Science and Engineering Ethics (2021) 27:61
https://doi.org/10.1007/s11948-021-00336-3

ORIGINAL RESEARCH/SCHOLARSHIP

Implementing Ethics in Healthcare AI-Based Applications: A Scoping Review

Magali Goirand[1] · **Elizabeth Austin**[1] · **Robyn Clay-Williams**[1]

Received: 21 January 2021 / Accepted: 4 August 2021 / Published online: 3 September 2021

Abstract

A number of Artificial Intelligence (AI) ethics frameworks have been published in the last 6 years in response to the growing concerns posed by the adoption of AI in different sectors, including healthcare. While there is a strong culture of medical ethics in healthcare applications, AI-based Healthcare Applications (AIHA) are challenging the existing ethics and regulatory frameworks. This scoping review explores how ethics frameworks have been implemented in AIHA, how these implementations have been evaluated and whether they have been successful. AI specific ethics frameworks in healthcare appear to have a limited adoption and they are mostly used in conjunction with other ethics frameworks. The operationalisation of ethics frameworks is a complex endeavour with challenges at different levels: ethics principles, design, technology, organisational, and regulatory. Strategies identified in this review are proactive, contextual, technological, checklist, organisational and/or evidence-based approaches. While interdisciplinary approaches show promises, how an ethics framework is implemented in an AI-based Healthcare Application is not widely reported, and there is a need for transparency for trustworthy AI.

Keywords Ethics · AI · Healthcare · Machine learning · Bioethics · Care robots · CDSS · IAT

Introduction

Artificial Intelligence (AI) emerged in 1955 and is recognised as a technology that can profoundly reshape healthcare and society at large (Crawford et al., 2019; Floridi & Cowls, 2019). While many definitions for AI have come into use, the 1955 definition "the science of making machines do things that would require intelligence if done by people" is still relevant for healthcare (Joshi & Morley, 2019, p. 15). Because of the nature of AI technology, large high-tech companies are entering

✉ Magali Goirand
magali.goirandampaire@hdr.mq.edu.au

[1] Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

the field of AI-based Healthcare Applications (AIHA). However, these companies do not share the same history and culture of ethics as the medical field. In 2016, Google DeepMind in collaboration with the Royal Free London NHS Foundation Trust developed an AI application for the management of acute kidney injury which raised ethical concerns because of a lack of transparency in their processes breaching the trust of patients (Powles & Hodson, 2017). It is in the interest of members of the public, and the healthcare institutions who served them to gain clarity on how ethics guidelines are implemented in AI-based Healthcare Application (AIHA).

AIHA take on many forms. Examples include sensors and wearables for diagnostics and monitoring, care robots, automated image interpretation, clinical decision support systems, and smartphone apps for mental health therapy (NHS, 2019). Views on the benefits of AI in healthcare span the whole spectrum from utopia to dystopia, and one subfield of AI in particular, Machine Learning (ML), holds a lot of promises but also raises the most concerns (Sparrow & Hatherley, 2019; Topol, 2019). In some pessimistic scenarios, medical care is mostly provided by AI replacing clinicians, the elderly monitored and tended to by robots. In some optimistic scenarios, AI promises a highly personalised medicine accessible even to remote locations, faster and more accurate medical diagnoses, to free clinicians from administrative tasks who would in return have more time for their patients and to allow the elderly to live safely and independently at home. There are multiple definitions of AI. The European Commission defines AI as "systems that display intelligent behaviour by analysing their environment and taking actions—with some degree of autonomy—to achieve specific goals" and can be software-based, or embedded in hardware, with the subsets of ML systems being characterised by their ability to learn from their environment (Bird et al., 2020). In healthcare, the environment extends from clinical points of care, nursing homes, to private homes where collected information and its usage are sensitive and protected by the culture of ethics of the medical fields. As illustrated by the Google DeepMind and the Royal Free London NHS Foundation Trust data privacy breach, traditional operationalisation of medical ethics frameworks may not be sufficient (Powles & Hodson, 2017). In this case, the technology company focused mostly on the performance of the app to the detriment of data privacy. Such ethical concerns emerging from the infusion of AI in various sectors including healthcare are being addressed by a range of initiatives to harness the power of AI to the benefit of people in the form of AI-specific ethics guidelines (Jobin et al., 2019).

There have been at least 84 AI ethics frameworks published in the last 6 years by various stakeholders involved in different aspects of AI development and implementation processes including technology providers (e.g., IBM), implementers (e.g., PricewaterhouseCoopers UK), non-governmental (e.g., IEEE) and governmental organisations (e.g., CSIRO) (Jobin et al., 2019). Ethics frameworks comprise ethics principles, recommendations, best practices and/or standards guiding the development of ethical AI systems. While the published AI ethics frameworks share some ethics principles, these are open to interpretation during operationalisation at the different stages of AI design, development and deployment (Jobin et al., 2019). For example, transparency is a principle widely present in ethics frameworks that can be implemented only to meet legal requirements, or as a participative process in the development of the application (Jobin et al., 2019). Consequently, concerns have been raised about the integrity of the translation of ethics guidelines and frameworks during implementation (Metzinger, 2019; Morley et al., 2020). In the context of healthcare, it is not yet clear what ethics frameworks have been adopted in AI development, how they have been implemented, and how successful these implementations have been. Gaining an understanding of the strategies and challenges will help gain clarity about the operationalisation of AIHA ethics frameworks.

In this review, we sought to survey how AI ethics frameworks have been implemented and evaluated in healthcare applications, and map the scope, challenges and practices of these initiatives. This scoping review aimed to examine the implementation of ethics frameworks in AI applications in healthcare. The questions addressed by the review were:

1. What ethics frameworks have been implemented in AI applications in healthcare?
2. What ethics frameworks does the existing literature report should be implemented (but have not yet been implemented) in AI applications in healthcare?
3. What are the challenges identified in the literature relating to the implementation of an ethics framework into an AI application in healthcare?
4. What strategies have been used to implement an ethics framework into AI applications in healthcare?
5. How have implementations of ethics frameworks been evaluated?

Method

We conducted a systematic scoping review of the peer-reviewed and grey literature related to the implementation of ethics frameworks in AI applications in healthcare published between 2015 and 2020. The protocol was prospectively registered in July 2020 in the Open Science Framework (osf.io/cb8nv). This scoping review followed the Joanna Briggs Institute's (JBI) guidance for systematic scoping reviews (Peters et al., 2015), and the preferred reporting items for systematic reviews and meta-analyses statement (PRISMA) (Moher et al., 2009). The process is designed to be reproducible. It consists of four steps: identification, screening, eligibility and data capture. The identification step consists in collecting the articles as described in the search strategy. During screening, articles are excluded based on information from their abstracts and keywords. The full texts of the articles remaining after the first screening are further scanned and if they meet the eligibility criteria are selected for the data capture. The data capture consists in extracting information from each article based on a pre-defined template.

Inclusion Criteria

Peer-reviewed articles, as well as grey literature articles in English and French language (languages spoken by the researchers) published between 2015 and 2020, were considered when meeting the following criteria: (1) the context of the discussion is an AI-based application in healthcare, (2) the AI-based application includes an ML component, (3) the discussion includes ethical considerations and how to implement or operationalise them into the AI-based application. Because AI technology evolves rapidly, we chose to limit our search to the past five years.

Search Strategy

Because of the cross-disciplinary nature of the review and the rapid pace of change in AI, we adopted a search strategy encompassing traditional medical publication databases and grey literature. The medical publications databases were Ovid MEDLINE, Ovid Embase, Scopus and Web of Science. Databases were queried on 16 July 2020. Databases were searched using medical subject headings and keywords representing the two main concepts of technology and ethics. The search strategy was developed in consultation with a research librarian and devised to capture intelligent assistive and decision support technologies based on machine learning. The full search strategy for all databases is available in "Appendix 2". For the grey literature, we used scienceresearch.com, a deep web technology search engine, as well as a targeted hand search of websites of technology providers, think tanks, and government bodies involved with AI-based applications for healthcare. The latter were selected from a list of organisations that released an AI ethics framework (Jobin et al., 2019) and had activities related to healthcare regardless of their geographic location. Key players such as Verily or McKinsey & Company, who did not appear in Jobin et al. (2019)'s list were also part of the search (full list available in "Appendix 1"). Websites were searched between 9 and 15 July 2020. Websites were searched using the local search function and a combination of the following keywords: "Artificial Intelligence", "ethics", and "healthcare" as systematically as possible within the capabilities of each website. When the search returned a large number of results (i.e., more than two pages of results), the first 20 items at a minimum were screened for relevance and recorded until a decline of relevance was observed (i.e., characterised by items becoming unrelated to healthcare). When a local search was not available but there was a list of publications available on the website, a search using the same keywords within each publication was performed to select or reject them. A log of the search strategies and results were kept in an Excel (MS Excel for Mac version 16, Microsoft, Seattle, WA) spreadsheet.

Study Selection

The search results from the databases were uploaded into EndNote citation management software (version X9.3; Thompson Reuters, New York, NY), and duplicates were removed. Then all references including grey literature results were uploaded into Rayyan web application software (Ouzzani et al., 2016) and duplicates further removed. The title and abstract of each article were screened by pairs of independent reviewers for inclusion based on pre-established criteria. Inclusion conflicts between reviewers were resolved through discussion. Articles were excluded if (1) they were press releases, promotional in nature, biographies, videos, podcasts, speech or meeting transcripts, education programmes, slides, or conference proceedings; (2) they were not healthcare specific, they did not mention any healthcare application or case study, or healthcare was mentioned only as an industry among others; (3) they did not feature any AI applications at all, or featured AI applications with no Machine Learning such as rule-based inference engines for example, or discussed AI as a technology among others, or discussed data only; (4) they mentioned ethics as only a review by an ethics committee, or featured technology as a solution to an ethical problem such as the usage of AI for data privacy preservation, or discussed only ethical challenges without specifying any strategy, or discussed data ethics only, or featured no operationalisation of an ethics framework.

Data Extraction and Quality Assessment

Pairs of reviewers independently extracted the data of the eligible articles based on a data extraction form developed for this purpose. A pilot of the data extraction form was conducted for usability prior to the full extraction. The data extracted were comprised of authors, title, information about the AI-based application, the characteristics of the ethics framework, the strategies for implementing the ethics framework and the challenges and their resolution if any, the assessment of the ethics framework implementation if any, and the type of publication (see Fig. 1). The quality assessment of the articles was performed by pairs of independent reviewers using

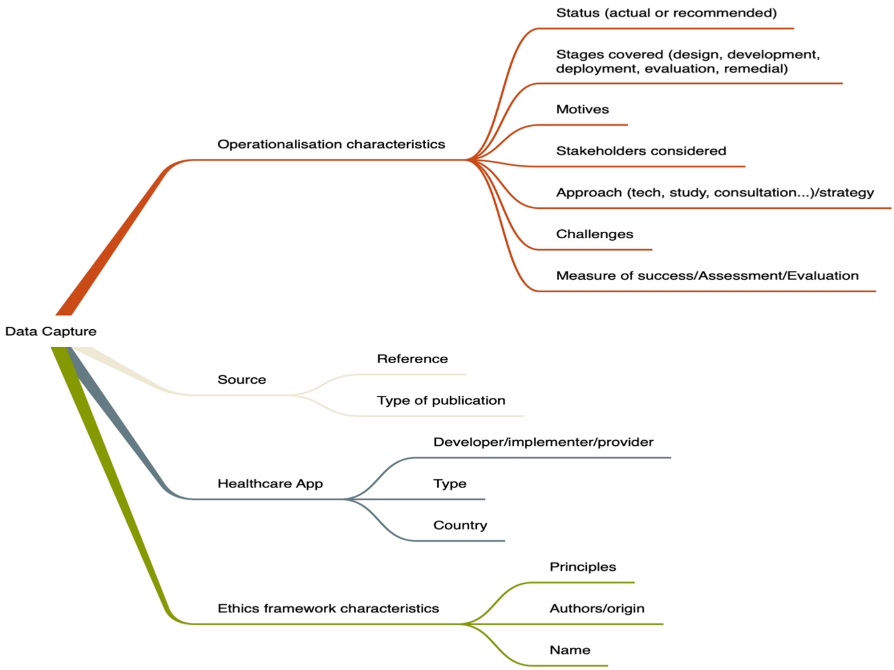

Fig. 1 Data capture model

JBI critical appraisal tools, selected depending on the type of article. Any disagreement was resolved through discussion. The outcomes were used to map the quality of the literature only, not as an exclusion criterion because of the exploratory nature of the scoping review.

Data Processing and Analysis

A narrative synthesis was performed for this review due to the qualitative and heterogeneity of the data. First, the characteristics of the applications were analysed to draw an overview of the data. Second, an analysis of the ethics frameworks was performed using the principles upon which they relied. Lastly, the characteristics of the operationalisation of the frameworks were analysed, compared and synthesised.

Results

The combined search of databases and targeted websites yielded a total of 8054 articles, which after removal of duplicates was reduced to 4444 items. Out of the 4444 title and abstract screened articles, 480 were flagged potentially relevant and underwent a full-text review. 33 met the inclusion criteria. Figure 2 illustrates the PRISMA diagram for the identification, screening, and inclusion processes.

80% (n = 26) of the articles dated from 2018 onward. 16 articles were theoretical, 17 featured actual applications among which 11 were about deployed applications, and 6 were about prototypes. The types of application covered by the articles included generic Machine Learning (ML) (n = 3), Clinical Decisions Support Systems (CDSS) (n = 8), drones (n = 2), Intelligent Assistive Technologies (IAT) which include generic IAT (n = 3), virtual bots (n = 6) and robots (n = 13) (Fig. 3).

Quality Assessment

Among the 33 articles, seven were appraised as qualitative studies, and 26 as text/opinion pieces (see "Appendix 3"). The qualitative studies varied in score, while the opinion/text pieces scored highly. While the process is imperfectly suited to the nature of the articles, which do not follow traditional medical review protocols, it establishes the credibility of the sources. Two articles were grey literature (Buston et al., 2019; CPAIS, 2019). Two articles had authors with corporate correspondents (Joerin et al., 2020; Rajkomar et al., 2018).

Ethics Frameworks that Have Been Implemented and/or Are Recommended to be Implemented in AIHA

The ethics principles categories identified by Jobin et al. (2019) were used to qualify the extent of coverage of each article's framework (see Fig. 4). The four principles, autonomy, non-maleficence, beneficence and fairness have been classified as bioethics principles following the standard Beauchamp and Childress (2013)'s framework

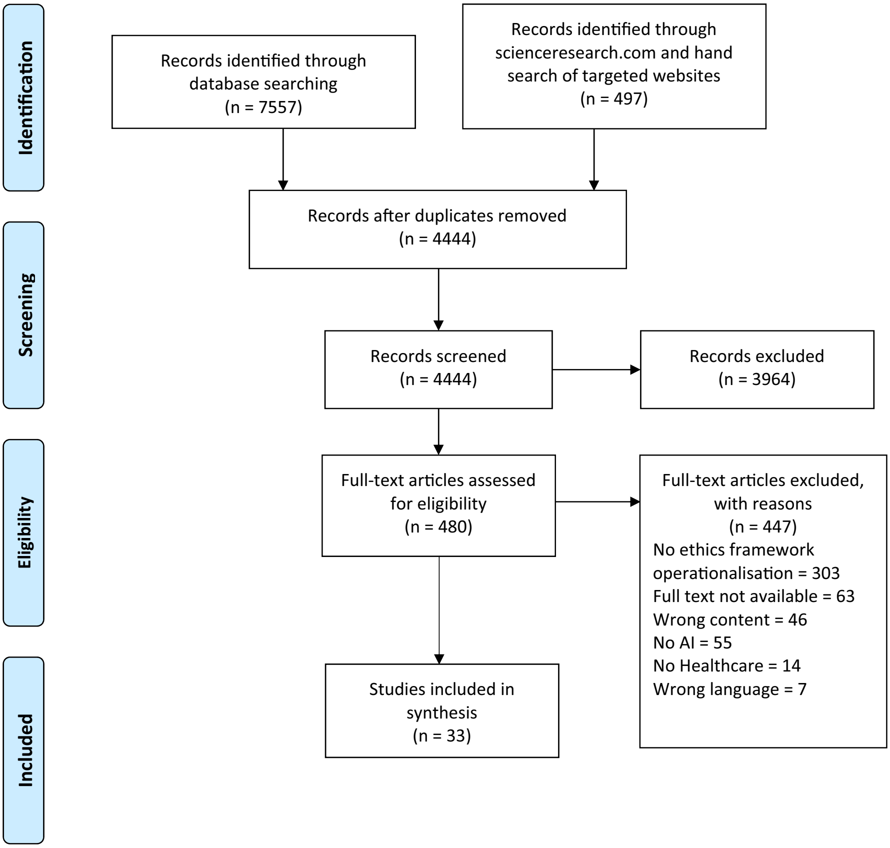

Fig. 2 PRISMA flow diagram of the article selection

widely used in the field of medical sciences. We found a strong representation of these bioethics principles with two thirds (n = 22) of articles covering at least three out of the four bioethics principles with non-maleficence the most considered (n = 27 or 82%). Privacy (n = 21) and transparency (n = 20) were considered in more than 60% of the articles. 24% (n = 8) of articles referenced an AI-specific ethics framework: AI4People (Cawthorne & Robbins-van Wynsberghe, 2020), American Medical Association Augmented Intelligence in Healthcare (Channa et al., 2020), UK House of Lords' AI guidelines (Joerin et al., 2020), Ethics guidelines for trustworthy AI by the European Commission (Beil et al., 2019), Australia's AI ethics framework (Milosevic, 2019), NHSX and Future Advocacy guidelines (Buston et al., 2019), Partnership on AI guidelines (CPAIS, 2019), and IEEE ethics framework (D. Peters et al., 2020) Other references were also made to machine ethics, robot ethics, care ethics, and diverse professional society ethical guidelines. Four articles (Beil et al., 2019; Cawthorne & Robbins-van Wynsberghe, 2020; Joerin et al., 2020; Poulsen et al., 2018) combined different sets of medical and technology ethical guidelines such as bioethics and European ethics guidelines for trustworthy AI (Beil et al.,

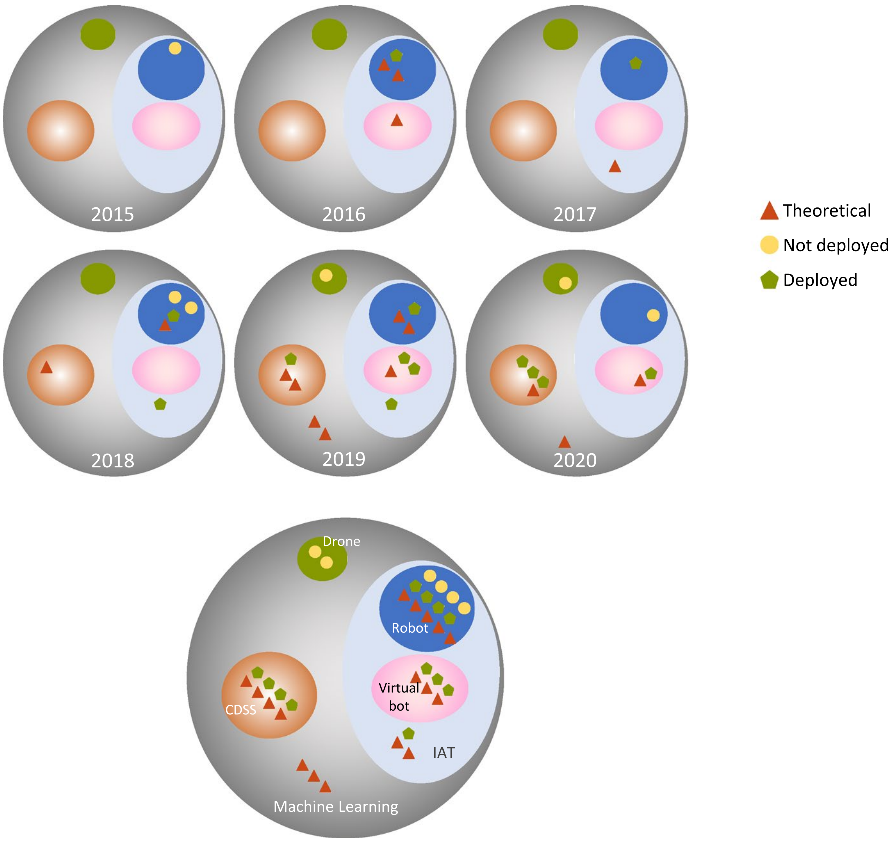

Fig. 3 Distribution of types of applications across articles and its evolution over time (one article covered three different types, all the others were monotype)

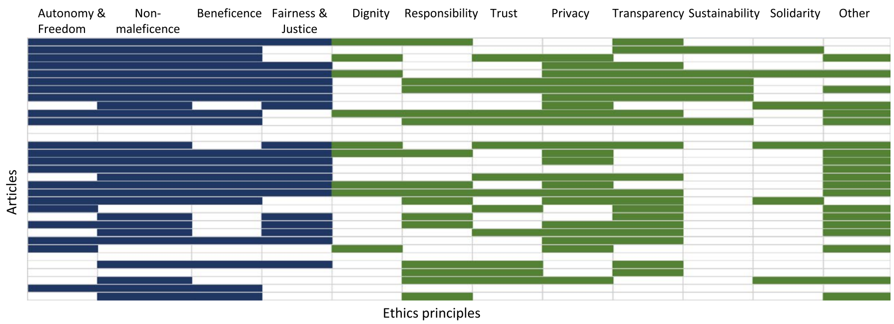

Fig. 4 Ethics principles coverage per article—in blue are the classic bioethics principles

2019). The list of ethics frameworks used in each article is available in "Appendix 4".

The process of categorising ethics principles laid out in the articles was not clear-cut, so we added an "other" category. Avoiding deception was reported for (n = 4) care robots shaped like animals or anthropomorphised to foster an emotional bonding (Battistuzzi et al., 2018; Cooney & Menezes, 2018; Kortner, 2016; Stahl & Coeckelbergh, 2016). The principle of participation or social interconnectedness was evoked in four articles (Battistuzzi et al., 2018; Hasenauer et al., 2019; Klein & Schlömer, 2018; Poulsen & Burmeister, 2019). For example, care robots for the elderly as a robot can become an enabler or an inhibitor of the ethics principle. Three articles mentioned the principle of interdependence between caregiver and patient as in the respect for the rapport existing between a caregiver and a patient (Ienca et al., 2018; Klein & Schlömer, 2018), and the danger of having the attention of the caregiver focused on the care robot rather than the patient (McBride, 2020). Two articles considered the impact of the AIHA on human skills, and human jobs (Cawthorne & Robbins-van Wynsberghe, 2020; Garner et al., 2016), and one recommended that a human therapist should always be prioritised over the AIHA (Cooney & Menezes, 2018). Efficacy as in the number of patients without the disease correctly identified as healthy was mentioned once (Channa et al., 2020). One article specified the relevance of an AIHA based on its context as part of ethics principles to follow (McCradden et al., 2020). Minimisation of power imbalances including in decision making was emphasised in five articles (Battistuzzi et al., 2018; Garner et al., 2016; Ienca et al., 2017; Joerin et al., 2020; Kortner, 2016). The transparency of the validation process and integration of the AIHA was cited once (Buston et al., 2019). One article included contestability, the ability to challenge a decision made by a CDSS in the set of principles (Milosevic, 2019).

Challenges Relating to the Implementation of an Ethics Framework into an AIHA

The articles identified challenges relating to the implementation of ethics frameworks. We grouped these into five levels: ethics principles, design, technology, organisational, and regulatory.

At the Ethics Principles Level

The relationships between ethics principles can be at odds. For example, in CDSS beneficence can compromise autonomy because of the lack of understanding of how decisions are made or managing biases in datasets can oppose justice and nonmaleficence (Beil et al., 2019). In the case of virtual bots for the elderly, trust may be compromised to ensure safety by being intrusive, and potentially also compromising privacy (Garner et al., 2016). Because ML systems learn from historical datasets, it embeds historical biases and disparities which means certain groups are favoured by default while others need to be protected by design to ensure fairness for all. Fairness in ML is multi-dimensional and can be achieved by optimising equality of outcomes, accuracy and allocation of resources for each group which requires compromises (Rajkomar et al., 2018). For example, optimal accuracy may translate into unequal outcomes for a protected group because the recommended treatment may be less effective due to lack of testing in this population (Rajkomar et al., 2018). The ranking of importance of each principle is context dependent as in the case of the elderly who may feel socially isolated, care robots can provide social connectedness to the detriment of privacy or independence (Poulsen & Burmeister, 2019). Patients, caregivers and clinicians can have divergent relationships with the ethics principles. While a caregiver would prioritise safety over privacy, a patient may prefer to preserve his/her privacy over safety when bathing with a shower robot for instance (Klein & Schlömer, 2018).

At a Design Level

Translating abstract principles into design requirements, features and functions of an AIHA creates conflict between technology and human values (Cawthorne & Robbins-van Wynsberghe, 2020). It is challenging for a machine to reproduce the complexity of human ethical decision making which is a desired feature when modelling an autonomous decision-making system whether it is part of a robot or software (van Rysewyk & Pontier, 2015).

At a Technology Level

Depending on its rigidity, the ethics framework implementation could impede continuous improvement or create blind spots to emergent issues (Cawthorne & Robbins-van Wynsberghe, 2020). Iterative processes built into an AIHA application, typical of AI models that keep learning to improve, can affect the relationship between the ethics principles (Battistuzzi et al., 2018).

At an Organisational Level

Implementing a framework requires interdisciplinary collaboration among groups with competing interests and disparate backgrounds ranging from clinicians, developers, legal counsellors to board members and investors of a technology provider. Developers tend to reduce ethics to fairness causing tensions with clinicians who have a broader scope for ethics (Abramoff et al., 2020). In the boardroom of an AIHA provider, there could be tensions between ethical, financial and legal considerations (Joerin et al., 2020). In the context of a research project, funding and expertise silos can impede collaboration across disciplines (Stahl & Coeckelbergh, 2016).

At a Regulatory Level

Regulatory and legal matters can hinder the application of ethics principles as, for example, seeking legal informed consent from every person entering in contact with a robot could prove impractical (Kortner, 2016). Lack of regulations of AI technology, and conflicting national and international legislations can pose a challenge (Dantas et al., 2017; Garner et al., 2016). Because outputs of an AIHA are not necessarily reproducible, and their uncertainty is difficult to quantify, it does not fit into existing medical regulatory ethical settings (Abramoff et al., 2020; Beil et al., 2019; Buston et al., 2019; Milosevic, 2019).

Strategies to Implement an Ethics Framework into AIHA

Implementation strategies from the surveyed articles covered the application lifecycle with the design stage the most represented (n = 27 or 82%), and commercialisation the least represented (n = 6 or 18%) (see Fig. 5). Only one strategy, which is targeted at the technology developers and is a grey literature article, spanned the entire lifecycle (Buston et al., 2019). The implementation strategies included proactive, contextual, technological, checklist, organisational and evidence-based approaches.

Being Proactive

One approach is to consider ethics at the design and requirement capture stage by embedding ethical values into the application using methodologies such as Value-Sensitive Design (VSD) (Battistuzzi et al., 2018; Cawthorne & Robbins-van Wynsberghe, 2019, 2020; Ienca et al., 2017; Ienca et al., 2018; D. Peters et al., 2020), Care-Centred Value-Sensitive-Design (CCVSD) (Van Wynsberghe, 2016), Values in Motion Design (Poulsen et al., 2018) and Attentive Framework (Poulsen & Burmeister, 2019). The ethical values can be captured or validated by consultation of stakeholders (Battistuzzi et al., 2018; Cawthorne & Robbins-van Wynsberghe, 2019, 2020; Fukuzumi et al., 2019; Garner et al., 2016; Ienca et al., 2017; Ienca et al.,

Fig. 5 Coverage of the implementation of the ethics framework into the life cycle of an application per article

2018; Klein & Schlömer, 2018; D. Peters et al., 2020; Poulsen et al., 2018; van Rysewyk & Pontier, 2015). Another proposed way is to translate a professional code of conduct into capabilities (Cooney & Menezes, 2018).

Being Contextual

Four articles stressed the importance of a deep understanding of the context of the AIHA (Ienca et al., 2017; McBride, 2020; Sendak et al., 2020; Van Wynsberghe, 2016). van Wynsberghe gives poignant situations where a care robot's purpose is to protect nurses from toxins when collecting urine samples from children undergoing chemotherapy or from injuries when lifting a patient off his/her bed. The rapport of care between nurse and patient needs to be respected when programming the robot.

Technological Approach

Another approach is technological and included coding ethics in the operational system (Anderson et al., 2019; Milosevic, 2019), embedding ethics principles in the algorithm, and the collection of data of the AIHA (Abramoff et al., 2020; Rajkomar et al., 2018), monitoring and evaluating the applications (Abramoff et al., 2020; Joerin et al., 2020; Kortner, 2016; McCradden et al., 2020; Rajkomar et al., 2018), and adopting a continuous improvement approach (Joerin et al., 2020; Kortner, 2016; McCradden et al., 2020). User eXperience (UX) techniques were recommended in particular for respecting privacy and transparency (Joerin et al., 2020; Kretzschmar et al., 2019).

Checking and Verifying

Checklists have been developed to guide the dialogues between stakeholders (CPAIS, 2019; Stahl & Coeckelbergh, 2016), raise awareness about potential conflicts of interest including funding (Morch et al., 2019), and focus on the patient, the system, the technical and the medical aspects of an AIHA (Beil et al., 2019). At the evaluation stage, a verification methodology of ethical compliance was recommended which involves case studies and experts' consultations (Anderson et al., 2019; Dantas et al., 2017).

Organisational Checks and Balances

The importance of the responsibility of the AIHA creators and the need for counselling through an advisory board was highlighted by authors from a private company (Joerin et al., 2020).

Evidence-Based

Another recommendation was to follow an evidence-based introduction of a complex intervention (Carter et al., 2020). AI systems trained within a context can perform very differently in another context. It is therefore important to prove the system before deploying it into a clinical setting.

Evaluation of the Implementations of Ethics Frameworks into AIHA

Less than a third of the proposed implementation strategies included specifics about the evaluations of the implementation (n = 9, 27%) (Beil et al., 2019; Cawthorne & Robbins-van Wynsberghe, 2020; CPAIS, 2019; Dantas et al., 2017; Hasenauer et al., 2019; Klein & Schlömer, 2018; Kretzschmar et al., 2019; McBride, 2020; Morch et al., 2019). Measures of success included better outcomes for the patients than without AI (Abramoff et al., 2020; Carter et al., 2020; Channa et al., 2020), efficacy (Abramoff et al., 2020; Kretzschmar et al., 2019; Poulsen et al., 2018), safety (Abramoff et al., 2020; Kretzschmar et al., 2019), equity (Abramoff et al., 2020; Carter et al., 2020), effectiveness (Poulsen et al., 2018), cost-effectiveness and reduction (Carter et al., 2020; Hasenauer et al., 2019), privacy (Kretzschmar et al., 2019), clear professional responsibilities, autonomy, usage in relevance settings, and that trust is maintained (Carter et al., 2020). The approach, whether a design framework, questionnaire or stakeholders' consultations, acts as an evaluation methodology. One article featuring an actual care robot application reported a successful evaluation as the robot behaved as expected (Anderson et al., 2019). One article concluded mixed results in their evaluations of different mental health mobile applications (Kretzschmar et al., 2019).

Discussion

Dearth of Studies About the Implementation of Ethics Frameworks in AIHA

In this review, we aimed to examine the implementation of ethics frameworks in AIHA. While 4444 articles were screened out of which 480 were eligible for full-text review, only 0.75% (n = 33) were ultimately included. The attrition of articles between screening, eligibility and inclusion indicates a dearth of studies on how ethics are implemented in AIHA. Additionally, only a third of the articles were about deployed applications. This lack of data could be due to the novelty of the ethics frameworks for AI as according to Jobin et al. (2019), 88% of them have been issued after 2016, the culture of non-disclosure of business practices in the private sector, and/or indicative of the complexity of the task (assuming failures don't get reported).

Limited Uptake of AI-Specific Ethics Frameworks

The review reveals that while there have been at least 84 AI-specific ethics frameworks published in the last 6 years (see (Jobin et al., 2019) for full list), the adoption of these frameworks in healthcare has been limited. Eight or 24% of the articles (Beil et al., 2019; Buston et al., 2019; Cawthorne & Robbins-van Wynsberghe, 2020; Channa et al., 2020; CPAIS, 2019; Joerin et al., 2020; Milosevic, 2019; D. Peters et al., 2020) mention an AI-specific ethics framework. Of the eight, only three report on the adoption of an ethics framework for deployed applications (Channa et al., 2020; CPAIS, 2019; Joerin et al., 2020) and one on a pilot (Cawthorne & Robbins-van Wynsberghe, 2020). Among these four AI-specific frameworks, only one was specific to healthcare developed by the American Medical Association (Channa et al., 2020). The limited uptake of AI-specific ethics frameworks in AIHA may be due to the novelty of the frameworks, or their lack of healthcare specialisation. The limited uptake could also signal the complexity of operationalising ethics in AIHA and harmonising the cultures of the healthcare and high-tech industry.

Trust is a Wicked Problem

Trust is a wicked problem, one difficult to solve because of compounding factors such as the challenges around ML explicability, and reluctance to adopt because of ethical concerns resulting in unknown variables. Critically, technology performance is heavily dependent on data whose collection depends on adoption (Hasenauer et al., 2019; Kretzschmar et al., 2019; Sendak et al., 2020). For instance, mental health chatbots users need to trust the app to use it and understand its limitations for safety reasons. It means privacy, confidentiality but also transparency about the limitations of the app, and how the collected data are used need to be communicated, while the data are needed to increase the performance of the app, hence its safety. In the case of a CDSS whose users were clinicians, the trust in relationships between the different parties was more important than the trust in the technology (Sendak et al., 2020). Nonetheless, the challenge AI applications are facing with adoption is circular because of their large need for data to keep learning and improving, hence becoming safer, hence more trustworthy.

Need for Examining the Motives for Implementing an Ethics Framework

Applications involving IATs, and robots in particular are the most represented and spread over the five-year window. CDSSs are becoming more prominent only in the last two years and equally represented as IATs in 2020. Because IATs interface with patients directly, ethical issues may be more visible and pressing, prompting more studies. In one instance (Hasenauer et al., 2019), the implementation of an ethics framework for a care robot in a nursing home was motivated by the need for the consumers to adopt the robot after they expressed reservations and ethical concerns about it. In other words, the implementation of an ethics framework was motivated by the need to gain the trust of the consumers, and the assumption that the care robot is the adequate solution is not challenged. It is noteworthy that only two strategies (Carter et al., 2020; Van Wynsberghe, 2016) call for explicitly investigating whether an AIHA is even a good idea in the given context first.

Challenges Cannot be Overcome with Technology Alone

The five levels of challenges identified (ethics principles, design, technology, organisational and regulatory) reinforce the need for an interdisciplinary approach and indicate that it can't be fixed with technology alone. In one CDSS application (Sendak et al., 2020), the authors highlighted how trust in the AIHA was rooted in relationships, not the features and functions of the technology. Another CDSS article that was technology focused (Rajkomar et al., 2018) stressed the importance of educating clinicians on ML technology to be able to understand ML output and orient design choices. The Australian Human Rights Commission report on the usage of AI (Commision, 2019) advocates educating public servants and regulatory advisors involved in making decisions about the acquisition and deployment of AI systems for them to better understand the implications and consequences of ML technology. Conversely, people involved in the design and development of ML technology need to be equipped to assess the ethical and societal consequences of their development (Commision, 2019; Villani, 2018). There is a need to create a common language and understanding on both sides about what is at play among all stakeholders and to coordinate effort from researchers, developers, end-users, and implementers to ensure the gap between principles and practice is closed (Morley et al., 2020).

Balance of Power is Unclear in Strategies

Strategies to implement ethics frameworks were mostly proactive and involved some degree of stakeholders' consultation, yet it is unclear how decisions were made in particular in the case of conflicts between ethics principles and/or between stakeholders needs. It is also unclear how much education was being provided for stakeholders to understand the technology and its implications in these consultations. How well stakeholders' voices are being heard, and who has the last say on how ethics and values are operationalised are known issues in AI in general and not specific to AIHA (Crawford et al., 2019).

Some Strategies Strive to be Holistic

Most strategies were multi-pronged combining some of the identified approaches (contextual, technological, checklist-oriented, organisational and evidence-based). Only one spanned the entire lifecycle of an application, and in particular, the commercialisation/procurement stage is the least represented. This could be a concern given the importance of the economic interests at stakes. Two strategies included considerations about the consequences of the AIHA on its broader environment, or second and third-order consequences usually seen as unintended consequences (Cawthorne & Robbins-van Wynsberghe, 2020; Peters et al., 2020). Both applications are on the fringe of the healthcare system, as one is a drone delivering blood sample, and the other a mental health counselling app. It is possible that because they are not contained within a medical environment, they are better positioned to enquire about broader consequences of the deployment of such applications on society. One circles back to the effects on developers being exposed to possibly disturbing information when having to label training datasets (Peters et al., 2020). The other enquires about the impact on the environment, and societal issues such as job losses, but also whether the introduction of a "good drone" could pave the way for acceptance of drones which may not be as beneficial, lowering our guard in some ways (Cawthorne & Robbins-van Wynsberghe, 2019, 2020). Such thinking implies the need for vigilance and calls for ethics frameworks to adjust or evolve with societal context. An abductive usage of the framework is recommended. Not unlike the way an AI model works, it begs to keep learning, monitoring and improving to be able to respond adequately.

No Clear Measure of Success

While some strategies include ways of evaluating the implementation of the ethics framework, we could not find clear measures of success and few explicitly reported on whether these implementations were successful (Anderson et al., 2019; Kretzschmar et al., 2019). The absence of clearly defined measures of successful implementation of ethics frameworks could indicate a lack of maturity in this emerging field. It could also be another symptom of the complexity of operationalising an ethics framework and reinforce the need for a common language between the different stakeholders. Clear measures of success are critical to evaluate whether the selected framework is fit-for-purpose and whether the implementation of the ethics framework was effective.

Strengths and Limitations

We used a comprehensive search strategy including academic databases and grey literature, and followed a rigorous process to screen, and appraise articles. Nonetheless, private sector initiatives are not necessarily accessible through a public literature search, and likely underrepresented in this review. Due to the interdisciplinary nature of the topic, the data were heterogeneous and did not always lend itself neatly to the process of a systematic scoping review. As such, the screening and eligibility determination was not straightforward, and the data capture messy. Having two pairs of reviewers with different backgrounds helped alleviate possible biases and potential gaps in the data capture. Because of the limitations to English and French languages, China, Korea and Japan are likely underrepresented in the current review while being prominent players in the field of AI. Finally, a number of the authors within our data set operate within the same network. Out of 33 articles, one pair (Ienca et al., 2017, 2018) had the same first author, two pairs (Cawthorne & Robbins-van Wynsberghe, 2019, 2020; Poulsen & Burmeister, 2019; Poulsen et al., 2018) had the same first and second authors and studied the same application, and the second author of one of these pairs was the first author of another article (Van Wynsberghe, 2016). Another two articles shared the same author and drew from the same application (Abramoff et al., 2020; Channa et al., 2020). While it gives more perspectives and depth to these approaches, it is likely that the interconnectedness of the authors may overrepresent their strategies.

Conclusion

Operationalising ethics frameworks in AIHA is a complex endeavour, and not widely reported in the literature. Recommended and implemented ethics frameworks for AIHA draw from medical ethics, technology-specific ethics and professional ethics. Yet there has been a limited uptake of AI ethics frameworks in healthcare. One challenge when operationalising an ethics framework is that there is no one size fits all approach and a need for contextualisation. While an AI ethics framework for AIHA may be application-agnostic, the implementation of the framework needs to be adapted to the environment of the AIHA and the same AIHA may face different ethical issues when deployed into different environments with different populations. Ultimately, there is agreement on the need for proactive and inclusive approaches. Implementing an ethics framework should not be remedial and should take into consideration all stakeholders' needs. Embedding an ethics framework is particularly important due to the danger of power imbalance between the recipients or consumers and the AIHA providers. Measures of a successful implementation of AIHA ethics frameworks are beginning to emerge, bringing much needed clarity and transparency to the whole process. Cross-pollination between computer science, engineering, medical sciences, and social sciences promises to yield the most comprehensive ways to tackle the wicked issue of ethics in AIHA.

Appendix 1: Grey literature search strategy

- Scienceresearch.com limited to Food and Drug Administration using the following query: (artificial intelligence or machine learning or deep learning or robot* or chatbot* or intelligent assistive or decision support) AND (ethic* or privacy or fairness or equity).

- Manual search of the websites of the following organisations:

Organisation name	Organisation type
Accenture	Consultancy
ACM	Professional organisation
Amazon care	Technology company
Apple	Technology company
Bain & co	Consultancy
Cap Gemini	Consulting
CEPEJ (Council of Europe European Commission for the Efficiency of Justice)	Government
CNIL	Government
DeepMind	Technology company

Organisation name	Organisation type
Department of industry innovation and science Australia	Government
EP think tank	Government
European Parliament	Government
Future Advocacy	Consultancy
GE healthcare partners	Consultancy
Google research publication	Technology company
Gov of the republic of Korea	Government
IBM	Technology company
ICO	Government
IEEE	Professional organisation
Intel	Technology company
McKinsey	Consultancy
Microsoft	Technology company
Ministry of economic affairs and employment Finland	Government
Mission Villani, AI for humanity	Government
National Science and technology council (USA white house)	Government
NHSX	Government
NITI Indian Gov	Government
Novartis foundation	Technology company
OpenAI	Technology company
Partnership on AI	Think tank
Personal Data protection Commission Singapore	Government
PWC UK	Consulting
SAP	Technology company
Tata Consultancy Services	Consultancy
The Boston Consulting Group	Consultancy
The Norwegian Data Protection Authority	Government
Tieto	Technology company
UK Government	Government
UK Parliament	Government
Verily (Google life sciences)	Technology company

Appendix 2: Peer-reviewed literature search strategy

See Table 1.

Table 1 Bibliographic databases search strategy

		OVID - MEDLINE	OVID- EMBASE	SCOPUS	Web of Science
TECHNOLOGY	MeSH	artificial intelligence/ or machine learning/ or deep learning/ or supervised machine learning/ or support vector machine/ or unsupervised machine learning/	machine learning/ or learning algorithm/ or semi supervised machine learning/ or supervised machine learning/ or support vector machine/ or unsupervised machine learning/ or deep learning/ or artificial intelligence/	N/A	N/A
	OR	OR	OR		
	Keywords	(artificial intelligence or machine learning or deep learning or robot* or chatbot* or intelligent assistive or decision support).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]		TITLE-ABS-KEY ("artificial intelligence" OR "machine learning" OR "deep learning" OR robot* OR chatbot* OR "intelligent assistive" OR "decision support")	TOPIC:("artificial intelligence" OR "machine learning" OR "deep learning" OR robot* OR chatbot* OR "intelligent assistive" OR "decision support")
AND		AND	AND	AND	AND
ETHICS	MeSH	ethics/ or bioethics/ or ethics, clinical/ or ethical analysis/ or "ethical review"/ or ethics, medical/ or social responsibility/	ethics/ or bioethics/ or medical ethics/	N/A	N/A
	OR	OR	OR		
	Keywords	(ethic* or privacy or fairness or equity).tw.		TITLE-ABS-KEY (ethic* OR privacy OR fairness OR equity)	TOPIC:(ethic* OR privacy OR fairness OR equity)
AND				AND	AND
HEALTH	Keywords	N/A	N/A	TITLE-ABS-KEY (health* OR medic*)	TOPIC: (health* OR medic*)

Appendix 3: Quality assessment results

See Tables 2 and 3.

Table 2 Qualitative study appraisal tool results

Authors	Is there congruity between the stated philosophical perspective and the research methodology?	Is there congruity between the research methodology and the research question or objectives?	Is there congruity between the research methodology and the methods used to collect data?	Is there congruity between the research methodology and the representation and analysis of data?	Is there congruity between the research methodology and the interpretation of results?	Is there a statement locating the researcher culturally or theoretically?	Is the influence of the researcher on the research, and vice- versa, addressed?	Are participants, and their voices, adequately represented?	Is the research ethical according to current criteria or, for recent studies, and is there evidence of ethical approval by an appropriate body?	Do the conclusions drawn in the research report flow from the analysis, or interpretation, of the data?
Sendak et al. (2020)	Yes	Unclear	Unclear	Yes	Yes	Unclear	No	Yes	Yes	Yes
Anderson et al. (2019)	Unclear	Yes	Yes	Yes	Unclear	No	No	No	No	Yes
Klein and Schlömer (2018)	Unclear	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Morch et al. (2019)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Unclear	Yes
Battistuzzi et al. (2018)	Unclear	Yes	Yes	Yes	Yes	No	No	Unclear	N/A	Yes
Fukuzumi et al. (2019)	Unclear	Yes	Unclear	Unclear	Unclear	No	No	Unclear	Unclear	Unclear

Table 2 (continued)

Authors	Is there congruity between the stated philosophical perspective and the research methodology?	Is there congruity between the research methodology and the research question or objectives?	Is there congruity between the research methodology and the methods used to collect data?	Is there congruity between the research methodology and the representation and analysis of data?	Is there congruity between the research methodology and the interpretation of results?	Is there a statement locating the researcher culturally or theoretically?	Is the influence of the researcher on the research, and vice- versa, addressed?	Are participants, and their voices, adequately represented?	Is the research ethical according to current criteria or, for recent studies, and is there evidence of ethical approval by an appropriate body?	Do the conclusions drawn in the research report flow from the analysis, or interpretation, of the data?
Garner et al. (2016)	Unclear	Unclear	Unclear	Unclear	No	No	No	Unclear	Yes	Unclear

Table 3 Text and opinion appraisal tool results

Authors	Is the source of the opinion clearly identified?	Does the source of opinion have standing in the field of expertise?	Are the interests of the relevant population the central focus of the opinion?	Is the stated position the result of an analytical process, and is there logic in the opinion expressed?	Is there reference to the extant literature?	Is any incongruence with the literature/sources logically defended?
van Rysewyk and Pontier (2015)	Yes	Yes	Yes	Yes	Yes	Yes
Poulsen et al. (2018)	Yes	Yes	Yes	Yes	Yes	Yes
Cawthorne and Robbins-van Wynsberghe (2020)	Yes	Yes	Yes	Yes	Yes	Yes
Channa et al. (2020)	Yes	Yes	Yes	Yes	Yes	No
Kretzschmar et al. (2019)	Yes	Yes	Yes	Yes	Yes	Yes
Cooney and Menezes (2018)	Yes	Yes	Yes	Yes	Yes	Yes
Rajkomar et al. (2018)	Yes	Yes	Yes	Yes	Yes	Yes
Joerin et al. (2020)	Yes	Unclear	Yes	Unclear	Yes	N/A
Kortner (2016)	Yes	Yes	Yes	Unclear	Yes	Yes
Beil et al. (2019)	Yes	Yes	Yes	Yes	Yes	Yes
Ienca et al. (2018)	Yes	Yes	Yes	Yes	Yes	Yes
Milosevic (2019)	Yes	Yes	Yes	Yes	Yes	Yes
Stahl and Coeckelbergh (2016)	Yes	Yes	Yes	Yes	Yes	Yes
Buston et al. (2019)	Yes	Yes	Yes	Unclear	Yes	Unclear
Cawthorne and Robbins-van Wynsberghe (2019)	Yes	Yes	Yes	Yes	Yes	Yes

Table 3 (continued)

Authors	Is the source of the opinion clearly identified?	Does the source of opinion have standing in the field of expertise?	Are the interests of the relevant population the central focus of the opinion?	Is the stated position the result of an analytical process, and is there logic in the opinion expressed?	Is there reference to the extant literature?	Is any incongruence with the literature/sources logically defended?
Van Wynsberghe (2016)	Yes	Yes	Yes	Yes	Yes	Yes
CPAIS (2019)	Yes	Yes	Yes	Yes	No	N/A
Abramoff et al. (2020)	Yes	Yes	Yes	Yes	Yes	Yes
Hasenauer et al. (2019)	Yes	Yes	Yes	Yes	Yes	Yes
Poulsen and Burmeister (2019)	Yes	Yes	Yes	Yes	Yes	Yes
McCradden et al. (2020)	Yes	Yes	Yes	Yes	Yes	Yes
Ienca et al. (2017)	Yes	Yes	Yes	Yes	Yes	Yes
D. Peters et al. (2020)	Yes	Yes	Yes	Yes	Yes	Yes
McBride (2020)	Yes	Unclear	Yes	Yes	Yes	Yes
Carter et al. (2020)	Yes	Yes	Yes	Yes	Yes	Yes
Dantas et al. (2017)	Yes	Yes	Yes	Yes	Yes	Yes

Appendix 4: Article characteristics

See Tables 4, 5 and 6.

Table 4 Characteristics of articles

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Abramoff et al. (2020)	Clinical Decisions Support Systems (CDSS)	Autonomous CDSS for diabetic retinopathy screening	Deployed	Private and academic institute	Bioethics	Approaches: User eXperience (UX), ergonomics, emphasis on the AI model design to meet ethics principles, HIPAA/HITECH requirements for data stewardship. Clinical validation and comparison with a human system. Motives: Foster trust in the system. Stakeholders consulted: designers, developers Challenges: Competing interests between medical and engineering experts during design and development; tendency to reduce ethics to fairness Motives: Fostering trust Stakeholders consulted: Designers and developers

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Anderson et al. (2019)	Robot	Autonomous elderly care robot	Deployed	Open-Source Software academic project	Machine ethics based on cases from ethical dilemmas resolved by ethicists	Approaches: Robot ethics is embedded in the Operating System (OS) using case-supported principle-based behaviour paradigm. The robot behaviour is compared with the consensus of a group of ethicists for a give case Motives: humanistic Stakeholders consulted: ethicists and clients
Battistuzzi et al. (2018)	Robot	Social Assistive Robot for elderly	Not deployed	Academic European Union project	Alzheimer Europe's Guidelines and Position on the Ethical Use of Assistive Technologies for/by People with Dementia, Value-Sensitive Design	Approaches: Values Sensitive Design (VSD), iterative, mapping tasks from scenarios from optimal to possible with a technical description including ethical themes Challenges: tension between ethics principles of autonomy, safety, dignity, and the impracticality of ethics guidelines because of their lack of context Motives: humanistic Stakeholders consulted: ethicists and clients

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Beil et al. (2019)	CDSS	CDSS for Intensive Care patients based on predictions of the course of the illness	Theoretical	Not specified	Bioethics and European ethics guidelines for trustworthy Artificial Intelligence (AI)	Approaches: checklist with medical, technical, patient-centred, and system-centred sections, and dos and don'ts Challenges: Tension between justice, autonomy because of the futility of treatment, between autonomy and beneficence regarding the understanding of the decisions, between beneficence and non-maleficence with the pros and cons of predictive values, between justice and non-maleficence because of biases in datasets, and justice and beneficence regarding explicability. Lack of regulation to guide the decision-making process with AI Motives: Trust for adoption of the system Quote: "Of note, a fully autonomous decision-making system could serve justice best at the level of society. However, such an approach would violate the principle of human autonomy and be ignorant of empathy that is considered a fundamentally human trait." Stakeholders consulted: Clients, patients, medical staff, clinicians and developers

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Buston et al. (2019)	ML generic	Any Machine Learning (ML) Digital Health application	Theoretical	Not specified	NHSX and Future Advocacy guidelines	Approaches: Stakeholders analysis to be conducted by developers including direct and indirect users, how they will be consulted, the cultural context and the environment within which the app will operate. Value and consequences matrices to capture requirements and concerns. These two activities should be repeated at fixed times, and the process should be made transparent to the public, in an open-source manner Challenges: Not addressed Motives: Trust building, and also evolve the process into a regulatory framework Stakeholders consulted: direct and indirect users
Carter et al. (2020)	CDSS	CDSS for breast cancer detection	Theoretical	Not specified	e.g. IDEAL framework	Approaches: Evidence-based introduction of complex intervention such as IDEAL Challenges: Not addressed Motives: Not addressed Stakeholders consulted: clinicians, patients and health system stakeholders at large

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Cawthorne and Robbins-van Wynsberghe (2019)	Drone	Autonomous drone transporting blood samples between point of care and lab location	Not deployed	Academic, public and private partnership	Bioethics, VSD	Approaches: VSD and values hierarchy for a combination of bottom-up and to-down approaches plus abductive approach with an ongoing evaluation for cocreation, commodification as values evolve Challenges: Conflicts between technology and human values, translating abstract precepts into actionable items Motives: humanistic Stakeholders consulted: direct and indirect users, clinicians, developers, designers…
Cawthorne and Robbins-van Wynsberghe (2020)	Drone	Autonomous drone transporting blood samples between point of care and lab location	Not deployed	Academic, public and private partnership	Bioethics and AI explicability AI4People framework, Value-Sensitive Design (VSD)	Approaches: VSD Challenges: Manage unintended consequences Motives: Not addressed Stakeholders consulted: drones' community, critics, impacted stakeholders: Not addressed

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Channa et al. (2020)	CDSS	Autonomous CDSS for diabetic retinopathy screening	Deployed	Private and academic institute	Principles of Safety, Efficacy, and Equity (SEE) as described by the American Medical Association (AMA)	Approaches: Assess the performance of the AI system within the context of safety, efficacy, and equity as per the AMA Augmented Intelligence in Health guidelines Challenges: Not addressed Motives: FDA Approval Stakeholders consulted: Clinicians, policy makers, developers
Cooney and Menezes (2018)	Robot	Art therapy robot	Not deployed	Academic project	Code of Professional Practice by the Art Therapy Credentials Board	Approaches: Subject the robot to the same code of conduct as a therapist, translate the requirements into capabilities, then pilot it Challenges: No one way of doing art therapy as such difficult to codify, interface of multiple disciplines (engineering, art therapy, social science etc.), how to assess the success of the art therapy sessions Motives: Not addressed Stakeholders consulted: Not addressed

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
CPAIS (2019)	Virtual bot, Intelligent Assistive Technology (IAT), CDSS	Mental health chatbot, Brain Computer Interface (BCI) for Motor Neuronal Disease (MND) patients, Magnetic Resonance Imagery (MRI) CDSS	Deployed	Private companies and private and academic partnership	Partnership on AI	Approaches: Set of 36 questions designed to instigate a dialogue between stakeholders and expand the horizon of the developers to include the broader society implications Challenges: The answer to the questions can be binary Motives: Create a dialogue Stakeholders consulted: Not addressed Quote: "By thinking through this list, I will have a better sense of where I am responsible to make the tool more useful, safe, and beneficial for the people using it. The public can also be better assured that I took these parameters into consideration when working on the design of a system that they may trust and then embed in their everyday life." Software Engineer, PAI Research Participant

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Dantas et al. (2017)	Robot	Care robot for the elderly	Deployed	European project with private and academic partners	Verification Methodology for Ethical Compliance	Approaches: Verification of ethical compliance in three steps (1) legal analysis and principles review through literature review, (2) use cases and services built by collecting data on expectations, needs, habits, (3) verification matrix for ethical compliance used iteratively tracking each requirement from each use case completed by each partner. Additionally, there should be an oversight ethics committee of the project, Challenges: Lack of harmonisation among standards in different countries at times incompatible with European Union standards Motives: Safeguarding the teleology of the common good Stakeholders consulted: end-users, developers and researchers
Fukuzumi et al. (2019)	CDSS	CDSS for determining whether to see or not a doctor	Theoretical	Not specified		Approaches: Collect issues from private enterprise doing business plan, issue guideline, and verify validity of guidelines with end-users through a survey Challenges: Not addressed Motives: Not addressed Stakeholders consulted: Not addressed

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Garner et al. (2016)	Robot	Virtual care bot for elderly	Theoretical	Not specified	User-derived principles	Approaches: Literature review and focus groups to extract relevant ethical issues which are then mapped to functionality using UX and technology design approaches Challenges: Fostering trust to increase adoption of the technology, and lack of legislation to help guide safe design which could be alleviated through transparency of processes and goals Motives: Humanistic Stakeholders consulted: care givers, a reference group for end-users, service providers
Hasenauer et al. (2019)	Robot	Social Assistive Robots	Theoretical	Academic and mixed public/private research institute partnership	Empirical on-field research	Approaches: Ethical assessment performed before adoption in order to reconfigure the organisation management. First stage is assessed quantitively and qualitatively for the level of desired care, the cost/benefit analysis performed at a later stage once reorganisation done Challenges: Acceptance of the technology characterised by anxiety, attitude of the personal and patients, balancing pros and cons from an ethical perspective, Motives: For robots to be adopted, ethical concerns need to be addressed Stakeholders consulted: Not addressed

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Ienca et al. (2017)	IAT	Neuro-assistive and Rehabilitation Technology for patients with nervous system diseases or injuries	Theoretical	Various	Bioethics	Approaches: VSD with a purpose to improve quality of life, technological approach to address risk reduction, safety and privacy issues, though the latter are usually remedial Challenges: Affordability of the systems, interdependence and privacy Motives: Not addressed Stakeholders consulted: clients, informal carers

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Ienca et al. (2018)	IAT	IAT for dementia patients care	Deployed	Various as it is a review of IATs	Bioethics	Approaches: Proactive ethics framework to anticipate potential issues and unintended consequences through cooperative, participatory and contextual design and VSD. It is a process; values are context dependent. Use of CYBATHLON competition to promote and showcase the approach Challenges: Imbalance of power between designers and end-users, and competing goals between stakeholders, affordability which could be achieved through just reimbursement policies but also affordability to countries with lower level of economic development, translating technology from lab to bedside and community Motives: Humanistic Stakeholders consulted: Ethicists, developers, end-users, clinicians, and designers

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Joerin et al. (2020)	Virtual bot	Mental health chatbot	Deployed	Private company (X2AI Inc.)	APA General principles for Psychologists and UK House of Lords Report AI Code principles	Approaches: Combined ethics frameworks from psychology and AI fields. At organisational level, relies on advisory board, and advice from industry as well as academic experts. Alignment of purpose of harm reduction required from advisors. Transparency and accessibility of terms of use and privacy policy for end-users. Strong stewardship of data. Ethos is that accountability rests in the hands of the creator. Training in ethics part of onboarding procedure at the company, investment in training and dialogues fostering ethics. Monitoring for biases and collecting feedback from users and respect for cultural differences through customisation Challenges: Silos between fields of engineering and psychology and maintaining values alignment between advisory groups Motives: Humanistic, and trust to foster adoption Stakeholders consulted: core service experts, clients, research partners

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Klein and Schlömer (2018)	Robot	A robotic shower system to support independent living in older age	Deployed	Academic European Union project	MEESTAR (Model for the Ethical Evaluation of Socio-Technical ARrangements)	Approaches: MEESTAR methodology including interviews, focus groups, tools to guide questions and identify ethical concerns. It is a discursive approach. The process is iterative starting with small group and pilot Challenges: Dialogue with primary users can be challenged by their disabilities. Conflicts between caregivers and patient's wishes. Conflicts between ethical principles such as privacy and autonomy for example. Discussions around ethics in focus groups can be difficult due to language/concept/vocabulary issues Motives: Humanistic and technology adoption Stakeholders consulted: Patients, caregivers

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Kortner (2016)	Robot	Social Service Robot for the elderly	Deployed	European consortium of universities and one private tech company		Approaches: Continuous improvement of the system, close monitoring and data protection thanks to technological solutions. Experience gathered from the STRAND project Challenges: What is legal is not necessarily ethically sound or feasible such as the informed consent for any person entering in contact with a robot. Preserving dignity is a challenge, and ethics are contextual Motives: Humanistic Stakeholders consulted: Elderly, developers and other clients Quote: "Ultimately, ethical considerations in robotics for senior citizens should not only ask what view we have on technology but also what image of humanity prevails in robotics."

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Kretzschmar et al. (2019)	Virtual bot	Mental health chatbot	Deployed	Various	Oxford Neuroscience, Ethics and Society Young People's Advisory Group (NeurOx YPAG)	Approaches: UX, data stewardship asking for complete transparency about the usage of collected data Challenges: Tension between privacy and beneficence, non-maleficence because of the need for AI for data to become safer. Possibly a mental health chatbot could reinforce the sense of isolation of clients because of its lack of human contact. Ability of chatbot to understand subtle complex sentences from clients Motives: humanistic, trust and privacy for adoption Stakeholders consulted: Youth age 14–18
McBride (2020)	Robot	Social Assistive Robot for autistic children	Not deployed	Academic and private enterprise project	ACTIVE ethics: An information systems ethics for the Internet age	Approaches: ACTIVE ethics framework and deep understanding of the autistic child Challenges: Not addressed Motives: Humanistic Stakeholders consulted: client, autistic child, designers

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
McCradden et al. (2020)	ML generic	Any ML Digital Health application	Theoretical	Not specified		Approaches: Data-centred, use statistical analysis of datasets to help clinicians assess whether the AI model is suitable to their patients, create different models to adapt to different populations, train clinicians to understand CDSS results and how to communicate the results for decision making Challenges: When and by whom quality control should be performed Motives: Inequalities in healthcare pose an ethical threat to the core values of health institutions Stakeholders consulted: clinical staff, hospital decision makers, researchers
Milosevic (2019)	ML generic	Any ML Digital Health application	Theoretical	Not specified	Australia's ethics guidelines for AI applications	Approaches: Open Distributed Processing Enterprise Language to represent how ethics constraints are coded into the AI system. Ethics principles are considered as constraints of the system which are modelled as obligations, permissions or prohibition then translated into behavioural constraints. The concept of deontic token is used to model the dynamics between parties Challenges: Explicability Motives: Not addressed Stakeholders consulted:

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Morch et al. (2019)	Virtual bot	Mental health application	Theoretical	Not specified	Medical guidelines, the European Society of Human Reproduction and Embryology's Best Practices in the Design of Medical Guidelines	Approaches: Checklist comprised of series of questions about funding, conflict of interest, objectives, privacy, transparency, security, risks, biases Challenges: It is not a regulatory or a legal framework Motives: Humanistic Stakeholders consulted: Developers

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Peters et al. (2020)	Virtual bot	Mental health application	Theoretical	Not specified	IEEE Ethically Aligned Design, IEEE PS7000 and Association for Computing Machinery ethical standards	Approaches: A combination of Responsible Design Approach, VSD (based on UK Design Council guidelines) and Responsible Innovation embedding ethics and wellbeing, and Spheres of technology experience Challenges: Unintended consequences such as mental health of developers being exposed to disturbing data during labelling process. Potential lack of empathy from the system, privacy issues, and translating principles into practical functions Motives: Humanistic Stakeholders consulted: Multidisciplinary team of experts, striving to consult as many stakeholders as possible
Poulsen et al. (2018)	Robot	Care robots for the elderly	Theoretical	Academic project	Combination of Robot ethics, machine ethics and care ethics	Approaches: Values in Motion Design. There are some principled ethics embedded in the robot at design time, but the robot keeps learning and adapt to the values of the user over time Challenges: Dynamic values trade-offs performed at runtime Motives: Not addressed Stakeholders consulted: Robot experts, ethicists, caregivers and users

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Poulsen and Burmeister (2019)	Robot	Care robot for the elderly	Theoretical	Not specified	Care ethics, the attentive framework	Approaches: The attentive framework where decisions are made in context and not fixed by design. The values trade-offs are constantly being updated by the AI system through continuous learning Challenges: For such a framework to work, it requires computational consciousness, and intrinsic values are achieved through extrinsic values Motives: Humanistic Stakeholders consulted: clients
Rajkomar et al. (2018)	ML generic	ML application for healthcare providers	Theoretical	Private company (Google)	Bioethics	Approaches: Technological approach for equal outcomes, performance and allocation. Determine the goal of a ML system and review it with diverse stakeholders. Decide which groups are classified as protected. Check for historical care disparities in data. Careful evaluation of datasets and model for fairness within desired context Challenges: The three equity principles are at odds with each other by nature of the technology Motives: Not addressed Stakeholders consulted:

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Sendak et al. (2020)	CDSS	Assists clinicians in the early diagnosis and treatment of sepsis	Deployed	Not specified	The Technology Acceptance Model	Approaches: Build trust and accountability without resorting to ML explicability: define the problem in context, build relationships with stakeholders, respect professional discretion, and create ongoing feedback loops with stakeholders. It involves onsite observations and consultations of different stakeholders at development time and focus groups Challenges: Trust in technology is rooted in relationships not in technical features and functions, over-reliance and deskilling of the clinicians, augment rather than replace existing and treatment processes Motives: Not addressed Stakeholders consulted: nurses, physicians, developers, data engineers, statisticians, solution architects, UX designers, administrators

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Stahl and Coeckelbergh (2016)	Robot	Care robots for autistic children	Theoretical	Academic European Union project	Framework for responsible innovation	Approaches: Responsible Research and Innovation (RRI) which includes a set of questions designed to enquire about the purpose, the process, the product and people at different levels from anticipation to action, dialogues between stakeholders Challenges: Interdisciplinary collaboration, be less theoretical and closer to action when doing research, funding silos Motives: Responsible innovation needs to have ethics embedded into the design Stakeholders consulted: patients, caregivers, healthcare professionals, and community members, striving for a broad outreach of stakeholders

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
van Rysewyk and Pontier (2015)	Robot	Machine simulating a full ethical agent	Not deployed	Academic project	Hybrid "top-down" normative theories (rule-utilitarianism and prima-facie deontological ethics) and "bottom-up" (case-based reasoning) machine ethics	Approaches: CCVSD (Care-Centred Value Sensitive Design), understand care in context is the starting point and involves onsite observations and start with assessment of where is there a relevant need for robot assistance-consultation Challenges: How to manage the principle of responsibility when delegating a task/practice to a robot is challenging and can be alleviated by making the chain of responsibility explicit in the design Motives: Not addressed Stakeholders consulted: Ethicists assisting developers and designers with moral deliberations

Table 4 (continued)

Authors and year	Type of application	Short description	Status of the application	Technology providers	Ethics framework characteristics	Synopsis
Van Wynsberghe (2016)	Robot	Care robots for assisting nurses in paediatric oncology department	Theoretical	Not specified	Care-Centred Value Sensitive Design (CCVSD)	Approaches: Technology-based, coding consequentialist and deontological ethics using different types of AI, and continuously improving the model. Combination of rule-based, and case-based algorithms coding ethical behaviours which are then simulated with prototypical scenarios from medical ethics. The testing is done by comparing "textbook" behaviours with the resolution offered by the machine Challenges: The challenge is for this machine/technical ethical agent to match the complexity of the human ethical decisions Motives: Not addressed Stakeholders consulted: Not addressed

Table 5 Distribution of ethics principles – yes means it is included, while a blank means it is not

Authors and year	Autonomy & Freedom	Non-maleficence	Beneficence	Fairness & Justice	Dignity	Responsibility	Trust	Privacy	Transparency	Sustainability	Solidarity	Other
Abramoff et al. (202)0	Yes	Yes	Yes	Yes	Yes	Yes			Yes			
Anderson et al. (2019)	Yes	Yes	Yes						Yes	Yes	Yes	
Battistuzzi et al. (2018)	Yes	Yes	Yes		Yes		Yes	Yes				Yes
Beil et al. (2019)	Yes	Yes	Yes	Yes				Yes	Yes			
Buston et al. (2019)	Yes	Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes	Yes
Carter et al. (2020)	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes		
Cawthorne and Robbins-van Wynsberghe (2019)	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes		Yes
Cawthorne and Robbins-van Wynsberghe (2020)	Yes	Yes	Yes	Yes				Yes	Yes	Yes		
Channa et al. (2020)		Yes		Yes				Yes			Yes	Yes
Cooney and Menezes (2018)	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes			Yes
CPAIS (2019)	Yes	Yes	Yes			Yes	Yes	Yes	Yes	Yes		Yes
Dantas et al. (2017)												
Fukuzumi et al. (2019)												
Garner et al. (2016)	Yes	Yes		Yes	Yes		Yes	Yes	Yes		Yes	Yes
Hasenauer et al. (2019)	Yes	Yes	Yes	Yes	Yes	Yes		Yes				Yes
Ienca et al. (2017)	Yes	Yes	Yes	Yes				Yes				Yes
Ienca et al. (2018)	Yes	Yes	Yes	Yes								Yes
Joerin et al. (2020)		Yes	Yes	Yes			Yes	Yes	Yes			Yes
Klein and Schlömer (2018)	Yes	Yes	Yes	Yes	Yes	Yes		Yes				Yes
Kortner (2016)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes			Yes

Table 5 (continued)

Authors and year	Autonomy & Freedom	Non-maleficence	Beneficence	Fairness & Justice	Dignity	Responsibility	Trust	Privacy	Transparency	Sustainability	Solidarity	Other
Kretzschmar et al. (2019)	Yes	Yes	Yes			Yes		Yes	Yes		Yes	
McBride (2020)	Yes						Yes		Yes			Yes
McCradden et al. (2020)		Yes		Yes		Yes			Yes			Yes
Milosevic (2019)	Yes	Yes		Yes		Yes		Yes	Yes			Yes
Morch et al. (2019)		Yes		Yes			Yes	Yes	Yes			Yes
Peters et al. (2020)	Yes	Yes	Yes	Yes				Yes	Yes			
Poulsen et al. (2018)	Yes				Yes			Yes				Yes
Poulsen and Burmeister (2019)												
Rajkomar et al. (2018)		Yes	Yes	Yes		Yes	Yes		Yes			
Sendak et al. (2020)						Yes	Yes		Yes			
Stahl and Coeckelbergh (2016)		Yes				Yes	Yes	Yes			Yes	Yes
van Rysewyk and Pontier (2015)	Yes	Yes	Yes									
Van Wynsberghe (2016)		Yes	Yes			Yes						Yes

Table 6 Distribution of lifecycle stages—yes means it is included, while a blank means it is not

Authors and year	Requirements capture	Design	Development	Testing, Quality Assurance	Deployment	Evaluation	Commercialisation Procurement
Abramoff et al. (2020)	Yes	Yes		Yes		Yes	Yes
Anderson et al. (2019)	Yes	Yes	Yes	Yes			
Battistuzzi et al. (2018)	Yes	Yes				Yes	
Beil et al. (2019)		Yes		Yes	Yes	Yes	
Buston et al. (2019)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Carter et al. (2020)							
Cawthorne and Robbins-van Wynsberghe (2019)	Yes	Yes	Yes		Yes	Yes	
Cawthorne and Robbins-van Wynsberghe (2020)		Yes		Yes		Yes	Yes
Channa et al. (2020)		Yes	Yes	Yes		Yes	Yes
Cooney and Menezes (2018)	Yes	Yes					
CPAIS (2019)			Yes			Yes	
Dantas et al. (2017)	Yes	Yes		Yes		Yes	
Fukuzumi et al. (2019)						Yes	
Garner et al. (2016)	Yes	Yes					
Hasenauer et al. (2019)	Yes	Yes			Yes	Yes	Yes
Ienca et al. (2017)		Yes	Yes	Yes			
Ienca et al. (2018)	Yes	Yes	Yes	Yes	Yes	Yes	
Joerin et al. (2020)	Yes	Yes	Yes	Yes		Yes	
Klein and Schlömer (2018)	Yes	Yes	Yes	Yes		Yes	
Kortner (2016)				Yes	Yes		
Kretzschmar et al. (2019)	Yes	Yes	Yes	Yes		Yes	
McBride (2020)		Yes	Yes		Yes		
McCradden et al. (2020)					Yes	Yes	
Milosevic (2019)							

Table 6 (continued)

Authors and year	Requirements capture	Design	Development	Testing, Quality Assurance	Deployment	Evaluation	Commercialisation Procurement
Morch et al. (2019)		Yes					
Peters et al. (2020)	Yes	Yes	Yes	Yes		Yes	Yes
Poulsen et al. (2018)		Yes					
Poulsen and Burmeister (2019)	Yes	Yes	Yes	Yes			
Rajkomar et al. (2018)		Yes	Yes	Yes	Yes	Yes	
Sendak et al. (2020)	Yes	Yes	Yes		Yes		
Stahl and Coeckelbergh (2016)		Yes					
van Rysewyk and Pontier (2015)		Yes	Yes	Yes			
Van Wynsberghe (2016)	Yes	Yes		Yes		Yes	

Acknowledgements Ms Mary Simons contributed to the development of the search strategy and Professor Wendy Rogers and Associate Professor Farah Magrabi to its refinement. Dr Catalin Tufanaru provided guidance on the scoping review process and the selection of risk of bias tools for critical appraisal.

Author Contributions MG, EA, RC-W conceived the review, MG conducted the search. MG, EA, RC-W screened titles and abstracts and full texts. MG, EA, RC-W extracted data and undertook the critical appraisal. MG undertook the qualitative analysis. MG wrote the introduction and the discussion. EA, RC-W revised the first draft of the paper and the final draft. All authors approved the final version.

Funding None.

Availability of Data and Material All data relevant to the study are included in the article of uploaded as supplementary information.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Abramoff, M. D., Tobey, D., & Char, D. S. (2020). Lessons learned about autonomous AI: Finding a safe, efficacious, and ethical path through the development process. *American Journal of Ophthalmology, 214*, 134–142.

Anderson, M., Anderson, S. L., & Berenz, V. (2019). A value-driven eldercare robot: Virtual and physical instantiations of a case supported principle-based behavior paradigm. *Proceedings of the IEEE, 107*(3), 526–540.

Battistuzzi, L., Papadopoulos, C., Papadopoulos, I., Koulouglioti, C., Sgorbissa, A., Maciejewski, A. A., et al. (2018). Embedding ethics in the design of culturally competent socially assistive robots. *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2018*, 1996–2001.

Beauchamp, T. L., & Childress, J. F. (2013). *Principles of biomedical ethics* (7th ed.). Oxford University Press.

Beil, M., Proft, I., van Heerden, D., Sviri, S., & van Heerden, P. V. (2019). Ethical considerations about artificial intelligence for prognostication in intensive care. *Intensive Care Medicine Experimental, 7*(1), 70.

Bird, E., Fox-Skelly, J., Jenner, N., Larbey, R., Weitkamp, E., & Winfield, A. (2020). *The ethics of artificial intelligence: Issues and initiatives*. Panel for the Future of Science and Technology.

Buston, O., Strukelj, N., & Fenech, M. (2019). Explainability in Data-Driven Health and Care Technology. Future Advocacy.

Carter, S. M., Rogers, W., Win, K. T., Frazer, H., Richards, B., & Houssami, N. (2020). The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. *Breast, 49*, 25–32.

Cawthorne, D., & Robbins-van Wynsberghe, A. (2019). From healthdrone to frugaldrone: Value-sensitive design of a blood sample transportation drone. *2019 IEEE International Symposium on Technology and Society*.

Cawthorne, D., & Robbins-van Wynsberghe, A. (2020). An ethical framework for the design, development, implementation, and assessment of drones used in public healthcare. *Science and Engineering Ethics*.

Channa, R., Wolf, R., & Abramoff, M. D. (2020). Autonomous artificial intelligence in diabetic retinopathy: From algorithm to clinical application. *Journal of Diabetes Science and Technology*.

Commision, A. H. R. (2019). Human Rights and Technology: Discussion Paper. Australian Human Rights Commision.

Cooney, M. D., & Menezes, M. L. R. (2018). Design for an art therapy robot: An explorative review of the theoretical foundations for engaging in emotional and creative painting with a robot. *Multimodal Technologies and Interaction, 2*(3).

CPAIS. (2019). Human-AI collaboration framework & case studies. *Collaborations between People and AI Systems (CPAIS)*. USA: Partnership on AI.
Crawford, K., Dobbe R., Dryer, T., Fried, G., Green, B., Kaziunas E., et al. (2019). AI Now 2019 Report. AI Now Institute.
Dantas, C., Balhau, P., Jegundo, A., Santos, L., Christophorou, C., Wings, C., et al. (2017). Verification methodology of ethical compliance for users, researchers and developers of personal care robots. In *5th World conference on information systems and technologies, WorldCIST* (Vol. 570, pp. 754–762).
Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*.
Fukuzumi, S., Jinno, M., Inagaki, K., Maeda, H., Mizukami, T., Sakura, O., et al. (2019). Extraction of new guideline items from the view point of ELSI (Ethics, Legal, Social Issues) for service utilized AI–focus on healthcare area. In *Thematic area on human interface and the management of information, HIMI 2019, held as part of the 21st International conference on human-computer interaction, HCI International 2019* (Vol, 11570, pp. 482–491).
Garner, T. A., Powell, W. A., & Carr, V. (2016). Virtual carers for the elderly: A case study review of ethical responsibilities. *Digital Health, 2*.
Hasenauer, R., Belviso, C., & Ehrenmueller, I. (2019). New efficiency: Introducing social assistive robots in social eldercare organizations. In *2nd Annual IEEE international symposium on innovation and entrepreneurship of the IEEE technology and engineering management society, TEMS-ISIE 2019*.
Ienca, M., Kressig, R. W., Jotter, F., & Elger, B. (2017). Proactive ethical design for neuroengineering, assistive and rehabilitation technologies: The cybathlon lesson. *Journal of Neuroengineering and Rehabilitation, 14*.
Ienca, M., Wangmo, T., Jotter, F., Kressig, R. W., & Elger, B. (2018). Ethical design of intelligent assistive technologies for dementia: A descriptive review. *Science and Engineering Ethics, 24*(4), 1035–1055.
Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399.
Joerin, A., Rauws, M., Fulmer, R., & Black, V. (2020). Ethical artificial intelligence for digital health organizations. *Cureus, 12*(3), e7202.
Joshi, I., & Morley, J. (2019). Artificial Intelligence: How to get it right. Putting policy into practice for safe data-driven innovation in health and care. In I. Joshi, & J. Morley (Eds.). NHSX.
Klein, B., & Schlömer, I. (2018). A robotic shower system: Acceptance and ethical issues. *Zeitschrift für Gerontologie und Geriatrie, 51*(1), 25–31.
Kortner, T. (2016). Ethical challenges in the use of social service robots for elderly people. *Zeitschrift Für Gerontologie und Geriatrie, 49*(4), 303–307.
Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., Singh, I., & NeurOx Young People's Advisory, G. (2019). Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (Chatbots) in mental health support. *Biomedical Informatics Insights, 11*, 1.17822E+17815.
McBride, N. (2020). Robot enhanced therapy for autistic children: An ethical analysis. *IEEE Technology and Society Magazine, 39*(1), 51–60.
McCradden, M. D., Joshi, S., Anderson, J. A., Mazwi, M., Goldenberg, A., & Zlotnik Shaul, R. (2020). Patient safety and quality improvement: Ethical principles for a regulatory approach to bias in healthcare machine learning. *Journal of the American Medical Informatics Association : JAMIA*.
Metzinger, T. (2019). Ethics washing made in Europe. *Tagesspiegel*, Retrieved 9, Sep, 2020 from https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html
Milosevic, Z. (2019). Ethics in digital health: A deontic accountability framework. In *23rd IEEE international enterprise distributed object computing conference, EDOC 2019* (pp. 105–111).
Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097.
Morch, C. M., Gupta, A., & Mishara, B. L. (2019). Canada protocol: An ethical checklist for the use of Artificial Intelligence in suicide prevention and mental health arXiv. *arXiv*.
Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Science and Engineering Ethics, 26*(4), 2141–2168.
NHS. (2019). The Topol review: Preparing the healthcare workforce to deliver the digital future.
Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan: A web and mobile app for systematic reviews. *Systematic Reviews*.

Peters, D., Vold, K., Robinson, D., & Calvo, R. A. (2020). Responsible AI-two frameworks for ethical design practice. *IEEE Transactions on Technology and Society, 1*(1), 34–47.
Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Baldini Soares, C. (2015). Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation, 13*(3), 141–146.
Poulsen, A., & Burmeister, O. (2019). Overcoming carer shortages with care robots: Dynamic value trade-offs in run-time. *Australasian Journal of Information Systems, 23*.
Poulsen, A., Burmeister, O. K., & Tien, D. (2018). A new design approach and framework for elderly care robots. In *29th Australasian conference on information systems, ACIS 2018*.
Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and Technology, 7*(4), 351–367.
Rajkomar, A., Hardt, M., Howell, M. D., Corrado, G., & Chin, M. H. (2018). Ensuring fairness in machine learning to advance health equity. *Annals of Internal Medicine, 169*(12), 866–872.
Sendak, M., Elish, M. C., Gao, M., Futoma, J., Ratliff, W., Nichols, M., et al. (2020). "The human body is a black box": Supporting clinical decision-making with deep learning. In *3rd ACM conference on fairness, accountability, and transparency, FAT* 2020* (pp. 99–109).
Sparrow, R., & Hatherley, J. (2019). The promise and perils of Al in medicine. *International Journal of Chinese & Comparative Philosophy of Medicine, 17*(2), 79–109.
Stahl, B. C., & Coeckelbergh, M. (2016). Ethics of healthcare robotics: Towards responsible research and innovation. *Robotics and Autonomous Systems, 86*, 152–161.
Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56.
van Rysewyk, S. P., & Pontier, M. (2015). A hybrid bottom-up and top-down approach to machine medical ethics: Theory and data. *Machine Medical Ethics, 74*, 93–110.
Van Wynsberghe, A. (2016). Healthcare robots: Ethics, design and implementation. *Healthcare Robots: Ethics, Design and Implementation*, 1–152.
Villani, C. (2018). For a meaningful Artificial Intelligence: Towards a French and European strategy. Aiforhumanity.fr.