Medical artificial intelligence ethics: A systematic review of empirical studies

Lu Tang[1] iD, Jinxu Li[1] and Sophia Fantus[2]

Abstract

Background: Artificial intelligence (AI) technologies are transforming medicine and healthcare. Scholars and practitioners have debated the philosophical, ethical, legal, and regulatory implications of medical AI, and empirical research on stakeholders' knowledge, attitude, and practices has started to emerge. This study is a systematic review of published empirical studies of medical AI ethics with the goal of mapping the main approaches, findings, and limitations of scholarship to inform future practice considerations.

Methods: We searched seven databases for published peer-reviewed empirical studies on medical AI ethics and evaluated them in terms of types of technologies studied, geographic locations, stakeholders involved, research methods used, ethical principles studied, and major findings.

Findings: Thirty-six studies were included (published 2013-2022). They typically belonged to one of the three topics: exploratory studies of stakeholder knowledge and attitude toward medical AI, theory-building studies testing hypotheses regarding factors contributing to stakeholders' acceptance of medical AI, and studies identifying and correcting bias in medical AI.

Interpretation: There is a disconnect between high-level ethical principles and guidelines developed by ethicists and empirical research on the topic and a need to embed ethicists in tandem with AI developers, clinicians, patients, and scholars of innovation and technology adoption in studying medical AI ethics.

Introduction

Artificial intelligence (AI) has shown exceptional promise in transforming healthcare, shifting physician responsibilities, enhancing patient-centered care to provide earlier and more accurate diagnoses, and optimizing workflow and administrative tasks.[1,2] Yet, the application of medical AI into practice has also created a wave of ethical concerns that ought to be identified and addressed.[3] Global and regional legal regulations, including recent guidelines issued by the WHO[4] and European Union,[5] provide guidance on how to stay abreast of an ever-changing world and placate growing concerns and worries about the moral impacts of medical AI on the provision of healthcare delivery.

Legal regulations are informed by theoretical/conceptual ethics models that serve to facilitate decision-making when it comes to applying medical AI in education, practice, and policy.[6] Medical AI ethics has been supported by the four principles of biomedical ethics (autonomy, beneficence, nonmaleficence, and justice) and ethical values related to

[1]Department of Communication and Journalism, Texas A&M University, College Station, TX, USA
[2]School of Social Work, University of Texas at Arlington, Arlington, TX, USA

Corresponding author:
Lu Tang, Department of Communication and Journalism, Texas A&M University, College Station, TX 77843, USA.
Email: ltang@tamu.edu

fairness, safety, transparency, privacy, responsibility, and trust.[3,7–9] These principles and values inform legal regulations, organizational standards, and ongoing policies in terms of how to focus on and mitigate the ethical challenges that impede the use of medical AI.

However, ethical frameworks that support medical AI legal regulations and guidelines are often developed without dialogue between ethicists, developers, clinicians, and community members.[6,10] This may mean that regulations used to inform best practices may not align with community members' experiences as medical AI users. The ethical concerns identified by policymakers and scholars may not be the same ones identified by actual patients, providers, and developers. This creates a disconnect whereby ethical decision-making tools may not be perceived as applicable to actual users of medical AI. Regulative practices are often informed by higher-level ethical theories and principles rather than examining actual knowledge and practices of medical AI ethics with real-world examples.[9,11,12] Identifying empirical work on the knowledge, attitudes, and practices of medical AI ethics can lead to mechanisms that may assist educators, researchers, and ethicists in clarifying and addressing perceived ethical challenges.[13,14] Medical AI interactions involve patients, families, and healthcare providers, and it is, therefore, critical to understand individuals' awareness of the ethical issues that may impact their AI utilization in health delivery and determine how these results can inform medical AI development and research. In identifying what stakeholders (e.g. patients, clinicians, and developers) see as the ethical risks of medical AI, we can initiate dialogue and develop practice guidelines, organizational standards, and legal regulations to determine ethically informed AI interventions that are relevant and applicable to practice.

The goal of this systematic review is to assess and bridge the gap between the theoretical frameworks of AI ethics and the actual stated worries and concerns of present and future stakeholders of medical AI. To that end, we provide an overview of the existing empirical studies on medical AI ethics in terms of their main approaches, findings, and limitations. Findings from our study can inform medical education and lead to novel strategies to translate medical AI ethics into public discourse and legislation guidance.

Method

Eligibility criteria

This study was guided by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).[15] We examined empirical studies in medical AI ethics published in journals and conference proceedings after 2000. Specifically, the inclusion criteria for this study were the following six: (1) must be empirical studies, (2) must be published in English, (3) must be peer-reviewed, (4) must be published in a journal or conference proceeding, (5)

must be published from January 2000 to June 2022, and (6) must be a study on medical AI, where ethics must be a substantial part of the investigation. Empirical studies referred to qualitative or quantitative studies driven by scientific observation or experimentation. By substantial, we meant that either the study was solely focused on ethics or at least one RQ or hypothesis was focused on ethics. Studies that only mentioned ethics a couple of times in a discussion of the findings as an afterthought were excluded.

Data sources and search strategy

The articles for this study were retrieved in June 2022. We searched a total of seven databases in the fields of information science, medicine, sociology, and communication: IEEE Xplore, PubMed/MEDLINE, CINAHL Complete, JSTOR, Sociological Abstracts (ProQuest), ACM Digital Library, and Communication Source (EBSCOhost). Based on the exploratory literature search, we used three categories of search terms: medical terms (medical, health, healthcare, and medicine), AI (machine learning, algorithm, AI, robotics, pattern recognition, automated problem solving, and natural language), and ethics (ethics, bias, bioethics, ethical, bioethical, autonomy, beneficence, nonmaleficence, justice, and responsibility).

We limited the initial search to abstracts and obtained a total of 5854 articles from the seven databases. First, the second author checked all initial articles for duplicates and removed 708 duplicated titles. Second, articles written in a language other than English, not published (such as those from preprint repositories), and published before 2000 ($n = 372$) were removed. Next, the abstracts of all remaining articles were reviewed, and pieces that were nonempirical and unrelated to medical AI ethics were removed ($n = 4722$), yielding 52 valid abstracts. The first and second authors both read the 52 full articles and identified 30 articles that met the inclusion criteria. The third author served as a tie-breaker when an agreement was not reached. The references of the above 30 articles were further reviewed, and an additional six articles meeting the inclusion criteria were identified. In the end, 36 articles were retained for this systematic literature review. Figure 1 shows the article inclusion flow diagram.

Coding procedure

In reviewing these articles, we coded the demographic information of each article, including the year of publication, journal/conference of publication, and discipline of the said journal/conference. Next, we coded the stakeholders studied in each study, including AI developers, clinicians, patients, patient family members, and other stakeholders. The type of medical AI technology studied in each article was also coded. Ethical principles studied were another variable
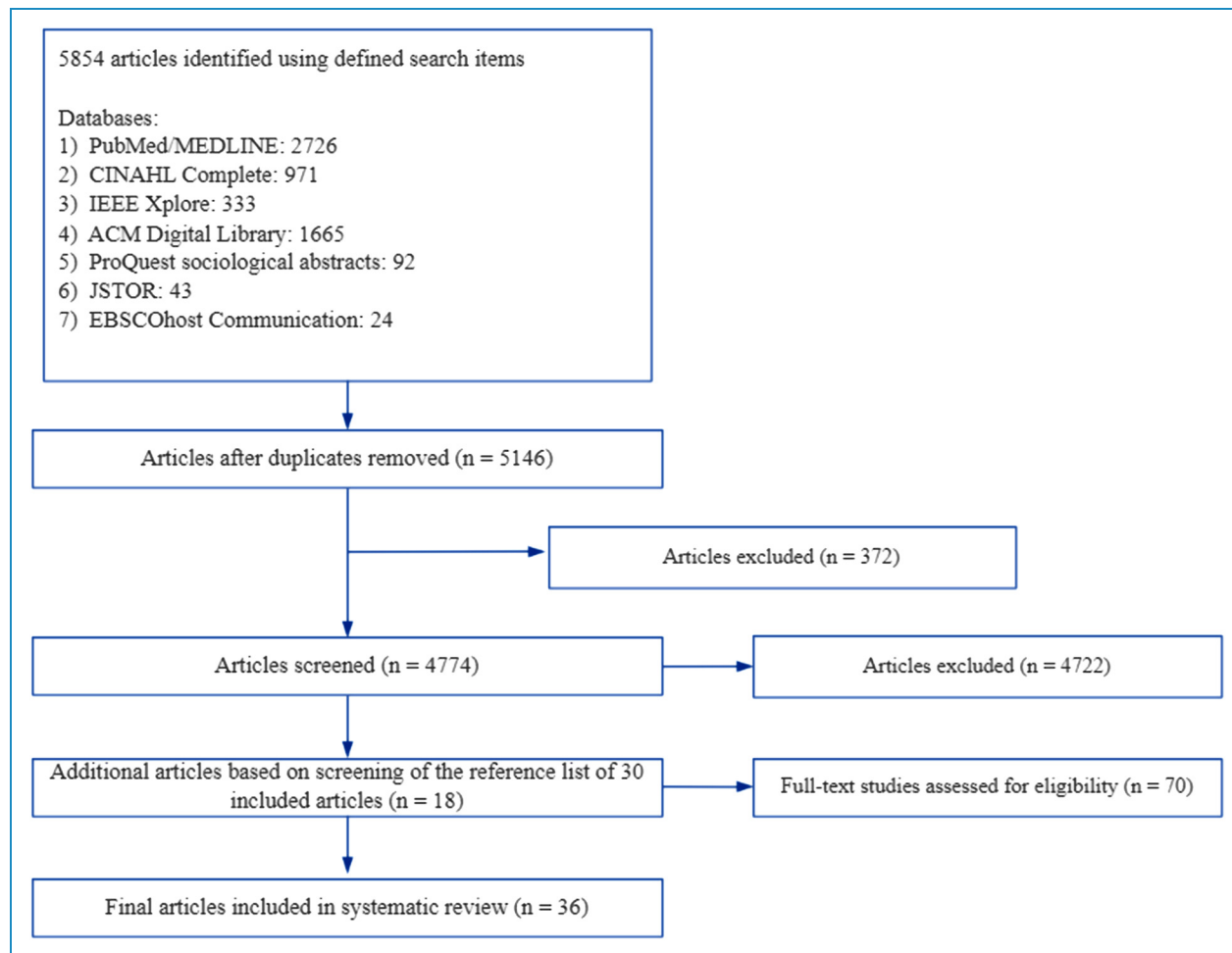
Figure 1. Systematic reviews flow diagram illustrating literature selection process.

coded. We started with five principles used in data collection, including autonomy, beneficence, nonmaleficence, justice, and responsibility. We coded whether each of the studies examined any of these principles. At the same time, we allowed additional principles to emerge from the data. Finally, we coded the research method used and the main findings of each article. The first two authors coded these articles independently and compared their results. They discussed the discrepancy to reach a consensus. When an agreement was not reached, the third author served as an arbitrator. See Table 1 for a summary of these studies.

Results

The 36 studies included in this systematic review were published between 2013 and 2022, with the most published during or after 2020 ($n = 27$), showing the topic's novelty. In terms of disciplines, the largest number of articles were published in journals of medicine ($n = 18$), followed by informatics ($n = 11$), science ($n = 2$), public health ($n = 2$), medical ethics (n = 2), and others (n = 1).

The three main types of AI technologies studied in these articles were risk prediction,[16–21] diagnosis and screening,[22–33] and intelligent assistive technology.[20,34–38]

Almost all studies were conducted based on data from Western countries, with the most from the USA. Other countries included Canada,[39,40] the UK,[22,34,41,42] France,[27,38,43] Denmark,[36,44] Germany,[33] Spain,[35] and Australia.[28] One study collected data from both North America and Asia,[45] and another study included data from three European countries (Switzerland, Germany, and Italy).[20] One study used data from nine Western countries.[34] Two non-Western countries/regions appeared in our collection of studies: sub-Saharan Africa[18] and India.[19,46]

These were categorized into three themes: (1) exploratory studies of stakeholder knowledge and attitude toward medical AI and ethics, (2) theory-building studies testing hypotheses regarding factors contributing to stakeholders' acceptance or adoption of medical AI, and (3) studies measuring and correcting biases in medical AI. We summarized the main findings of each of these categories of studies in Table 2.

Table 1. Overview of empirical studies on medical AI ethics.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Studies of stakeholder knowledge and attitude toward medical AI						
(1). Patients and family members						
Kendell et al., 2020	Using AL for early identification of rapidly deteriorating health and the need for advance care planning	Medicine	Privacy and access	Semi-structured interviews of patients with declining health and their caregivers ($n = 14$)	Canada	Patients and caregivers are mostly supportive of EMR-based algorithm. Their main ethical concerns are as follows: privacy, patients without primary care physicians, not all deaths can be predicted, and added workload for physicians.
Musbahi et al. 2021	Medical AI in general	Medicine	Security, bias, data quality, human interaction, and responsibility	Nominal group technique (four groups with seven participants each)	The UK	80% participants felt AI shouldn't be used to manage health without the involvement of a doctor. Top 5 benefits are faster health services, greater accuracy in management, available 24/7, reduced workforce burden, and equality in healthcare decision-making. Top 5 concerns: data cybersecurity, bias and quality of AI data, less human interaction, algorithm error, and limitation in technology
Nelson et al. 2020	Using director-to-patient and clinician decision-support AI tools to classify skin lesions	Medicine	Autonomy, human interaction, and privacy	Semi-structured interviews of skin cancer patients ($n = 48$)	The USA	Patients identified several major benefits: diagnostic speed, healthcare access, reduced cost, reduced patient anxiety (less waiting), increased triage efficiency, fewer unnecessary biopsies, increase patient self-advocacy, stimulate technology, and higher patient autonomy. Patients also identified several major risks: increase patient anxiety, loss of social interaction, and patient loss of privacy.

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Ongena et al. 2020	AI in radiology	Medicine	Accountability, patient engagement, human interaction, and patient autonomy	Survey and cognitive interviews of patients scheduled for CT, MRI, and/or conventional radiography ($n = 155$)	The USA	Exploratory factor analysis identified five factors in patients' attitude toward AI in radiology: 1. distrust and accountability (overall, patients were moderately negative on this subject) 2. procedural knowledge (patients generally indicated the need for their active engagement) 3. personal interaction (overall, patients preferred personal interaction) 4. efficiency (overall, patients were ambiguous on this subject) 5. being informed (overall, scores on these items were not outspoken within this factor).
Tran et al. 2019	Wearable device and AI in healthcare	Medicine	Security and privacy	Survey and case experiment with patients with chronic conditions ($n = 1183$)	France	20% patients considered that the benefits of these technologies greatly outweighed the dangers. Three percent considered the risks greatly outweighed the benefits. The main risks included: inadequately replace human intelligence in care, risk of hacking, and misuse of private patient data.

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Valles-Peris et al. 2021	AI and robotic system in healthcare	Public health	Human interaction and responsibility	Interviews with patients who have been hospitalized for COVID 19 (n = 13)	Spain	Interviews identified several issues: the empirical effects of imagined robots, the vivid experience of citizens with the care crisis, the discomfort of the ineffective, the virtualized care assemblages, the human-based face-to-face relationships, and the automatization of healthcare tasks. Patients used a "well-being repertoire" in talking about their personal experiences, which treated robots as a threat to good care due to a lack of human interaction. In contrast, they used the "responsibility repertoire" to discuss the health system, which considered automatization via robots and AI as positive and desirable.
Clinicians						
Blease et al. 2019	Future technology, including AI	Informatics	Accountability and safety	Survey of general practitioners (n = 720)	The UK	GPs expressed three types of comments regarding future technology: the limitations of future technology (such as lack of empathy and communication, clinical reasoning, and patient-centeredness), potential benefits of future technology (e.g. improved efficiencies), and social and ethical concerns. Specifically, social and ethical concerns encompass many divergent themes such as the need to train more doctors in response to medical workforce shortage; the uncertainty about whether such technologies will be accepted by patients, safety and accountability.

(continued)

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Bourla et al. 2018	AI-based CDSSs	Medicine	Security and privacy	Mixed method (focus group and survey of psychiatrists ($n = 515$))	France	Psychiatrists studied expressed moderate acceptance of CDSSs and demonstrated a lack of knowledge about such technologies. They raised many medical, forensics, and ethical issues, including therapeutic relationship, data security, data storage, and privacy risk.
Chen et al. 2021	AI innovation in radiology	Medicine	Autonomy	Interviews of radiologists ($n = 12$) and radiographer ($n = 16$) Focus group: radiographer ($n = 8$)	The UK	Radiologists believe that AI has the potential to take on more repetitive tasks and allow them to focus on more interesting and challenging work. They are less concerned that AI technology might constrain their professional role and autonomy. Radiographers showed greater concern about the potential impact that AI technology could have on their roles and skills development. They were less confident of their ability to respond positively to the potential risks and opportunities posed by AI technology.
Hendrix et al. 2021	AI in breast cancer screening	Informatics	Responsibility (against AI autonomy)	Experiment on primary care provider ($n = 91$)	The USA	The study identified three classes of respondents based on their views of medical AI: "Sensitivity First" (41%) found sensitivity to be more than twice as important as other attributes; "Against AI Autonomy" (24%) wanted radiologists to confirm every image; "Uncertain Trade-Offs" (35%) viewed most attributes as having similar importance. A majority (76%) accepted the use of AI in a "triage" role that would allow it to filter out likely negatives without radiologist confirmation.

(continued)

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Lim et al. 2022	AI in diagnostic medical imaging reports	Medicine	Responsibility, privacy, and security	Survey of nonradiologist clinicians ($n = 88$)	Australia	Nonradiologist clinicians believed that medicolegal responsibility with errors in AI-issued reports mostly lay with hospitals or health service providers (65.9%) and radiologists (54.5%). Regarding data privacy and security, nonradiologist clinicians felt significantly less comfortable with AI issuing image reports instead of radiologists ($P < 0.001$).
Morgenstern et al. 2021	AI in general	Public health	Bias and fairness and equity	Semi-structured interviews of experts in public health and AI ($n = 15$)	North America and Asia	Experts are cautiously optimistic about AI's effect on public health, especially on disease surveillance. Barriers to adoption such as confusion regarding AI's applicability, limited capacity, and poor data quality and risks such as propagation of bias, exacerbation of inequity, hype, and poor regulation.
Nichol et al. 2021	Machine learning–based prediction of HIV/AIDS	Medicine	Privacy, power disparities, alignment and conflicts of interests, benefit-sharing, stigma, and bias	Interviews of experts in informatics, African public health, and HIV/AIDS and bioethics ($n = 22$)	Sub-Saharan Africa	Sub-Saharan experts identified not only general ethical concerns regarding AI such as bias but also region-specific ethical concerns related to history of human rights abuses, lack of trust in government and non-African researchers, misuse of research funding, and US researchers' duty to help build research capacity in Sub-Saharan Africa. Stigma related to HIV/AIDS was highlighted as an ethical challenge.

(continued)

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Parikh et al. 2022	ML prognostic algorithms in the routine care of cancer patients.	Medicine	Bias and human interaction	Interviews of medical oncologists and advanced practice providers (*n* = 29)	The USA	Clinicians were concerned about biases in the application of predictions from ML prognostic algorithms. Notably, clinicians worried that such algorithms could reinforce confirmation bias (only applying the information when it confirms their intuition) or automation bias (overly trusting the mortality prediction while ignoring other sources of information, including their own intuition). They also identified ethical challenges in sharing mortality predictions with patients, especially automatic sharing of predictions based on EHR without contexts provided by clinicians.
Peronard 2013	Robot technology	Computer science	Equality and autonomy	Survey of clinicians (*n* = 200)	Denmark	This article is a comparative analysis between workers in healthcare with high and low degree of readiness for living technology such as robotics. The results showed important differences between high- and low-readiness types on issues such as staff security, documentation, autonomy, and future challenges.

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Wangmo et al. 2019	Intelligent assistive technology and wearable computing	Medical ethics	Autonomy, affordability, human interaction, and privacy	Interviews of researchers and health professionals ($n = 20$)	Switzerland, Germany, and Italy	Within the framework of autonomy, participants discussed informed consent and deception. Access to data and data sharing were discussed in relation to the principle of privacy and autonomy. Affordability is discussed as a concern with fair access.
(3) Multiple stakeholders						
Cresswell et al. 2018	Healthcare robotics	Informatics	Autonomy	Interviews with stakeholders from a wide range of backgrounds ($n = 21$)	Nine Western countries	The study identified four major barriers in the application of healthcare robotics: lack of clear incentive from users (clinicians and patients), the appearance of robots (too robotic or too human), changes to the structures of healthcare work, and new ethical and legal challenges. The absence of existing ethical framework is a barrier to the adoption.
Lai et al. 2020	AI in healthcare	Medicine	Social justice, responsibility, and privacy	Interviews with diverse stakeholders ($n = 40$)	France	Healthcare professionals concentrated on providing the best and safest care for their patients and are not always seeing the use of AI tools in their practice. Healthcare industrial partners consider AI to be a true breakthrough but recognize that legal difficulties to access individual health data could hamper its development. Institutional players emphasize their responsibilities on regulating the use of AI tools in healthcare. From an external point of view, individuals without a conflict of interest have significant concerns about the sustainability of the balance between health, social justice,

(continued)

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
						and freedom. Health researchers specialized in AI have a more pragmatic point of view and hope for a better transition from research to practice.
McCradden et al. 2020	AI in healthcare	Medicine	Transparency, fairness, privacy, and human involvement in decision-making	Interviews with meningioma patients ($n = 18$), caregivers ($n = 7$), and healthcare providers ($n = 5$)	Canada	The majority of participants endorsed nonconsented use of health data but advocated for disclosure and transparency. Few patients and caregivers felt that allocation of health resources should be done via computerized output, and a majority stated that it was inappropriate to delegate such decisions to a computer. Almost all participants felt that selling health data should be prohibited, and a minority stated that less privacy is acceptable for the goal of improving health. Certain caveats were identified, including the desire for deidentification of data and use within trusted institutions.
Muller et al. 2021	AI for dental diagnostics	Medicine	Responsibility	Interviews with dentists ($n = 8$) and patients ($n = 5$)	Germany	Both dentists and patients expressed hope for AI. Factors enabling the implementation of AI in dental diagnostics include higher diagnostic accuracy, a reduced workload, more comprehensive reporting, and better patient-provider communication. Barriers included over reliance on AI and responsibility for medical decisions and the explainability of AI.

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Thenral 2020	AI-enabled telepsychiatry platforms for clinical practice	Medicine	Privacy/confidentiality, ethical violations, security/ hacking, and data ownership	Interviews with psychiatrists ($n=$ 14), their patients ($n=14$), technology experts ($n=13$), and CEOs ($n=5$) of health technology incubation centers	India	Almost all respondents cited ethical, legal, accountability, and regulatory implications as challenges. The major issues stated by patients were privacy/ confidentiality, ethical violations, security/ hacking, and data ownership. Psychiatrists cited lack of clinical validation, lack of established studies or trials, iatrogenic risk, and healthcare infrastructure issues as the main challenges. Technology experts stated data-related issues as the major challenge. The CEOs quoted the lack of interdisciplinary experts as one of the main challenges in building deployable AI-enabled telepsychiatry in India.
Studies creating theoretical models of medical AI adoption						
Alaiad and Zhou 2014	HHR	Informatics	Privacy and autonomy	Survey of patients and healthcare professionals from home healthcare agencies ($n=108$)	The USA	Usage intention was a function of social influence, performance expectancy, trust, privacy concerns, ethical concerns, and facilitating conditions. Among them social influence was the strongest predictor. Ethical concerns were measured with three questions: I am concerned that (1) HHRs would change their behaviors over time, (2) HHRs would replace doctor's work in the future, (3) about the errors/ mistakes that would be made by HHRs.

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Esmaeilzadeh 2020	AI medical devices with clinical decision support features	Informatics	Privacy fairness	Survey of potential users (patients, $n = 307$)	The USA	The study examined how technological, ethical, and regulatory concerns predict perceived risks and intention to use AI-based tools. Among them, technological concerns (such as AI performance) are the strongest predictor of perceived risks. The only ethical concern significantly predict perceived risks is "perceived mistrust in AI."
Esmaeilzadeh et al. 2021	AI application in healthcare	Informatics	Privacy and fairness	Experiment on potential users (patients, $n = 634$)	The USA	The interactions between the types of healthcare service encounters (AI only, hybrid of AI and physicians, and physicians only) and health conditions (acute or chronical) significantly influenced individuals' perceptions of privacy concerns, trust issues, communication barriers, concerns about transparency in regulatory standards, liability risks, benefits, and intention to use. Patients facing acute illnesses are more concerned about the privacy of their personal health information, when AI and physician were involved, compared to physician only situations. There were no significant differences among scenarios regarding perceptions of performance risk and social biases.
Ploug et al. 2021	AI decision-making in healthcare	Informatics	Responsibility, explainability, and bias (discrimination)	Focus group and survey/ experiment of representative adult population ($n = 1027$)	Denmark	Responsibility, explainability and discrimination (tested for bias) are the three most important predictors of preference for AI use in hospitals.

(continued)

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
Sisk et al. 2020	AI-driven precision medicine technologies in pediatric care	Medicine	Privacy, human contact and fairness (social justice)	Survey of parents of pediatric patients ($n = 804$)	The USA	Factor analysis was used to identify seven concerns of parents in considering these technologies: quality/accuracy, privacy, shared decision-making, convenience, cost, human elements of care, and social justice. Among these issues, quality, convenience, and cost were positively associated with openness to AI technologies, while perceived importance of shared decision-making was negatively associated with openness. Mothers were less likely to be open to these technologies.
Measuring and identifying bias						
Borgese et al. 2021	Unhealthy alcohol use screening among hospitalized patients	Informatics	Fairness/bias	NLP	The USA	The study reports underestimation among Hispanics compared to non-Hispanic White patients, 18- to 44-year-old compared to 45 years and older medical/surgical inpatients, and non-Hispanic Black compared to non-Hispanic White medical/surgical inpatients.
Chen et al. 2019	Prediction of ICU mortality and 30-day psychiatric readmission	Medical ethics	Fairness/bias	Case study using ML	The USA	ICU model: female and public insurance patients have a higher model error rate than male and private insurance patients. Psych readmission: Black patients have the highest error rate. Private insurance patients have higher error rate.
Estiri et al. 2022	Prediction of COVID hospital admission, ICU	Informatics	Fairness/bias	Machine learning pipeline for modeling health outcomes	The USA	On the model level, inconsistent instances of bias were found. On the individual

(continued)

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
	admission, ventilation use, and death					level, all models reported higher error rates for older patients. The models marginally performed better for female and Latinx patients and worse for male patients.
Larrazabal et al. 2020	Using X-ray for diagnosis	Science	Fairness/bias	Deep neural network	The USA	A consistent decrease in performance for underrepresented genders when a minimum balance is not fulfilled.
Noseworthy et al. 2020	Diagnosis of low LVEF using CNN	Medicine	Fairness/bias	Retrospective cohort analysis	The USA	This study used a CNN model developed based on non-Hispanic White on a diverse population and found that it performed equally well in other racial and ethnic groups.
Obermeyer et al. 2019	Commercial prediction algorithms for identifying patients with complex health needs	Science	Fairness/bias	Compare predicted risk score to insurance claim data	The USA	At the same risk score, Black patients have more illness burdens than Whites. This is because risk score is based on cost prediction instead of illness prediction but unequal access to care means that we spend less money caring for Black patients than for White patients.
Seyyed-Kalantari et al. 2021	Underdiagnosis in chest X-ray pathology	Medicine	Fairness/bias	Deep learning model	The USA	This study focuses on algorithmic underdiagnosis in chest X-ray pathology classification across three large chest X-ray data sets, as well as one multisource data set. It finds that these technologies consistently and selectively underdiagnosed under-served patient populations and that the underdiagnosis rate was higher for intersectional under-served

Table 1. Continued.

Author and year	AI tech studied	Discipline	Ethical concerns	Method	Country	Results
						subpopulations, for example, Hispanic female patients.
Straw and Wu 2022	Liver disease prediction	Informatics	Fairness/bias	Experiments measuring the bias caused by sex-unbalanced data in four ML models (random forests, support vector machines, Gaussian naive Bayes, and logistic regression)	India	The study showed performance disparity where women suffer from higher false negative rates.
Other studies						
Frost and Carter 2020	Screening and diagnostic AI	Informatics	Morality, scientific uncertainty, and governance	Content analysis of media articles about AI (n = 136)	Study conducted in Australia but source of data unknown	Media tend to frame medical AI in terms of social progress, economic development, and alternative perspectives. Most of the coverage of AI is positive with very few articles discussing the ethical, legal, and social implications.
Yi et al. 2020	Data set used for deep learning models	Medicine	Fairness	An analysis images in publicly available chest radiograph data sets ($n = 23$)	The USA	Most data sets reported demographics in some form (19 of 23; 82.6%) and on an image level (17 of 23; 73.9%). The majority reported age (19 of 23; 82.6%) and sex (18 of 23; 78.2%), but a minority reported race or ethnicity (2 of 23; 8.7%) and insurance status (1 of 23; 4.3%). Of the 13 data sets with sex distribution readily available, the average breakdown was 55.2% male subjects, ranging from 47.8% to 69.7% male representation. Of these, 8 (61.5%) overrepresented male subjects and 5 (38.5%) overrepresented female subjects.

Abbreviations: AI, artificial intelligence; CDSSs, clinical decision support systems; CEOs, Chief Executive Officers; CNN, convolutional neural network; CT, computed tomography; EHR, electronic health record; EMR, electronic medical records; HHR, home healthcare robot; ICU, intensive care unit; LVEF, ventricular ejection fraction; ML, machine learning; MRI, magnetic resonance imaging; NLP, Natural language processing.

Studies on stakeholder knowledge and attitude toward medical AI and ethics (n = 21)

The largest group of empirical studies on medical AI ethics explored different stakeholders' attitudes toward medical AI and their ethical concerns through an inductive approach. Very few of the studies[20,39] were focused exclusively on ethics. The remaining studies included ethics as one dimension in stakeholder attitudes toward medical AI in general and often framed ethics broadly in terms of concerns, barriers, and potential challenges. Participant stakeholders included clinicians (such as physicians, radiologists, and dentists), patients, patient family members, and occasionally researchers and governmental, NGO, or industry leaders. Data collection methods included administered surveys, focus groups, or interviews. Approaches to data analysis were informed by either descriptive statistics (surveys) or qualitative data analysis methods (focus groups, interviews), including grounded theory and thematic analysis to identify patterns and themes associated with stakeholders' ethical concerns.

Overall, stakeholders in these studies were most concerned with the ethical impacts of medical AI on human connection, autonomy, privacy, data security, and fairness. Clinicians generally reported more ethical concerns related to the use of medical AI than patients. Human connection, predominantly in the form of the clinician–patient relationship, was the most frequently identified ethical issue by both clinicians and patients. Many studies showed that clinicians and patients believed that personal touch and human relationships facilitated ethical communication. This was especially true when explaining medical diagnoses and treatment options. The stated ethical concerns with medical AI from stakeholders are that it will depersonalize medicine, eliminate the need for human interaction, and reduce investment in patient-centered care.

Overall, six studies examined the attitudes of patients and their family members toward medical AI ethics. Surveys and interviews relying primarily on open-ended questions were administered to examine patients' or family members' attitudes toward medical AI and the perceived benefits and risks of incorporating AI. These studies typically used different medical scenarios to understand stakeholder perspectives, including advanced care planning,[40] skin cancer diagnosis,[32] radiology,[31] and COVID-19 hospitalization.[35] Overall, patients and family members expressed moderate support for medical AI, while identifying ethics as a major barrier/concern in accepting medical AI. Major ethical concerns included responsibility,[22,35] privacy,[32,38,40] data security,[22] bias and accuracy,[22] and lack of human interactions.[22,31,32,35] Overall, patients and family members expressed concern that AI technologies would disengage physicians from the healthcare process, demonstrating a preference for physician involvement in diagnosis, decision-making, and

Table 2. Main findings from the three themes of empirical studies on medical AI ethics.

Three types of studies	Main findings
Stakeholders' knowledge and attitude toward medical AI and ethics (*n* = 21)	• Patients and family members expressed concern that AI technologies would disengage physicians from the healthcare process. • Clinicians shared concerns about responsibility, data security, privacy, bias, human interaction/effective clinician–patient communication. • Autonomy was another common ethical concern identified by clinicians including patient autonomy and clinician autonomy. • Different stakeholders held diverse views toward medical AI and had different ethical concerns.
Creating theoretical models of medical AI adoption (*n* = 5)	• Ethics have a small effect on medical AI adoption. • Ethical concerns included: privacy, mistrust of AI, transparency, responsibility, fairness, autonomy/shared decision-making, and explainability.
Identify and correct bias in medical AI (*n* = 8)	• Most studies identified AI underperformance for underserved populations, such as women, racial and ethnic minorities, and patients on public insurance, but the results were inconsistent overall. • Highlighted the importance of testing the AI performance across different demographic groups during each step of algorithm development in future projects.

clinical communication. These results demonstrate that both clinicians and patients have stated ethical concerns about the impacts of medical AI on trustworthy relationships and transparent and informed communication that facilitate healthcare decisions.

Ten articles studied how clinicians, including physicians, psychiatrists, radiologists, radiographers, and other experts, perceived the use of AI technology in medicine and their ethical concerns. Like patients and family

members, these professionals shared concerns about responsibility,[28,29,43] data security,[20,27,28] privacy,[18,27,28] bias,[18,45,47] and human interaction/effective clinician–patient communication.[20,36,41,47] In addition, clinicians were also concerned with equality,[36] equity,[45] safety and accountability,[41] stigma and discrimination,[18] and affordability (distributive justice).[20] These findings may demonstrate that clinicians have a deeper understanding of the ways in which medical AI may have disproportionate impacts on social justice. This may mean that ethical concerns related to social justice, discrimination, and equity may arise from more knowledge on the uses of and impacts of medical AI in practice. Autonomy was another common ethical concern identified by clinicians, including patient autonomy and clinician autonomy. Medical professionals often worried that AI would infringe on their professional autonomy in making decisions or even the future of their specialty.[42] Patient autonomy referred to whether patients were mentally capable of giving informed consent to utilize medical AI in their care.[20] Occasionally, the term autonomy was used without any definition or explanation.[36]

Five studies examined multiple stakeholders simultaneously.[33,34,39,43,46] They often sought to provide an overview of the public opinion toward medical AI. Stakeholders studied included patients, caregivers, clinicians, academic researchers, industry leaders, and journalists. These studies typically found that different stakeholders held diverse views toward medical AI and had varied ethical concerns. For instance, a study of multiple stakeholders in Canada found that patients were more concerned with fairness in resource allocation, while all stakeholders considered privacy a major concern.[39] Similarly, a study of multiple stakeholders in India showed that patients were more concerned with ethical challenges such as privacy, confidentiality, ethical violations, security, and data ownership.[46] In contrast, CEOs were mostly concerned with the human resource challenges in developing and deploying AI-enabled telepsychiatry, and psychiatrists were more concerned with technical challenges.

Studies creating theoretical models of medical AI adoption (n = 5)

Five studies used surveys or experiments to test theoretically driven hypotheses that explained or predicted the adoption of medical AI. Theories such as technology acceptance theories[37] and the value-based AI acceptance model[48,49] were used to create hypotheses. These studies often utilized fictional vignettes to elicit participants' attitudes toward medical AI and their behavioral intentions. They found that several factors, including social influence,[37] technological concerns,[37,48,50] ethical concerns,[37,44,48–50]

regulatory concerns,[49] and situational factors such as the nature of the illness and tasks,[44,48] were likely to affect the perceived risks of medical AI, preference, and/or intention to adopt. While ethical concerns influenced the acceptance of medical AI or adoption intention, its effect was smaller than other factors, such as social influence[37] and technology concerns.[48] Among the ethical concerns examined, privacy,[37,48] mistrust of AI,[49] transparency,[48] responsibility,[44] fairness (i.e. free from discrimination),[44] autonomy/shared decision-making,[50] and explainability[44] were found to be significantly or nearly significantly related to adoption intention or preference. Overall, these studies suggested that ethics has a small effect on medical AI adoption, and few ethical issues stood out as more important.

Identify and correct bias in medical AI (n = 8)

The last category of studies focused on the ethical principle of fairness by measuring the biases in medical AI algorithms and some experimented with ways of correcting such biases. Even though this group of studies did not collect data from AI researchers per se, their research provided insights into ethical issues that AI researchers perceive as most important to address and mitigate. This line of research also provides empirical validation for the presence of AI biases. Recent research has shown that AI-based predictions and diagnoses were often plagued with biases against certain demographic groups, such as women and racial and ethnic minorities and such biases were likely to perpetuate existing health disparities.[51] Systematic biases and missing data in training data sets (such as electronic health records (EHRs) and insurance claims) contributed to the disparity in AI performance among different demographic groups. Eight articles in the sample tested the performance of AI prediction and diagnosis algorithms across different races/ethnicity, sex, and insurance status by comparing the error rates of diagnoses among different groups or comparing predictions with actual patient data. Most studies identified AI underperformance for underserved populations, such as women,[16,19,23] racial and ethnic minorities,[16,21,23,26] and patients on public insurance.[16,23] However, not all studies found underdiagnosis for underserved populations. For instance, one study showed that a convolutional neural network (CNN) model trained using data from non-Hispanic White patients performed equally well for other racial and ethnic groups in ECG analysis.[24] Another study using EHR data to predict COVID-19 hospital admission, ICU admission, ventilation use, and death showed that their algorithms performed marginally better for female and Latinx patients but worse for older patients and male patients.[17] These results demonstrated the unpredictability in the relationships among demographic variables, biophysical data, and algorithm performance and highlighted the importance of testing the

AI performance across different demographic groups during each step of algorithm development in future projects.

Other

Two studies did not fit into any of the previous three groups. One was a content analysis of media coverage of the application of AI in medical screening and diagnostics. It found that media coverage was overwhelmingly positive, with a very small percentage (<10%) incorporating the ethical, legal, and social implications of such technologies.[25] The other outlier was an examination of the reporting of demographic information in 23 publicly available chest radiograph data sets, as earlier research suggested that unbalanced training data sets might lead to bias in predictions.[52] The study found that while most data sets reported demographic data regarding age and sex, few reported race, ethnicity, and insurance status. Among those data sets that did report such data, certain groups (such as female patients) were often underrepresented. Such underreporting was likely to contribute to the bias in the algorithms trained using such data sets.

Discussion

Our systematic literature review of published empirical studies of medical AI ethics identified the three main approaches taken in this line of research and the major findings in each approach. The largest group of studies examines the knowledge and attitudes of medical AI ethics across various stakeholders. These studies demonstrate that the primary concern shared among patients and clinicians is the impact of medical AI on reducing human interaction and trustworthy healthcare communication. Although patients showed concerns with other ethical issues, clinicians identified principles of autonomy and justice as major considerations in the application of medical AI. Their shared concerns of autonomous decision-making as clinicians with respect to patient care and their responsibility of attaining full informed consent from their patients was an important finding to incorporate into medical education. Additionally, clinicians voiced concerns with respect to equity, discrimination, stigma, and distributive justice. These reported findings indicate that providers are concerned about patients' access to and affordability of healthcare and how this may influence treatment options and care plans.

Another group of studies utilized theoretical models to examine the adoption of medical AI technologies. These empirical findings demonstrate that ethical issues have a significant yet small effect on acceptance and adoption of medical AI in practice. The last group of studies was mathematical and technical in nature to identify and correct biases in the prediction by medical AI based on race, gender, and insurance status. These findings showed that although there are medical AI biases that may perpetuate

existing health disparities, there may be novel ways in which to reduce and mitigate these biases in medical AI algorithms.

Our review of these 36 studies identified a few concerns or limitations in the current academic research on medical AI ethics: a disconnect between high-level ethical principles and the concerns reported by different stakeholders, a lack of research that focuses on ethics of medical AI per se, and a need for consistent definition and operationalization of ethics-related concepts and terminologies. In this section, we will further discuss these limitations and call for truly interdisciplinary collaboration in medical AI ethics research.

Overall our study shows the urgent need to connect high-level ethical principles and guidelines developed by ethicists and international agencies and the empirical research on the topic. There is no empirical research on AI developers' knowledge and attitudes toward ethical issues. We can only deduce that bias and addressing fairness are the concerns of AI developers, as evidenced by the large number of studies aiming at measuring and correcting biases in algorithms. However, there is no existing research on AI developers' knowledge, attitude, and practices regarding AI ethics other than their research to repair AI biases, connected to the ethical principles of fairness and justice. The lack of inclusion of AI developers as integral stakeholders in empirical research could be also caused by the fact that this systematic review focused on published empirical studies in peer-reviewed journals and studies done on AI developers could potentially be found in the grey literature or technical publication. Nevertheless, future research needs to include AI developers as participants to understand their knowledge and practices of medical AI ethics and their ability to connect fears of bias with principles of justice, fairness, and social responsibility. Understanding the perspectives of AI developers can provide insights into the potential ethical challenges in the development of medical AI and create interdisciplinary collaborations to create truly ethical medical AI.

The systematic review also reveals several additional limitations in this body of literature. First, researchers' primary aims are not focused on ethics. Instead, in qualitative, open-ended interviews and focus group studies, ethics emerged naturally in dialogue with stakeholders, showing its importance in stakeholders' evaluation of medical AI. In quantitative surveys, ethics is one of the variables measured in addition to technical, social, legal, and regulatory factors. This can be tentatively attributed to the fact that most of these studies have been conducted by medical researchers who want to explore stakeholder opinions about medical AI in their respective subfields. Ethics scholars are surprisingly absent in the research teams conducting empirical studies of medical AI. Only three authors (1.8%) of these studies are affiliated with an ethics-related department or institution.

Second, we observed significant divergence in the definition and measurement of ethics terminology across studies. Researchers seldom define ethics or ethical principles in their research design, administered surveys, or interview guides with participants. Many ethical principles were mentioned in researchers' analyses without being accompanied by specific conceptual and operational definitions. In interviews and focus group studies, stakeholders (patients or clinicians) mention concerns or challenges of utilizing medical AI technologies that researchers then label as "ethics." In administered surveys, there is no consistency in the items used to measure different ethical concerns and no efforts to validate the scales used. Most of the existing studies rely on open-ended questions, and only a few studies used scales to quantitatively measure these variables[31,38,48–50]; however, the scales used were not validated. Researchers sometimes label items as ethics without context, and consequently, it is not clear whether these items were ethical issues, rooted in value conflict between two ethical courses of action, or whether these were more practice-based issues. For instance, a study of home healthcare robot (HHR) adoption measured ethical concerns using three items: HHR would change their behaviors over time, HHR would replace doctors, and HHR would make mistakes.[37] Furthermore, the same study also measured common ethical concerns such as privacy as a separate category. Another study on patients' perceived risks of AI medical devices measured ethical concerns in three categories: privacy concerns, fairness concerns (social bias), and "mistrust in AI," and finds that only mistrust in AI predicts perceived risks.[49] Sometimes, ethical concerns are intermingled with technical, legal, or regulatory challenges of AI. The lack of universal or shared ethics definition may be the result of limited collaboration with ethicists or lack of intention to focus on ethical concerns. Rather, without proper definitions or categorizations of ethical issues, researchers may be unknowingly confuse ethical issues with practice-based issues that may arise from AI utilization. Without context as to what makes these ethical issues (as opposed to development, practice, or legal issues), it is up to the readers to make sense of these and connect them to ethical principles and values in order to develop ethically informed care. Rather, concerns and challenges are often simply thematically organized as ethics rather than potentially seeing these responses as a comprehensive review of overall risks. This type of analysis may lead to a dearth of standardized and validated measurement tools to capture stakeholders' knowledge and attitude about medical AI and ethics. The lack of validated and common measurement scales prevents the comparison of different studies and reduces the generalizability of their findings and creates difficulty in assessing the stated ethical concerns as opposed to practice-based or generalized risks of incorporating medical AI into practice.

Our findings highlight the need to foster collaborations between ethicists, AI researchers, medical professionals, and scholars of innovation and technology adoption in studying medical AI ethics. Ethics scholars should inform AI developers, clinicians, and patients through each step of the medical AI pipeline. In other words, ethicists should be members of interdisciplinary and interprofessional research groups. We adopt the idea of "embedded ethics" developed by McLennan et al. (2022) to describe the importance of incorporating ethicists and ethical principles in the development, utilization, and understanding of medical AI during ongoing, interactive conversations.[6] The framework proposed by McLennan and colleagues (2022) asserts that there ought to be reciprocal interactions between ethicists and AI developers: ethicists educate AI developers on ethical principles and guidelines, while AI developers provide feedback to ethicists about feasibility, practicality, and process of development. While McLennan's approach focuses on the collaboration between medical AI developers and ethicists, we argue that it is imperative to broaden this model to include reciprocal dialogue with medical AI users, such as clinicians and patients. Medical AI technology is different from other medical technologies; neither AI developers nor clinicians alone can predict or explain the predictions made by medical AI and explain the results to patients.[44] Consequently, to inform AI development and ethics, there ought to be reciprocal dialogue throughout the entirety of the development and implementation process. In turn, we envision three components: (1) medical AI users need to actively participate in AI development conversations along with ethicists and developers, (2) AI developers ought to be included in practice outcomes of medical AI through conversations with patients and clinicians after development, and (3) ethicists need to be involved in the design, development, implementation, and interpretation of medical AI products from development to practice.

In our expanded model, we envision an ongoing and long-term collaboration to inform medical AI utilization and data interpretation. The integration of clinicians and patients into medical AI conversations can contribute to the development of transparent communication regarding informed consent, human interaction, and fairness. Collaborative dialogue between ethicists and clinicians can involve issues related to the decision to use medical AI technologies in the diagnosis and treatments of different types of patients, informed consent, and communication of AI-based diagnoses with patients. Ethicists' interactions with patients may include mitigating issues related to the protection of informed consent or refusal. Ethicists can play the additional role of mediating the conversation between AI developers, clinicians, and patients to create genuinely open and equitable discussion, especially when patients from historically marginalized groups are involved. Finally, by providing their feedback to ethicists, AI

developers, clinicians, and patients can influence organizational ethical guidelines to address reported ethical concerns and obstacles in employing and utilizing medical AI.

Finally, while medical AI has the potential to benefit low- and middle-income countries as well as less developed regions in high-income countries by enabling better public health surveillance and providing more and better healthcare to underserved populations,[4] all except three empirical studies on medical AI ethics were conducted in high-income countries, and most based in large urban hospital settings. It is imperative to engage stakeholders in low- and middle-income countries and regions in developing and implementing medical AI, recognizing the diversity in stakeholder opinions and the existing digital divide as medical AI becomes a global phenomenon.

ORCID iD: Lu Tang https://orcid.org/0000-0002-1850-1511

References

1. Rigby MJ. Ethical dimensions of using artificial intelligence in health care. *AMA J Ethics* 2019; 21: E121–E124.
2. Svensson AM and Jotterand F. Doctor ex machina: a critical assessment of the use of artificial intelligence in health care. *J Med Philos* 2022; 47: 155–178.
3. Gerke S, Minssen T and Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. In: Bohr A and Memarzadeh K (eds) *Artificial intelligence in healthcare*. London, UK: Academic Press, 2020, pp.295–336.
4. *Ethics and governance of artificial intelligence for health: WHO guidance*. World Health Organization, 2021, 165 p.

5. Cannarsa M. Ethics Guidelines for Trustworthy AI [Internet]. The Cambridge Handbook of Lawyering in the Digital Age. 2021:283–297. DOI: 10.1017/9781108936040.022.
6. McLennan S, Fiske A, Tigard D, et al. Embedded ethics: a proposal for integrating ethics into the development of medical AI. *BMC Med Ethics* 2022; 23: 6.
7. Francesca R. *Artificial intelligence: potential benefits and ethical considerations*. Brussels: European Parliament, 2016, 8 p.
8. Keskinbora KH. Medical ethics considerations on artificial intelligence. *J Clin Neurosci* 2019; 64: 277–282.
9. Jobin A, Ienca M and Vayena E. The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 2019; 1: 389–399.
10. Crigger E and Khoury C. Making policy on augmented intelligence in health care. *AMA J Ethics* 2019; 21: E188–E191.
11. Holmes W, Porayska-Pomsta K, Holstein K, et al. Ethics of AI in Education: Towards a Community-Wide Framework. *Int J Artif Intell Educ [Internet]*. 2021. DOI: 10.1007/s40593-021-00239-1.
12. Ryan M and Stahl BC. Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications. *J Inf, Commun Ethics Soc* 2020; 19: 61–86.
13. Borenstein J and Howard A. Emerging challenges in AI and the need for AI ethics education. *AI Ethics* 2021; 1: 61–65.
14. Morley J, Kinsey L, Elhalal A, et al. Operationalising AI ethics: barriers, enablers and next steps. *AI Soc [Internet]*. 2021. DOI: 10.1007/s00146-021-01308-8.
15. Moher D. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement [internet]. *Ann Intern Med* 2009; 151: 64.
16. Chen IY, Szolovits P and Ghassemi M. Can AI help reduce disparities in general medical and mental health care? *AMA J Ethics* 2019; 21: E167–E179.
17. Estiri H, Strasser ZH, Rashidian S, et al. An objective framework for evaluating unrecognized bias in medical AI models predicting COVID-19 outcomes. *J Am Med Inform Assoc* 2022; 29: 1334–1341.
18. Nichol AA, Bendavid E, Mutenherwa F, et al. Diverse experts' perspectives on ethical issues of using machine learning to predict HIV/AIDS risk in sub-Saharan Africa: a modified Delphi study. *BMJ Open* 2021; 11: e052287.
19. Straw I and Wu H. Investigating for bias in healthcare algorithms: a sex-stratified analysis of supervised machine learning models in liver disease prediction. *BMJ Health Care Inform [Internet]* 2022; 29. http://dx.doi.org/10.1136/bmjhci-2021-100457.
20. Wangmo T, Lipps M, Kressig RW, et al. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. *BMC Med Ethics* 2019; 20: 98.
21. Obermeyer Z, Powers B, Vogeli C, et al. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019; 366: 447–453.
22. Musbahi O, Syed L, Le Feuvre P, et al. Public patient views of artificial intelligence in healthcare: a nominal group technique study. *Digit Health* 2021; 7: 20552076211063682.
23. Seyyed-Kalantari L, Zhang H, McDermott MBA, et al. Underdiagnosis bias of artificial intelligence algorithms

applied to chest radiographs in under-served patient populations. *Nat Med* 2021; 27: 2176–2182.

24. Noseworthy PA, Attia ZI, Brewer LC, et al. Assessing and mitigating bias in medical artificial intelligence: the effects of race and ethnicity on a deep learning model for ECG analysis. *Circ Arrhythm Electrophysiol* 2020; 13: e007988.

25. Frost EK and Carter SM. Reporting of screening and diagnostic AI rarely acknowledges ethical, legal, and social implications: a mass media frame analysis. *BMC Med Inform Decis Mak* 2020; 20: 325.

26. Borgese M, Joyce C, Anderson EE, et al. Bias assessment and correction in machine learning algorithms: a use-case in a natural language processing algorithm to identify hospitalized patients with unhealthy alcohol use. *AMIA Annu Symp Proc* 2021; 2021: 247–254.

27. Bourla A, Ferreri F, Ogorzelec L, et al. Psychiatrists' attitudes toward disruptive new technologies: mixed-methods study [internet]. *JMIR Ment Health* 2018; 5: e10240.

28. Lim SS, Phan TD, Law M, et al. Non-radiologist perception of the use of artificial intelligence (AI) in diagnostic medical imaging reports. *J Med Imaging Radiat Oncol [Internet]*. 2022. https://onlinelibrary.wiley.com/doi/abs/10.1111/1754-9485.13388

29. Hendrix N, Hauber B, Lee CI, et al. Artificial intelligence in breast cancer screening: primary care provider preferences. *J Am Med Inform Assoc* 2021; 28: 1117–1124.

30. Larrazabal AJ, Nieto N, Peterson V, et al. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. *Proc Natl Acad Sci U S A* 2020; 117: 12592–4.

31. Ongena YP, Haan M, Yakar D, et al. Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire. *Eur Radiol* 2020; 30: 1033–1040.

32. Nelson CA, Pérez-Chada LM, Creadore A, et al. Patient perspectives on the use of artificial intelligence for skin cancer screening: a qualitative study. *JAMA Dermatol* 2020; 156: 501–512.

33. Müller A, Mertens SM, Göstemeyer G, et al. Barriers and enablers for artificial intelligence in dental diagnostics: a qualitative study. *J Clin Med Res [Internet]* 2021; 10. DOI: 10.3390/jcm10081612.

34. Cresswell K, Cunningham-Burley S and Sheikh A. Health care robotics: qualitative exploration of key challenges and future directions. *J Med Internet Res* 2018; 20: e10410.

35. Vallès-Peris N, Barat-Auleda O and Domènech M. Robots in healthcare? What patients say [internet]. *Int J Environ Res Public Health* 2021; 18: 9933.

36. Peronard JP. Readiness for living technology: a comparative study of the uptake of robot technology in the danish healthcare sector. *Artif Life* 2013; 19: 421–436.

37. Alaiad A and Zhou L. The determinants of home healthcare robots adoption: an empirical investigation. *Int J Med Inform* 2014; 83: 825–840.

38. Tran VT, Riveros C and Ravaud P. Patients' views of wearable devices and AI in healthcare: findings from the ComPaRe e-cohort [Internet]. *npj Digital Medicine* 2019; 2. DOI: 10.1038/s41746-019-0132-y.

39. McCradden MD, Baba A, Saha A, et al. Ethical concerns around use of artificial intelligence in health care research from the perspective of patients with meningioma, caregivers and health care providers: a qualitative study [internet]. *CMAJ Open* 2020; 8: E90–E95.

40. Kendell C, Kotecha J, Martin M, et al. Patient and caregiver perspectives on early identification for advance care planning in primary healthcare settings. *BMC Fam Pract* 2020; 21: 136.

41. 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.

42. Chen Y, Stavropoulou C, Narasinkan R, et al. Professionals' responses to the introduction of AI innovations in radiology and their implications for future adoption: a qualitative study. *BMC Health Serv Res* 2021; 21: 813.

43. Laï MC, Brian M and Mamzer MF. Perceptions of artificial intelligence in healthcare: findings from a qualitative survey study among actors in France. *J Transl Med* 2020; 18: 14.

44. Ploug T, Sundby A, Moeslund TB, et al. Population preferences for performance and explainability of artificial intelligence in health care: choice-based conjoint survey. *J Med Internet Res* 2021; 23: e26611.

45. Morgenstern JD, Rosella LC, Daley MJ, et al. "AI's gonna have an impact on everything in society, so it has to have an impact on public health": a fundamental qualitative descriptive study of the implications of artificial intelligence for public health [Internet]. *BMC Public Health* 2021; 21. DOI: 10.1186/s12889-020-10030-x.

46. Thenral M and Annamalai A. Challenges of building, deploying, and using AI-enabled telepsychiatry platforms for clinical practice among urban Indians: a qualitative study. *Indian J Psychol Med* 2021; 43: 336–342.

47. Parikh RB, Manz CR, Nelson MN, et al. Clinician perspectives on machine learning prognostic algorithms in the routine care of patients with cancer: a qualitative study. *Support Care Cancer* 2022; 30: 4363–4372.

48. Esmaeilzadeh P, Mirzaei T and Dharanikota S. Patients' perceptions toward human-artificial intelligence interaction in health care: experimental study. *J Med Internet Res* 2021; 23: e25856.

49. Esmaeilzadeh P. Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. *BMC Med Inform Decis Mak* 2020; 20: 170.

50. Sisk BA, Antes AL, Burrous S, et al. Parental Attitudes toward Artificial Intelligence-Driven Precision Medicine Technologies in Pediatric Healthcare. *Children [Internet]* 2020; 7. http://dx.doi.org/10.3390/children7090145.

51. Parikh RB, Teeple S and Navathe AS. Addressing bias in artificial intelligence in health care. *JAMA* 2019; 322: 2377–2378.

52. Yi PH, Kim TK, Siegel E, et al. Demographic reporting in publicly available chest radiograph data sets: opportunities for mitigating sex and racial disparities in deep learning models. *J Am Coll Radiol*. 2022; 19: 192–200.