

## Secondary Publication



Möllmann, Nicholas R.J.; Mirbabaie, Milad; Stieglitz, Stefan

### Is it alright to use artificial intelligence in digital health? : A systematic literature review on ethical considerations

Date of secondary publication: 02.10.2025

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-110614x

#### Primary publication

Möllmann, Nicholas R.J.; Mirbabaie, Milad; Stieglitz, Stefan (2021): Is it alright to use artificial intelligence in digital health? : A systematic literature review on ethical considerations, in: Health Informatics Journal, London [u.a.]: Sage, Vol. 27, Nr. 4, pp. 1–17, doi: 10.1177/14604582211052391.

#### Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by/4.0/legalcode>



# Is it alright to use artificial intelligence in digital health? A systematic literature review on ethical considerations

Health Informatics Journal  
2021, Vol. 27(4) 1–17  
© The Author(s) 2021  
Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/14604582211052391  
[journals.sagepub.com/home/jhi](https://journals.sagepub.com/home/jhi)



**Nicholas RJ Möllmann** 

Research Group Digital Communication and Transformation, University of Duisburg-Essen, Duisburg, Germany

**Milad Mirbabaie** 

Faculty of Business Administration and Economics, Paderborn University, Paderborn, Germany

**Stefan Stieglitz** 

Research Group Digital Communication and Transformation, University of Duisburg-Essen, Duisburg, Germany

## Abstract

The application of artificial intelligence (AI) not only yields in advantages for healthcare but raises several ethical questions. Extant research on ethical considerations of AI in digital health is quite sparse and a holistic overview is lacking. A systematic literature review searching across 853 peer-reviewed journals and conferences yielded in 50 relevant articles categorized in five major ethical principles: beneficence, non-maleficence, autonomy, justice, and explicability. The ethical landscape of AI in digital health is portrayed including a snapshot guiding future development. The status quo highlights potential areas with little empirical but required research. Less explored areas with remaining ethical questions are validated and guide scholars' efforts by outlining an overview of addressed ethical principles and intensity of studies including correlations. Practitioners understand novel questions AI raises eventually leading to properly regulated implementations and further comprehend that society is on its way from supporting technologies to autonomous decision-making systems.

## Keywords

artificial intelligence, ethics, digital health, healthcare, systematic literature review

---

### Corresponding author:

Nicholas RJ Möllmann, Research Group Digital Communication and Transformation, University of Duisburg-Essen, Forsthausweg 2, Duisburg 47057, Germany,

Email: [nicholas.moellmann@uni-due.de](mailto:nicholas.moellmann@uni-due.de)



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

## Introduction

“A robot may not injure a human being or, through inaction, allow a human being to come to harm.” The first of the Three Laws of Robotics expressed in 1942 by the science fiction author Isaac Asimov<sup>1</sup> is certainly known by many of us from various cinema and television films about robots and their artificial intelligent behavior. Despite the age of these laws, their relevance seems more present than ever if we consider current developments in the area of artificial intelligence (AI) in digital health as systems are used in various medical disciplines.

There is an ongoing excitement towards AI,<sup>2</sup> and both researchers and practitioners equally focus on related concepts. Although early studies can be traced back to the 1950s,<sup>3</sup> there is still no widely accepted definition of the term AI.<sup>4</sup> AI can generally be seen as “*the ability of a machine to perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting with the environment, problem solving, decision-making, and even demonstrating creativity.*”<sup>5</sup> The application of AI within healthcare yields numerous benefits for hospitals and patients. For instance, AI is being used for assisting medical specialists within the diagnostic process and deriving proper treatment strategies<sup>6–8</sup> as well as managing health-related information provided by patients during a self-anamnesis.<sup>9</sup>

The major aim of AI in digital health lies in generating advantages for involved individuals and organizations. AI is capable of collecting, processing and analyzing health-related information more quickly<sup>10,11</sup> and provide decision support for multiple stakeholders within healthcare, that is, patients, doctors, or hospitals.<sup>12</sup> However, the recent AI developments and applications in digital health come along with certain issues. Above all, recent literature poses various ethical questions which have not been considered sufficiently, especially how to prevent inequitable usage<sup>13,14</sup> or provide fair information practices.<sup>15</sup> There are no unified legal regulations, even though multiple governments are involved.<sup>16–19</sup> The rapid development of AI<sup>4</sup> hinders to regulate new advancement in a timely manner and further, there is currently no consensus on what an AI is allowed to decide, nor is there social acceptance for AI decisions.

Among others, AI poses uncertainties regarding bias, discrimination, autonomies, control, opacity, rational concerns, or dystopias.<sup>20,21</sup> Moreover, it is rather indistinct when AI should be used as decision-support system that assists medical specialists with specific tasks, for example, within the diagnostic process.<sup>7,8</sup> An advancement could be the application of AI as decision-making tool which takes over responsibilities of individuals and derives conclusions on its own. In this case, it has not yet been defined who actually is responsible for decisions in the event of an error. If, for example, a patient dies in a hospital due to an AI-based decision or recommendation, who is then liable for any compensation payments? Or, to reverse the example, what happens if a clinician does not follow a recommendation of an AI which leads to death?

The examples provide evidence that ethical concerns and the application of AI in digital health are closely linked to each other. Research on ethical considerations is no longer a mere part of science fiction, but a real-world concern affecting healthcare as AI is already used in digital health.<sup>22,23</sup> Medical practitioners are still at the beginning of understanding ethical implications related to AI as well as the extent to which systems can be applied to generate improvements within digital health, always keeping the overall outcome for patients in mind. We argue that this research is of great interest to scholars and practitioners alike as its relevance will increase even further. There is thus an urgent demand to provide an overview of ethical research approaches regarding AI in digital health to understand the intricacy of this emerging topic.<sup>24</sup> To address these pressing issues, our study is guided by the following research question:

**RQ:** Which ethical considerations on artificial intelligence in digital health are currently being studied, and what are the pertinent future research directions?

A systematic literature review (SLR) suits as starting point to provide an overview on ethical considerations in the context of AI within digital health and identifies which aspects have not yet been adequately covered in existing literature. We therefore explicitly focused on research articles that deal with ethical questions regarding the application of AI for digital health purposes. Our objective is to portray the ethical landscape of AI in digital health and provide a snapshot to guide future development. We present a review of the status quo based on the existing literature and scientific findings that highlight potential research areas where little empirical research has been carried out.

This paper structures our understanding of ethical challenges that have been researched in the context of AI as component of digital health. Researchers will find the overview of ethical considerations beneficial for validating less explored areas in which ethical issues still remain unanswered. To guide scholars' efforts, we outline an overview of which ethical principles are addressed by recent research and highlight the intensity of the studies reviewed including their correlated areas. Practitioners will be able to assess which ethical aspects are relevant when applying AI and what novel questions AI, unlike other technologies, raises.

### *Related work*

Research on ethical considerations and related consequences is primarily concerned with three types<sup>25</sup>: First, the moral development of future physicians throughout their medical training.<sup>26,27</sup> Second, the evaluation of ethical values of medical students and differentiations in their professions.<sup>28,29</sup> Third, the evaluation of certain ethical principles.<sup>26,30</sup> However, ethical considerations are becoming more eminent and profound in digital health due to increasing capabilities and application of AI.<sup>31,32</sup>

Martinez-Martin and Kreitmair<sup>33</sup> investigated ethical issues for digital psychotherapy apps and discovered that lack of adequate regulation, especially regarding safety and privacy, is a major concern. They discuss that appropriate regulatory oversight is needed for the safety and protection of data. Research is also concerned with AI as potential blackbox since the results of underlying algorithms, that is, neural networks or deep learning approaches, are not always comprehensible.<sup>34</sup> Park et al.<sup>35</sup> examined transparency regarding data, test data, and results of AI algorithms. The results show that clarifying any bias or errors in the data prior the application is crucial. The authors argue that transparency is a key ethical issue for AI in healthcare. This goes in line with a recent study of Reddy et al.<sup>36</sup> who developed a governance model for the application of AI in healthcare. The study identified transparency of AI for decision making as one of four essential elements connected to ethical considerations. Besides transparency, fairness, trustworthiness, and accountability are additional components that need to be considered before integrating AI in digital health.

Alongside the model of Reddy et al., other frameworks have emerged in recent years to categorize ethical concerns. For example, Recht et al.<sup>37</sup> developed the FAT (fairness, accountability, and transparency) paradigm for the integration of AI into clinical practice. They argue that most ethical concerns with AI in healthcare can be classified in one of the three groups. Moreover, countries and organizations have also realized that ethical frameworks must be created due to the increasing spread of AI. The European Union established a high-level expert group on AI in 2019 and developed the "European Ethics Guidelines for Trustworthy AI."<sup>16</sup> Trustworthy AI should follow three guidelines: (1) lawful, complying laws and regulation, (2) ethical, obeying ethical

principles and values, and (3) robust, not causing any harm. Furthermore, the guideline defines various key requirements, that is, technical robustness and safety or societal and environmental well-being. Besides the European Union, the World Health Organization is currently working on reviewing existing ethics principles to derive recommendation on how to assure ethical use of AI,<sup>38,39</sup> and even the Australian government seeks to develop an ethics framework.<sup>17</sup>

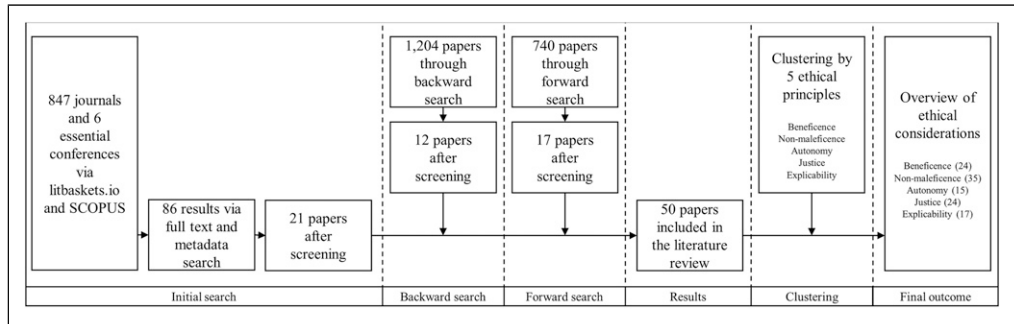
Despite the growing numbers of frameworks, guidelines, principles, or paradigms, there is no uniform ethical strategy to apply AI in digital health.<sup>40</sup> However, existing frameworks frequently cover similar circumstances. We therefore follow a more holistic approach and use the framework of good AI society as proposed by Floridi et al.<sup>41</sup> as theoretical lens. The authors created a “*synthesis of existing sets of principles produced by various reputable, multi-stakeholder organizations and initiatives.*”<sup>41</sup> They assessed various documents, for example, the asilomar AI principles,<sup>42</sup> the Montreal declaration for responsible<sup>43</sup> AI or the five overarching principles for an AI code,<sup>44</sup> which yielded in 47 principles. Floridi et al. identified a high degree of overlaps compared to the four core principles of bioethics,<sup>45</sup> that is, *beneficence, non-maleficence, autonomy, and justice*. Combined, they form the core of modern bioethics discussion,<sup>46</sup> determining an optimal course of action in healthcare but fit ethical challenges posed by AI.<sup>41</sup> *Beneficence* generally describes that healthcare specialists should act in a way that the patient benefits from it. Thus, AI needs to foster the well-being of living beings.<sup>41,43</sup> *Non-maleficence* aims at avoiding any harm, for example, to patients or physicians. In the context of AI, negative consequences must be avoided, for example, misuse regarding privacy or security.<sup>41</sup> *Autonomy* respects the capabilities of individuals to make independent decisions. AI therefore needs to be designed to endorse this demand.<sup>41</sup> *Justice* describes that every patient should be treated similar. AI can thus be applied to correct previous non-justice decisions, that is, unfair discrimination.<sup>41</sup> Despite significant overlaps and suitability for AI, the four core principles are not exhaustive.<sup>41</sup> The authors thus added a fifth principle *explicability* incorporating intelligibility and accountability which defines the “*need to understand and hold to account the decision-making processes of AI.*”<sup>41</sup>

## Research approach

For providing an overview on ethical considerations, that is, which attempts scholars have been undertaken, we executed as SLR to identify relevant articles and examine them according to the ethical framework for AI. Since we were interested in providing a holistic picture, we decided to start our literature search within the information systems (IS) discipline. We argue that IS, as interdisciplinary community addressing dynamic and evolving sociotechnical challenges, must promote further research guiding medical specialists with ethical considerations on AI in digital health. Our SLR is depicted in [Figure 1](#).

Generally, an SLR is a structured approach to determine recent knowledge about a specific topic, which benefits from the disclosure to potential theoretical aspects.<sup>47</sup> SLRs assist scholars in highlighting their contribution to theory and practice by searching for relevant literature within scientific databases using dedicated keywords.<sup>47</sup> Our research can be classified as descriptive approach since we examined existing knowledge describing the current circumstances based on scientific facts and further transferring it to an emerging topic.<sup>48</sup> Leaning on Bem, who suggested a “*coherent conceptual structuring.*”<sup>49</sup> we analyzed the literature according to the five categories (beneficence, non-maleficence, autonomy, justice, and explicability) within the ethical framework for AI. This classification assisted us to interpret the results in a more holistic way.

Following the literature review approach provided by vom Brocke et al.,<sup>50,51</sup> we defined our search scope relying on the taxonomy of literature reviews.<sup>52</sup> The literature search was performed



**Figure 1.** SLR procedure.

using litbaskets.io, an IT artefact developed to assist scholars in retrieving relevant literature from the most recognized databases within the IS discipline.<sup>53</sup> Litbaskets.io converts the given search string to be used by Scopus's advanced search via ISSN numbers of selected outlets, thus providing the possibility to search across indexed scientific sources.<sup>53</sup> Scholars can further apply different filters and restrictions, that is, selecting specific outlets or limiting by a year, to narrow down the search results based on certain criteria.<sup>53</sup>

Since we aimed at obtaining a comprehensive overview of the current research on ethics in the context of AI and digital health, we deliberately selected the largest filter containing 847 outlets (for a full list of the included outlets, please see <https://www.litbaskets.io/>). Furthermore, we did not limit our search starting with a specific year. Since litbaskets.io is focusing on journals but does not provide the option to include conferences, we extended the literature search by selecting the elementary IS conferences manually via Scopus. Our SLR acknowledged peer-reviewed articles only. Less relevant sources, for example, editorials or commentaries, were excluded. We performed full text and metadata search to avoid the possibility of overlooking relevant publications as metadata of articles is not equally indexed by Scopus. The final query used for the SLR looks as follows:

*(AI or "artificial intelligence" or "chatbot\*" or "chat-bot\*" or "conversational agent\*" or "digital assistant\*" or "virtual assistant\*" or "personal assistant\*" or "virtual agent\*" or "ai-based system\*") AND (health or healthcare or "digital health" or "hospital\*" or medicine or medical) AND ("ethic\*" or "moral\*").*

For obtaining thorough results and providing an overview on the current state of research, we carried out several exemplary literature searches prior the execution of our SLR. Our final search string covers three major areas: First, since there are various systems that are using AI and frequently examined in current research, we linked the most common ones, that is, "chatbot," "chat-bot," "conversational agent," "digital assistant," "virtual assistant," "personal assistant," "virtual agent," and "ai-based system." Second, we used various terms for the topic area of digital health, that is, "health," "healthcare," "hospital," "medicine," and "medical." Third, we applied the terms "ethic" and "moral" to cover ethical and moral issue in the context of our research which are used interchangeably.<sup>54</sup> Parentheses nest clauses or group phrases within the search string. Separate statements are linked by Boolean expressions, quotation marks specify terms which must appear next to each other, and wildcards (e.g., "\*") cover different spellings or the plural of expressions.

Once the initial search was conducted, we carefully read the title, abstract and keywords of each publication to validate its relevance according to our research question. Since we wanted to provide an overview on the current state of research, we focused on articles which are largely concerned with ethical aspects. We excluded articles, for example, if they only dealt with technical implementations of AI within the realm of healthcare. Next, we categorized the articles according to the five principles on ethics within the theoretical framework. Since a basic search is not able to provide a comprehensive review, and we were also interested in retrieving relevant literature outside the IS discipline, we further conducted a backward search. We therefore collected every reference of the bibliographies of each identified article. Comparable to the initial search, we included references to other conference and journal publications and excluded non-scientific sources, such as web pages or business reports. To identify relevant literature according to our research question, we followed the equal procedure as performed within the initial search: Reading the title, keywords and abstract, and categorizing the articles according to the theoretical framework. The final step of our SLR was the forward search. To identify even further literature, we considered every paper, which was found during the initial and backward search, that has been cited by other research after its initial publication. Yet again, we read title, abstract, and keywords and performed the categorization process as outlined above.

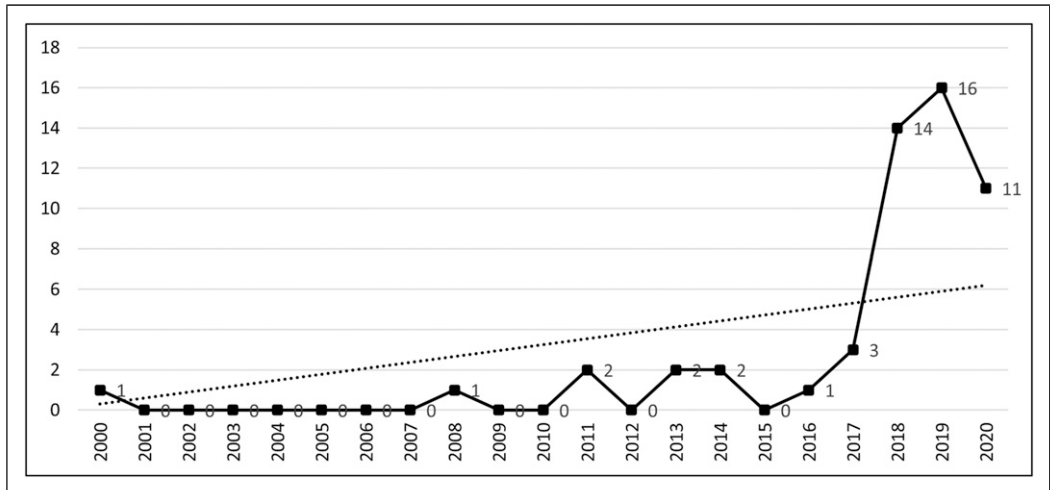
## Results and analysis

In total, we screened 2,106 articles of which 50 are relevant with respect to our research question. The execution of the SLR resulted in 21 relevant results within the initial search (of 86), 12 via backward search (of 1,204) and, finally, 17 relevant publications were identified through the forward search (of 2,106). We analyzed the publications according to their distribution of year, their outlet and, as outlined above, their distribution by their ethical principle. We thereby provide an overview on ethical concerns related to the application of AI in digital health. We further show in which area extensive research has been undertaken and where little or less research was conducted yet but seems promising for future studies.

### *Distribution of articles by year and outlet*

No articles were found by our SLR before the year 2000. It is noteworthy that there are only few or even no publications up until 2016. But from 2018 onward, the number of publications has risen considerably with 14 in 2018, 16 in 2019, and already 11 in this year. If we consider 2020, the work at hand presents only the first 9 months. It can be assumed that the total number of publications might be similar to the 2019, if not greater. However, the undeniable trend of ethical concerns in the context of AI within digital health, represented by the rapid increase in publications over the past 3 years, symbolizes the salience and legitimacy of this research area. [Figure 2](#) provides the distribution by years.

From the total of 50 articles, 44 have been published in journals (88%) and only 6 (12%) in conference proceedings. Almost all articles were published in different outlets. Only three outlets published more than one article: Journal of Medical Internet Research (6;12%), Bulletin of the World Health Organization (2;4%) and Proceedings of the AAMAS Workshop on Intelligent Conversation Agents in Home and Geriatric Care Applications (2;4%). This finding is surely related to the rapid increase of publications in recent years and indicates that ethical aspects, in the context of AI in digital health, are addressed from various research streams.



**Figure 2.** Total number of articles per year.

### *Distribution of articles by ethical principle*

The classification, with respect to the ethical framework for AI and its five principles as, is outlined in [Table 1](#). Overall, non-maleficence (avoiding harms) clearly stands out as the most heavily researched ethical principle with 35 articles (70%). Beneficence (creating benefits), as well as justice (avoid discrimination), follow with each of 24 (48%) publications. The ethical principles of explicability (17;34%) (understanding the results), and autonomy (15;30%) (fostering autonomous decisions), have undergone least research. This illustration aids in validating whether a research area has received much attention and if some areas are highly correlated. However, this is just an overview of the status quo. Areas with less research do not indicate that further studies are inevitably required.

The papers as retrieved by our literature search deal with a variety of different approaches and diverse ethical viewpoints in the context of AI in digital health. A majority of the research tries to develop models, frameworks, guidelines, or principles<sup>31,36,41</sup> to use them for various purposes in healthcare or suggest that they shall be used in practice but without providing clinical evidence or proof of their validity. Occasionally, new categories or names for ethical constructs are created, for example, “trustworthiness,”<sup>36,59</sup> but without really differing from the four core principles commonly used in bioethics (or the additional fifth category that we also used to categorize our results). Furthermore, ethical aspects in the context of AI in healthcare are being studied in a wide range of research areas, for example, technical implementations,<sup>96</sup> the use of robotic systems in operating theaters,<sup>95</sup> care robots for the elderly<sup>93</sup> or in the area of caregivers and geriatric patients.<sup>69</sup>

## **Discussion**

Recent development in the field of AI in digital health come along with certain ethical issue. There are challenges regarding bias, discrimination, opacity, rational concerns, or dystopias.<sup>20</sup> In addition, there are various questions that try to answer in what way AI is allowed to interact or intervene or if systems should be decision-supporting or decision-making. Even though AI is frequently being

**Table I.** Relevant articles found in the SLR assigned to the five theoretical foundations.

Article	(1)	(2)	(3)	(4)	(5)
Alami et al <sup>55</sup>	—	x	—	x	x
Alexiou et al <sup>56</sup>	x	—	—	—	—
Alexiou et al <sup>57</sup>	x	—	—	—	—
Avila-Tomas et al <sup>58</sup>	x	—	—	—	—
Bærøe et al <sup>59</sup>	x	x	—	x	x
Bauer et al <sup>60</sup>	x	—	—	—	—
Bezemer et al <sup>61</sup>	x	—	x	—	—
Bjerring and Busch <sup>62</sup>	—	x	—	—	x
Blease et al <sup>63</sup>	—	—	x	—	—
Braun et al <sup>64</sup>	—	x	—	x	x
Cresswell et al <sup>65</sup>	—	—	—	x	x
Dalton-Brown <sup>66</sup>	—	x	—	x	x
Eichenberg et al <sup>67</sup>	—	—	—	x	—
Ellahham et al <sup>68</sup>	—	x	—	x	—
Eschweiler and Wanner <sup>69</sup>	—	x	—	—	—
Fiske et al <sup>70</sup>	—	x	—	—	x
Floridi et al <sup>41</sup>	x	x	x	x	x
Furey and Blue <sup>71</sup>	—	x	—	—	—
Gamble <sup>72</sup>	—	x	x	—	—
Garner et al <sup>73</sup>	x	x	—	—	—
Geis et al <sup>74</sup>	x	x	x	x	—
Gill <sup>75</sup>	x	—	—	—	—
Goldhahn et al <sup>76</sup>	x	x	—	x	—
Graham et al <sup>77</sup>	x	—	x	x	—
Horgan et al <sup>78</sup>	x	x	x	—	—
Ienca et al <sup>79</sup>	x	x	x	x	—
Kostopoulos <sup>80</sup>	x	x	—	x	—
Kovach et al <sup>81</sup>	—	x	—	—	x
Kretzschmar et al <sup>82</sup>	x	x	—	—	—
Lai et al <sup>83</sup>	x	—	—	—	—
Li <sup>84</sup>	—	x	x	—	—
Ludwin and Murray <sup>85</sup>	—	x	—	—	—
Lupton <sup>86</sup>	x	—	—	x	—
Luxton <sup>22</sup>	—	x	—	—	—
Luxton <sup>87</sup>	x	x	x	x	—
Luxton <sup>88</sup>	—	—	—	—	x
Luxton <sup>38</sup>	x	x	—	x	—
Martinez-Martin et al <sup>89</sup>	—	x	—	—	—
Martinez-Martin and Kreitmair <sup>33</sup>	—	x	x	—	x
Milosevic <sup>31</sup>	x	x	x	x	x
Ngiam and Khor <sup>90</sup>	—	x	—	x	x
Park et al <sup>35</sup>	x	x	—	x	x
Pontier and Widdershoven <sup>91</sup>	—	—	x	—	—
Pouloudi and Magoulas <sup>92</sup>	x	x	—	—	—

(continued)

**Table I.** (continued)

Article	(1)	(2)	(3)	(4)	(5)
Poulsen and Kreps <sup>93</sup>	—	X	x	—	—
Recht et al <sup>37</sup>	—	X	—	x	x
Reddy et al <sup>36</sup>	—	X	—	x	—
Schönberger <sup>94</sup>	—	X	—	x	x
Steil et al <sup>95</sup>	x	—	—	x	—
Vayena et al <sup>96</sup>	—	X	x	x	x
Total	24 (48%)	35 (70%)	15 (30%)	24 (48%)	17 (34%)

(1) Beneficence; (2) Non-maleficence; (3) Autonomy; (4) Justice; and (5) Explicability.

used in digital health,<sup>24,97</sup> ethical considerations have only recently started to move into researchers' attention.

Our results reveal that there has been an enormous increase in publications from 2018 onwards. This could be related to the fact that AI is increasingly being applied within digital health. Recent research provides evidence that AI assists physicians in the diagnostic process,<sup>98,99</sup> improves collaboration among healthcare workers<sup>100–102</sup> and further aids patients in their course of disease.<sup>103,104</sup> Besides, technology which is needed for implementing AI-based systems improved and simultaneously became more affordable.<sup>105</sup> However, the development of a supporting technology is usually in the foreground, ethical considerations are frequently considered after systems are already in use. This may be the reason why ethics research on AI in digital health only recently begun although systems are already used for various productive purposes.

The three ethical principles that have received the most research yet, that is, non-maleficence, beneficence, and justice, are somewhat most likely to be compared to the Three Laws of Robotics.<sup>1</sup> If we would need to jointly describe these categories, they basically suggest that AI should yield in positive outcomes for individuals involved. We thus argue that the increased research attempts seem obvious. However, autonomy and explicability draw a different picture. AI is more likely to threaten the physicians' independence compared to other technologies. Currently, AI assists healthcare specialists with certain tasks, for example, within the anamnesis.<sup>9</sup> However, the capabilities of AI continue to rise. Systems are already capable of gathering information quicker,<sup>11</sup> are able to better interpret complex interdependencies<sup>106</sup> and derive assumptions without any subjective bias, for example, based on missing experience, pressure, or simply tiredness. There is an increasing risk that the results will resemble or even be superior to those of medical experts. Therefore, autonomy is surely restricted when systems gain consecutive advantages over a human being. As we are on the way from a supporting technology to autonomous decision-making, research in the field of autonomy was not yet very widespread but is becoming vital. This is also supported by the results of our SLR which shows that almost all articles have been published in the last 3 years (12 out of 15). Furthermore, AI-based decisions tend to become in-transparent for individuals. Existing programs in digital health are developed to follow certain logics or processes. Similarly, AI-based systems build upon basic machine learning algorithms,<sup>32</sup> using statistical calculations. However, current scientific developments indicate greater results when using neural networks or deep learning approaches.<sup>107,108</sup> These much more complex methods are, however, a black box to humans.<sup>34</sup> It is usually possible to assess which data AI is processing, but the outcomes are not always comprehensible to humans. Thus, there are AI approaches making it impossible to transfer insights to other circumstances due to the potential unexplainability of algorithmic outcomes.<sup>55</sup> Furthermore,

AI algorithms are trained using an exemplary dataset which contains certain features (i.e., patient characteristics).<sup>23</sup> Incomplete or incorrect data pose a high risk of generating incorrect results which may even no longer be recognizable due to the complexity of underlying algorithms. And, the data which is used to train an algorithm, is currently still classified by humans and somewhat might contain subjective bias. Neural networks or deep learning approaches are increasingly being used since they offer greater benefits compared to conventional methods. On the one hand, we argue that the increasing studies on explicability is related to the fact that neural networks and deep learning approaches are becoming more common.<sup>109</sup> On the other hand, the resulting black box explains why papers explicitly deal with this problem and the extension of the four core principles of bioethics is reasonable. There are thus several ethical considerations in the area of AI in digital health that need to be addressed by future research endeavors. Table 2 outlines exemplary research questions within each ethical principle guiding scholars' efforts.

While we provide an overview of current ethical research on AI in digital health, this research is not free of limitations. We linked the most common notations for systems using AI in our search term, that is, conversational agent, digital assistant, and virtual assistant. However, we did not use technical terms such as machine learning, neural network, or deep learning. Even though this helped us to focus on relevant literature and excluded technical implementations, we may have missed the retrieval of studies dealing with ethical consideration of AI in digital health. Furthermore, there are many different types and applications of AI in digital health, for example, decision-support

**Table 2.** Exemplary future research questions for each ethical principle.

Ethical principle	Exemplary research questions
Beneficence	<ul style="list-style-type: none"> <li>•What medical data/patient information should be used by AI in digital health to achieve the best possible outcomes?</li> <li>•How can physicians' expertise be combined with the capabilities of AI-based systems?</li> <li>•To what extent can medical knowledge be enriched when using AI in digital health?</li> <li>•How can any subjective bias (e.g., individuals belonging to a certain minority) be identified and prevented?</li> </ul>
Non-maleficence	<ul style="list-style-type: none"> <li>•Which are preventable impairments caused by using AI in digital health?</li> <li>•To what extent are reprehensible decisions/recommendations recognizable?</li> <li>•What might be a potential control mechanism for AI applied in digital health?</li> </ul>
Autonomy	<ul style="list-style-type: none"> <li>•To what extent feel healthcare professionals threatened by AI?</li> <li>•What are the differences between AI and other technologies potentially harming healthcare professional's autonomy?</li> <li>•What factors contribute to the perception of losing the individual autonomy?</li> <li>•Who should be accountable for algorithm-based decisions derived by AI?</li> </ul>
Justice	<ul style="list-style-type: none"> <li>•How can medical experts be psychologically supported when AI derives morally reprehensible decisions?</li> <li>•To what extent can AI be applied to improve the doctor-patient relationships in digital and/or virtual surroundings?</li> <li>•Which guidelines should be followed to develop AI systems when being applied in digital health?</li> </ul>
Explicability	<ul style="list-style-type: none"> <li>•To what extent can certain algorithms be applied to provide explainable and understandable results for healthcare professionals?</li> <li>•How can in-transparent results derived by AI be presented in a comprehensible manner?</li> <li>•What data is particularly suitable for the development of AI systems generating understandable decisions?</li> </ul>

system<sup>7,8</sup> or virtual assistants.<sup>114</sup> Thus, ethical considerations might alter and need to be addressed separately.<sup>69</sup>

Besides the exemplary research questions, future scientific endeavors should provide a more granular picture of ethical considerations considering the specific type of system and area of application. Differences might arise between decision-support or decision-making tools depending on whether a healthcare specialist must approve a decision made by an AI. Distinctions might also be needed depending on the type of interaction. For example, systems can be addressed via a text- or speech-based input.<sup>110–112</sup> Furthermore, there are also variations as to who or where an AI supports. Systems today assist in specific health-related processes like diagnostics including the anamnesis,<sup>9,113</sup> helping patients to cope with their disease<sup>103,104</sup> or improve collaboration among healthcare employees.<sup>100–102,112</sup> Moreover, researchers should address issues related to retraceability, explainability, and causability contributing to the trustworthiness of AI.<sup>116</sup> Long-term and thorough trust in AI is only achievable if generated outcomes are retraceable, hence replicable in case an incident happens.

## Conclusions

In this article, we illustrated ethical consideration of AI in digital health in current academic research. We presented our findings from the SLR, in which we classified the relevant literature according to the five principles on ethics: beneficence, non-maleficence, autonomy, justice, and explicability. The results are helpful for practitioners as well as for scholars from a variety of research disciplines. The main theoretical contribution of this paper lies in providing an overview of the status quo covering ethical considerations on AI in digital health. The review is beneficial for validating less explored areas in which ethical issues still remain less researched or even unexplored. By highlighting the intensity of studies, we seek to assist researchers' effort for future studies in the field of AI in digital health. On a practical level, practitioners understand what ethical aspects are relevant when using AI for health-related purposes, which might result in properly regulated implementations. The application of AI is beneficial for digital health, but a few things must be considered in advance, that is, how transportable results are achievable, how far AI is allowed to be integrated in processes, and to what extent humans can be held accountable.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen. Please see to the open access contract here: [https://www.uni-due.de/ub/en/ea\\_verlage.php](https://www.uni-due.de/ub/en/ea_verlage.php)

## ORCID iDs

Nicholas RJ Möllmann  <https://orcid.org/0000-0002-9075-4433>

Milad Mirbabaie  <https://orcid.org/0000-0002-9455-5773>

Stefan Stieglitz  <https://orcid.org/0000-0002-4366-1840>

## References

1. Asimov I. *'Runaround' I, Robot*. New York City: Doubleday, 1950.
2. Selz D. From electronic markets to data driven insights. *Electron Mark* 2020; 30: 57–59.
3. Moor J. The dartmouth college artificial intelligence conference: the next fifty years. *AI Mag* 2006; 27: 87–91.
4. Duan Y, Edwards JS and Dwivedi YK. Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *Int J Inf Manage* 2019; 48: 63–71.
5. Rai A, Constantinides P and Sarker S. Next-generation digital platforms: toward human-AI hybrids. *MIS Q* 2019; 43: iii–ix.
6. Li H, Wang X, Liu C, et al. Dual-input neural network integrating feature extraction and deep learning for coronary artery disease detection using electrocardiogram and phonocardiogram. *IEEE Access* 2019; 7: 146457–146469.
7. López-Martínez F, Núñez-Valdez ER, Lorduy Gomez J, et al. A neural network approach to predict early neonatal sepsis. *Comput Electr Eng* 2019; 76: 379–388.
8. De Ramón Fernández A, Ruiz Fernández D and Prieto Sánchez MT. A decision support system for predicting the treatment of ectopic pregnancies. *Int J Med Inform* 2019; 129: 198–204.
9. Denecke K, Lutz Hochreutener S, Pöpel A, et al. Talking to Ana: a mobile self-anamnesis application with conversational user interface. In: International Conference on Digital Health, Lyon, France, 2018. Epub ahead of print 2018. DOI: [10.1145/3194658.3194670](https://doi.org/10.1145/3194658.3194670).
10. Dellermann D, Lipusch N, Ebel P, et al. Design principles for a hybrid intelligence decision support system for business model validation. *Electron Mark* 2018; 29: 423–441.
11. Nasirian F, Ahmadian M and Lee O-K. AI-based voice assistant systems: evaluating from the interaction and trust perspectives. In: Twenty-third Americas Conference on Information Systems, Boston, MA, USA, 2017.
12. Guha S and Kumar S. Emergence of big data research in operations management, information systems, and healthcare: past contributions and future roadmap. *Prod Oper Manag* 2018; 27: 1724–1735.
13. Sonja M, Ioana G, Miaoqing Y, et al. *Understanding value in health data ecosystems: a review of current evidence and ways forward*. *Rand Heal Q*; 7.
14. Salerno J, Knoppers BM, Lee LM, et al. Ethics, big data and computing in epidemiology and public health. *Ann Epidemiol* 2017; 27: 297–301.
15. Libaque-Sáenz CF, Wong SF, Chang Y, et al. The effect of fair information practices and data collection methods on privacy-related behaviors: a study of mobile apps. *Inf Manag* 2021; 58(1): 103284. DOI: [10.1016/j.im.2020.103284](https://doi.org/10.1016/j.im.2020.103284).
16. EU. *Ethics guidelines for trustworthy AI*, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (2020).
17. Dawson D, Schleiger E, Horton J, et al. *Artificial Intelligence: Australia's Ethics Framework*. Data61 CSIRO, Australia.
18. O'Sullivan S, Nevejans N, Allen C, et al. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. *Int J Med Robot Comput Assist Surg* 2019; 15: e1968.
19. Schneeberger D, Stöger K and Holzinger A. The European legal framework for medical AI. In: Machine Learning and Knowledge Extraction, pp. 209–226.
20. Ploug T and Holm S. The right to refuse diagnostics and treatment planning by artificial intelligence. *Med Heal Care Philos* 2020; 23: 107–114.
21. Turel O, Matt C, Trenz M, et al. Panel report: the dark side of the digitization of the individual. *Internet Res* 2019; 29: 274–288.

22. Luxton DD. Artificial intelligence in psychological practice: current and future applications and implications. *Prof Psychol Res Pract* 2014; 45: 332–339.
23. Mirbabaie M, Stieglitz S and Frick NRJ. Artificial intelligence in disease diagnostics: a critical review and classification on the current state of research guiding future direction. *Health Technol (Berl)* 2021; 11: 693–731.
24. Mirbabaie M, Hofeditz L, Frick NRJ, et al. Artificial intelligence in hospitals: providing a status quo of ethical considerations in academia to guide future research. *AI Soc* 2021 Epub ahead of print 28 June 2021. DOI: [10.1007/s00146-021-01239-4](https://doi.org/10.1007/s00146-021-01239-4).
25. Page K. The four principles: can they be measured and do they predict ethical decision making? *BMC Med Ethics* 2012; 13: 10. DOI: [10.1186/1472-6939-13-10](https://doi.org/10.1186/1472-6939-13-10).
26. Price J, Price D, Williams G, et al. Changes in medical student attitudes as they progress through a medical course. *J Med Ethics* 1998; 24: 110–117.
27. Bore M, Munro D, Kerridge I, et al. Selection of medical students according to their moral orientation. *Med Educ* 2005; 39: 266–275.
28. Rezler AG, Schwartz RL, Obenshain SS, et al. Assessment of ethical decisions and values. *Med Educ* 1992; 26: 7–16.
29. Rezler AG, Lambert P, Obenshain SS, et al. Professional decisions and ethical values in medical and law students. *Acad Med* 1990; 65: 31–32.
30. Hebert PC, Meslin EM and Dunn EV. Measuring the ethical sensitivity of medical students: a study at the University of Toronto. *J Med Ethics* 1992; 18: 142–147.
31. Milosevic Z. Ethics in digital health: a deontic accountability framework. In: 2019 IEEE 23rd International Enterprise Distributed Object Computing Conference (EDOC), Paris, France. IEEE, pp. 105–111.
32. Knijnenburg B and Willemsen M. Inferring capabilities of intelligent agents from their external traits. *ACM Trans Interact Intell Syst* 2016; 6: 1–25.
33. Martinez-Martin N and Kreitmair K. ethical issues for direct-to-consumer digital psychotherapy apps: addressing accountability, data protection, and consent. *JMIR Ment Heal* 2018; 5: e32. DOI: [10.2196/mental.9423](https://doi.org/10.2196/mental.9423).
34. Knight W. *The dark secret at the heart of AI*. MIT Technol Rev, 2017.
35. Park SH, Kim Y-H, Lee JY, et al. Ethical challenges regarding artificial intelligence in medicine from the perspective of scientific editing and peer review. *Sci Ed* 2019; 6: 91–98.
36. Reddy S, Allan S, Coghlan S, et al. A governance model for the application of AI in health care. *J Am Med Inform Assoc* 2020; 27: 491–497.
37. Recht MP, Dewey M, Dreyer K, et al. Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations. *Eur Radiol* 2020; 30: 3576–3584.
38. Luxton DD. Ethical implications of conversational agents in global public health. *Bull World Health Organ* 2020; 98: 285–287.
39. WHO. *Big data and artificial intelligence*, <https://www.who.int/ethics/topics/big-data-artificial-intelligence/en/> (2020, accessed 19 September 2020).
40. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019; 25: 1337–1340.
41. Floridi L, Cowsls J, Beltrametti M, et al. AI4People—An ethical framework for a good ai society: opportunities, risks, principles, and recommendations. *Minds Mach* 2018; 28: 689–707.
42. Benevolent. *Principles developed in conjunction with the 2017 Asilomar conference*, <https://futureoflife.org/ai-principles> (2017, accessed 19 September 2020).
43. de Montréal U. *The montreal declaration for the responsible development of artificial intelligence*, <https://www.montrealdeclaration-responsibleai.com/> (2020, accessed 19 September 2020).

44. UK Parliament. AI in the UK: ready, willing and able? Select Committee on Artificial Intelligence. Report of Session 2017 -19. 16 April 2018. Authority of the House of Lords.
45. Beauchamp TL and Childress JF. *Principles of Biomedical Ethics*. 7th ed. Oxford: Oxford University Press, 2012.
46. Lawrence DJ. The Four principles of biomedical ethics: a foundation for current bioethical debate. *J Chiropr Humanit* 2007; 14: 34–40.
47. Webster J and Watson R. Analyzing the past to prepare for the future: writing a literature review. *MIS Q* 2002; 26: 13–23.
48. Bear A and Knobe J. Normality: part descriptive, part prescriptive. *Cognition* 2016; 167: 25–37.
49. Bem DJ. Writing a review article for Psychological Bulletin. *Psychol Bull* 1995; 118: 172–177.
50. vom Brocke J, Simons A, Riemer K, et al. Standing on the shoulders of giants: challenges and recommendations of literature search in information systems research. *Commun Assoc Inf Syst* 2015; 37: 205-224. DOI: [10.17705/ICAIS.03709](https://doi.org/10.17705/ICAIS.03709).
51. vom Brocke J, Simons A, Niehaves, et al. Reconstructing the giant: on the importance of rigour in documenting the literature search process. In: European Conference on Information Systems, Verona, Italy, June 2009. DOI: [10.1108/09600031211269721](https://doi.org/10.1108/09600031211269721).
52. Cooper HM. Organizing knowledge syntheses: a taxonomy of literature reviews. *Knowl Soc* 1988; 1: 104–126.
53. Boell SK and Blair W. [www.litbaskets.io](http://www.litbaskets.io)an IT artifact supporting exploratory literature searches. In: Australasian Conference on Information Systems, Perth, Western Australia, Australia, 2019.
54. Jones TM. Ethical decision making by individuals in organizations: an issue-contingent model. *Acad Manag Rev* 1991; 16: 366–395.
55. Alami H, Lehoux P, Auclair Y, et al. Artificial intelligence and health technology assessment: anticipating a new level of complexity. *J Med Internet Res* 2021; 22: e17707. DOI: [10.2196/17707](https://doi.org/10.2196/17707).
56. Alexiou A, Psixa M and Vlamos P. Artificial intelligence applications and innovations. In: IFIP International Federation for Information Processing, Neos Marmaras, Greece. Springer, 2011, pp. 297–302.
57. Alexiou A, Theocharopoulou G and Vlamos P. Ethical issues in neuroinformatics. In: IFIP International Federation for Information Processing, Paphos, Cyprus, 2013, pp. 700–705.
58. Avila-Tomas JF, Olano-Espinosa E, Minué-Lorenzo C, et al. Effectiveness of a chat-bot for the adult population to quit smoking: protocol of a pragmatic clinical trial in primary care (Dejal@). *BMC Med Inform Decis Mak* 2019; 19: 249. DOI: [10.1186/s12911-019-0972-z](https://doi.org/10.1186/s12911-019-0972-z).
59. Bærøe K, Miyata-Sturm A and Henden E. How to achieve trustworthy artificial intelligence for health. *Bull World Health Organ* 2020; 98: 257–262.
60. Bauer M, Glenn T, Monteith S, et al. Ethical perspectives on recommending digital technology for patients with mental illness. *Int J Bipolar Disord* 2017; 5: 6.
61. Bezemer T, de Groot MC, Blasse E, et al. A human(e) factor in clinical decision support systems. *J Med Internet Res* 2019; 21: e1132. DOI: [10.2196/11732](https://doi.org/10.2196/11732).
62. Bjerring JC and Busch J. Artificial intelligence and patient-centered decision-making. *Philos Technol* 2020; 34: 349–371. DOI: [10.1007/s13347-019-00391-6](https://doi.org/10.1007/s13347-019-00391-6).
63. Blease C, Kaptchuk TJ, Bernstein MH, et al. Artificial intelligence and the future of primary care: exploratory qualitative study of UK general practitioners' views. *J Med Internet Res* 2019; 21: e12802. DOI: [10.2196/12802](https://doi.org/10.2196/12802).
64. Braun M, Hummel P, Beck S, et al. Primer on an ethics of AI-based decision support systems in the clinic. *J Med Ethics* Epub ahead of print 2020. DOI: [10.1136/medethics-2019-105860](https://doi.org/10.1136/medethics-2019-105860).
65. Creswell JW and Creswell DJ. *Research design: qualitative, quantitative, and mixed methods*. SAGE Publications, 2018.

66. Dalton-Brown S. The ethics of medical ai and the physician-patient relationship. *Cambridge Q Healthc Ethics* 2020; 29: 115–121.
67. Eichenberg C, Khamis M and Hübner L. The attitudes of therapists and physicians on the use of sex robots in sexual therapy: online survey and interview study. *J Med Internet Res* 2019; 21: e13853. DOI: [10.2196/13853](https://doi.org/10.2196/13853).
68. Ellahham S, Ellahham N and Simsekler MCE. Application of artificial intelligence in the health care safety context: opportunities and challenges. *Am J Med Qual* 2019; 35: 341–348.
69. Eschweiler GW and Wanner L. *How can intelligent conversational agents help? the needs of geriatric patients and their caregivers*, ICAHGCA@AAMAS, 2018.
70. Fiske A, Henningsen P and Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res* 2019; 21: e13216. DOI: [10.2196/13216](https://doi.org/10.2196/13216).
71. Furey E and Blue J. She just doesn't understand me! Curing Alexa of her Alexithymia. In: 26th Irish Conference on Artificial Intelligence and Cognitive Science, Dublin, Ireland, 2018.
72. Gamble A. Artificial intelligence and mobile apps for mental healthcare: a social informatics perspective. *Aslib J Inf Manag* 2020; 72: 509–523. DOI: [10.1108/AJIM-11-2019-0316](https://doi.org/10.1108/AJIM-11-2019-0316).
73. Garner TA, Powell WA and Carr V. Virtual carers for the elderly: A case study review of ethical responsibilities. *Digit Heal* 2016; 2: 1–14.
74. Geis JR, Brady AP, Wu CC, et al. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. *Can Assoc Radiol J* 2019; 70: 329–334.
75. Gill SP. Socio-ethics of interaction with intelligent interactive technologies. *AI Soc* 2008; 22: 283–300.
76. Goldhahn J, Rampton V and Spinaz GA. Could artificial intelligence make doctors obsolete? *BMJ* 2018; 363: 4563. DOI: [10.1136/bmj.k4563](https://doi.org/10.1136/bmj.k4563).
77. Graham S, Depp C, Lee EE, et al. Artificial intelligence for mental health and mental illnesses: an overview. *Curr Psychiatry Rep* 2019; 21: 116.
78. Horgan D, Romao M, Morrè SA, et al. Artificial intelligence: power for civilisation – and for better healthcare. *Public Health Genomics* 2019; 22: 145–161.
79. Ienca M, Wangmo T, Jotterand F, et al. Ethical design of intelligent assistive technologies for dementia: a descriptive review. *Sci Eng Ethics* 2018; 24: 1035–1055.
80. Kostopoulos L. The emerging artificial intelligence wellness landscape: benefits and potential areas of ethical concern. *Calif West Law Rev* 2018; 55: 235–254
81. Kovach KA, Aubrecht JA, Dew MA, et al. Data safety and monitoring for research involving remote health monitoring. *Telemed e-Health* 2011; 17: 574–579.
82. Kretzschmar K, Tyroll H, Pavarini G, et al. Can your phone be your therapist? young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomed Inform Insights* 2019; 11: 1178222619829083. DOI: [10.1177/1178222619829083](https://doi.org/10.1177/1178222619829083).
83. Lai M-C, Brian M and Mamzer M-F. Perceptions of artificial intelligence in healthcare: findings from a qualitative survey study among actors in France. *J Transl Med* 2020; 18: 14.
84. Li H. Impact of artificial intelligence based on big data on medical care. *J Phys Conf Ser* 2020; 1533: 032077. DOI: [10.1088/1742-6596/1533/3/032077](https://doi.org/10.1088/1742-6596/1533/3/032077).
85. Ludwin SK and Murray TJ. Dilemmas in medical ethics in the age of big data. *Mult Scler J* 2017; 23: 1306–1308.
86. Lupton M. Some ethical and legal consequences of the application of artificial intelligence in the field of medicine. *Trends Med* 2020; 18: 2-7. DOI: [10.15761/TiM.1000147](https://doi.org/10.15761/TiM.1000147).
87. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. *Artif Intell Med* 2014; 62: 1-10. DOI: [10.1016/j.artmed.2014.06.004](https://doi.org/10.1016/j.artmed.2014.06.004).

88. Luxton DD. Should Watson be consulted for a second opinion? *AMA J Ethics* 2019; 21: 131–137.
89. Martinez-Martin N, Insel TR, Dagum P, et al. Data mining for health: staking out the ethical territory of digital phenotyping. *Npj Digit Med* 2018; 1: 68.
90. Ngiam KY and Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol* 2019; 20: 262–273.
91. Pontier MA and Widdershoven GAM. Robots that stimulate autonomy. In: 9th IFIP WG 12.5 International Conference, Paphos, Cyprus, 2013, pp. 195–204.
92. Pouloudi A and Magoulas GD. Neural expert systems in medical image interpretation: development, use, and ethical issues. *Neural Expert Syst Med Image Interpret* 2000; 10: 451–472.
93. Poulsen A and Kreps D. This changes everything –ICT and climate change: what can we do? In: IFIP International Federation for Information Processing 2018, Zurich, Switzerland, 2018, pp. 314–328.
94. Schönberger D. Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. *Int J Law Inf Technol* 2019; 27: 171–203.
95. Steil J, Finas D, Beck S, et al. Robotic systems in operating theaters: new forms of team–machine interaction in health care. *Methods Inf Med* 2019; 58: e14–e25.
96. Vayena E, Blasimme A and Cohen IG. Machine learning in medicine: addressing ethical challenges. *PLOS Med* 2018; 15: e1002689. DOI: [10.1371/journal.pmed.1002689](https://doi.org/10.1371/journal.pmed.1002689).
97. Lu J, Song E, Ghoneim A, et al. Machine learning for assisting cervical cancer diagnosis: an ensemble approach. *Futur Gener Comput Syst* 2020; 106: 199–205.
98. Parisi L, RaviChandran N and Manaog ML. Feature-driven machine learning to improve early diagnosis of parKinson’s disease. *Expert Syst Appl* 2018; 110: 182–190.
99. Vasanthselvakumar R, Balasubramanian M and Palanivel S. Detection and classification of kidney disorders using deep learning method. *J Mech Contin Math Sci* 2019; 14: 3045–3051.
100. Georgiadis DT. Dynamic mobile agents for collaboration systems. In: Mediterranean Conference on Information Systems, Naples, Italy, 2011.
101. Just JE, Cornwell MR and Huhns MN. Facilitating Human Collaboration with Agents. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences, Big Island, Hawaii, 2005. Epub ahead of print 2005. DOI: [10.1109/HICSS.2005.270](https://doi.org/10.1109/HICSS.2005.270).
102. Seeber I, Waizenegger L, Seidel S, et al. Collaborating with technology-based autonomous agents: Issues and research opportunities. *Internet Res* 2020; 30: 1–18.
103. Bickmore T, Kimani E, Trinh H, et al. Managing chronic conditions with a smartphone-based conversational virtual agent. In: International Conference on Intelligent Virtual Agents, Sydney, Australia, 2018, pp. 119–124.
104. Shaked N. Avatars and virtual agents – relationship interfaces for the elderly. *Healthc Technol Lett* 2017; 4: 83–87.
105. Spohrer J and Banavar G. Cognition as a service: an industry perspective. *AI Mag*; 36: 71–86.
106. Dellermann D, Ebel P, Söllner M, et al. Hybrid Intelligence. *Bus Inf Syst Eng* 2019; 61: 637–643.
107. Goodfellow I, Bengio Y and Courville A. *Deep Learning*. 1st ed. Cambridge, MA: MIT Press, 2016.
108. Fogassi L, Ferrari PF, Gesierich B, et al. Neuroscience: parietal lobe: from action organization to intention understanding. *Science* 2005; 308: 662–667.
109. Nguyen P, Tran T, Wickramasinghe N, et al. DeepPr: a convolutional net for medical records. *IEEE J Biomed Heal Inform* 2017; 21: 22–30.
110. Brachten F, Brünker F, Frick NRJ, et al. On the ability of virtual agents to decrease cognitive load: an experimental study. *Inf Syst E-bus Manag* 2020; 18: 187–207. DOI: [10.1007/s10257-020-00471-7](https://doi.org/10.1007/s10257-020-00471-7).
111. Mirbabaie M, Stieglitz S, Brünker F, et al. Understanding collaboration with virtual assistants – the role of social identity and the extended self. *Bus Inf Syst Eng* 2021; 63: 21–37. DOI: [10.1007/s12599-020-00672-x](https://doi.org/10.1007/s12599-020-00672-x).

112. Mirbabaie M, Stieglitz S and Frick NRJ. Hybrid intelligence in hospitals: towards a research agenda for collaboration. *Electron Mark* 2021; forthcoming.
113. Frick NR, Brünker F, Ross B, et al. Comparison of disclosure/concealment of medical information given to conversational agents or to physicians. *Health Inform J* 2021; 27: 146045822199486.
114. Holzinger A, Langs G, Denk H, et al. Causability and explainability of artificial intelligence in medicine. *WIREs Data Min Knowl Discov* 2019; 9: e1312. DOI: [10.1002/widm.1312](https://doi.org/10.1002/widm.1312).