support can hinder the adoption of AI (Morrison, 2021, p. 650).

Government regulation

The third component of the environmental context is government regulation. Government regulation can have both positive and negative effects on innovation (Tornatzky & Fleischer, 1990, p. 173). When governments impose new restrictions on industries or apply pressure to find technical alternatives to current practice, innovation is essentially mandated or encouraged for these organisations (Tornatzky & Fleischer, 1990. P. 173). Similarly, strict safety and testing regulations can delay or inhibit innovation in many organisations as the cost of innovation can

be quite high (Tornatzky & Fleischer, 1990, p. 173). Thus, there is either a positive or negative relationship between government regulation and innovation adoption (Tornatzky & Fleischer, 1990, p. 173).

In the context of AI adoption in healthcare, government regulations may impact not only AI adopting organisations, but also AI technology providers and, in turn, organisations which indirectly adopt AI. In Morrison's study (2021, p. 650), eight out of twelve experts highlighted the current regulatory landscape as a barrier to AI adoption. According to these experts, the regulation is confusing for developers to navigate and the roles and remits of regulators are unclear (Morrison, 2021, p. 650). In this context, several participants expressed concerns about legal liability, as no case law on AI exists (Morrison, 2021, p. 650). There is great uncertainty around the question who can be held liable, the HIT provider, the clinic, or the physician (Al Badi et al., 2022, p. 203; Pumplun et al., 2021a, p. 10; Pumplun et al., 2021b, p. 6322). The interviewees in Hofmann et al. (2019, p. 10) mentioned another two key legal challenges, namely regulatory approval, and data protection law. The fact that ML systems have not yet undergone any approval processes is a barrier to their adoption (Hofmann et al., 2019, p. 10; Morrison, 2021, p. 10; Sun & Medaglia, 2019, p. 375). However, regulatory approval remains a challenge due to the so-called 'black box' problem and the lack of a validation framework (Hofmann et al., 2019, p. 10). There is no existing legal basis or agreed gold standard that specifies the necessary approval requirements, such as the accuracy values or the size of the test population for testing the safety of a ML product (Hofmann et al., 2019, p. 10; Morrison, 2021, p. 650; Sun & Medaglia, 2019, p. 375). As ML systems can learn from new experiences and adapt themselves, legal approval is not trivial (Pumplun et al., 2021b, p. 6322; Morrison, 2021, p. 650; Sun & Medaglia, 2019, p. 375). Moreover, the debate about whether privacy laws pose a challenge to ML is contentious in the study by Hofmann et al. (2019, p. 10); two radiologists argued that data protection laws are not a significant obstacle and even facilitate the set-up of ML projects. Another radiologist and an IT expert, however, claimed the opposite (Hofmann et al., 2019, p. 10). Nevertheless, most experts agreed that strict requirements for the protection of sensitive patient data impede the adoption of AI (Morrison, 2021, p. 651; Pumplun et al., 2021a, p. 10; Pumplun et al., 2021b, p. 6322). Some respondents mentioned that strict privacy laws hinder the willingness of clinics to adopt ML systems as clinics are concerned to not obtain the necessary patient data to train the ML system (Pumplun et al., 2021a, p. 10; Pumplun et al., 2021, p. 6322).

Optimal data

Another important factor that is not included in the TOE framework, but should be added in the context of AI adoption in healthcare, is access to great amounts of *optimal data*, which is patient data that is digitized, of high-quality, available in uniform formats and unbiased. Access to such data, however, is not easy. Lack of digitized patient data was mentioned as an obstacle to AI adoption by many respondents in different studies (Pumplun et al., 2021a, p. 11; Pumplun et al., 2021b, p. 6323; Morrison, 2021, p. 651; Hofmann et al., 2019, p. 9). Respondents argued that digital patient data does not exist in every clinic, as many processes in clinics are still paper based; the integration of an electronic medical record system, however, represents a prerequisite to the application of ML systems (Pumplun et al., 2021a, p. 11; Pumplun et al., 2021b, p. 6324). Further, digitized patient data is often stored in unstructured file types, such as images, texts, or videos (Pumplun et al., 2021a, p. 11; Pumplun et al., 2021b, p. 6323; Watson et al., 2020,p. 170). According to experts, the quality of the unstructured data is highly dependent on the particular clinic where the data is generated and its clinical staff, as, for example, doctors' letters are often written in free-text formats which are filled with synonyms and can be interpreted differently by whoever reads them (Pumplun et al., 2021a, p. 11; Pumplun et al., 2021b, p. 6323). Indeed, this lack of data quality has been mentioned as a barrier to AI adoption by several interviewees (Pumplun et al., 2021a, p. 11; Pumplun et al., 2021b, p. 6324, Watson et al., 2020, p. 170; Weinert et al., 2022, p. 6). Data can partially lose quality if it must first be anonymised, as relevant correlations can be lost in this process (Pumplun et al., 2021a, p. 11; Pumplun et al., 2021b, p. 6324). It is also problematic when data must be transferred into a machine-readable format for which there are no uniform quality standards (Pumplun et al., 2021a, p. 11; Pumplun et al., 2021b, p. 6324). However, one physician explained that analysing the quality of patient data is difficult anyway because there is no "ground truth" for a healthy patient, as the human body is a highly complex system (Pumplun et al., 2021b, p. 6324). Biased algorithms, pose another challenge to AI adoption (Pumplun et al., 2021, p. 6324; Morrison, 2021, p. 651; Hofmann et al., 2019, p. 9; Hercheui et al., 2021, p. 84; Watson et al., 2020, p. 170). If ML systems are trained on the basis of a demographically or regionally biased database, the system could draw incorrect, non-generalisable conclusions (Pumplun et al., 2021b, p. 6324). Indeed, AI tools' outputs are only as good as the input of data it relies on (Sun & Medaglia, 2019, p. 373). In addition, as mentioned earlier, ML technologies require huge data sets to achieve the most accurate outputs; the lack of interoperability of different data formats can therefore complicate massive data collection (Hercheui et al., 2021, p. 81; Pumplun et al., 2021a, p. 10; Pumplun et al., 2021b, p. 6323; Sun & Medaglia, 2019, p. 376; Watson et al., 2020, p. 5; Weinert et al.,

2022, p. 5). Patients' medical data is usually provided in a variety of 'proprietary data formats' (Pumplun et al., 2021a, p. 10, Pumplun et al., 2021b, p. 6323). Different clinical legacy systems from different providers need to work together, but they are often difficult or impossible to convert for ML systems, making it very difficult to create uniform formats. However, in order to be able to use as much data as possible, interoperability between different data formats is key (Pumplun et al., 2021a, p. 10; Pumplun et al., 2021b, p. 6323).

Funding

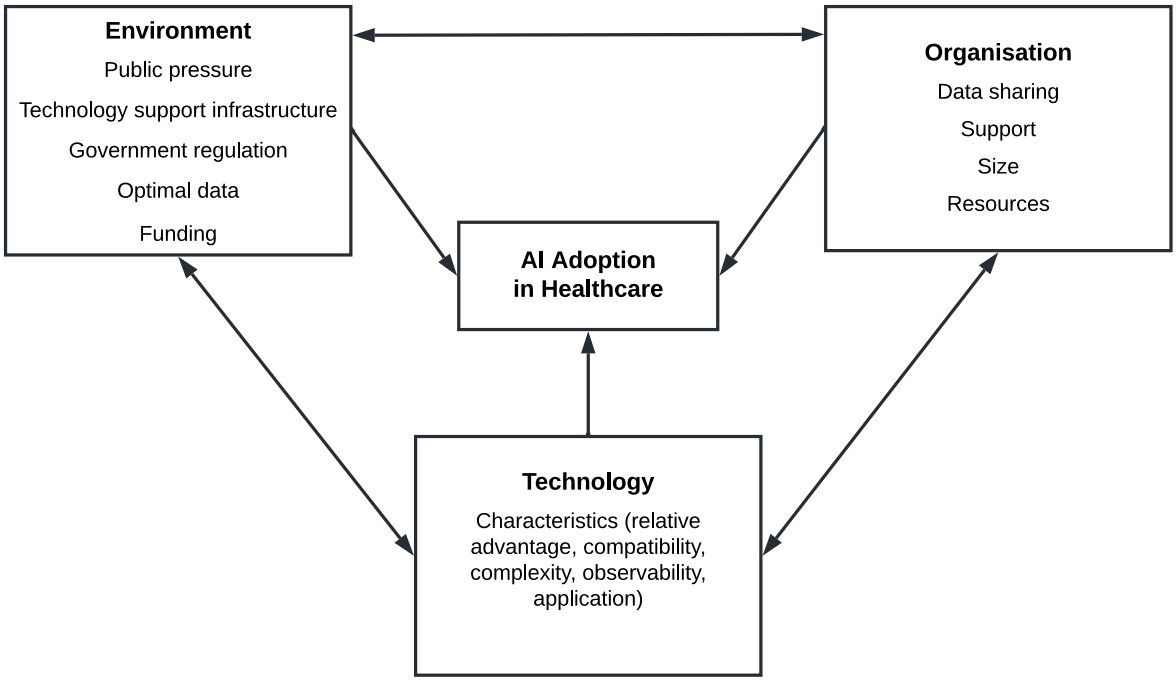
Another important factor that is omitted in the TOE framework, but should be added in the context of AI adoption in healthcare, is *funding*. In the context of ML adoption for medical diagnosis of rare diseases, funding may impact not only AI adopting organisations, but also AI technology providers and, in turn, indirectly AI adopting organisations. Out of all twelve interviewees in the study by Morrison (2021, p.651), six participants, across all four key informant groups, argued that the lack of funding is a barrier to AI adoption in the NHS. According to these respondents, NHS organisations often cannot look beyond the initial start-up costs to future benefits. Watson et al. (2020, p. 170) found in their interview analysis that financial concerns varied across institutions. Several AMCs stated that large parts of their work were not funded, while other respondents had problems with model development stalling or being abandoned altogether due to lack of funding (Watson et al., 2020, p. 170). Only a few institutions that reported having the most models in development stated that they were not concerned about funding (Watson et al., 2020, p. 170). Funding difficulties for technology providers have a negative impact on AI adoption, as this can prevent the development of qualitative AI technologies and thus, logically, also their adoption. In addition, the adoption of AI technologies is resource-intensive, and resources in healthcare organisations are often limited, which is why substantial funding is of significance for healthcare organisations to be able to adopt AI technologies (Pumplun et al., 2021a, p. 9; Pumplun et al., 2021b, p. 6321).

3.2 Summary: TOE framework for AI adoption in healthcare

In the previous chapter, the TOE framework was adapted to the context of AI adoption in healthcare using findings from existing literature. The adapted TOE framework is shown in Figure 4 on the next page; all the factors shown in Figure 4 are potentially relevant in the adoption of ML for medical diagnosis of rare diseases, either by hindering or supporting adoption.

The factor *availability*, which is originally part of the TOE framework, and the factor *trialability*, which originates from the DOI theory, have not proven to be relevant factors in the context of AI adoption in healthcare and have therefore been excluded. Further, the three factors *application*, *optimal data* and *funding* are not part of the original TOE framework. However, as they have been identified in empirical research as influential for the AI adoption in healthcare organisations, they are also considered as potentially relevant factors in the adoption process of ML for rare disease diagnosis.

Figure 4: *TOE framework adapted to the context of AI adoption in healthcare*



Source: Own representation based on Tornatzky & Fleischer, 1990, p. 153

Table 1 on the next page provides an overview of the literature on AI adoption in healthcare addressed in this thesis. All studies are recent, with the oldest dating from 2019. This further underlines that the adoption of AI technologies is still in its infancy. Researchers conducted interviews and surveys with medical and IT experts and managers to find out more about the factors influencing AI adoption in healthcare. In some studies, the respondents also included researchers, policy decisionmakers or employees at regulatory bodies (see Table 1). The interviews were generally analysed using qualitative methods, while quantitative methods were used for the surveys. Moreover, the studies were conducted in different regions, some in European countries such as in Switzerland, Germany and the UK, as well as out of Europe, in the United Arab Emirates, China and the United States. As shown in Table 1, most researchers referenced

a theoretical framework in their study. For example, Pumplun et al. (2021b) and Weinert et al. (2022) also applied the TOE framework.

Table 1: *Literature review*

Literature	Innovation	Data collection	Research method	Country	Theoretical foundation
Al Badi et al. (2022)	AI in healthcare	IT managerial-level executives (n = 27)	Survey and in-depth interviews; qualitative and quantitative analysis with the analytic hierarchy process (AHP) method	United Arab Emirates	-
Fan et al. (2020)	AI-based medical diagnosis support system (AIMDSS)	Healthcare professionals (n= 191)	Paper & Internet survey; quantitative PLS analysis	China	Unified theory of user acceptance of technology and trust theory (UTAUT)
Hemmer et al. (2022)	AI-based clinical decision support system (CDSS)	Physicians (n = 5) and AI experts (n = 5)	Semi-structured in-depth interviews; inductive qualitative content analysis	unknown	Technology acceptance model (TAM) and UTAUT
Hercheui et al. (2021)	AI in healthcare	NHS Clinicians (n = 22)	Semi-structured interviews; deductive qualitative content analysis	United Kingdom	TAM
Hofmann et al. (2019)	ML in radiology	Radiologists (n = 3) and IT/Computer Science experts (n = 3)	Expert interviews; deductive qualitative content analysis	Europe	Radiology business models of Enzmann und Schomer (2013)
Morrison (2021)	AI in healthcare	Key informants (n = 12) in the UK healthcare AI ecosystem (doctors, managers, researchers, personnel at regulatory bodies)	Semi-structured interviews; thematic analysis	United Kingdom	Diffusion of Innovations (DOI) theory
Pumplun et al. (2021a)	ML systems for medical diagnostics in clinics	Medical experts from clinics and their suppliers (n = 22)	Semi-structured interviews; deductive qualitative content analysis	Germany and Switzerland	Framework of non-adoption, abandonment, scale-up, spread and sustainability (NASSS)
Pumplun et al. (2021b)	ML systems for medical diagnostics in clinics	Medical experts from clinics and their suppliers (n = 15)	Expert interviews; iterative multi-cycle coding process	Germany	TOE and NASSS
Sun & Medaglia (2019)	AI system (IBM Watson) in the public healthcare	Hospital managers/doctors, IT firms, and government policy-makers (n = 19)	Semi-structured interviews & analyses of policy documents and secondary data; inductive qualitative content analysis	China	Construct of framing
Watson et al. (2020)	Predictive modelling (PM) and ML in clinical care	Leaders from AMCs with medical and IT background (n = 33)	Semi-structured interviews; inductive qualitative content analysis using the grounded-theory approach	United States	-
Weinert et al. (2022)	AI technologies in hospitals	Chief information officers (CIOs) from hospitals (n = 40)	Web-based surveys; quantitative study design (descriptive analyses)	Germany	Model on AI readiness by Jöhnk et al. (2021) and TOE

Source: Own representation

As most studies analysed interviews, one must be aware that the answers obtained are subject-driven. Interviews with experts from different regions can lead to different findings, as the health systems and government regulations of the individual countries can differ a lot. In addition, Sun and Medaglia (2019) found in their study that different stakeholders sometimes needed very diverse or even contradictory framings of the challenges regarding the adoption of AI technologies in the healthcare sector. For example, government policymakers did not mention economic challenges. IT managers were silent about technological challenges but mentioned social challenges. Policy decision-makers, however, mentioned technological challenges

and hospital managers and doctors commented on economic challenges and were the only group silent on social challenges. However, such extreme differences in opinion depending on the interviewed stakeholder did not occur in any other study. While other studies did not present such extreme differences in the mentioned challenges, different studies have occasionally raised different factors. Further, different opinions were sometimes expressed about the influence of certain factors within a study or between different studies. Nevertheless, all the factors shown in Figure 4 have empirical evidence.

In summary, the literature review demonstrates that a number of studies have examined factors that influence AI adoption in healthcare, either positively or negatively. However, given the media hype around AI in healthcare there is still very little research on the topic. Based on the research findings to date, no hypotheses can be formulated on the adoption process of AI in healthcare. Moreover, no study has yet explicitly investigated the adoption of ML systems for medical diagnosis in the rare disease setting. It is therefore necessary to investigate whether the factors defined in Figure 4 have an influence on ML adoption for rare disease diagnosis and, if so, what influence, whether that is positive or negative, each of these factors has.

4. Research methodology

After completing this theoretical foundation of the TOE framework and the literature review, the following chapter is dedicated to the research methodology of this thesis. First, the research setting will be presented. Then the data collection will be explained and finally, the third sub-chapter with the applied data analysis will follow.

4.1 Research setting: ML product for rare disease diagnosis in the UK

In order to answer the research question - *Why is the adoption of ML-based diagnostics increasing slowly in healthcare organisations?* - this paper looks at the adoption process of a particular ML product for the diagnosis of rare diseases in the UK.

This particular ML product is provided by a private technology start-up company located in London. For reasons of anonymity, the name of the technology provider is not mentioned in this paper. The technology provider was founded in 2015 and has 16 employees. With its ML scan, the company aims to bring early disease identification to healthcare providers in the primary care setting in the NHS. For this purpose, the technology provider has coded the diagnostic criteria for more than 100 rare diseases.

The ML product of the use case is subject to registration and was approved by the Medicines and Healthcare products Regulatory Agency (MHRA). The ML product of the use case is a class one medical device according to the EU Medical Device Directive (MDD). A class one medical device is associated with the lowest risk, meaning that the device has a low to moderate risk to the patient and user. As the ML product is a clinical case finding decision support tool, it is always the doctor who ultimately decides on clinical action and not the ML product itself.

The function of the ML product can be explained in five steps:

1) The particular ML algorithm or software captures disease characteristics from EHRs in a patient population.
2) Patients are matched against published diagnostic criteria for hundreds of rare diseases (and growing).
3) The clinical team and disease specialists of the technology provider collect a detailed medical history for the patients that are flagged by the ML product because they may potentially have a rare disease. This is also called "case finding".
4) Healthcare providers, such as GP practices, receive a report from the technology provider in their inbox describing the patients' details, such as age, NHS number as well as the suspected disease, the reasons for suspicion and the diagnostic pathway, including external sources such as the NHS website or other websites so that the healthcare providers can learn more about the particular rare disease.
5) Finally, healthcare providers decide how best to help each patient by combining their clinical expertise with new findings from the ML product.

According to current data, about 22% of patients flagged by the ML product are placed by GPs on a diagnostic pathway within the NHS.

Currently, the technology provider is helping the Genomic Medicine Service Alliance (GMSA) in England in some pilot projects to find patients in primary care, i.e. GP practices, who should be put forward for genomic testing. Genetic examinations study a person's genetic material to see if there are any changes in genes. They provide information about an existing or assumed risk for the occurrence of a genetic disease or developmental disorders (NHS website, 2021). Genomic testing is mainly used to diagnose rare and inherited health conditions and some types of cancer. Indeed, as already mentioned in Chapter 2.2, rare diseases are predominantly genetic. Returning to the GMSA: The GMSA is a series of seven new facilities in different areas of England, where genetic services are redesigned. This aims to ensure equitable access to both clinical and laboratory services. The project, which involves the NHS North East and Yorkshire

GMSA, concerns the search for cases of rare diseases, i.e. aims to improve and shorten the process of diagnoses for patients with rare diseases. For this purpose, the technology provider of this paper's use case approached various GP practices in the area to take part in this pilot project. Nevertheless, in this pilot project, only half of the GP practices chose to participate. Hence, for half of the practices there must have been decisive reasons not to participate. The screening of patients with the ML product in this particular pilot was planned towards the end of 2021, but was postponed and finally took place from March until April 2022.

4.2 Data collection

For the empirical data collection, a case-study approach was used. As the study is guided by a more general research question and previous research findings are insufficient to formulate hypotheses (Yin, 2018), a case study seemed appropriate to capture the challenges of ML adoption in the rare disease setting. As explained in chapter 2.2, this study is embedded in work package one of the Screen4Care research project. The use case of this thesis was therefore identified in the context of the Screen4Care research project.

Data was obtained through semi-structured expert interviews. This is a method of data collection in which experts are asked a series of open-ended questions followed by probe questions to further explore their responses and the topic of interest (Galletta, 2013). Semi-structured interviews in qualitative research, such as this one, are a mixture of structured and unstructured interviews where some questions are predetermined while others are not (Galletta, 2013).

For this thesis, two-way semi-structured interviews were conducted with key stakeholders, which were experts. They were questioned on the expectations, experiences, and challenges of adopting the ML product in the field of rare diseases. The experts included the CEO of the technology provider of the ML product presented in the previous chapter and a GP partner of a GP practice, located in Northumberland, which has adopted the technology provider's ML product within the pilot of the NHS North East and Yorkshire GMSA. Two GP partners and one salaried GP work in the respective adopting GP practice. Again, for reasons of anonymity, the name of the GP practice is not mentioned in this thesis.

The semi-structured interviews in this paper were based on two interview guides from the Screen4Care research project (see Appendix 1), one for the managing director of the technology provider and one for the GP of the adopting organisation. However, there are no significant differences which can be found between the two interview guides, as they both address concepts

such as perceived ease of use, perceived benefits of the technology, funding, the role of public policy etc.

The CEO of the technology provider served as an early product manager for the tool and was thus instrumental in the development of the ML system. As shown in Table 2 on the next page, the managing director has a background in human genetics, clinical medicine and an MBA. The interviewed GP of the adopting organisation has a degree in medicine. Both interviews were conducted virtually by MS Teams due to travel limitations. The interview with the CEO of the technology provider took place on April 14, 2022 and the interview with the GP of the adopting organisation took place on June 17, 2022.

Table 2: *Overview of experts interviewed*

Interviewee	Background information	Interview details
CEO (male) of the technology provider	Human genetics, clinical medicine, MBA	April 14, 2022 / MS Teams / 31:03 minutes
GP (female) of the GP practice/adopting organisation	Medicine	June 17, 2022 / MS Teams / 22:19 minutes

Source: Own representation

The interviews lasted about 22-31 minutes (see Table 2). The managing director of the technology provider was sent the interview guide in advance, although he claimed that he did not have time to review it before the interview. The GP, however, did not receive the interview guide before the interview. Both interviews were recorded with the consent of the interviewees, the audio file was then imported into MAXQDA and later transcribed intelligently verbatim. For the transcription, the transcription rules according to Kuckartz and Rädiker (2020, p. 2-3) were followed. The interviews were anonymised directly in MAXQDA. Finally, the transcripts of these two interviews formed the basis for the data analysis.

4.3 Data analysis

The interviews were analysed using qualitative content analysis. They were coded directly with the qualitative analysis software MAXQDA. The coding was done according to the deductive category assignment method according to Mayring (2014). Consequently, the category system with its categories was theoretically derived from the adapted TOE framework shown in Figure 4 (see coding scheme in Appendix 3). Thus, the topic areas or concepts of the interview guides (see Appendix 1) were not adopted one-to-one as classification categories. In addition, an inductive component (Mayring, 2014) was included in the analysis during coding by adding a

further statement relevant to the research question, which could not be assigned to the previous category system, to the code *'Overarching challenge'*. In total, 15 codes were defined and 62 codings were made. Since the coded text sections of a category were sufficiently comprehensible and the content captured with them could be directly transferred into a result text (see Appendix 4), no further fine coding was conducted.

5. Findings

This chapter presents the findings of the data analysis. They are presented factor by factor, based on the potential influence of the factors for the adoption of ML for rare disease diagnosis defined in chapter three. The findings of the case study show which of the potential influencing factors hinder and thus slow down the widespread adoption of the considered ML product.

5.1 Findings regarding the technological context

Characteristics

- *Relative Advantage*

According to the CEO of the technology provider, the relative advantages of the ML product are manifold. For example, GPs are educated about rare diseases, as in some cases they've never heard of the diseases that the ML product identifies. Moreover, the CEO added that the ML product can sometimes support GPs in solving long, complex patient cases. So far, 22% of the patients identified by the ML product have been referred to a next stage by GPs, and among them, only a very small proportion of patients have already been diagnosed. In addition, the GP mentioned that the ML product has already been used in a pilot project in another part of the country. The data from this pilot project revealed that about 50% of the patients which the ML product suggested for review were actually referred to the next step on the diagnostic pathway by the reviewing GPs. The CEO of the technology provider further argued that a small economic benefit can be made with early diagnosis, as patients undergo fewer unnecessary investigations. The managing director, however, added that therapies for rare diseases are also very expensive: "But of course, when these patients are put forward into a rare disease like specialty or anything, the dimension gets diverted and often they end up in a far better care position" (TP, 104). As the actual results of the pilot are not known yet, the interviewed GP was still cautious about naming relative advantages of the ML product:

> I think it's very early in its stages to say whether or not it has added public benefit, because the evaluation stage hasn't been completed, so we don't know whether we have found new diagnoses or whether those diagnoses have been reached more rapidly. (AO, 471)

However, according to the GP, if the desired added value of the ML product, namely speeding up the diagnosis of rare diseases and thus improving the treatment of patients with rare diseases, occurs, the ML product can help save money, albeit only a small amount, in the healthcare system by avoiding unnecessary investigations for patients with rare diseases. Another potential benefit of the ML product, as the GP explained, is that the ML product could improve the patchy care in disadvantaged areas: "It was very interesting, it found many more people in much more deprived areas, so whether it might be a way of trying to level up the care that's provided in more deprived areas, I don't know" (AO, 481). Moreover, even though this was not the aim of the pilot project, the GP mentioned that income generation does clearly not belong to the relative advantages of the ML product: "For us as an organisation, it has not produced any more income at all. (...) It was not done as an income generating process" (AO, 399). As rare diseases will always affect only a minority of the population, the additional benefit of the ML product for the diagnosis of rare diseases for the entire health system is also limited. The GP underlined this:

> If you look at this in terms of population health improvement, it's probably a small add on. If you addressed all the things that we already know about, levels of exercise, obesity, smoking, if you were able to reduce all those, you would probably get far greater health gains than using this sort of technology. (AO, 496)

In summary, the key stakeholders certainly see potential in the ML product for faster diagnosis of rare diseases. However, even with an effective and efficient ML technology for rare disease diagnosis, the added value for the entire healthcare system will be limited, as rare diseases affect only a minority of the population. This perceived small added value could be one reason for the slow adoption of ML technologies for rare disease diagnosis.

Compatibility

The technology provider's ML scan is used exclusively at primary care level, which is where the technology provider believes it is best suited:

> The GPs have a good longitudinal view of these patients. They've got a good and a broad view from the different areas of disease within this patient. Rare diseases are

> often multi-systemic, so if you find these patients in secondary care or tertiary care, they will be at a specialist level, but there often the information is silent. So, you might have information from ophthalmologists one day, and you might have to look quite hard to find out that that infectious patient is also being treated by renal physicians, but that information flows back into the GP primary care record. (TP, 148)

However, the technology provider also works closely with the specialised centres, as the pathways on which the rare disease patients are placed are dictated by the specialised centres. In order to make the ML technology as compatible as possible with the GPs' workflow, the CEO of the technology provider highlighted that the GPs will receive reports from the technology provider only once a month or perhaps only once a year, depending on the GPs' available time. According to the CEO of the technology provider, this high compatibility has paid off positively:

> I think what's allowed us in other pilots is that clinicians quite quickly realised that we're not flagging a huge number of patients. So, the clinicians themselves may review five of these reports a month. But in those five, there are one or two cases that they think, wow, this is great, I can take this forward. (TP, 114)

In summary, no problems or challenges were identified in the case study regarding the compatibility of the ML product. Thus, compatibility does not seem to be an obstacle to the adoption of the ML technology.

- *Complexity*

When asked if any special training is required for the GP wanting to use the ML product, the CEO of the technology provider answered now. No special type of training for reading the reports, using the results or using the product at all is necessary. According to him, the tool is designed in a way that any GP can use it. Also, the GP does not perceive the ML product as complex per se, but what makes the reading of the reports complex is poor quality data being fed into the ML product:

> Then it's just time for me to sit down and go through the patients' notes. It varied a lot as to how easy that was, because again it comes down to how well all the notes are being recorded, how well the coding is added and how well structured the data is. (AO, 444)

Thus, the stakeholders interviewed do not associate complexity with the ML product itself. Complexity does consequently not seem to hinder the adoption of the ML technology.

- *Observability*

In the interview with the GP, it showed that the GP did not have much knowledge about the ML product until she heard about the experiences that were made with the ML scan in a pilot in another part of the UK. The managing director of the technology provider knows that the company needs to prove the benefits of the ML product in order to gain trust of potential adopting organisations. However, as a prerequisite for this, the technology provider needs to have partnerships with healthcare organisations to do various pilots: "We need to be flagging thousands of patients in a week, a month or a day if we get lucky, but we need to have the partnerships set up to deliver all of that" (TP, 210). The technology provider has already tried to provide evidence of the benefits of the ML product through publications on pilot successes as well as external evaluations of some of the pilots conducted by the company, which were then submitted to the NHS. In addition, according to the managing director, the company is in the process of setting up a new study to determine patients' opinions about the ML product. The CEO is aware that gathering evidence of the added value of the ML product is resource-intensive, but of utmost importance, especially as Covid has pushed the issue of rare diseases even further back. According to the managing director, the technology provider will continue to invest in evidence gathering and look for bigger and better ways to find proof, such as with the use of the ML product to diagnose hard-to-diagnose diseases. However, according to him, it is not easy to create evidence when it comes to rare diseases:

> One of the difficulties in creating evidence in this space is that, because we're working with rare diseases, it takes so long to get these diagnoses done. Even if we flag a patient today, maybe in the old days it might have taken seven years to get that diagnosis, but it still takes maybe a year in some of these cases to get them referred all way through to the point of the diagnosis. (TP, 123)

Thus, findings of the case study demonstrate that it is difficult to provide evidence of the added value of ML technologies for rare disease diagnosis, such as for the ML product of the use case. This lack of evidence could be one reason why the adoption of such technologies in healthcare organisations is only increasing slowly.

- *Application*

According to the managing director of the technology provider, the ML scan can be used not only for the diagnosis of rare diseases, but also for the diagnosis of hard-to-diagnose diseases such as familial hypercholesterolaemia for example: "Familial hypercholesterolaemia (…) it's a hard-to-diagnose disease and our system is pretty good at taking these patients up" (TP, 328). However, according to the interviewed CEO of the technology provider, the potential lack of data interoperability is a challenge to the widespread adoption and use of the ML product. As the CEO explained, data interoperability is running in the UK. At primary care level, everything is aligned with SNOMED CT, which is currently the most comprehensive health terminology. In addition, the CEO elaborated that there is also interoperability between ICD-10 and ICD-11. ICD is the International Classification of Diseases (ICD), the international standard diagnostic classification which organises content into meaningful standardised criteria and enables the storage and retrieval of diagnostic information for epidemiological and research purposes (*International Classification of Diseases (ICD)*, 2022). From the CEO's point of view, however, data interoperability is a challenge if the technology provider wants to export its ML product:

> There will be a problem when you start exporting med technology to other countries. Certainly, our focus is currently in the UK. If we start exporting new med technology into other countries, we need to figure out how we adapt for different coding ontologies. (TP, 178)

In summary, the findings of the case study underline that the ML product can be used not only for the diagnosis of rare diseases, but also for diseases that are difficult to diagnose. However, the lack of data interoperability seems to be an obstacle to the widespread application of the ML product abroad.

5.2 Findings regarding the organisational context

Data sharing

To gain access to as many EHRs as possible, the managing director of the technology provider explained that the company always works with several GP practices and other administrative structures that adopt the ML technology as a group:

> So it's not done on a case-by-case basis by the clinicians saying, "Can you help us with a particular case?". So, we have contracts either with the GP practices themselves or with bigger umbrellas, potentially super group practices. (…) they adopt

> technology as a group or even above that within the NHS, there are bigger structures of management just above GP practices and beyond. These bodies will decide that they would like to run a quality improvement process of trying to identify patients that may have rare diseases or in fact hard-to-diagnose diseases. And they will ask us to run the scan across the entire electronic record data set. (TP, 61)

However, as the GP claimed in the interview, many GP practices in the area did not agree to adopt the ML product, because they did not want to share their data with an external company:

> One of the big barriers to using this product was the agreement of the other GPs in the area, because many people were very suspicious of the new technology and very reluctant to allow other organisations access to their data. A lot of concerns about how the data was going to be used, whether it was going to remain confidential to that patient, whether the data was safe. So, I think that was a big barrier. (AO, 437)

Thus, as these findings of the case study demonstrate, organisational unwillingness to share data hinders the adoption of the ML product.

Support

According to the CEO, it is usually the proactive and innovative GP practices, which are those that perform well on NHS metrics, that are most likely to support and thus adopt the ML scan. In addition, the managing director noted that GPs interested in genomic medicine and rare diseases are increasingly and better adopting the ML product. Hence, it is no coincidence that the GP interviewed has a particular interest in genomic medicine:

> Over the last two years, I wanted to take a masters in genomic medicine, because I have a particular interest in this area, and they asked me if I would be prepared to review some patients. So, I was happy to help them with that work. (AO, 344)

Indeed, the GP's interest in rare diseases seems to be one of the reasons why her organisation adopted the ML product, as she does not see improving rare disease diagnosis as an urgent priority in healthcare: "I think I'm afraid, I would have to say that it would not be the most pressing priority at this point in time, because I feel there's a lot of unmet care needs already" (AO, 495). Moreover, the GP emphasized repeatedly that her expectations of the ML product were very low at the beginning and that she is still sceptical about its benefits now. Therefore, this underlines that evidence for the ML product is of significance to gain people's trust. Moreover, the GP interviewed is not the only one who does not fully trust the ML product. She

emphasized that she has learnt much about how to persuade people to engage with the technology. In addition, the fact that half of the GP practices in her area have chosen not to participate in the pilot and thus not to adopt the ML product also demonstrates the mistrust that other GPs had or still have towards this technology. Indeed, she underlined this by stating: "Within our area, only half of practices chose to take part in it. There was a great deal of suspicion" (AO, 489). Further, GPs have reservations about the ML product, because they fear that patients may feel that their privacy is being violated by a third party, in this case the technology provider. However, according to the CEO of the technology provider, these fears are unfounded:

> One, we don't have access to the identifiable data. Second thing is, we find that the patient population is incredibly recipient, because these patients have issues that they are trying to find diagnosis for and things like that. (...) So, there's a mismatching what the medical fraternity expects or is worried about and what the patients actually feel. So, hopefully that can be cleared up in the long run. (TP, 317)

Thus, these findings of the case study show that various reasons can lead to a lack of an organisation's support for the ML technology and thus hinder its adoption. These reasons include lack of interest in rare diseases, lack of trust in the technology, and reservations about patient responses.

Size

The GP highlighted that GP practices in the North East of England are all rather small. It was not found that larger GP practices were more likely to participate in the pilot than smaller ones:

> At an individual practice level, we still have quite small practices who are led by the GPs concerned, as GP partners. So, it would be very much up to each set of GPs within the practice to make those decisions. (AO, 487)

Thus, the findings of this case study suggest that organisational size was not a relevant factor in the adoption of the ML product in the pilot in question. Other factors seem to influence GPs' opinion on whether to adopt the ML product or not.

Resources

The technology was purchased by the GMSA as part of the pilot project. It was thus funded by the UK government. However, the GP does not know how much money was spent by the UK government on the pilot. The adoption did not generate any income for the organisation, but as

the pilot was largely funded by the UK government, the adopting GP practice itself had to spend only minimal resources, such as for putting messages in connection with the pilot on their website. Moreover, the GP added that the GP practice did not have to take care of the necessary IT infrastructure either:

> Logistically then, the computer system of the practice concerned, you know the data holders, was something, I think, that [name of the technology provider] had to take into account. That wasn't really my problem, I didn't get involved with that part of it. (AO, 441)

In terms of specialized and skilled HR, the CEO of the technology provider confirmed that any GP with a good clinical background is able to use the ML product. However, as mentioned previously when discussing the factor *support*, according to the managing director, GPs with an interest in genomic medicine and rare diseases have adopted the ML product more. In summary, the lack of financial or technical resources within an organisation were apparently not the reasons for half of the GP practices not participating in the pilot and not adopting the ML product. Findings of the case study show that the adopting organisation did not have to make a large financial investment and did not have to take care of the technical infrastructure. However, according to the findings, it could be that a possible lack of HR with low interest and knowledge about genomic medicine and rare diseases impeded the widespread adoption of the ML product in the pilot.

5.3 Findings regarding the environmental context

Public pressure

According to the managing director of the technology provider, there are many new developments regarding policy support for rare diseases in the UK:

> Within the UK, there's a strengthening political agenda on rare diseases (…) with some of the rare disease frameworks and rare disease action plans and all of these things (…). So, in the last six years, I've seen quite a good evolution of these things and what we're doing at [name of the technology provider] falls directly in line with some of the priorities that they outline specifically on the early diagnostic side. (TP, 265)

In addition, the CEO of the technology provider does not only perceive political support for rare diseases in the UK: "I think there is additionally quite a lot of political support outside of

direct NHS and medical device regulation" (TP, 295). From the managing director's point of view, the political pressure generated is very helpful:

> One, it does ultimately, potentially lead to funding from within the NHS, because they set up initiatives that will lean on private providers to assist the NHS and delivering those goals, but secondly, it creates a bit of vacuum and traction within the NHS, where that filters all the way down to the bottom levels when people start realising that rare diseases are a focus within the NHS at the top level. (TP, 270)

The GP of the adopting organisation also sees an important role for public policy in the process of adopting the ML product. Hence, she believes that political pressure will increase in the future: "I think we are in a very early stage of that process" (AO, 480). As already mentioned when discussing the results for the factor *support*, the managing director of the technology providing company feels great support from patients: "The patient population really loves the idea of what we do. (…) So, they love the aim of this project" (TP, 314). However, the CEO of the technology provider claimed that there is a lack of incentivisation structures from the NHS for rare disease diagnoses:

> The NHS itself is really, really poorly set up to facilitate any extra time spent on rare diseases. So the NHS at ground level is just completely response to the incentivisation structures that the NHS pairing structures create. And there are no incentivisation structures for rare disease diagnoses. (TP, 323)

Thus, to further encourage adoption, the CEO of the technology providing company is calling for incentivisation structures from the NHS for rare disease diagnoses. Moreover, according to the interviewed CEO, Covid has not really helped to raise public awareness about rare diseases. Consequently, the manging director emphasised that the technology provider will continue to gain more political support for its ML product: "I think looking at our expansion fronts that we focus on right now (…) one is continue the expansion of political front within the NHS on our side" (TP, 274). Thus, findings of the case study show that the ML product is supported by the public, but there is room for improvement, for instance in creating incentivisation structures for rare disease diagnoses by the NHS.

Technology support infrastructure

According to the CEO interviewed, the technology company takes care of the technology support infrastructure for the ML product: "We look after all of that internally. (...) we supply full up to date software with all of the overhead set that go with that directly into the NHS at a very,

very reduced rate for them at the moment" (TP, 219). Indeed, the GP confirmed that they did not have to take care of the technology infrastructure. Since adopting organisations are not responsible to take care of the technical infrastructure themselves, it can be assumed that this factor does not hinder the adoption of the ML product and was therefore not decisive for the fact that half of the GP practices did not participate in the pilot project.

Government regulation

Since Brexit, which happened on January 1, 2021, the Medicines and Healthcare products Regulatory Agency (MHRA) became the UK's independent regulator for medicines and medical devices. Medical devices are nowadays subject to registration with the MHRA in the UK (*Regulating Medical Devices in the UK*, 2022). Since the UK's withdrawal from the EU, Northern Ireland is following the EU timetable for the implementation of the EU MDR and In Vitro Diagnostic Medical Device Regulation (IVDR) (*Regulating Medical Devices in the UK*, 2022). Great Britain, however, continues to apply the UK MDR of 2002 until the end of June 2023. The general approach of the future regulation from July 1, 2023, aims to be aligned more closely with the EU MDR and IVDR, but simultaneously deviate in a way that makes sense for the UK government, industry and, most importantly, patient safety (*Chapter 1 – Scope of the Regulations,* 2022). The analysis of the interview with the CEO of the technology providing company suggests that the technology provider welcomes the transition from the MDD to the EU MDR, although this is more of an issue in the EU than in the UK post-Brexit. He stated:

> When we need to upgrade into MDR, the new device regulation, which is already active, I see a much clearer pathway for us then within the MDD. It's going to take some work for us to get into the position to take advantage of that, but in fact I think, it will allow us to be much more innovative and to roll out much more progress of AI-driven or ML-driven algorithms to identify patients. We've got a lot of this work on our background, and we know that it performs far better than some of the algorithms that we have in play to date already. (TP, 289)

Hence, from a technical point of view, the EU MDR will contain stricter rules on software which is used as a medical device, such as the ML product. On the other hand, however, the MDR is an improvement to the MDD, as the MDR emphasizes patient safety and transparency more. Further, the CEO added that the company benefits not only from the NHS's legal approaches to AI algorithms, but also from the US Food and Drug Administration (FDA):

> So, the NHS has released the latest update of the way, they evaluate digital technologies and by falling in line with these types of things, we really allow ourselves to take advantage of this. I think even looking at the FDA's approach to AI algorithms and how they have got to a more mature state what you can easily understand from the documentation that's out there and from other examples that are happening on how the FDA and the medical device service things look at regulating these AI-driven algorithms. (TP, 296)

Finally, the CEO of the technology provider emphasised that the fact that the company he works for is developing a case finding clinical decision support tool gives the company a lot of legal leeway:

> We are not a diagnostician. (…) we are not a screening tool. (...) So, we're technically a case finding clinical decision support tool and that allows us to do quite a lot because it shifts the responsibility of the decision making on the clinician as the final decision point. (TP, 302)

Overall, in the case study, no obstacles related to government regulations were identified. Thus, this factor does not seem to be an obstacle to the adoption of the ML product.

Optimal data

The managing director argued that access to big amounts of data is essential for the use of the ML product:

> Within the rare disease space, there's a saying that I like to use, 'to find rare disease patients is like finding a needle in the haystack, but in order to find it, you need the haystack'. So, you need a massive, massive data set. (TP, 181)

However, to obtain as much data as possible, healthcare organisations must be willing to share their data. As already mentioned when discussing the factor *data sharing*, this willingness is not always given. Since Brexit, the UK no longer follows the EU GDPR, but the UK GDPR, which is almost identical to the EU GDPR (*Using Personal Data in Your Business or Other Organisation, 2021*). This means that GP practices in the UK can, in principle, share their patient data anonymously with technology providers if they want. The managing director explained that the technology provider is working with data from primary and secondary care in the UK to get as much data as possible: "We started to do a bit of shared care records, so primary and secondary care, which allows us a much better search in a facility, because we got so much

more data to work with" (TP, 166). In addition, as already mentioned when discussing the factor *application*, in order to be able to work with data from different GP practices and different levels of care, the interoperability of different data formats and sources must be given in each case. Indeed, the CEO of the technology provider explained:

> So, interoperability becomes really important, because even at a single GP practice where you've got five, ten thousand patients, you're not going to find huge amounts of rare disease patients in there. I mean, you will find a lot of patients with rare diseases, but the diseases that we particularly are looking for, you may flag ten cases in five thousand at the moment. (TP, 183)

Further, as already mentioned when discussing the factor *complexity*, the GP highlighted that the review of the referral letters from the technology provider was sometimes difficult due to insufficiently recorded and structured data. The CEO of the technology company also pointed out that the data is not always of good quality, which can lead to the ML product incorrectly flagging patients: "In some cases, when the GPs review that flag, they will know that the patient is already diagnosed, but it's not well coded in the electronic record. Of course, we have that problem of data. It doesn't always match reality" (TP, 79). Overall, it can be argued that obtaining large datasets of good quality required for the ML product is not easy and this seems to be a major challenge for the widespread adoption of the technology.

Funding

In the interview, the CEO of the technology company explained how the ML product is funded:

> So, for the development of the tool as a start-up, we've been largely funded by Angel investors to date. We've got one or two small venture capital firms that have invested in some of our early rounds. Beyond that, we have actually won quite a lot of grant support from the UK government, the innovate UK side of things. So, between those two funding sources, we're formally in an excess of pounds. We have a number of grant applications going forward for some bigger AI-orientated evidence generation funding that I've been talking about. But beyond that, we are generating revenues. We generate revenues from two main sources. The first source is directly from the NHS state payers. Certain bodies pay us to help them find patients and then we generate revenue from pharma. Our pharma revenues come from working with pharma and building algorithms within and beyond the pharma industry. (TP, 192)

As a start-up, according to the managing director, the goal is clearly to continue to generate revenue from pharma and the NHS. When the benefit of the ML product becomes obvious, the GP also sees a taxpayer funding responsibility for it: "I think ultimately that if it is found to be of benefit and certainly in the UK, obviously we have a centrally funded healthcare system, then yes, it would be the taxpayers who would be responsible" (AO, 473). However, according to the CEO, the revenue the company generates from the NHS does not even come close to covering the cost of maintaining the service. In contrast, the technology provider's partnerships with pharmaceutical companies contribute significantly to funding the company's work. Indeed, the managing director of the technology provider has found that the real revenue generator in the rare disease setting, is the pharma industry:

> So, I've looked at various forms of funding in the past and I've realised that some of the rare disease advocacy groups themselves are not well funded, but where the real money-spinner is all within the setting, is certainly within pharma. So, in my opinion, we could set up a few consortiums where funding is supplied by some of the big pharma companies that are doing rare disease things. Ultimately, for them there's a direct selfish thing, that improved diagnostic rare diseases in the future can lead to increased revenues for them. (…) they could shift money into consortiums where many of them get to as a group and then those consortiums could be run as similar to most non-profit organisations, and they could decide where they fund into different projects. It does happen to some degree, but it's been incredibly difficult for us as a small start up to get into the radar of some of these things. (TP, 246)

Moreover, in addition to consortium-based and private healthcare efforts, big government funding, such as UK or European level funding, would also be beneficial according to the managing director. He added that it is not easy to justify large sums of money for funding in the field of rare diseases:

> It's incredibly difficult to find justification for large amounts of funding in this area for lots of reasons. One being that the outcomes and the impact is almost hidden. You know the cost of rare diseases is incredibly difficult thinking quantified on society and on health systems, although we've done some work on this topic ourselves to try and show some of the numbers. But secondly, this is quite a long-term claim and there's not a lot of easy funding for things that take a long time to settle in and become part of mainstream medicine. (TP, 231)

In sum, there is private and public funding for the development and provision of the ML product. However, findings of the case study show, it is challenging to justify and obtain large funding sums in the field of rare diseases, even more so as a start-up. More private and public funding could help to further promote the widespread adoption of the ML product.

5.4 Other findings

Overarching challenge

When asked about the most relevant lessons he has learned from his work on the ML product, the CEO of the technology provider answered as follows:

> So, learnings to date are, that it's really quite difficult to push against the prevailing winds of a big healthcare system and it's really difficult to bring something that's quite innovative into play. I mean, we've been going for a long time and the last thing is that working in a rare disease space is incredibly tricky. It means that you need to spend much more time, spend much more resources and it's just far more difficult to actually get anywhere with it. So, it's been quite a hard learning. (TP, 331)

This statement underlines that it is generally difficult to bring innovative solutions to the market in the health sector. Further, the development of innovative technologies in the area of rare diseases is particularly challenging.

6. Discussion

The findings of the case study demonstrate which factors hinder the adoption of the ML product. In this chapter, the findings of the case study will be related to the findings of the existing literature on AI adoption in healthcare, outlined in chapter three.

Among the factors, which were not identified as obstacles in the case study, is the factor *compatibility*. The managing director of the technology provider claimed that the ML product is positively received by GPs due to the reasonable amount of work it generates. In one of the earlier studies, it was argued that AI technologies are better suited for secondary care than for primary care, partly because secondary care is digitally more advanced and focuses more on medical specialties. However, the ML product in the case study is used only at primary care level and the technology provider has had positive experiences with it. Thus, it can be concluded that it depends on the use of the respective AI technology in each case, whether it fits better

into primary or secondary care level. In addition, the factor *complexity* of the technology could not be identified as a barrier to the adoption in this case study. This finding does not correspond to most previous studies, which have considered the complexity of AI technologies as a barrier to the adoption in healthcare. However, the influence on the adoption of this factor also seems to depend strongly on the AI algorithms in question. As already mentioned at the beginning of this paper, neural networks are more difficult to understand than other methods, such as decision trees for instance. Another factor that could not be identified as an obstacle in the case study is the factor of the organisational *size*. Previous studies have argued that larger healthcare organisations are more likely to adopt AI technologies, because they have more financial resources and data. However, as the ML product in the use case was largely funded by the pilot in question, the GP practice did not have to invest much of its own financial resources in adopting the ML product. In addition, to identify as many patients with rare diseases as possible, a lot of data is needed. Even in a large healthcare organisation, many patients with rare diseases are unlikely to be identified. Moreover, some previous studies also highlighted that a lack of organisational resources in terms of financial resources and IT infrastructure, represents a major barrier to AI adoption in healthcare. However, the factor organisational *resources* regarding financial resources and IT infrastructure was not identified as an obstacle for the adoption of the ML product. The adopting GP practice did not have to invest much of its own financial resources and did not have to take care of the necessary IT infrastructure. As the technology provider is responsible for the follow up and updating of the software, the factor *technology support infrastructure* could not be identified as an obstacle to the adoption in the case study either. Indeed, one of the previous studies has claimed that technical support from the technology provider is of significance for the adoption of AI technologies in healthcare organisations. Further, in contrast to some previous studies, the case study did not identify any problems related to the factor *government regulation* in the adoption of the ML product. Presumably, however, this factor depends heavily on the regulations in place and the AI technology in question. This means that a decision support tool for case finding, as in the use case, is subject to less strict regulations than, for example, a screening tool. When analysing the factor *public pressure* in terms of patient support, the case study has shown, in line with previous studies, that patient support for the technology should be given in order to achieve adoption. The technology provider argued that patients strongly support the ML product, although the medical community does not seem to perceive this. To really find out to what extent the patient population supports the ML product, one would have to ask the patients themselves. However, it can be assumed that patients with rare diseases who have been waiting for a diagnosis for years see great

potential in technologies such as the ML product. Therefore, patient support in the rare disease setting is unlikely to be a barrier to adoption.

Factors which were identified as barriers to the adoption of the ML product, both in previous research and in the case study, include all factors related to data. Lack of willingness for *data sharing* poses a major problem to the adoption of AI technologies in general. Moreover, lack of data interoperability limits the *application* of AI technologies and thus their widespread adoption. Further, less data and data of poor quality, also called lack of *optimal data*, leads to ineffective and inefficient AI technologies. AI technologies with higher error rates in turn bring fewer *relative advantages*. Consequently, all these hurdles are responsible for the slow increase in the adoption of AI technologies in healthcare organisations. Moreover, also in line with previous studies, the case study identified the lack of organisational *support* for the ML product as a barrier to its adoption. However, the case study has shown that it is particularly difficult to gain support for technologies for rare diseases, as the added value, thus the *relative advantage* of such technologies for the health system is very low compared to technologies for common diseases. The same applies to the factor *observability*. In line with previous studies, the case study confirmed that the demonstration of the added value of a technology is significant to promote its adoption. However, as highlighted in the interview analysis, providing evidence for technologies for rare disease diagnosis is a particularly big hurdle, as diagnosing a rare disease, even with the help of the ML product, can take up to a year. Regarding organisational *resources* in terms of skilled and specialized HR, previous studies argued that an insufficient number of personnel with expertise in the field of medicine and data science hinders the adoption of AI technologies in healthcare. However, according to the case study findings, it is not that an insufficient number of staff with expertise in medicine and data science is hindering the adoption of the ML product, but rather the low level of interest in genomic medicine and rare diseases among GPs. In addition, some of the existing literature claimed that lack of *funding* is a barrier to widespread adoption of AI technologies in healthcare. The case study has shown that while there is private and public funding for the ML product, it is indeed very difficult to justify large amounts of money for rare disease technologies, even more so as a start-up. Thus, more consortium-based, private and public funding would be beneficial to further promote the adoption of the ML product. Regarding the factor *public pressure* in terms of policy support, the interview analysis revealed, corresponding to the findings of previous studies, that policy support for the technology is essential to achieve widespread adoption. Even though the case study suggested that public policy tries to support patients with rare diseases through action plans and so on, according to the managing director, there is still room for improvement, for instance by

creating incentivisation structures in the NHS for rare diseases or by providing more public funding. As rare diseases only affect a minority of the population, it is all the more important to raise public awareness and encourage society to benefit from possible new innovations for the diagnosis of rare diseases. Finally, the one statement of the managing director summarizes the problem of widespread adoption of ML for rare disease diagnosis well: Bringing AI technologies to market is challenging in general, and when it comes to AI technologies for the rare disease field, it is even more challenging.

7. Conclusion

After discussing the findings of the case study, this chapter presents the principal findings of the thesis and then highlights the limitations of the paper and opportunities for future research.

7.1 Principal findings

As outlined in the beginning of this paper, ML technologies have great potential to significantly improve the diagnosis of patients with rare diseases. Nevertheless, the adoption of ML technologies in the practice of medical diagnostics is only slowly increasing. Thus, this paper tried to answer the following research question: *Why is the adoption of ML-based diagnostics for rare diseases increasing slowly in healthcare organisations?* As a basis for answering the research question, this paper drew on the TOE framework. The TOE framework defines factors in the technological, organisational and environmental context which influence innovation adoption in organisations. The framework was adapted with the findings from existing literature on AI adoption in healthcare according to the context of this paper. The following potential influencing factors for the adoption of ML-based diagnostics for rare diseases were thus identified: *Characteristics of the technology (relative advantage, compatibility, complexity, observability, application)* in a technological context, *data sharing, support, size* and *resources* in an organisational context and *public pressure, technology support infrastructure, government regulation, optimal data* and *funding* in an environmental context. In order to examine whether and to what extent these factors actually have an influence on ML adoption for rare disease diagnosis, a case study, embedded in the Screen4Care research project, was conducted. The use case dealt with the adoption of an ML product, hence, a case finding clinical decision support tool for rare diseases, developed by a start-up company in London and used in different pilot projects in primary care in the UK. Two expert interviews have been conducted: one with the CEO of the technology provider of the ML product and one with a GP of a GP practice that has adopted the

technology as part of a pilot. For the analysis of these interviews, a theoretically derived deductive and inductive qualitative content analysis was carried out. The interview analysis demonstrated that factors related to data such as lack of willingness for *data sharing*, lack of *optimal data* as well as lack of data interoperability which limits the *application* of AI technologies, were clearly perceived as barriers to the adoption of the ML product. These findings correspond to findings of previous studies on AI adoption in healthcare. Thus, data issues are reasons for the slow growth of AI adoption in healthcare organisations in general, not only for ML-technologies for rare disease diagnosis. Further, corresponding to previous studies, the case study has demonstrated that an *innovation-supportive* culture in adopting organisations, high levels of *observability* of the results and experiences with the technology, sufficient *funding* and *public pressure* in the form of public policy awareness are key to achieving widespread adoption of the ML product. However, as outlined in the case study, it is particularly difficult to gain support for a technology that is used for rare disease diagnosis such as the ML product, as the relative advantage or added value of such a technology to the overall health system is relatively low compared to technologies which are used for common diseases. Evidence gathering is also challenging for rare disease technologies, as diagnosing rare diseases can take a considerable amount of time, even with ML technologies, such as the ML product of the use case. Unsurprisingly, justifying large amounts of funding for technologies used to diagnose rare diseases, hence for a minority of the population, is also difficult. This makes the role of public policy for rare disease technologies all the more important. Finally, the case study of this article has shown that the adoption of AI technologies is a challenge in itself and an additional hurdle in the rare disease setting. It is therefore not surprising that early diagnosis and thus the provision of effective treatments for people with rare diseases is considered one of the greatest global health challenges of our time.

7.2 Limitations and future research opportunities

Some studies investigating AI adoption in healthcare already exist. However, no study has investigated the adoption process of ML technologies for the diagnosis of rare diseases. To overcome this shortcoming, this paper validated the literature review findings with a case study approach to gain a deeper understanding about the challenges of ML adoption in the rare disease setting. A case study allows for an in-depth analysis of the case, but this also means that these findings cannot be generalised with a similar certainty as quantitative analyses can. The case study looked at one particular ML product, which was a case finding clinical decision support tool. Findings might differ for another technology. In addition, the case study examined a use

case in the UK. The UK has a public health system and follows the UK MDR, apart from Northern Ireland, which follows the EU MDR. However, there are different health systems in different countries as well different regulations in place and so on. Therefore, findings may vary depending on the country. Moreover, for the case study of this paper, two expert interviews were conducted. One must be aware that answers in interviews are always subject-driven. In addition, more interviews, especially with actors on the adopting side of the ML product, such as adopting GP practices, would have led to more generalisable findings. As far as this work is concerned, further research on the adoption of different ML technologies for rare disease diagnosis in different countries should provide more information on the general challenges of adoption. It would also be relevant to conduct interviews with non-adopting organisations to find out their reasons for not adopting a particular technology. Moreover, as the development of AI technologies is very fast-moving, it consequently remains ambiguous how certain factors will impact the adoption in the future. For example, initial research has already been conducted to enable the conversion of different types of medical data into a uniform format. Hence, it is possible that data interoperability may soon no longer be an obstacle to widespread adoption. Thus, future research is of utmost importance. Moreover, this work has focused only on the obstacles which explain the slow adoption growth of ML-based diagnostics for rare diseases in healthcare organisations. It would also be worthwhile for future research to look closely at how these obstacles could now be overcome and what measures must be taken to do so. Finally, although the findings of this case study are not generalisable, they provide a first insight into the application of ML in the field of rare diseases and in this way attempt to draw the attention of policy makers to the urgent need to develop appropriate policies to facilitate market access for technology providers and strengthen the medical market through AI-assisted timely diagnosis for rare diseases.

References

Ahmadi, H., Nilashi, M., Ibrahim, O., Ramayah, T., Wong, M. W., Alizadeh, M., Jafarkarimi, H., & Almaee, A. (2015). Exploring potential factors in total hospital information system adoption. *Journal of Soft Computing and Decision Support Systems*, *2*(1), 52-59.

Alanazi, A. (2022). Using machine learning for healthcare challenges and opportunities. *Informatics in Medicine Unlocked*, 100924.

Al Badi, F. K., Alhosani, K. A., Jabeen, F., Stachowicz-Stanusch, A., Shehzad, N., & Amann, W. (2022). Challenges of AI Adoption in the UAE Healthcare. *Vision*, *26*(2), 193-207.

Areas of Focus | Screen4Care. (n.d.). Retrieved June 22, 2022, from https://screen4care.eu/en/the-project/area-of-focus

Baker, J. (2012). The technology–organisation–environment framework. *Information systems theory*, 231-245.

Benbya, H., Davenport, T. H., & Pachidi, S. (2020). Artificial intelligence in organisations: current state and future opportunities. *MIS Quarterly Executive*, *19*(4).

Bini, S. A. (2018). Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care?. *The Journal of arthroplasty*, *33*(8), 2358-2361.

Brasil, S., Pascoal, C., Francisco, R., dos Reis Ferreira, V., A Videira, P., & Valadão, G. (2019). Artificial intelligence (AI) in rare diseases: is the future brighter?. *Genes*, *10*(12), 978.

Chapter 1 - Scope of the Regulations. (2022, June 26). GOV.UK. Retrieved August 22, 2022, from https://www.gov.uk/government/consultations/consultation-on-the-future-regulation-of-medical-devices-in-the-united-kingdom/outcome/chapter-1-scope-of-the-regulations

Decherchi, S., Pedrini, E., Mordenti, M., Cavalli, A., & Sangiorgi, L. (2021). Opportunities and challenges for machine learning in rare diseases. *Frontiers in Medicine*, *8*.

Dhruv, S. (2021, March 15). *Why Machine Learning is Important in Education Sector.* Aalpha. Retrieved July 6, 2022, from https://www.aalpha.net/blog/why-machine-learning-is-important-in-education-sector/

Döbel, I., Leis, M., Vogelsang, M. M., Neustroev, D., Petzka, H., Rüping, S., Voss, A., Wegele, M., & Welz, J. (2018). Maschinelles Lernen–Kompetenzen, Anwendungen und Forschungsbedarf. *Fraunhofer IAIS, Fraunhofer IMW, Fraunhofer Zentrale. Zugriff am*, *21*, 2020.

Enzmann, D. R., & Schomer, D. F. (2013). Analysis of radiology business models. *Journal of the American College of Radiology*, *10*(3), 175-180.

Evans, W. R., & Rafi, I. (2016). Rare diseases in general practice: recognising the zebras among the horses. *British Journal of General Practice*, *66*(652), 550-551.

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*, *294*(1), 567-592.

Galletta, A. (2013). *Mastering the semi-structured interview and beyond: From research design to analysis and publication* (Vol. 18). NYU press.

Groft, S. C., Posada, M., & Taruscio, D. (2021). Progress, challenges and global approaches to rare diseases. *Acta Paediatrica*, *110*(10), 2711-2716.

Hemmer, P., Schemmer, M., Riefle, L., Rosellen, N., Vössing, M., & Kühl, N. (2022). Factors that influence the adoption of human-AI collaboration in clinical decision-making. *arXiv preprint arXiv:2204.09082*.

Hercheui, M., & Mech, G. (2021). Factors Affecting The Adoption Of Artificial Intelligence In Healthcare. *Global Journal of Business Research*, *15*(1), 77-88.

Hofmann, P., Oesterle, S., Rust, P., & Urbach, N. (2019). Machine learning approaches along the radiology value chain–Rethinking value propositions.

Holzinger, A., Kieseberg, P., Weippl, E., & Tjoa, A. M. (2018, August). Current advances, trends and challenges of machine learning and knowledge extraction: from machine learning to explainable AI. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction* (pp. 1-8). Springer, Cham.

IMI Innovative Medicines Initiative | Screen4Care. (2021, October 1). IMI Innovative Medicines Initiative. Retrieved June 22, 2022, from https://www.imi.europa.eu/projects-results/project-factsheets/screen4care

International Classification of Diseases (ICD). (2022, August 31). Retrieved July 22, 2022, from https://www.who.int/standards/classifications/classification-of-diseases

Jöhnk, J., Weißert, M., & Wyrtki, K. (2021). Ready or not, AI comes—an interview study of organisational AI readiness factors. *Business & Information Systems Engineering*, *63*(1), 5-20.

Kuckartz, U., & Rädiker, S. (2020). Fokussierte Interviewanalyse mit MAXQDA. *Schritt Schritt*, 55-74.

Lebcir, R., Hill, T., Atun, R., & Cubric, M. (2021). Stakeholders' views on the organisational factors affecting application of artificial intelligence in healthcare: a scoping review protocol. *BMJ open*, *11*(3), e044074.

L'heureux, A., Grolinger, K., Elyamany, H. F., & Capretz, M. A. (2017). Machine learning with big data: Challenges and approaches. *Ieee Access*, *5*, 7776-7797.

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

Mayring, P. (2014). Qualitative content analysis: theoretical foundation, basic procedures and software solution.

Melnikova, I. (2012). Rare diseases and orphan drugs. *Nature Reviews Drug Discovery*, *11*(4), 267.

Morrison, K. (2021). Artificial intelligence and the NHS: a qualitative exploration of the factors influencing adoption. *Future Healthcare Journal*, *8*(3), e648.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. b*, *4*, 51-62.

NHS website. (2021, November 18). *Genetic and genomic testing.* nhs.uk. Retrieved July 22, 2022, from https://www.nhs.uk/conditions/genetic-and-genomic-testing/

Pumplun, L., Fecho, M., Wahl, N., Peters, F., & Buxmann, P. (2021a). Adoption of Machine Learning Systems for Medical Diagnostics in Clinics: Qualitative Interview Study. *Journal of Medical Internet Research*, *23*(10), e29301.

Pumplun, L., Fecho, M., Wahl-Islam, N., & Buxmann, P. (2021b). Machine learning systems in clinics–how mature is the adoption process in medical diagnostics?.

Regulating medical devices in the UK. (2022, January 1). GOV.UK. Retrieved July 22, 2022, from https://www.gov.uk/guidance/regulating-medical-devices-in-the-uk

Rogers, E.M. (1995). Diffusion of Innovations. 4th Edition, the Free Press, New York.

Ronicke, S., Hirsch, M. C., Türk, E., Larionov, K., Tientcheu, D., & Wagner, A. D. (2019). Can a decision support system accelerate rare disease diagnosis? Evaluating the potential impact of Ada DX in a retrospective study. *Orphanet journal of rare diseases*, *14*(1), 1-12.

Schaefer, J., Lehne, M., Schepers, J., Prasser, F., & Thun, S. (2020). The use of machine learning in rare diseases: a scoping review. *Orphanet journal of rare diseases*, *15*(1), 1-10.

Schieppati, A., Henter, J. I., Daina, E., & Aperia, A. (2008). Why rare diseases are an important medical and social issue. *The Lancet*, *371*(9629), 2039-2041.

Sernadela, P., González-Castro, L., Carta, C., Van Der Horst, E., Lopes, P., Kaliyaperumal, R., Thompson, M., Thompson, R., Queralt-Rosinach, N., Lopez, E., Wood, L., Robertson, A., Lamanna, C., Gilling, M., Orth, M., Merino-Martinez, R., Posada, M., Taruscio, D., Lochmüller, H., Robinson, P., Roos, M., & Oliveira, J. L. (2017). Linked registries: connecting rare diseases patient registries through a semantic web layer. *BioMed Research International*, *2017*.

Sitem Insel. (n.d.-b). Retrieved August 2, 2022, from https://sitem-insel.ch/en/info/news/new-eu-research-project-screen4care-accelerating-diagnosis-for-rare-disease-patients-through-genetic-newborn-screening-and-artificial-intelligence

Stuart, J. (2010). Artificial Intelligence A Modern Approach Third Edition.

Sukardi, ., Hasyim, ., & Supriyantoro, . (2021). Using the Technology-organisation-environment Framework Approach in the Acceptance of Telemedicine in the Health Care Industry. *European Journal of Business and Management Research*, *6*(3), 47–54. https://doi.org/10.24018/ejbmr.2021.6.3.837

Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, *36*(2), 368-383.

Supervised Machine learning - Javatpoint. (n.d.). www.javatpoint.com. Retrieved August 22, 2022, from https://www.javatpoint.com/supervised-machine-learning

Tornatzky, L. G., & Fleischer, M.. (1990). *The process of technological innovation*. Lexington books.

Using personal data in your business or other organisation. (2021, June 28). GOV.UK. Retrieved August 14, 2022, from https://www.gov.uk/guidance/using-personal-data-in-your-business-or-other-organisation

Verma, V. K., & Verma, S. (2021). Machine learning applications in healthcare sector: An overview. *Materials Today: Proceedings*.

Watson, J., Hutyra, C. A., Clancy, S. M., Chandiramani, A., Bedoya, A., Ilangovan, K., Nderitu, M., & Poon, E. G. (2020). Overcoming barriers to the adoption and implementation of predictive modeling and machine learning in clinical care: what can we learn from US academic medical centers?. *JAMIA open*, *3*(2), 167-172.

Weinert, L., Müller, J., Svensson, L., & Heinze, O. (2022). Perspective of Information Technology Decision Makers on Factors Influencing Adoption and Implementation of Artificial Intelligence Technologies in 40 German Hospitals: Descriptive Analysis. *JMIR medical informatics*, *10*(6), e34678.

What is Screen4Care? | Screen4Care. (n.d.). Retrieved August 22, 2022, from https://screen4care.eu/en/the-project/screen4care

Wiens, J., & Shenoy, E. S. (2018). Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clinical Infectious Diseases*, *66*(1), 149-153.

Wolverton, C. C., & Lanier, P. A. (2019). Utilizing the Technology-organisation-environment framework to examine the adoption decision in a healthcare context. In *Handbook of Research on the Evolution of IT and the Rise of E-Society* (pp. 401-423). IGI Global.

Yin, R. K. (2018). Case Study Research and Applications: Design and Methods (6th ed.). Thousand Oaks, CA: Sage.

Zhang, H., Xiao, Y., Zhao, X., Tian, Z., Zhang, S. Y., & Dong, D. (2022). Physicians' knowledge on specific rare diseases and its associated factors: a national cross-sectional study from China. *Orphanet journal of rare diseases*, *17*(1), 1-13.

Declaration of independence

Selbstständigkeitserklärung für die Masterarbeit

Ich erkläre hiermit, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche kenntlich gemacht. Mir ist bekannt, dass andernfalls der Senat gemäss dem Gesetz über die Universität zum Entzug des auf Grund dieser Arbeit verliehenen Titels berechtigt ist.

Bern, 23.09.22 Nicole Widmer

Declaration of publication

Einverständniserklärung zur Veröffentlichung der Masterarbeit

Ich erkläre hiermit, dass ich der Veröffentlichung der von mir verfassten Masterarbeit im Falle einer Benotung von 5.0 oder höher auf der Homepage des KPM zustimme. Die Arbeit ist öffentlich zugänglich.

Bern, 23.09.22 Nicole Widmer