Chapter 10
Barriers of Artificial Intelligence in the Health Sector

Laura Beltempo, Jasmin Zerrer, Ralf-Christian Härting, and Nathalie Hoppe

Abstract Demographic change, shortage of qualified employees and increasing cost pressure—the healthcare sector has to deal with various challenges. Coping with the current COVID-19 pandemic is an additional issue. All these barriers contribute to the fact that digitalization in the healthcare sector is moving forward more and more. Without the application of advanced technologies, healthcare organizations would reach their limits. In this context, the use of AI is becoming increasingly important. The potentials are wide-ranging and include applications in diagnostics and therapy, as well as the development of pharmaceuticals. But what challenges are associated with the use of AI in healthcare? Within the framework of a qualitative empirical study according to Mayring, this question has been investigated. Based on a systematic literature review, the following barriers of AI in healthcare have been identified and examined: Disagreement in data protection, lack of compatibility with ethical aspects, quality of training data, knowledge, and trust of physicians in AI-supported systems. The next step in the research design have been expert interviews among medical staff as well as AI developers with focus on AI in the healthcare sector mainly in Germany. According to these interviews, the data are analyzed and evaluated. Based on the results of the study, potential activities have been derived in order to be able to successfully overcome the barriers of AI in the healthcare sector in the future. Finally, the opinions of physicians and developers on the identified barriers are compared and discussed.

L. Beltempo · J. Zerrer · R.-C. Härting (✉) · N. Hoppe
Aalen University of Applied Sciences, Aalen, Germany
e-mail: ralf.haerting@hs-aalen.de

L. Beltempo
e-mail: laura.beltempo@kmu-aalen.de

J. Zerrer
e-mail: jasmin.zerrer@kmu-aalen.de

N. Hoppe
e-mail: nathalie_vanessa.hoppe@kmu-aalen.de

C. P. Lim et al. (eds.), *Artificial Intelligence and Machine Learning for Healthcare*,
Intelligent Systems Reference Library 229,
https://doi.org/10.1007/978-3-031-11170-9_10

Keywords Artificial intelligence · AI · Healthcare · Barriers · Digitalization · Trust · Data privacy · Expert interviews · Systematic literature review · Qualitative study

10.1 Introduction

Digitalization became one of the most important economic topics of the future [1]. Big Data technologies have the potential to process large amounts of data from different origins and various source structures with high speed [2]. This capability is also used by artificial intelligence systems to generate far-reaching changes in treatment methods [3]. Machine Learning, Deep Learning, Natural Language Processing, Big Data Analytics, and a lot more, are terms that are summarized as artificial intelligence (AI). But what exactly is meant by the term AI? Since there is currently no general accepted definition, AI can be defined as an empirical discipline of computer science that explores the mechanisms of intelligent human behavior [4]. With the help of computer program simulations, algorithms are programmed that resemble the functionality of a human brain. In this way, artificial neurons are used to form networks that enable machine learning. The two main requirements that an AI system must fulfill are the ability to learn and the ability to deal with uncertainty and probabilities [2].

In order to strengthen Germany's position in the development, research and application of AI in international competitions, the German government adopted the "Artificial Intelligence" strategy in November 2018. The goal is to promote the trademark "AI Made in Europe" to further establish and expand responsible and public welfare-oriented application and development of AI systems in Germany and Europe. The funding is aimed to advance the application areas of AI in all parts of society, including medicine [5].

Automating human or manual decision-making behaviors using AI makes it possible to improve patient care while reducing costs and long waiting times. In some medical fields, AI is even better than humans. Currently, the most prominent field of application of AI-based technologies is radiology, followed by oncology and pathology [6]. However, despite many opportunities, the use of AI also creates some hurdles. Besides trust and privacy issues, ethics is a controversial subject, especially autonomous decision-making [7]. Particularly in times of COVID-19, the need for immediate ethical consideration of medical actions become the focus of current debates [8]. The main ethical implications are negative consequences, such as damaging things or, in the worst case, harming people. Since an algorithm is not a legal entity, the question of who can be liable for an AI-based failure in general needs to be clarified [7]. Unfortunately, so far, there are no specific liability rules for any damage resulting from AI application. Therefore, human and machine actions must not be directly equated [9].

Digitalization and the use of AI in healthcare is an inevitable process that many people cannot fully grasp and appreciate [7]. Therefore, this study uses a qualitative

research approach to explore the impact of AI, focusing on barriers in the healthcare sector.

10.2 Empirical Investigation

This section describes the applied strategy step by step and presents the results of the systematic literature review. The literature-based hypotheses build the foundation for the data collection using expert interviews. The process of data collection and the data analysis are also explained in detail.

10.2.1 Research Design

The empirical investigation is based on a qualitative research design. Due to the fact that the research subject is very current and that there is plenty of literature on the topic, a deductive-inductive method according to Philipp Mayring was used.

Figure 10.1 provides an overview of the applied research design. The aim of the research is to investigate the barriers of AI in the health sector. For this purpose, a systematic literature review was conducted in the first step. Several challenges were identified, which were subsequently formulated as a hypothesis. In order to substantiate the generated hypotheses, semi-structured expert interviews have been con ducted for the data collection. Afterwards, the interviews were transcribed and coded with the software MAXQDA. The results from the analysis of the interviews are interpreted together with the findings from the literature research and finally lead to answer the research question. The original hypotheses were extended by new hypotheses, which resulted from the analysis of the interviews. Based on that, a conceptual model was developed, which can be used as a basis for a quantitative research work in the future.

The individual steps of the empirical investigation are described more detailed in the following sections.

10.2.2 Systematic Literature Review and Generation of Hypotheses

The development of a hypothesis framework provides the basis for a deductive investigation according to Mayring, with the aim of answering the research question. For the literature review, the following databases were accessed: ScienceDirect, Google Scholar, Springer and IEEE Xplore. Because the topic AI in healthcare is a current

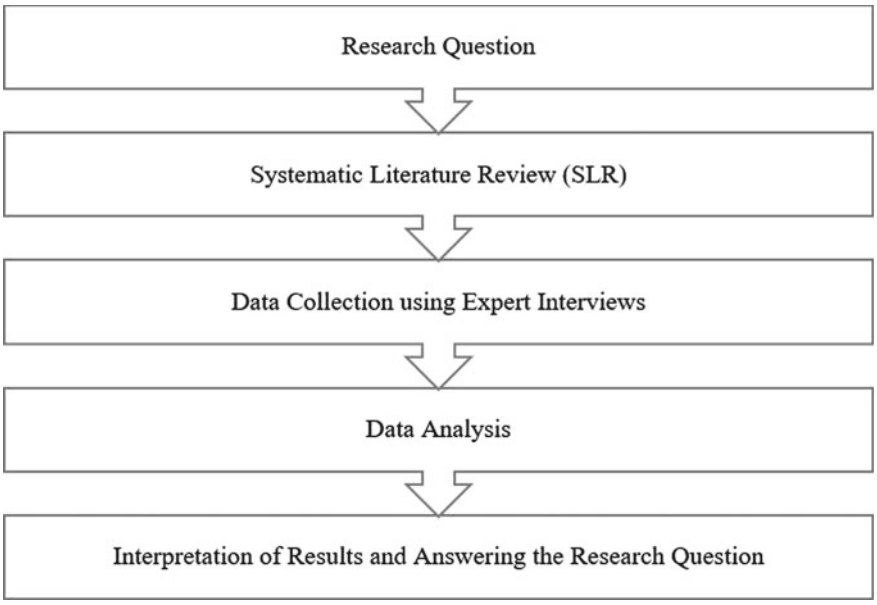

Fig. 10.1 Research design

issue, the selection of journals focused on the most recent ones available. In addition, only journals with a rating of A and B were used to ensure a good quality of the articles. The two ranking portals CORE Rankings Portal [10] and VHB-Jourqual3 [11] were used to check and evaluate the scientific quality. The following section presents the results of the SLR and the hypotheses that were derived based on these findings. A total of six hypotheses were formulated in order to identify barriers for the application of AI in healthcare.

Quinn et al. (*in 2021*) mention that the recommendations and actions of AI systems cannot be explained for proprietary or technical reasons. In this context, the "Black Box" phenomenon is often discussed. The lack of transparency can diminish autonomy and affect trust in the recommendations of an AI mechanism [12].

Cheng et al. (*in 2021)* describe explainability and interpretability as a key for transparent AI algorithms. They distinguish between local and global transparency. While local transparency focuses on explaining decisions in a specific context, global transparency aims to understand the entire AI System. At the same time, total transparency can also create potential risks. For example, disclosing additional data and information can make AI more vulnerable to hacker attacks. Therefore, a balance needs to be found between necessary transparency and the associated risks [13]. As a lack of transparency is assumed to influence the use of AI in healthcare, the following hypothesis is derived:

H1: Lack of transparency impairs the use of AI.

Kokciyan et al. (*in 2021*) describe that the misbehavior of AI technologies is not always clear or statistically detectable. For this reason, regulatory structures are needed to constrain these behaviors without compromising autonomy [14]. The focus is placed on decision algorithms such as classification or evaluation to minimize undesired misbehavior and to achieve fairness. Another cause emerges as a result of failures in measuring algorithm performance. When scientists present their research results, they usually argue that algorithms have a certain level of accuracy. However, this assumption depends on whether the training and testing samples are representative of the target group and whether their distributions are similar enough. In practice, it has been demonstrated that the learned AI algorithm achieves zero training bias for unrepresentative samples in the beginning phase. Due to the fact that new added data increases the probability of AI-misbehavior [13], the following hypothesis is generated:

H2: Misbehavior of systems complicates the use of AI.

Khullar et al. (*in 2021*) focus their study on the issue of liability when using AI in healthcare. This includes the question of who should be responsible for medical failures resulting from joint treatment between physicians and algorithms. In their study, they concluded that the public is significantly inclined to believe that physicians should be liable. However, physicians believe that healthcare providers and organizations should be responsible for AI-related medical failures. Because the use of AI in healthcare is new, there is currently not much established jurisprudence on how liability issues should be decided. So far, it is also unclear how these issues will be resolved in the future, because no major cases involving the use of AI in medicine have been decided by the courts. It is imaginable that physicians will take responsibility in situations where AI is recommended, instead of the machine or its AI algorithm. In addition, the liability issue may depend on who developed the used algorithm. If a physician uses an AI system developed by their own healthcare organization, the organization may be assigned a higher liability risk due to "Enterprise liability." In this case, hospitals and medical groups are blamed for failures rather than individual physicians [15]. This leads to the following hypothesis:

H3: Liability issues complicate the use of AI.

Bartoletti et al. (*in 2019*) mention data privacy as a challenging factor for the use of AI in healthcare. In this study on ethical and data protection aspects, it was clearly shown that data protection laws are a hurdle to achieving medical innovations. Privacy is tied to cultural norms and evolves over time. Therefore, it is also conceivable that data protection will change along with technological progress. However, this is always dependent on the respective country and generation. Data can also be quickly misused by hackers and thus increase distrust in the release of data to collection points. In addition, many people are insecure about the lack of transparency in data protection. This is because data protection notices are often too long and incomprehensible to the layperson. When it comes to health data, this is the most private information a person has, so there must be a high level of trust. On the one hand, health data is very sensitive and requires special protection. On the other hand,

the use of such data in a research context can improve health conditions enormously. Therefore, finding a good balance is a major challenge, especially in the development of AI algorithms in healthcare [16]. Due to the challenging factors on data privacy, the following hypothesis is generated:

H4: Uncertainty in data privacy hampers the use of AI.

Straw (*in 2020*) describes that bias in health care can lead to data-driven, algorithmic, and human errors. Prejudices in AI data can bring terrible consequences in healthcare. This can put certain groups of people at a disadvantage when it comes to medical care. The reason for this usually lies in deep-rooted inequalities between ethnicities and genders, which are now reflected in the underrepresented datasets. Accordingly, AI systems can reinforce discrimination based on the trained dataset. Therefore, the challenge of choosing the right dataset should be countered by using standardized training specifications and continuous development of datasets. However, there is still a risk that unconscious bias may appear in the data, putting certain groups of people at a disadvantage through the use of AI [17, 18]. This leads to the following hypothesis:

H5: Discrimination has a negative impact on the use of AI.

Boddington et al. (*in 2017*) focus their study on answering ethical questions in relation to AI. In society, artificial intelligence can scare people through various levels of unpredictability. This lack of knowledge about the degree of autonomy of systems makes the public skeptical and distrustful. Machines are often ascribed the same kind of autonomy as humans and consequently there is concern about how to control them [19]. Today, AI systems are used in hospitals, law enforcement, and other sensitive areas. Therefore, it is highly desirable that AI behaves correctly in these areas. Doctors, therefore, do not find it easy to rely on AI systems and perhaps transfer their decision-making authority to the machine in the future. In most cases, physicians are not properly informed about how such an AI algorithm works and about its strengths and weaknesses. The lack of knowledge among physicians leads to a negative attitude toward the use of AI systems [20]. Due to the misunderstanding in society that AI is at risk of taking over the world the following hypothesis was formulated:

H6: Decision-making authority impairs the use of AI.

10.2.3 Data Collection

The data collection includes sixteen semi-structured interviews with nine physicians, who have experience with AI in the healthcare sector, as well as seven AI-researchers and AI-developers. The interviews were conducted between December 2021 and February 2022 and the main focus has been on the German healthcare sector. Only one expert was from Austria. Participants from different functional areas were interviewed to gain a broad understanding of the situation and minimize individual bias.

Table 10.1 Interviewee overview

Number	Gender	Professional field	Position
1	Female	Surgery	Senior dental radiologist
2	Male	Cardiology	Specialist for cardiac surgery
3	Male	Radiology	Radiologist
4	Male	Researcher	Manager digital health & innovation
5	Male	Researcher/developer	Research assistant/Ph.D. student
6	Male	Radiology	Radiologist
7	Female	Researcher/developer	Senior AI and FEM Expert
8	Male	Researcher	Research group leader
9	Female	Surgery	Dentist
10	Male	Researcher/developer	Research group leader
11	Male	Researcher/developer	Developer and project manager
12	Male	Researcher/developer	Research assistant
13	Male	Surgery	Hand surgeon
14	Male	Radiology	Radiologist
15	Male	Radiology	Radiologist
16	Male	Internal medicine	Doctor, program director telemedicine, E-health and artificial intelligence

The interviewees in this study are between 25 and 64 years old and have an average of 8.3 years of experience with AI in healthcare, so a high level of expertise in this research topic is assumed. Table 10.1 gives an overview about all participants.

The questionnaire was created based on the hypothesis model and includes eighteen questions. The questionnaire consists of six parts. Part A includes socio-demographic questions, such as age, gender, or academic background. Part B includes questions about the experience with AI in the healthcare sector and part C to F contain one or two questions about each hypothesis. The questionnaire was distributed a few days before the interview, so that the experts had the opportunity to prepare the topics. The interviews were conducted and digitally recorded via the platform Zoom with a duration of 30–60 min. Afterwards, the answers were transcribed in the original German language, coded, and subsequently analyzed. The results are presented in detail within Sect. 10.3.

After the preparation of the questionnaire, a pre-test was carried out together with a physician in order to check the comprehensibility of the questions and to eliminate overlaps in content.

10.3 Results

This section presents the results of the expert interviews. First, the process of data analysis is presented. Then the conceptual model is introduced by the results from the expert interviews, followed by the derived hypotheses.

10.3.1 Data Analysis

After the semi-structured interviews were completely transcribed, the data analysis process started. The data was analyzed by using the Philipp Mayring coding guideline. For this process, the software MAXQDA was used. In the first step, individual categories were created by generating inductive and deductive codes. The inductive codes consist of keywords that can be derived from the structured literature research and were searched in the interviews. Deductive codes emerged from the analysis of the interviews and were generated for topics that were mentioned by many experts. At the end, a combination of inductive and deductive codes was formed. The individual categories were substantiated during the analysis with the help of anchor examples, so that the assignment of codes was transparent and comprehensible.

In this way, subcategories of related categories could be formed in the next step. For example, the codes "Black Box", "Misbehavior/wrong decisions" and "Lack of Empathy" build the third-order category, which is combined into the second-order category "Trust of Physicians". The second order categories are divided into "Data Privacy", "Ethical Principles", "Knowledge of Physicians", "Training Data" and "Trust". In this study, the core category represents the barriers of AI in the health sector.

A total of 623 codes were generated from all 16 interviews. These were subcategorized and grouped into 22 categories. To decide which of the codes would ultimately be included in the conceptual model, it was determined which codes were most frequently addressed by the experts. The remaining codes were not included in the analysis because they were only mentioned by a few experts. In this way, it was possible to identify the most important aspects that have a negative influence on AI in healthcare.

10.3.2 Empirical Findings and Model Conceptualization

To understand the method, a conceptual model was developed, which is presented in Fig. 10.2. It is based on the coded expert interviews and the previous SLR. A total of five main hypotheses and seventeen sub-hypotheses were identified, which were explained in detail in the following.

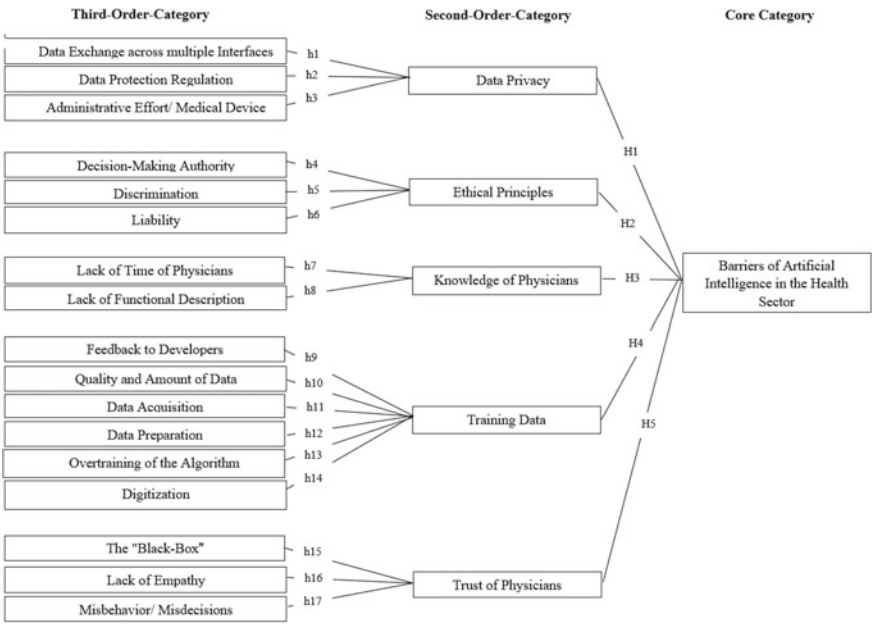

Fig. 10.2 Conceptual model

10.3.2.1 Data Privacy

To minimize the risk of hacking, hospitals avoid exchanging data between multiple interfaces. Each hospital has its own data center that is completely isolated from the outside. Because of sensitive data in medicine, the IT security level is very high. Therefore, the use of medical data is associated with a lot of effort for AI developers. To guarantee a secure transmission of the data, the encryption and anonymization of the data is essential. In this context, the permission of the patients is required for the use of data in the research framework.

[…] Clearly, if data is exchanged with other university hospitals via interfaces or even uploaded to some cloud, there is a huge risk. […]

On the one hand, IT security is a hurdle, because data collection is made more difficult. On the other hand, personal data are particularly worth protecting.

[…] In other words, this is not something that is thought up for fun; it is about handling data sensibly. And that is why it is good that something like this exists. And it has to be balanced somehow. […]

In Germany, there are doubts about data being exchanged between interfaces. In other countries, this is seen as an opportunity to provide better and faster medical care in emergencies. Therefore, data transfer between hospitals should be made possible in general. The framework of encryption and anonymization of the data must be observed at all costs. Sometimes it can be difficult to anonymize data when they are

needed from a specific application area, e.g., from the head area. It should not be possible to draw conclusions about individuals.

[…] But I think if the data is all anonymized and encrypted, it should actually already be possible to transfer it with a relatively high level of security. […]

h1: Data exchange across multiple interfaces has a significant impact on data privacy.

The data in Germany must be processed in a DSGVO-compliant manner. Patients and participants in studies should be informed immediately about what happens to their data and that the data can be withdrawn at any time. This should be followed without exception in order to avoid uncertainties and complications with data protection. A way should be found that supports the publication of knowledge and not the privatization of knowledge or algorithms.

[…] A European path that is linked to the data protection regulation. I think this is really important, even if it's tiresome, even if it slows down. […]

For many researchers and developers, the effort to follow the requirements of the DSGVO and other ethics committees is too high. Therefore, they have less motivation to research AI-innovations in connection with patients. To sum it up, data protection is a major barrier to the use of AI because developers don't get enough data. In addition, ethics committees are becoming stricter, leading to even higher hurdles for physicians in acquiring data.

[…] Sorry, but I don't feel like doing that. It's probably incredibly difficult to find a balance between the need for data and the need to protect data. […]

Compared to other countries, such as the USA, Germany is already lagging in terms of data collection and computing power. It can be seen as a risk but also as an opportunity to say that if we can't centralize the data, then we can develop methods to achieve good quality without centralizing the data. This would be a locational advantage, to build AI in Europe that is compliant with data protection. The reason why other countries are more advanced in AI development is because of less regulation in terms of data protection, less complicated paperwork, and more freedoms. This leads to the risk that qualified specialists will prefer to deploy their expertise abroad. However, the management of the Corona pandemic in Germany shows how damaging strict data protection laws are for an efficient IT infrastructure in healthcare. It can be concluded that if Germany fails to catch up with other countries by adapting the data protection regulation, the country runs the risk of buying expensive AI algorithms from abroad instead of selling themselves.

[…] We are good at basic AI research, but we are losing a lot of ground in developing products because it is really difficult in Germany at the moment. […]

Another problem is that politicians in Germany are making demands that contradict each other. They would prefer all data to be centrally available and visible to every-one, but totally privacy-compliant and totally private. These goals are not easy to combine. In addition, the German healthcare system is one of the most complex in the world. Because of this great interaction between and against each other of the most diverse institutions, it is extremely inhibiting to innovation. This does not just affect AI; it affects the entire digitalization and innovation readiness in healthcare. In

this context, the difficult structures but also the convenience of doctors, who do not want to change anything in their paper documentation, are major inhibiting factors.

[…] And that is also the biggest barrier, that this structure of the health care system is reorganized a bit, so that it becomes a bit more fluid. […]

h2: Data protection regulation has a significant impact on data privacy.

Practically, AI systems have a supportive role in medical treatments. Therefore, there is an increasing development towards structured reporting, especially in radiology. Currently, AI is implemented to support faster decision-making and minimize risks. Physicians report that this makes their daily work easier.

[…] To facilitate administrative activities, so that time is again made possible for the encounters between people, doctors, and patients on both sides. […]

You need standardized data to train an AI. The problem is that the documentation must be personalized. So, standards are actually not correct from a legal point of view.

h3: Administrative effort has a significant impact on data privacy.

In summary, it was found that physicians and developers see a major hurdle to the use of AI in data privacy laws. Compared to other countries, the many administrative requirements imposed by ethics committees and approval applications slow down technological progress in healthcare.

H1: Uncertainty in data privacy hampers the use of AI.

10.3.2.2 Ethical Principles

Doctors can quickly be offended in their position as physicians if you want to replace their work with AI. However, this often depends on the gender, age, area of application and perception of the person. In this context, today's AI systems are purely for support, and they are an aid to decision-making. Nevertheless, there is a high level of skepticism on the part of physicians about the replacement of medical decisions through AI systems. However, some experts have the opinion that these concerns will become less pronounced in the future or will be ignored. Interviewee 4 mentioned in this regard:

[…] The question is merely how quickly and to what extent will certain physicians and areas of application be replaced. Perhaps not 100 percent, but I can very well imagine that in the medium to long term a 90 percent replacement rate could prevail. […]

The majority of respondents believe that complete replacement by AI will never be possible. Medicine is too complex and multi-layered for that. There are always special cases. In addition, talking to patients is an important part of medical treatment. There are ethical questions that cannot be taken over by artificial intelligence. The living is just one example where the human being must make the decision.

[…] And in the end, it is always very important that the human factor does not disappear completely, because computers can calculate well and calculate quickly, but they can't replace the human factor. […]

On the other hand, standardized treatment decisions can be replaced by AI soon. Ultimately, this transfer tends to give doctors more time for each patient rather than replacing them. It is also possible that staff will be eliminated by assistance systems, thereby somewhat counteracting the shortage of personnel in the healthcare sector. This always depends on the situation and the area of application. For example, radiology with its imaging processes is already very advanced with the use of AI. Interviewee 6 explained:

[…] So we do individual CT examinations in radiology today. We have 3000 images for one examination. And that is of course time-consuming to look at. People are not always one hundred percent concentrated. We get tired, and of course these systems can help us. […]

In the end, doctors will see AI systems as an aid, but they would never allow them to replace them. Interviewee 1 said the following statement:

[…] We have different roles and different stakeholders. And I don't think they're going to be able to replace that because we're probably not going to allow it to happen in the first place, which is to be replaceable. After all, we want to be loved. […]

In summary, taking over medical decisions is going to take a long time, if AI ever catches on. This will be difficult, as the use of AI is contrary to the preservation instinct of physicians. However, it is to be expected, that in the future the demographic change and the shortage of staff will continue to come to a head and at some point, there will be no other option for the widespread use of AI systems in healthcare.

h4: Decision-making authority has a significant impact on ethical principles.

When AI is used, certain groups of people may be favored. This is not intentional but happens unintentionally and depends on the data set used. Researchers often exercise a training set in good faith and do not detect bias that may occur. If the data set contains more information about men than about women, the AI will work well with men and worse with women. To prevent this, attention must be paid to the quality of the data. If it is a specific disease, then discrimination would make sense. There is a chronic disease that is more prevalent in dark-skinned people. If the disease is detected early, deadly consequences can be prevented. That means, in this case, you want a discriminatory model that is specifically positive for dark-skinned people. In healthcare, AI applications must pass a high study setting before they are allowed to use. The bottom line is that it must be defined in advance where and for which group of people the system is to be used. From this point of view, data should be categorized so that a qualitative evaluation in medicine is possible.

[…] AI systems show discriminatory behavior, but they are not proactive agents. It is a tool. It is not an autonomous system in its own right. […]

Often, incorrect data sets are discovered only when the AI is used. Then it is important to report the mistakes to a central location and to reprogram the AI. The patient makes a contract with the doctor. Therefore, the doctor will always remain responsible for his actions. In this context, physicians need to be informed about the training data so that they do not exceed the target group. In addition, AI can be used for social scoring systems, especially in other countries. Social scoring would

be used to decide whether a person is preferred for medical treatment or not. That would be very discriminatory and could be abused by AI.

[…] And that just shows that there is definitely, […], a big risk of discrimination. Also, scientifically, so not only social discrimination. […]

On the one hand, AI is more objective and neutralizes certain prejudices that people can have. On the other hand, AI generates a new form of discrimination that is not caused by emotions, but by errors in the input of basic data. Often this happens un-consciously and should be closely examined by ethics committees before AI is introduced in healthcare.

h5: Discrimination has a significant impact on ethical principles.

Experts have different opinions on the liability situation regarding AI. Who is liable in the context of damage caused by faulty measurement of AI? Five interviewees can imagine that the developer can also be held liable in the future. If AI should make decisions on its own at some point and a guarantee for safe results is given by the developer, he will be liable. If no guarantee is given, the hospital must expect possible errors and learn to deal with them. In general, it is also fact that developers or the company are always responsible for a defect in the product.

[…] And that would also be difficult in the development of such products if the user were then held liable, for a product problem. […]

Other 7 interviewees see the liability rather with the physicians. If the medical device is not defective, the user is responsible. In the case of assistance systems, it is difficult not to hold the physician liable because the developer only offers a recommendation with the AI. In the end, the physician must make the final decision and is therefore responsible. The risk for medical technology manufacturers to spend a lot of money on liability cases is too high. That is why the companies will not be liable. From this point of view, the liability will probably remain with the physicians, because they are expected to check the results of the AI system with their expertise and then make a decision. The opinion of interviewee 1 was:

[…] The physician performing the action is the one who makes the decisions and the one who is liable for them. […]

Another option would be to have the liability run through insurance companies in the future. In principle, customers would then pay part of the insurance when they purchase the product, and this would be balanced out somehow. This is how malpractice is handled in hospitals. If it can be proven that the doctor was in fault, the hospital is liable for the doctor via an insurance policy.

[…] I could imagine that a hospital would then also be liable via *insurance for an error by the AI system. […]*

In addition, it is very important for the development of AI algorithms that the liability does not belong to the producers. Otherwise, no scientist would take the risk and develop an AI. To sum it up, an AI software ends up being a medical device and has to go through a very strict approval process. So, when the software comes onto the market, it can be considered that the quality is of a very high standard, especially in Europe. Since AI has no identity, it cannot be considered a legal entity. Therefore, the solution to the liability issue via insurance should be the closest.

h6: Liability has a significant impact on ethical principles.

To conclude, all mentioned aspects, decision-making authority, discrimination and liability issues complicate the use of AI in healthcare. This leads to the following hypothesis:

H2: Ethical principles hampers the use of AI.

10.3.2.3 Knowledge of Physicians

For doctors, lack of time is always present. That is why AI systems that make a pre-selection are a time-saving help. In this way, the workload can be minimized, and output can be increased at the same time. In addition, AI helps to make decisions faster and reduce errors in medical treatments.

[…] It was actually a great relief to have this kind of AI support, so that you didn't have to do the paperwork 20 times again but could check it off quickly with this kind of AI support. […]

The time problem also means that physicians don't even find the time to inform about AI and to implement AI systems in their practices. This is because they have to ensure patient care and are thereby focused on medical care. It is a dilemma because in the end, AI would save them time again. So, a big challenge is to even realize AI in the first place and then implement it despite the lack of time and complex insurance policies.

[…] but it, just makes life difficult for doctors because they have to decide between two things. But their main job is to care for patients, so that's the main hurdle. […]

h7: Lack of time of physicians has a significant impact on the knowledge of physicians.

There is the problem that we use AI systems in practice that have theoretical errors that we do not even know about. This means that if doctors use such an assistance system, it should only work in the background and provide suggestions. No decisions should be imposed on the physician. An experienced physician should still be able to estimate the results and determine the level of confidence. A broad knowledge transfer of AI will not take place due to time constraints. There is always an explanation of possible risks when using a device, but there is no time or desire for more in most cases. Nevertheless, doctors should be made aware of AI tools. It is important that developers communicate exactly what AI can and cannot do, so that doctors can interpret the results of AI correctly.

[…] I mean doctors specialize in a thousand things, have to learn so much, have such a hard job. I don't think they want to deal with the theory of deep learning. […]

For most physicians, AI is incomprehensible, especially if they are not familiar with machine learning. As an external user, decisions made by AI systems are difficult to understand. In this context, it would be useful to provide appropriate and easy-to-understand instructions for the use by the developer or via user manual. Depending on the requirements, instructions may not be sufficient in some cases. Therefore, good training by the developer and information about the risks should not be neglected in any case.

[…] I don't think you just have to specify that in a user manual, but you also have to clearly train the doctor. […]

h8: Lack of functional description has a significant impact on the knowledge of physicians.

Mainly, the time limitation and the recording capacity of physicians prevent the flow of knowledge about AI. In addition, in many cases, there is insufficient training on how to operate AI systems. Physicians should be made more aware of the benefits of AI implementation, and sensitization to the topic should be promoted.

H3: Lack of Knowledge of Physicians hampers the use of AI.

10.3.2.4 Training Data

AI is a learning system. In practice, learning ends at a certain point and the system is integrated into practice. However, learning is actually not completely finished, because the same learning data can be used again and improved if necessary. This continuous learning process is very important for an AI system to improve its quality in the long-term. The interaction between physicians and developers plays a major role in this process because developers and manufacturers depend on feedback from physicians so that they can optimize the algorithm. After all, the accuracy of the AI results must still be determined with a specialist in the end, and this determination can minimize undesirable misbehavior. It is also conceivable to develop a feedback system that directly informs the AI system of the failure. In this way, the algorithm in the AI system would be trained further by telling it:

[…] Hey, here is another example of a situation where there is actually no tumor. But you said there is one, just take a look at it again. Try to update your model again. Hoping that if the same picture or a similar picture comes in, this failure will not happen again. […]

The challenge is that once a system is approved, it can't be changed. In this case, the AI system would have to go through the approval process again due to the strict regulatory requirements, which involves a lot of time and effort.

h9: Feedback to developers has a significant impact on training data.

A large amount of data that has simultaneously a high quality is important for an effective and specific training of AI. This is the only way to ensure that AI algorithms are provided with one hundred percent quality to the users. Interviewee 6 mentions in this regard:

[…] one of the main problems is the lack of well-adopted datasets. We require them to train the AI systems, so that they all perform a super learning. We need large amounts of annotated, and qualitatively well-annotated datasets! […]

h10: Quality and amount of data has a significant impact on training data.

In medicine, a lot of data is generally available and well documented, but the challenge is to gain access to the data. Interviewee 1 commanded:

[…] My hurdle right now is getting data out of these thousands of x-rays from the hospital. So, I wish that we could get better access to research data. […]

Because we are dealing with sensitive patient data, it is usually stored on local servers in hospitals and doctors' offices. The data pool is highly protected, so it is not possible to simply pull out and merge the data needed to train the algorithm. Currently, there is no suitable tool in Germany to exchange and merge data, that is why interviewee 5 said:

[…] It would be a dream if there is a cloud or something like that, where all data is stored completely anonymized, with the same prerequisites as I just described. And that all research institutes would have access to the data, possibly also German companies, if they pay a bit of money for it. […]

h11: Data acquisition has a significant impact on training data.

Data need to be specially prepared for the training of AI, because they are not always structured in the same way. A typical example of this is marking a tumor in a CT scan. Some physicians mark the tumor with a cross, encircle it, or mark nothing. As every doctor has his or her own procedure, it is very time-consuming for the developer of an AI algorithm to prepare the dataset in such a way that it can be used for training. Interviewee 8 explains that:

[…] the heterogeneity of the data, i.e., how the data look, is another problem. You have to prepare it so that you can use them, for example for machine learning or other AI methods. […]

In most cases, a physician must be consulted to explain the data in detail to the developer. This is very time-consuming for both parties, so standardized documentation criteria in Germany can help to minimize the effort involved in data preparation.

h12: Data preparation has a significant impact on training data.

AI systems are trained with plenty of data before they are approved. During this process, it is important to ensure that the system is not overtrained. This means that a part of the available data must always keep aside, which will be used for the final validation. In this way, it is possible to check whether the AI only performs well based on training data or also based on validation data. If it comes to the realization at this point that there are certain misfitting during the training, it is probably an over fitting.

[…] If you have over-fitted the AI on training data, then you have to either increase the amount of data or just adjust the targeting of the data even more to match the target population. […]

Overtraining can also result because AI algorithms need to be adjusted and revised based on the feedback from the physicians. The training process is triggered again and, ideally, the algorithm learns this special case. However, a challenge is also, that retraining can lead to the AI forgetting what it has already learned. This means that whenever the AI system is extended, the complete validation process must be repeated to ensure that the AI can still do what it has already learned. This means it must be guaranteed by the developer that the algorithm is at least as good as before, so that no undesirable overtraining effects occur.

h13: Overtraining of the algorithm has a significant impact on training data.

Digitization plays a major role for the application of AI, because only where digitization is well-advanced, AI can be applied. Compared to other countries, like China

or the USA, Germany is far behind in the healthcare sector. The entire digital areas are comparatively weak in the German healthcare system. Interviewee 10 mentions the Corona pandemic as an example for this:

[…] that fax machines are still used in Germany today, even in clinics or in administration. This is a shame. We are lagging in terms of digitization in general. The processes should be digital with reasonable interfaces that reach the target quickly […]

The complication that emerges especially for developers is that it is hard for them to get the data they need for the AI training. This is because one physician stores the data online, while another physician still writes the reports with a pen and a paper. To successfully integrate AI into the German healthcare system, according to interviewee 7.

[…] Germany really first requires the data as a digital document. That is the biggest hurdle. […]

h14: Digitization has a significant impact on training data.

AI systems are only as good as the training data they have used to learn specific functions. Raw data or test data are crucial to ensure that the system knows exactly what it is supposed to find during the application. If only an insufficient, perhaps unrepresentative, selection of training data is used, there is a risk that maybe wrong diagnoses may occur.

[…] Very important is that you have to know what kind of goal you have with AI, because if you do not define this goal exactly, then you are off the track! […]

Therefore, the developers of AI algorithms must pay close attention during training to guarantee that the system is trained on the correct data, which are relevant to the topic. The experience and expertise that a developer contributes also plays a major role, because not everyone can recognize patterns, for example in the field of imaging. If people with limited knowledge are hired in advance for AI training, the probability is high that the programmed algorithm will have weaknesses in the end.

H4: Training Data has a significant influence on AI.

10.3.2.5 Trust of Physicians

The term "Black Box" is a big problem for users, as described earlier in the SLR. For the experts, the challenge is to understand where the AI results come from. Only if the physician was directly involved in the technical development, there is a chance to understand the system better. Especially in the medical context, algorithms are comprehensible to the user from a certain perspective, or not at all. The lack of transparency naturally also has a negative impact on the trust of users. In the initial phase, every physician will only trust an AI system if it is possible to understand the decisions. Interviewee 3 has the following opinion:

[…] I think that it is still a issue, that people are very skeptical, also because the decisions or the results of the AI systems are often not completely comprehensible. And that creates a bit of a lack of trust. […]

To reduce this skepticism, the research area "Explainable AI" is currently working on making AI algorithms more transparent. For example, there are methods that are being used in the field of imaging. With some of these, an attempt is finally made to visualize which areas were relevant for the result and where the algorithm gained the most information. After all, when users trust in AI, patients' confidence will probably increase as well.

h15: The "Black Box" has a significant influence on the trust of physicians.

One challenge of using AI is the lack of empathy towards patients. This affects the trust of physicians, because they have to show empathy to their patients in certain situations. AI will probably never be able to replace empathy from a doctor. Interviewee 6 comments on this aspect:

[…] I do not think that AI will be able to replace the doctor or other medical professionals with all the humanity, with the attention as well as with the empathy. That is an important function. […]

As an example, serves a patient with a malignant tumor. There is the possibility of an operation, which is not promising. In this case, the situation may arise that it is not useful or that the patient does not want the surgery. Since a human being is capable of empathy, it can better assess the situation. An AI would most probably try to save the patient's life, regardless of the patient's individual wishes. This contributes to the fact that a lack of empathy currently affects the confidence of users.

h16: Lack of empathy has a significant influence on the trust of physicians.

AI-based misbehavior or wrong decisions will probably always happen, the same way as human misbehavior exists. Because the current state of development in Germany is not yet advanced enough to use AI as an autonomously acting system in medicine, the term wrong decision is used more frequently. This is because misbehavior would imply that AI has a will of its own, which is currently not the case. Basically, it can be said that AI systems that make serious mistakes are not used in medicine at all. However, to give the physician more confidence and safety, the AI developer can indicate the probability that a mistake may occur. Then the users can decide for themselves whether it is acceptable to utilize the AI system or not. Interviewee 11 says:

[…] I don't want to go to a hospital where the algorithm might be able to detect a disease with 80 percent accuracy. That would be a disaster. That is why it is good that there is a regulatory framework for this. […]

This means that trust in AI can be increased if the reliability of detection is more exactly defined.

[…] the good thing is that when a wrong decision is made, the algorithms are supposed to learn from it. In the context of technical development, it is something very good, this trial-and-error of learning AI systems […]

Since AI is currently only used as a support system in Germany, a specialist always takes a final look at the AI results. In this way, the physician always has the final responsibility and can minimize possible errors.

h17: Misbehavior/ wrong decisions have a significant influence on the trust of physicians.

Humans are usually reluctant and skeptical towards new technologies at the beginning. But once they have been implemented for a period of time, they are accepted. Trust towards AI may not yet be completely established, for example, because physicians are afraid of being unconsciously influenced by AI. Interviewee 4 thinks:

[…] in reality, a person is influenced by these unconscious decision-making aids. And I think there is currently still a huge skepticism, if you directly confront someone as a medical professional. […]

One way to eliminate skepticism and build more trust is to convince physicians, for example, through valid certifications and high-quality results and studies. Only in this way, AI can gradually build trust. But of course, trust can also be quickly destroyed if mistakes happen. Creating more transparency plays a crucial role in this regard, because the user will only trust such a system, especially in the initial phase, if it can understand the decision.

[…]What creates trust is when everything that happens is very transparent and the highest security standards are really maintained, so that there are no data scandals in this area. And people should be educated about what possibilities there are through the application of AI […]

H5: Trust of physicians has a significant influence on AI.

10.4 Discussion

The barriers identified in this study are data privacy and ethical aspects such as the physicians' decision-making authority. In addition, it was figured out that the knowledge level of physicians has a significant impact on the use of AI in healthcare. Moreover, the quality and amount of training data as well as the trust of physicians turned out to be moderating factors. This study shows which influencing factors slow down the use and development of AI systems in healthcare. If this conceptual model will be considered, these issues can be targeted to drive forward the implementation of AI in Germany. One of the biggest barriers in this context is that Germany has stricter legislation and more requirements to meet than other countries. It is important to protect personal data, but there should be a balance between sharing data for research and data protection. Otherwise, the gap in technological progress will grow and may lead to a competitive disadvantage in the future. In addition, the influence of the various ethical aspects for the use of AI should be considered. A significant hurdle in decision-making authority is that the use of AI is contrary to the preservation instinct of physicians. It is to be expected, that in the future the demographic change and the shortage of medical staff will increase the need for AI in healthcare. Subsequently, attention must be directed to the selection of training data and the area of application to prevent discriminatory results of AI. In addition, uncertainty regarding the liability situation poses another challenge to the use of AI. Different opinions hold that in the case of AI errors, either the physicians, the developers or an insurance company is liable. The third option is most likely because the physician

would not want to be liable for the use of a defective product and the developers would not take the risk of developing something for which they could be held liable. Therefore, impartial insurance companies are the best way to overcome this barrier. Moreover, the lack of knowledge of physicians about AI is an additional hurdle. This is mainly connected to the time limitation and the recording capacity of physicians. Nevertheless, physicians should receive extensive training before using AI systems and be made aware of potential risks. The topic of digitization and AI systems should be promoted more for physicians. Another influence on the research question is the quality and quantity of training data. In addition, the selected data must be relevant to the topic being researched. Therefore, people with good knowledge of AI systems are highly required. Otherwise, it could lead to weaknesses in the algorithm. The last major impact on the use of AI is the trust of physicians. To overcome physicians' skepticism and build more trust, it takes valid certifications and high-quality results and studies. More transparency also leads to greater acceptance and understanding for the decision of an AI. Even in the experience of errors, trust is not immediately lost if the AI is comprehensible to physicians.

Subsequently, the statements of the physicians and developers are contrasted and compared with each other. On the subject of data protection, a clear difference can be seen between the opinions of engineers and physicians. From the point of view of developers, data protection is a major hurdle. They also agree that data protection is holding back the development process and research into AI, particularly in Germany. Compared to the U.S., developers see a significant competitive disadvantage and fear that if regulations do not adapt, technology will have to be acquired from abroad in the future. The physicians' side mainly does not see the progress of AI threatened by data protection laws, since only anonymized data sets are used.

On the topic of decision-making authority, physicians and developers also differ in their opinions. On the one hand, the doctors see AI as a purely supportive activity whose task is to simplify the doctor's daily work. In addition, due to its lack of emotional and empathic capacity, AI does not have the potential to consider a treatment decision holistically and to place it in the overall context of the patient. In addition, doctors would never let technology get this far because they do not want to have the feeling of being replaceable. On the other hand, for some physicians it is conceivable that in the future AI will automate certain treatment decisions and the physician will merely act as confirmation of the measure. In individual cases, this could be feasible in the future, especially in the fields of radiology, pathology, and ophthalmology.

The ethical aspects also include discrimination. The majority of doctors agree with the hypothesis that discrimination has a negative impact on the use of AI. But there are also three participants who said that the AI system is only as discriminating as the data set on which the algorithms are based and trained. The developers are determined that AI, on the one hand, neutralizes certain prejudices that humans may have, on the other hand, new forms of discrimination can be generated. In general, discrimination already exists to some extent in the testing of drugs. These are often tested only on men. However, it is known that drugs frequently have a different effect

on women. Discrimination in healthcare is therefore a general problem that should be combated.

Compared to the uncertainty in liability, the most doctors and developers think doctors are responsible when AI makes a wrong decision. The executing physician is the one who makes the decisions and the one who is liable for them. It would make the development of an AI much more difficult and not allow any further development if the one who offers the software is responsible for it. Because the claims for damages are so high that no manufacturer will take over. Therefore, it may be likely in the future that insurance companies will pay for such damage. Compared to the study mentioned in Sect. 10.2.2, the liability issue may also depend on who developed the used algorithm. In this case, hospitals and medical groups are blamed for failures rather than individual physicians.

On the subject of doctors' lack of knowledge and training data, the experts are agreed, regardless of whether they are doctors or developers. The time limitation and the recording capacity of physicians prevent the flow of knowledge about AI. In connection with insufficient training on how to operate these systems, this hampers the use of AI. Moreover, the continuous learning process is very important for an AI system to improve its quality in the long term. The interaction between the physician and the developer plays a major role, because the developers depend on feedback from the physicians so that they can optimize the algorithm. After all, the accuracy of the AI results must still be determined with a specialist in the end, and this determination can minimize undesirable misbehavior. It is also conceivable to develop a feedback system that directly informs the AI system of the error.

Regarding the last hypothesis about trust, the most physician's experts referred to AI as a black box. AI systems cannot be understood at the moment of decision because they are very complex. But, since they are based on algorithms, any decision can be decoded or unlocked in retrospect. Also, it's a support for the physicians, they always have a final look over it anyway. So, the physicians can understand decisions to a certain extent, but not always. If physicians were involved in the development and know how the algorithm is structured, then they can understand relatively well what it has done. In this context, physicians then also have more confidence in AI systems. In contrast, the developers are of the opinion that AI is a black box for the physician. AI is only used as a support and the doctors always have to look over it themselves. That is why it is not so bad if not everything is comprehensible for doctors. There is a field of research called "Explainable AI" that is trying to make AI algorithms a bit more comprehensible to increase the level of trust.

10.5 Limitations and Further Research

As a result of the qualitative study, expert interviews were used to investigate the barriers of AI in the healthcare sector. Five influencing factors were identified with their associated indicators, and possible coping strategies were presented. A conceptual model was created for a further quantitative research. Future research should

quantitatively verify or refute the established hypotheses of this qualitative approach. The study provides added theoretical value by providing initial insights into a very current area of research in medicine.

It must be considered that this study has limitations. It focuses exclusively on the aspects mentioned in the qualitative expert interviews. The sample represents the knowledge of physicians, researchers, and developers. By interviewing a targeted group, the respective barriers in the individual medical fields can be addressed more precisely. Radiology appears to be well suited for this purpose, because AI is most advanced in the field of imaging.

A total of 16 experts were interviewed. This means that the sample size was higher than expected, which is why the validity of this study can be classified as good. Furthermore, the sample was relatively homogeneous. Out of the sixteen experts, four worked in radiology and seven have a background in research and/or development. Since 13 of the sample were male, an interview with female experts could therefore lead to discrepancies or possibly different results. Furthermore, a heterogeneous group should be interviewed to achieve higher validity and representativeness.

Additionally, it seems promising to conduct an international study, e.g., with physicians and AI developers from the U.S., to get a different perspective on the topic. During the interviews, it became obvious that many countries are already more advanced in the area of AI. It would also be interesting to compare the AI challenges or potentials between Germany and other countries.

Since digitalization and the new technologies that come along with it are advancing faster and faster, it can be expected that the topic AI will play an increasingly important role in the healthcare sector in the next years. When using well-functioning AI algorithms, there is the chance to support doctors and medical professionals considerably and to relieve them of time. In that context, the shortage of specialists can possibly be compensated in the future and doctors can focus more on the individual patient.

Acknowledgements We thank Professor Lakhmi C. Jain and all reviewers for supporting our research. Thank you for your valuable and helpful comments.

References

1. Härting, R.-C., et al.: The Potential Value of Digitization for Business: Insights from German-speaking Experts (2017) https://doi.org/10.18420/IN2017_pp.165
2. Dai, D., Boroomand, S.: A review of artificial intelligence to enhance the security of Big Data systems: state-of-art, methodologies, applications, and challenges. Arch. Computat. Methods Eng. **29**, 1291–1309 (2022). https://doi.org/10.1007/s11831-021-09628-0
3. Gjellebaek, C., et al.: Management challenges for future digitalization of healthcare services (2020). https://doi.org/10.1016/j.futures.2020.102636
4. Helm, J.M., et al.: Machine learning and artificial intelligence: definitions. Appl., Future Dir. (2020). https://doi.org/10.1007/s12178-020-09600-8

5. Federal Ministry for Economic Affairs and Climate Action: Digital Summit Event: Strengthening AI and Trust in Digital Technologies (2021). https://www.bmwi.de/Redaktion/EN/Pressemitteilungen/2021/05/20210518-digital-summit-event-strengthening-aI-and-trust-in-digital-technologies.html (accessed on 04/03/2022).
6. Mintz, Y., Brodie, R.: Introduction to artificial intelligence in medicine. Minim. Invasive Ther. Allied Technol. **28**, 73–81 (2019). https://doi.org/10.1080/13645706.2019.1575882
7. Siau, K., Wang, W.: Artificial intelligence (AI) ethics: ethics of AI and ethical AI. J. Database Manag. 14 (2020). https://doi.org/10.4018/JDM.2020040105
8. Saxena, S.: Evolving uncertainty in healthcare service interactions during COVID-19: artificial intelligence—a threat or support to value cocreation? Cyber-Phys. Syst. (2022). https://doi.org/10.1016/B978-0-12-824557-6.00014-5
9. Ludvigsen, K.R., Nagaraja, S.: Dissecting liabilities in adversarial surgical robot failures: a national (Danish) and EU law perspective. Comput. Law Secur. Rev. **44** (2022). https://doi.org/10.1016/j.clsr.2022.105656
10. Core: http://portal.core.edu.au/jnl-ranks/. Accessed on 02/07/2022
11. VHBonline (2022) https://vhbonline.org/vhb4you/vhb-jourqual/vhb-jourqual-3/gesamtliste. Accessed on 02/07/2022
12. Quinn, T.P., et al.: Trust and medical AI: the challenges we face, and the expertise needed to overcome them. J. Am. Med. Inform. Assoc. **28**(4), 890–894 (2021). https://doi.org/10.1093/jamia/ocaa268
13. Cheng, L., et al.: Socially responsible AI algorithms: issues, purposes, and challenges (2022). http://arxiv.org/abs/2101.02032
14. Kokciyan, N., et al.: Sociotechnical perspectives on AI ethics and accountability, IEEE Internet Comput. **25**(06), 5–6 (2021). https://doi.org/10.1109/MIC.2021.3117611
15. Khullar, D., et al.: Public versus physician views of liability for artificial intelligence in health care. J. Am. Med. Inform. Assoc. **28**(7), 1574–1577 (2021). https://doi.org/10.1093/jamia/ocab055
16. Riaño, D., et al.: Artificial Intelligence in Medicine: 17th Conference on Artificial Intelligence in Medicine, AIME 2019, Poznan, Poland, Proceedings, vol. 11526. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-21642-9
17. Straw, I.: The automation of bias in medical artificial intelligence: de-coding the past to create a better future. Artif. Intell. Med. (2020). https://www.sciencedirect.com/science/article/pii/S0933365720312306
18. Norori, N., et al.: Addressing bias in big data and AI for health care: a call for open science. J. Patterns Cell Press. 1–9 (2021). https://doi.org/10.1016/j.patter.2021.100347
19. Boddington, P., et al.: Minds and machines special issue: ethics and artificial intelligence. Minds Mach. **27**(4), 569–574 (2017). https://doi.org/10.1007/s11023-017-9449-y
20. Lokhorst, G.-J.C.: Computational meta-ethics. Minds Mach. (2011). https://doi.org/10.1007/s11023-011-9229-z