Impact of Artificial Intelligence on Human Experience in Health Care: A Systematic Review

Tahani Ali Salem Maharem
Medical- Surgical Nursing Department, Faculty of Nursing, Alexandria University, Egypt.
Assistant professor of Medical Surgical Nursing, Nursing Department, Faculty of Applied Medical Sciences, Al Baha University, Saudi Arabia.
E- mail: med.sur87@gmail.com or tahani@bu.edu.sa
https://orcid.org/0000-0002-7134-8261
Mobile:966549779639

ABSTRACT

Aim: This systematic review seeks to identify the impact of artificial intelligence on human experience in health care. **Background**: The application of artificial intelligence (AI) in healthcare has given rise to a number of ethical questions as the technology advances and AI becomes more useful. Over the past ten years, there have been notable attempts to reconcile ethical issues with the AI-driven revolution in health. It's still difficult to deploy AI-related technologies and initiatives in healthcare settings ethically, despite the growing interest in AI ethics. **Methods**: A broad range of databases within the fields of legal sciences, social sciences, health-care sciences and the more general sciences practitioner base "Web of Science" were explored. Articles were selected according to strict inclusion/exclusion criteria and systematically analyzed regarding their content and authorship. To address this current issue, we examined 253 papers that summarised the consistent themes of ethical AI initiatives and were published between 2000 and 2020 and related to AI ethics in healthcare.. The systematic literature review was carried out using a hermeneutic approach, and articles were screened and chosen using the Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) approach. **Results**: Through the integration of pertinent insights from the fields of AI governance and ethics, we put forth a framework for responsible AI initiatives that comprises five fundamental themes in AI for policy makers, healthcare practitioners, and AI solution developers. These themes are summarized as following: *Sustainability, Human centeredness, Inclusiveness, Fairness, and Transparency*. **Conclusion**: Identifying the impact of artificial intelligence on human experience in health care.

Keywords: artificial intelligence, human experience, health care, systematic review

Introduction

The field of artificial intelligence (AI) is based on algorithmic computing technology. Its applications include data mining, robotic process automation, machine learning, and artificial neural networks, and self-learning algorithms that enable it to make intelligent predictions and decisions in real time **(Chen and Asch, 2017; Davenport and Kalakota, 2019).** In the healthcare industry, artificial intelligence (AI) is described as the use of smart, data-driven technologies to support and speed up decision-making by more effectively utilising healthcare resources and data, ultimately resulting in more individualised, high-quality healthcare services. To effectively inform healthcare decision-making,, Machine learning algorithms are commonly utilised by AI technologies to conduct "intelligent" analytical and inferential tasks on health data. These tasks include infodemiology, which is the study and prediction of diseases and pandemics, diagnosis and treatment of neurological and chronic illnesses, interpretation of radiological images and medical scans, provision of healthcare services and treatments, drug discovery, and matching eligible patients to clinical trials. Furthermore, AI has the ability to address societal global health-specific challenges **(Mehta et al., 2020)** and expedite the attainment of health and well-being-related Sustainable Development Goals **(Vinuesa et al., 2020).** However, a number of ethical concerns, such as algorithmic bias that results in inconsistent Questions concerning AI's effectiveness in the healthcare sector are being raised by discriminatory outcomes, privacy violations, disputes over data ownership, and a lack of transparency in data use **(Vayena et al., 2018).**

Because of this, Professor Stephen Hawking advised supporters of AI to proceed with caution, stating that the creation of AI would be the greatest event in human history. Regretfully, if we don't figure out how to reduce the risks, it might also be the last. The following are current examples of unethical healthcare AI use: serve as examples of the ethical concerns:

Claimed to be a game-changer in cancer treatment, IBM's Watson supercomputer bases its treatment recommendations on a small number of fictitious cancer cases. As a result, patients may receive dangerous and erroneous medical advice, putting their health and safety at risk **(Ross and Swetlitz, 2018)**. This instance highlights the detrimental effects of *ambiguous processes in data access and use, as well as algorithmic design and explanation.*

According to the UK Information Commissioner's Office (ICO), Google's DeepMind received data from 1.6 million patients without providing them with the appropriate notice or getting their consent **(Powles and Hodson, 2017)**. In this instance, the institutions' and regulators' lack of resilience and adaptability in their responses to data policies has jeopardised patients' autonomy.

Algorithms that forecast future healthcare requirements using surrogate health status metrics run the risk of perpetuating errors and inequities in treatment. For instance, for For a given degree of illness burden, Black patients typically spend, $1100–1800 less annually than White patients. Black patients required just as much coordinated healthcare as White patients, but the algorithm misinterpreted their needs as being less. This algorithm does not account for certain significant a few crucial elements include employment status, hospital accessibility, insurance status, and household incom. **(Ober- meyer et al. 2019).** *Because of "label choice bias," medical decisions in this situation can be difficult and possibly unfair.*

The past ten years have seen a number of discussions on finding a balance between ethical issues and digital transformation in an effort to address the ethical challenges and worries raised by artificial intelligence **(Culnan and Williams, 2009; Zhang and Hon, 2020).** Researchers should examine the effects of According to **Newell and Marabelli (2015),** there is "irresponsible" use of AI-powered analytics on individuals, organisations, and society at large (to inform algorithmic decisions).and communities. This request for research has been addressed by two recent studies: In order to address the ethical challenges of AI, A framework for ethics AI that combines stakeholder theory and social contracts theory has been put forth by **Wright and Schultz (2018).** It develops several best practises, including minimising disruptions, minimising social inequalities, and acknowledging the transition.

Flyverbom et al. (2019) have since emphasised the relationship between digital transformations and responsible business practises. They propose that businesses should use digital technologies like big data analytics to expand their scope beyond profit-making and to proactively address societal issues.

In the healthcare sector, awareness of responsible methods for AI development, application, management, and governance has increased. **(Morley et al., 2020; Peters et al., 2020). Arrieta et al. (2020,)** define responsible AI as "a methodology for the large-scale implementation of AI methods in real organisations with fairness, model explainability, and accountability at its core." This brings us to the idea of responsible AI. Several AI ethics frameworks and guidelines are available, but they are not specifically designed with healthcare in mind **(Morley et al., 2021**) or are very abstract **(Jobin et al., 2019).** Furthermore, reviews that have already been written about responsible AI concentrate on examining the primary developments in AI for healthcare while taking ethical considerations into account using a bibliometric approach **(Wamba and Queiroz, 2021)** or using a thematic review approach to identify ethical issues unique to the application of AI in healthcare **(Morley et al., 2020; Trocin et al., 2021).** Our review goes one step further by offering helpful recommendations on how healthcare organisations can develop an ethical AI project.

Developing AI systems that are safe and morally sound for healthcare organisations can be a difficult task that involves large investments in algorithm management, data governance, and careful consideration of the social and environmental effects of AI. According to **Almai et al. (2020),** Initiating responsible AI requires cross-disciplinary and cross-sectoral collaboration among data scientists, healthcare providers, and policy makers in order to create a sustainable AI ecosystem that benefits all stakeholders and society. Numerous studies from the social sciences, medicine, and healthcare information systems have examined these issues, and computer science disciplines. We aim to provide researchers and practitioners with an understanding of responsible AI initiatives by analysing the existing literature on responsible AI in healthcare. They may encounter when implementing responsible AI.

In conclusion, our goal is to provide a response to the the following research question: What ethical applications for AI-related projects and technologies are there in the healthcare sector? Our review aims to enumerate the critical issues and methodically identify significant themes of responsible AI initiatives related to responsible AI that have not yet received enough attention in order to respond to this research question. We suggest directions for further study on responsible AI in healthcare, based on the conclusions from our review.

Theoretical background

Using virtue ethics theory as a lens to understand responsible AI

Research on AI ethics in healthcare has become more and more popular in recent years. According **to Floridi et al. (2018),** *a number of ethical precepts have been recognised as suitable options for the creation of AI systems. Nevertheless, a lack of a single framework for regulating AI means that a large portion of current AI-driven research lacks the ethical, legal, and practical considerations for widespread adoption* **(Schwalbe and Wahl, 2020). (Morley et al., 2020).**

Even though frameworks for AI ethics have been revised multiple times to account for the complexity of these issues, they still don't offer much guidance on what programmes should be put in place to encourage the ethical use of AI **(Jobin et al., 2019; Morley et al., 2021).** The literature has frequently demanded that an ethical framework for AI governance in healthcare be defined and standardised in order to enable the adoption of responsible AI-based health systems **(Morley et al., 2021).**

Rather than emphasising the appropriateness of an action or its consequences, virtue ethics theory places more emphasis on the virtue or moral character of the person performing the action in a particular circumstance (what will a virtuous person do in a given situation) **(Chatterjee et al., 2009; Audi, 2012).** Various explanations of the virtues that characterise a virtuous person have been offered by virtue ethicists. Virtue ethics can assist managers in making more moral decisions when it comes to management practises **(Audi, 2012).** A firm's daily activities and operations can also be guided by virtue ethics, which can also be used to lower service-related risks and "increase a firm's reputation" and moral standing in the community in which it operates **(Chakrabarty and Bass, 2015, p. 497).**

This idea can be expanded to include health service providers using AI. Because of this, virtue ethics is a suitable theoretical framework for creating a responsible AI initiative framework in the healthcare industry. Previous research indicates that virtues like fairness and honesty can have a positive effect on an agent's actions, whether it be a person or an organisation **(Audi, 2012; Chun, 2005).** Therefore, by conducting a comprehensive literature review, we are able to examine healthcare organisations' ethical AI practises and how those practises relate to AI ethical principles that might be backed by virtue ethics. We concentrate on the six core ethical qualities - fairness, transparency, trustworthiness, accountability, privacy, and empathy - that have been identified in the literature on AI ethics and are thought to be significant and frequently mentioned in the healthcare industry, despite the lack of a universally recognised ethical framework.

Our first understanding of responsible AI was developed through AI research **(Blobel et al., 2020; Bukowski et al., 2020; Floridi et al., 2019; Reddy et al., 2020).** We examine how these six ethical qualities arose in earlier healthcare research before discussing them in the context of AI in healthcare (see Table 1):

Methodology

This work used a systematic literature review (SLR) approach to address topical research questions by the literature to identify, select, critically evaluate, and compile findings in order to gain a deeper understanding of responsible AI applications in healthcare. The research is a methodical literature review. The four steps of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) model, a well-established methodology, serve as the foundation for the employed search strategy. Twenty ● First, a search strategy incorporating the identified key words is formulated. ● Second, a list of key words covering the two research areas is identified, along with their synonyms and related concepts. ● Third, this search strategy is applied, making any necessary modifications to the modalities or requirements of each database. ● Fourth, a single file that can be used for in-depth analyses is created by pooling the articles and publications that were retrieved from each database query.

Search and acquisition

Finding databases, journals, keywords, and a time frame is the first stage in SLR. To assist us in screening and choosing articles, we used the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) approach, as shown in Fig. 1.

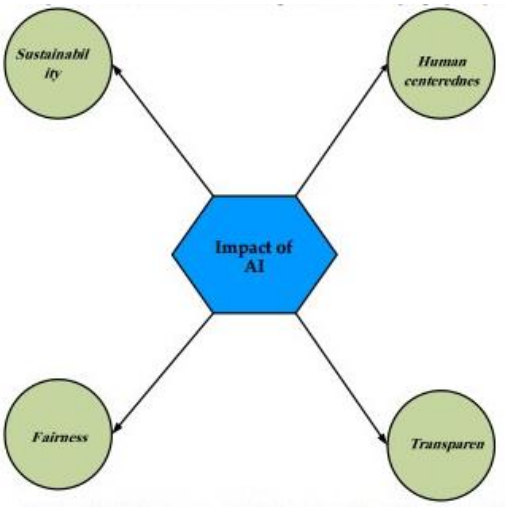


Fig. 1: The five main themes of AI

Two important digital bibliographical databases, PubMed and Google Scholar, were searched for research articles.. For our review, we have selected articles that were released within the last year, that is, from January 1, 2000, to December 31, 2020, including articles that are still in press because artificial intelligence is developing quickly. We searched for pertinent articles that addressed ethical concerns related to AI in healthcare or provided examples of responsible AI deployment. We therefore created a list of keywords based on the definition of AI as proposed by Davenport and **Kalakota (2019), Toh et al. (2019),** and **Morley et al. (2020).** The first group of search terms is associated with AI-related medical technology. Since the second set of keywords related to terms related to AI ethics, we developed a list of search terms based on the AI ethical principles mentioned in the AI governance frameworks **(Jobin et al., 2019; Reddy et al., 2020).** As previously indicated, we go over the most significant and frequently mentioned ethical traits that have come to light in the literature on AI ethics and are thought to be crucial for fostering and deepening our understanding of responsible AI. We also list some other recently developed ethical principles (e.g., explicability and sustainability) in addition to these most frequently cited ones. When describing ethical principles, different fields may use different terminology, even though some phrases are interchangeable. In order to

include a larger range of articles for our review, the goal of our keyword search strategy is to cover as many keywords related to AI ethics as possible. Afterwards, we used the single search term "health" to narrow down the results to studies that were related to healthcare. To get a wide variety of articles about health, we avoided using niche terms like "medicine" or "healthcare" and instead used a genetic term, "health.. Our application of Significant validity and consistency were present. Excerpts from articles mentioning the chosen keywords were examined and categorised. The codes were then expanded upon to produce themes and subthemes, which were eventually combined to produce overarching themes.

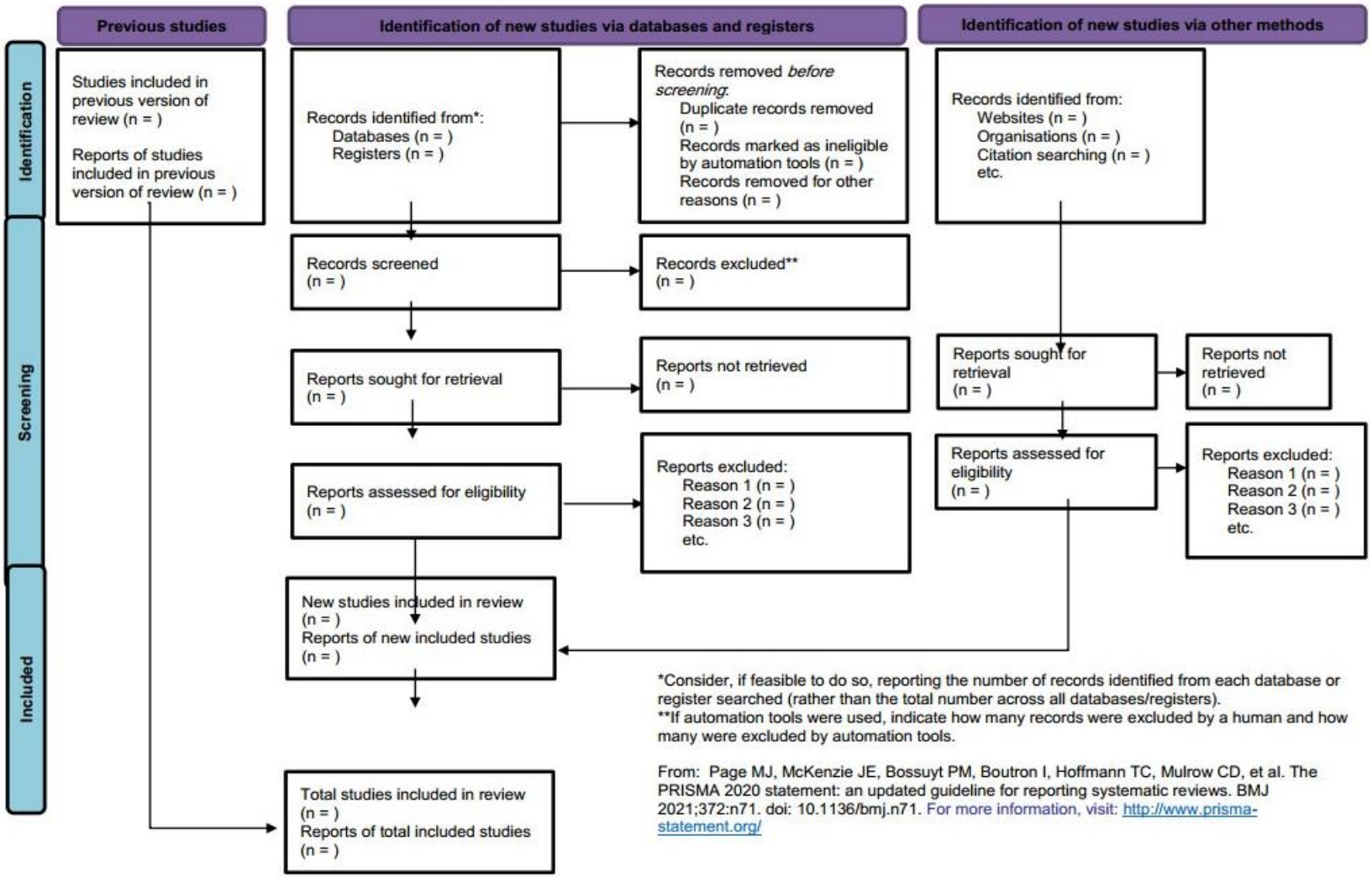
By enlisting an independent coder with an academic background who was not involved in the research project, intercoder reliability was examined **(Lombard et al., 2002).** In order to produce an intercoder reliability statistic, the coder was instructed to code a subset of a randomly chosen number of distinct articles. The percentage agreement of intercoder reliability was 94.85%.

Critical assessment - the framework

Research on AI ethics in healthcare has yielded five themes: trans-parent AI, human-centric AI, inclusive AI, sustainable AI, and fair AI. Fig. 3 illustrates the shift in the themes' frequency across the articles. It is possible for an article under one category, like Sustainable AI (Theme A), to overlap with themes from other categories due to the similarities between topics and research areas. According to the study pattern, these themes have been investigated at a faster rate since 2015. Less attention has been paid to inclusivity-related transparency issues, which are the most often discussed issues related to AI governance and ethics. To show the extent of each theme and subtheme, a thematic map was made.

Sustainable AI

We identified two subthemes within the category of sustainable AI: (1) developing accountable local leadership to increase the sustainability of AI solutions, and (2) AI for social sustainability.


Previous studies
Identification of new studies via databases and registers
Identification of new studies via other methods
Identification
Screening
Included
Studies included in previous version of review (n =)
Reports of studies included in previous version of review (n =)
Records identified from*:
Databases (n =)
Registers (n =)
Records removed *before screening*:
Duplicate records removed (n =)
Records marked as ineligible by automation tools (n =)
Records removed for other reasons (n =)
Records identified from:
Websites (n =)
Organisations (n =)
Citation searching (n =)
etc.
Records screened (n =)
Records excluded** (n =)
Reports sought for retrieval (n =)
Reports not retrieved (n =)
Reports sought for retrieval (n =)
Reports not retrieved (n =)
Reports assessed for eligibility (n =)
Reports excluded:
Reason 1 (n =)
Reason 2 (n =)
Reason 3 (n =)
etc.
Reports assessed for eligibility (n =)
Reports excluded:
Reason 1 (n =)
Reason 2 (n =)
Reason 3 (n =)
etc.
New studies included in review (n =)
Reports of new included studies (n =)
Total studies included in review (n =)
Reports of total included studies (n =)
*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).
**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.
From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/

Responsible local leadership

In order to develop AI healthcare technologies that are beneficial to the local population and adapt to local contexts, responsible local leadership is essential for making AI more sustainable **(Alami et al., 2017).** The development and application of AI technologies require the participation of local governments, educational institutions, research centres, international agencies, non-governmental organisations (NGOs), business, and civil society **(Alami et al., 2020). Alami et al. (2020)** Provide a framework that consists of four components to direct the application of sustainable AI healthcare technologies: (1) strong monitoring systems, (2) local expertise training and retention, (3) a systems-based approach to implementation, and (4) accountable, inclusive local leadership that involves all stakeholders. Encouraging cooperation amongst stakeholders and empowering local actors are essential to advancing sustainable AI. Local AI experts need to be developed and retained in order to guarantee that AI technology satisfies industry standards. To assist experts and decision-makers in better understanding and adhering to AI standards, local stakeholders ought to allocate a portion of their resources to providing consulting services **(Hosny and Aerts, 2019).** Furthermore, in order to prevent the deployment of disjointed AI-driven solutions that are unsupported by local infrastructure and expertise, international organisations can assist in establishing appropriate governance strategies and identifying crucial areas for investment and interventions **(Hosny and Aerts, 2019).**

Social sustainability

In recent years, interest in AI for sustainability has grown among academics and practitioners alike. The application of AI should carefully consider how it will affect human and environmental well-being **(World Health Organisation, 2021).** AI must be used ethically in healthcare organisations to create long-term profits while balancing the needs of stakeholders and minimising ethical concerns. If a healthcare organisation creates AI algorithms that inadvertently or intentionally violate human rights and wellbeing, it could seriously damage its credibility and reputation. For example, it has been noted that misusing AI, such as replacing established health services with smart technology, raises ethical questions **(Carter et al., 2020; Powell, 2019)** which might make already-existing health disparities worse **(Abramoff et al., 2020).** Thus, it makes sense for healthcare companies to create AI solutions that promote and prioritise social and economic sustainability. It is specifically necessary for them to create ethical governance policies that take into account socially beneficial methods, deal with ethical concerns during the AI systems' initial design and after launch, and integrate AI ethics into their social responsibility plan **(McCall, 2020).**

Beyond the fields of computer science and information technology, all healthcare personnel should receive training on the responsible use of AI-based technology in healthcare. To guarantee the complementary application of expertise in AI technology, pedagogy, ethics, healthcare policy, and clinical practise, partnerships between academic institutions and health service providers should be established. Some have suggested that in order to prepare future medical and healthcare students for the ethical issues surrounding data collection and AI use, medical schools should incorporate these programmes into their curricula **(McCoy et al., 2020).**

Human-centric AI

Embedding humanness in AI agents to meet ethics of carerequirements

Since the current ethical standards, laws, and guidelines are general in nature and lack a central ethical framework and specific principles that apply in the AI-based healthcare environment, Ethics of Care has been proposed as the primary ethical framework for developing AI-based health systems. A few advantages of the Ethics of Care approach are that it provides sufficiently specific principles, that it embodies values relevant to the design of AI-based health systems within the framework of compassionate practises, and that it is closely aligned with and related to the preservation and upkeep of a trustworthy

relationship between patients and AI agents **(Dalton-Brown, 2020).**

According to **Dalton-Brown (2020),** AI agents must possess human qualities like empathy, recognition, and human experiences in order to comply with care ethics. The application of humanised AI in healthcare has been found to have several advantages. For example, clinical reasoning and value-based care are decided by a doctor's judgement, and humans are open to exclusively human abilities. Furthermore, although machines are not susceptible to human carelessness or forgetfulness, fatigue, or cognitive errors, healthcare professionals or carers may experience burnout **(LuXton, 2014).** This could lead to more precise diagnosis and focus when interacting with patients. Furthermore, because humanised AI applications, like AI carebots, are not susceptible to the same personal prejudices that human therapists might, their advice may be more objective and detached. Because care seekers would feel less nervous discussing private matters with an AI carebot, some patients might therefore prefer interacting with one over a human care provider **(Kan- dalaft et al., 2013).** For AI developers, the British Standard's BS8611 for carebots offers guidelines on how to evaluate and minimise societal risks like betrayal, privacy invasion, security breaches, safety concerns, and loss of trust **(ISO, 2014).** According to **Wangmo et al. (2019) and Yew (2020),** patients view AI carebots' ability to mimic humans or animals as deception in certain situations, even though using carebots in place of Humans could reduce the possibility of moral blunders that human carers might encounter. **Wangmo et al. (2019)** discovered that a few interviewees voiced concerns about AI carebots appearing as humans or pets, which could unintentionally mislead elderly individuals suffering from dementia by mistakenly believing them to be actual people or animals.

Health professionals' function in upholding public confidence

Healthcare professionals are essential in the delivery of AI-enabled care because AI will never be able to completely replace cognitive abilities and human trust **(Mesko´ et al., 2018; Wangmo et al., 2019)**. Actually, the ability of medical professionals to uphold widespread public trust is what determines the integrity of AI **(Deo, 2015).** Three crucial ethical roles exist for health professionals utilising AI technologies in their practises: (1) as medical domain experts who should give computer and the clinical context needed by data scientists; (2) as gatekeepers for data quality, ensuring that data inputs are relevant, accurate, and properly sourced; and (3) as interpreters of AI black-BoX solutions who offer recommendations in real-time and after the fact to **(Miller, 2020).** Healthcare professionals should collaborate with industry leaders to develop new ethical approaches to address emerging ethical issues that may arise in the future, as suggested by **Darcy et al. (2016).** This will help ensure that AI technology advances in a way that maintains the public's trust in medicine. **Poulsen et al. (2020)** argue that while AI robots in healthcare assist medical professionals in providing more services, it is unclear how these robots influence codes of conduct, particularly in the area of cybersecurity. They emphasise the significance of incorporating cybersecurity concerns in codes of conduct for robot developers and carers in their conclusion since the onus is on humans and not the machine to ensure that an AI system is secure and safe to use.

Developing artificial wisdom through interdisciplinary collaboration **Jeste et al. (2020)** argue that artificial wisdom (AW), not artificial intelligence (AI), will be more important in the future as AI technology will not be primarily needed for advancing society's technological needs. Rather, wisdom will be more closely linked to people's longevity, well-being, and happiness than intelligence. The creation of AW requires the careful and close cooperation of computer scientists, neuroscientists, mental health professionals, and ethicists, all of whom together offer the greatest advantages to humanity **(Jeste et al., 2020).** Even when doctors are aware that they are working with machines, **Powell (2019)** reinforces Jeste's argument by saying that many medical decisions require both ethical judgements and the doctor-patient relationship for practical functionality **(Eduard and Jordi,**

2019). Hoorn and Winter (2018) conclude robot doctors are better than human doctors at communicating bad news or unfavorable information to patients, but **Blease et al. (2019)** emphasize that communication and empathy are interacting with AI agents on an emotional level, rather than just relying tionship **(Luxton, 2014),** interdisciplinary collaborations (Littmann et al., 2020) and empathy and understanding **(Wangmo et al., 2019)** to arrive at a shared decision, often handling large areas of uncertainty and balancing competing risks.

Inclusive AI

Inclusive communication and involvement in AI governance

The communication dynamic between patients and providers is changing as a result of digitalization. According to **LuXton (2014),** patients are worried that AI-based health systems could alter the way they interact with their doctors, potentially affecting the quality and cost of medical care. In order to tackle this issue, it can be greatly beneficial for scientific communities and public agencies to make sure that there is inclusive communication between healthcare providers and the general public **(LuXton, 2014; Noorbakhsh-Sabet et al., 2019; Poulsen et al., 2020).**

Furthermore, the design of AI solutions should involve stakeholders (e.g., AI companies, healthcare organisations, regulatory agencies and policy makers, and patients) from a variety of fields and cultures, with varying languages and communication styles. This will help to reduce unintentional biases **(Aitken et al., 2019; Char et al., 2020; World Health Organization, 2021). Horgan et al. (2019) and Vayena et al. (2015) advocate that legislation (e.g., inclusive impact assessment; World Health Organization, 2021)** can be implemented to make AI applications more inclusive and ensure legal certainty and clarity.

Fair AI

Alleviating algorithmic and data bias

According to **Rajkomar et al. (2018) and (2019),** biases are generally present in the design of AI models (e.g., label bias and cohort bias), training data (e.g., minority bias, missing data bias), interactions with clinicians (e.g., automation bias and feedback loops bias), and interactions with patients (e.g., privilege bias and agency bias). According to **Wang and Kosinski's 2017** study, for instance, computer vision algorithms were utilised to analyse thousands of facial photos from public dating website profiles in order to determine the sexual orientation of the subjects. Another study, which had more significant ethical ramifications, documented several instances in which AI algorithms have discriminated against people from underprivileged communities, ethnic minorities, and certain groups in contexts like health insurance and credit ratings **(Ienca and Ignatiadis, 2020; O'Neil, 2016).**

Scholars have proposed a number of solutions to reduce the impact of this kind of bias **(e.g., Eaneff et al., 2020; Shameer et al., 2018).** The influence of algorithmic bias on clinical decisions can be reduced in clinical settings by expanding the prediction task and developing various predictive models by taking contextual variables into account **(Shameer et al., 2018).** Furthermore, in order for the performance of AI models and algorithms to be deemed acceptable, algorithmic stewardship programmes need to regularly assess them (Eaneff et al., 2020; Vollmer et al., 2020). In order to guarantee safety and equity in the creation of algorithms for credit scores and health insurance, algorithm stewardship programmes are made to keep an algorithm inventory under the direction of a centralised therapeutics committee **(Eaneff et al., 2020). Obermeyer et al. (2019)** also suggest that identifying label choice bias in algorithms could potentially address structural inequalities.

Data representation and equality

There could be significant socio-political and ethical repercussions from the bias and limitations of the data used to train AI **(Strydom and Strydom, 2018).** Unrepresentative and unequal datasets used to train biassed AI algorithms may result in biassed medical diagnoses and treatment decisions, or even worse, discriminatory profiling of individuals living in low-resource environments **(Ienca and Ignatiadis, 2020;**

Mittelstadt et al., 2016; Powell, 2019). According to **Faraj et al. (2018),** AI algorithms are political by design because their creators' and the people who put together the datasets' values, preferences, and standards are ingrained in them. For instance, if applied in certain sub-Saharan African populations, an AI system trained on data skewed towards an overdiagnosis of schizophrenia in African Americans may have detrimental effects **(Vayena et al., 2018).** If algorithms trained with unrepresentative datasets are adopted in healthcare, they have the potential to exacerbate health disparities and may lead to underestimation or overestimation of risks in certain patient populations **(Reddy et al., 2020; Vayena et al., 2018).**

To mitigate AI bias, a data governance panel comprising a representative patient sample, clinical professionals, and experts in AI, ethics, and law should be established **(Char et al., 2018; Reddy et al., 2020).** The panel would oversee and assess datasets and algorithms used to train AI to make sure the data is representative and the algorithms are impartial and sufficient to inform necessary model outcomes. In order to uphold ethical obligations to vulnerable individuals and prevent discrimination in the use of AI, stakeholders must be open and honest about which communities and individuals are being monitored. Community leaders should also be involved to enable them to identify and report any adverse incidents affecting members of their community **(Reddy et al., 2020).**

Health disparity in low resource settings

To prevent the socio-economic digital health divide from getting worse, all socio-economic classes must have fair and equal access to affordable AI health technologies **(Alami et al., 2020; Horgan et al., 2019; Mehta et al., 2020).** Healthcare disparities between developed and developing nations may be reduced by AI-powered health systems **(Panchmatia et al., 2018).** However, the majority of AI-based health applications are created and used in high-income nations, raising doubts about their ability to improve the quality of healthcare in low-to-middle-income countries (LMICs) **(Alami et al., 2020).** Developing nations face financial difficulties and are dependent on aid for development, which makes investing in public health difficult. Additionally, in low-income countries (LMICs), companies may be able to commercialise solutions that would not receive regulatory approval in high-income countries due to a lack of governance **(Christie, 2018).** Previous research revealed that the majority of medical supplies donated to low- and middle-income countries (LMICs) were of poor quality, occasionally breaking down, or lacking user manuals and training for local staff **(Martinez-Martin and Kreitmair, 2018).** Certain nations might not be able to implement AI technology past the pilot stage due to the exorbitant cost and investment required. The requirements for responsible innovation in AI-powered health technologies go far beyond following and adhering to morally and ethically charged frameworks **(Blobel et al., 2020).** They also include making sure that a wide range of users are involved. Along with the sociodemographic diversity of the condition itself, which includes behavioural, cognitive, and emotional variations, users may differ in terms of race, ethnicity, and socioeconomic status. Health professionals have frequently been urged to take a more proactive approach by participating in the co-design and development of cutting-edge AI health technologies **(Miller and Polson, 2019; Panchmatia et al., 2018).**

There have been several examples given of how AI can improve healthcare equity. AI automated translation solutions, for instance, could make healthcare services more accessible to people in areas where language barriers exist, and interactive chatbots powered by AI could enhance patient care by assisting patients in receiving care and follow-up services on time **(LuXton, 2014).** AI, according to **Alami et al. (2020),** may be able to foresee and predict the spread of diseases or vulnerabilities within particular populations or groups, enabling more successful interventions in low- and middle-income countries **(Hosny and Aerts, 2019).**

Additionally, **Alami et al. (2019)** suggest that an over-reliance on AI may lead to the loss of clinical critical thinking and local practise skills, and there is a possibility that funding

for AI could be diverted from social and health budgets and resources. Furthermore, there's a chance that artificial intelligence (AI) diagnostic systems developed in wealthy countries will recommend medical interventions (such as surgery or medication) that are unavailable or prohibitively expensive in low-income countries (LMICs) **(Hosny and Aerts to Price and Cohen, 2019).** Conversely, AI-based health applications may offer several benefits to low- and middle-income countries (LMICs) where resources and expertise are limited, and they may also serve as a means of enabling universal access to affordable, high-quality healthcare for all. **Alami et al. (2020)** assert that public-private partnerships can offer smart health solutions to improve the health outcomes of those at risk of non-communicable diseases (NCDs) by using AI to intervene at various points in the patient journey (from health literacy and awareness to diagnosis and treatment). In order to stop the spread of pandemics and non-communicable diseases (NCDs) in low- and middle-income countries (LMICs), a coordinated, strategic, and focused approach involving all stakeholders in the healthcare value chain - including governments, academia, healthcare providers, civil society organisations, and the private sector - must be taken.

Transparent AI

Safeguarding personal privacy

Patients may experience identity theft, fraud, algorithmic bias, privacy invasion, information leakage, or other severe problems as a result of the gathering and use of personal health data by AI and analytical algorithms **(Toh et al., 2019; Wearn et al., 2019).** According to a survey conducted in the UK in 2018, 49% of adults expressed reluctance to share their personal health data in order to develop algorithms that could potentially improve the quality of care. The primary reason for this reluctance to share health data is the possibility that the data may be compromised or unintentionally leaked **(McNair and Price, 2019).** Specifically, data breaches may result in criminal activity or discrimination against marginalised or vulnerable populations **(Xafis et al., 2019).**

Patients may experience identity theft, fraud, algorithmic bias, privacy invasion, information leakage, or other severe problems as a result of the gathering and use of personal health data by AI and analytical algorithms **(Toh et al., 2019; Wearn et al., 2019).** According to a survey conducted in the UK in 2018, 49% of adults expressed reluctance to share their personal health data in order to develop algorithms that could potentially improve the quality of care. The primary reason for this reluctance to share health data is the possibility that the data may be compromised or unintentionally leaked **(McNair and Price, 2019).** Specifically, data breaches may result in criminal activity or discrimination against marginalised or vulnerable populations.

The potential to overcome the constraints of conventional face-to-face psychological interventions has been made possible by AI advances in machine learning. These advances have created new avenues for decoding and analysing neural data to deliver customised neurointerventions. It is crucial to assess the ethical ramifications of neural data analytics' increasing integration into the healthcare ecosystem and create a roadmap for ethical innovation in this area while considering a range of privacy concerns, such as mind readiness **(Benke and Benke, 2018).**

Explainability of AI-driven models and decisions

The field known as explainable AI (XAI) has emerged as a result of this important issue **(Rai, 2020).** Knowing an algorithm's inner workings and who is in charge of putting it into practise are two crucial components of XAI **(Floridi et al., 2018).** According to **Norgeot et al. (2020),** responsible AI usage in healthcare requires that patient data access, analysis, and interpretation, as well as Holding AI model development responsible. In addition to providing intelligent and personalised explanations of the results generated by algorithms, an effective AI-enabled healthcare system should be able to demonstrate the stability and dependability of the AI models, according to **Davenport and Kalakota (2019) (Norgeot et al., 2020).** The Baig group (2020) **Zitnik et al. (2019)** argue that the

explainability of clinical decisions can be improved by merging patient data from various sources, while others recommend setting up a central review board to look at the architecture and processes of machine learning as well as the methods used to interpret data in AI healthcare projects.

Furthermore, if medical professionals can justify the efficacy of a particular AI-driven treatment for a patient, the public will have more faith in the clinical judgements made. According to **Shaw et al. (2019),** decision support in the healthcare industry can only be beneficial if its results can be integrated with the human decision-making processes that are fundamental to the provision of health services. Additionally, **Cabitza et al. (2017)** propose that doctors could be able to investigate the consequences by fusing machine learning and visualisation, physicians may be able to explore the implications of outputs in rich interactive ways, alleviating the tension between accuracy and interpretability.

Addressing the loss of confidentiality by legislation

Big data storage and collection have grown rapidly, and advances in AI health technologies have created previously unanticipated challenges. For example, medical data is now accessible through shared networks, mobile devices, and even body-attached sensors **(Wang et al., 2018).** The public's awareness of data confidentiality has increased with the advancement of information retrieval and storage technologies. Concerns about how public trust in the healthcare system might be damaged by confidentiality violations are becoming more widespread **(Ahmed et al., 2020).** Patients may self-medicate, see another doctor, give erroneous information, or decide not to seek treatment if these worries are present, which could lower the actual quality of care given **(Yüksel et al., 2017).** According to research **(Ienca and Ignatiadis, 2020; Price and Cohen, 2019),** the most frequent argument against sharing data with a third party is the possibility of discrimination by insurance companies. Continuous medical monitoring and privacy violations involving medical devices can make more marginalised people feel more stigmatised and may even make it more difficult for them to get health insurance and care if they are unable to adopt new, healthier lifestyle standards **(Brig- anti and Le Moine, 2020; Mittelstadt, 2017).** Thus, legislation should be implemented to protect patients' confidentiality by creating a unique cause of action for those who wish to sue AI agents or healthcare organizations for misuse of their data **(Lupton, 2018).**

User empowerment

New methods of facilitating the sourcing of patient personal health data have been investigated with the introduction of new technologies intended to collect data voluntarily submitted by patients. It is advised that patients keep track of their symptoms and treatments, monitor their progress towards fitness goals, view test results, and communicate more conveniently with medical professionals by using online personal health repositories like Microsoft HealthVault. The degree to which privacy and ethical issues are addressed from a technical (e.g., digital key cryptography) and transparent (firm-generated reassurances) perspective, as well as by encouraging people to share and distribute their personal data more actively, will determine whether the public accepts AI-powered digital health technology. Making end users prosumers of AI-driven digital health systems can empower them to take control of their own data and protect their privacy **(Peters et al., 2020).** This can result in the deployment of AI-powered health systems being more ethically aligned **(Benke and Benke, 2018; Chadwick et al., 2014).**

Concerns among users regarding the adoption of AI-powered digital health systems have given rise to a recurring theme of empowerment **(Manrique de Lara and Pela´ez-Ballestas, 2020).** Users who are reluctant to disclose certain types of health data typically hold information about pregnancy, contraception, sexual health, mental health, and other sensitive, private, and potentially stigmatising topics **(Powell et al., 2006).** In order to ascertain whether AI models and their outcomes will be approved by clinicians who will use the model to make clinical decisions

and by users (like patients) whose data was collected, assessment mechanisms should be developed **(Vollmer et al., 2020).**

Furthermore, the legislature must intervene in AI-related medical disputes by extending the common law's definitions of malpractice and negligence to AI agents **(Lupton, 2018).**

Informed consent for data use

Data privacy and ethical concerns regarding the use and storage of patient data are major roadblocks to data ownership in AI-powered digital health ecosystems **(Bukowski et al., 2020).** Healthcare stakeholders have also emphasised patient autonomy and informed consent as ethical priorities **(Wangmo et al., 2019).** This draws attention to the numerous and important ethical problems that come up when attempting to obtain informed consent from patients. About clinical trials, **Pfister and Jung (2020)**

Furthermore, they put forth a safe framework to support the creation of moral AI applications in the healthcare industry. These applications would entail overseeing WIC documentation throughout the whole data value chain, from gaining consent to publishing research findings in an academic journal to profitably using clinical study outcomes. They emphasised that voluntary participation in a clinical trial requires the introduction of a written informed consent (WIC) procedure. In a counterargumentative position, **Larson et al. (2020)** controversially claimed that in exceptional circumstances such as emergencies (e.g., acute life-threatening situations), patient consent is not required before the data is used. The GDPR specifically allows competent public health authorities to process personal data lawfully for purposes of substantial public interest without obtaining informed consent **(Holub et al., 2020).**

For instance, in order to conduct a clinical study during the COVID-19 pandemic, consent from the patient or their legal guardian may be obtained subsequently **(European Medicines Agency, 2020).** Although the Royal Free London NHS Foundation Trust shared patient data without consent in order to develop a clinical application, recent incidents have raised concerns about privacy violations even though patients' explicit consent is expected when sharing their data **(Carter et al., 2020; Reddy et al., 2020).** A cautious policy intervention **(Panch et al., 2019)** and an informed consent model (such as Ploug and Holm, 2016's Meta Consent) that considers practical implementation might be useful to allay the privacy concerns mentioned above.Developing arguments is one of the most important problems in ethical AI.

The argument development section, in contrast to the previous section's examination of initiatives pertaining to responsible AI, focuses on identifying the most urgent problems and obstacles associated with responsible AI that have not been sufficiently addressed in the body of existing literature, thus inspiring additional study **(Boell and Cecez-Kecmanovic, 2014).** More precisely, we list the major problems and obstacles of responsible AI that are unique to the SHIFT themes along with a list of possible fixes that have been suggested by earlier studies.

Sustainable AI

The medical community is still not fully aware of the ways that new AI technologies can bring ethical challenges to the practise of medicine **(Rigby, 2019).** To prepare and educate upcoming healthcare professionals in the use of AI technologies, it appears unclear what kind of AI There ought to be ethics education **(Combs and Combs, 2019).** To address the ethical issues raised by the AI revolution in healthcare, some attempts have been made to incorporate AI courses into curricula **(McCoy et al., 2020; Park et al., 2019)** and integrate ethical decision-making training into clinical training **(Combs and Combs, 2019).** For example, the introduction of virtual patients (VP) in medical education enables students to learn clinical and ethical decision making through interactions between practitioners and patients.

Table 2
Challenges, issues and proposed solutions on responsible AI implementation.

Responsible AI themes	Challenges	Issues challenging responsible AI	Proposed solutions
Sustainable AI	What kind of AI ethics training should be integrated into medical school curriculums?	• Healthcare practitioners are unaware that new developments in AI technology could complicate providing care in ethical ways (Rigby, 2019). • There is uncertainty about the specific AI ethics training that future healthcare professionals should receive to get ready to use AI technologies (Combs and Combs, 2019).	• Using AI-enabled virtual patients to incorporate ethical decision-making training into clinical-based training (Combs and Combs, 2019) • Integrating rigorous courses on data science into the foundational curriculum for health research (McCoy et al., 2020; Park et al., 2019).
Human-centric AI	How can a right balance be struck between delivering individualized care based on AI while attaining long-term profitability for healthcare providers?	Patients being treated as commodities rather than individuals (Quinn et al., 2021) Loss of interpersonal processes of responding to clinical problems in a way that prioritizes the needs and preferences of patients.	Using a patient-centered approach in designing medical AI that promotes informed choices aligned with patient values and respects patient autonomy (Sarkar et al., 2021; Quinn et al., 2021) Delivering value with a patient-centric process (Agarwal et al., 2020)
Inclusive AI	How can the pursuit of commercialization of AI be inclusive to the broad public?	Lack of inclusiveness policy for AI governance in relation to access and use of health data (Aitken et al., 2019) Health data is commercialized for AI solutions without considering the voices of diverse patient groups (Rickert, 2020)	A continuous dialogue among authorities, technology giants and healthcare service providers to resolve potential ethical complexities of AI commercialization in healthcare (Rickert, 2020) Engage diverse patient groups and broad stakeholders in AI governance (Chen et al., 2020)
Fair AI	***Challenge 4***: How can representative data be created and utilized to address patients' needs fairly?	Contextualized factors such as sociodemography, health state, and social culture are not understood adequately to develop AI-based solutions that can cater for patients' needs (Obermeyer et al., 2019) Training AI algorithms with unrepresentative dataset limits the ability to provide meaningful assessments or predictions (Carter et al., 2020; Ienca and Ignatiadis, 2020)	Allow for value pluralism and ensure that no protected characteristics are enforced (Morley et al., 2021) Explore local datasets or local checklist for AI training to promote equal patient outcome, equal performance and equal allocation (Rajkomar et al., 2018, 2019; Vollmer et al., 2020) Avoid label choice bias by conducting further validation studies (Char et al., 2020; Obermeyer et al., 2019)
Transparent AI	***Challenge 5***: To what extent should AI-led data use need to be transparent to all stakeholders?	The fuzziness of AI governance in addressing the questions of how to interpret predictions properly (Benke and Benke, 2018) The opaque design of explainable AI results from its ambiguous definition (Jiménez-Luna et al., 2020)	Higher education institutions and funding bodies should maintain, curate, and promote open-science repositories, with clear incentives for compliance (Blobel et al., 2020) The ex-post reviews could be conducted by a multi-disciplinary committee to evaluate the quality of AI-driven explanations (Baric-Parker and Anderson, 2020; Blobel et al., 2020)

Communication **(Combs and Combs, 2019).** However, VP has been criticized for its representation of diversity in a population and non-transparent algorithms for providing patient feedback.

Human-centric AI

According to some, AI will revolutionise healthcare and allow patients to be treated more like commodities than like unique individuals **(Quinn et al., 2021).** For instance, health insurance companies may force patients to use AI-powered health systems without giving them the option to see a human doctor instead, motivated by financial incentives and the belief that these systems are better than traditional methods. Another illustration is the Uberization of mental health services in the UK as a result of the cheaper and more easily accessible AI-powered health systems (such as chatbots) **(Cotton, 2021).** This could lead to a shift in emphasis away from patient-centered approaches to treating clinical issues and the interpersonal processes involved in psychotherapy that place the needs and preferences of individuals first.

Medical AI has to change its focus from a problem-oriented approach to a patient-centered one in order to overcome this challenge **(Quinn et al., 2021). Agarwal et al. (2020)** take things a step further and propose a patient value-centered approach that considers three essential aspects of value: process, preferences, and precision. Whereas the latter is largely motivated by value creation, the former approach is founded on the ideas of patient-centeredness. Both methods centre clinical decision-making around patient rights as a moral requirement.

Inclusive AI

According to a number of studies, the healthcare industry is about to be colonised and commercialised by large technology companies like Google, Apple, and IBM as well as top healthcare providers like the NHS. These companies use the vast amounts of digital data that are being amassed online to make money **(Downey, 2019; Larson et al., 2020).** The commercial goals include selling users goods and services, advertising, and the on-sale of archived data to other parties, like health and pharmaceutical companies. For instance, the NHS patient data is thought to be worth a staggering £10 billion annually **(Downey, 2019).** Management consultants McKinsey proposed commercial models that range from the NHS receiving a curated dataset free of charge to a royalty fee and shared ownership of products or discounts on products created through the partnership **(Downey, 2019).** The reciprocal agreement between IBM and the Italian government, which was signed in early 2016, is another notable example of patient data commercialization. IBM committed to investing $150 million in a health centre to be used for developing e-health applications. IBM will receive access to Lombardy residents' important health data in exchange **(Monegain, 2016).** Under these circumstances, health data are commercialised for AI applications without taking inclusive policies into account **(Aitken et al., 2019)** or various patient viewpoints **(Rickert, 2020).** Regular communication and collaboration among government authorities, technology giants, and healthcare service providers is imperative to regulate the commercialization of AI **(Rickert, 2020).** In the meantime, broad stakeholders and a variety of patient groups should be included in AI governance. The All OF Us Research Programme (All of Us Research Programme In- vestigators, 2019) is one of the first initiatives that have been introduced to address this challenge.

Fair AI

It's possible that the dataset used to train AI algorithms is not representative, which would make it more difficult to make insightful evaluations or predictions **(Carter et al., 2020; Ienca and Ignatiadis, 2020).** Lack of knowledge about contextualised factors like sociodemography, health status, and social culture, which must be carefully taken into account when developing AI solutions, results in unrepresen- tative datasets **(Obermeyer et al., 2019).** According to **Vollmer et al. (2020),** artificial intelligence algorithms ought to be trained on regional datasets and ought to exercise caution when aiding or deciding on medical matters when faced with scientific ambiguity.

Furthermore, while some AI systems can infer ethnicity, which is important in some clinical situations, if not properly governed and managed, this function could be exploited for racial profiling or discrimination, and it could be used to marginalise people on the basis of their gender,

ethnicity, socioeconomic status, sexual orientation, or pathology **(Alami et al., 2020; Ienca and Ignatiadis, 2020; LuX- tonne, 2014; Price and Cohen, 2019).** This problem appears to continue despite the development of some viable solutions, such as investigating local datasets or local checklists for AI training **(Rajkomar et al., 2018, 2019; Vollmer et al., 2020)** and carrying out validation checks to reduce label choice bias **(Char et al., 2020; Obermeyer et al., 2019).** We therefore call for more research in developing a method to ensure that representative datasets are used to avoid structural inequalities in AI development.

Transparent AI

Population health research is seeing a rise in the use of AI-based techniques. A large portion of the data gathered for this type of research comes from anonymous secondary health data or public social media, exempting it from ethics committee review **(Samuel and Derrick, 2020).** In addition to reducing the risk of health inequity or over-surveillance, governance guides academics' ethical decision-making and reassures the public that researchers are operating morally. There are still issues regarding what an ethics governance framework is, how to handle, process, and interpret data predictions in an ethically responsible way, and how, in some jurisdictions, such a system would stop data from being exported to nations with weak research ethics oversight **(Benke and Benke, 2018). Blobel et al. (2020)** highlighted the need for higher education institutions and funding agencies to oversee, curate, and steer open-science repositories, providing clear incentives for adherence as a viable solution to this ethical conundrum. In order to create best practises, for instance, it ought to be mandatory that algorithms and related data be stored in repositories to which only specific stakeholders have access. With the help of these controlled repositories, other researchers and interested parties can test the algorithms using their own data, looking for anomalous predictions and drawing attention to any problems or concerns that might exist in the AI prediction models.

Furthermore, **Blobel et al. (2020)** proposed an ex-post review of novel prediction algorithms applied in particular industries as an additional layer of ethical regulation. A multidisciplinary committee made up of academics and stakeholders (such as professionals or technology users) from a variety of disciplines, such as health and medicine, artificial intelligence, social science, and ethics, could conduct ex-post reviews in the field of public health. The committee's objectives would be to reduce the likelihood of potential harm by examining scientific inquiries into the source and calibre of data, algorithms, and artificial intelligence; verifying the steps taken to validate the prediction models; and requesting additional validation when necessary. By taking on a regulatory function, the committee might take on the responsibilities of an AI ombudsman or supervisor. The best people for this kind of position are probably already in regulatory bodies; in fact, regulatory bodies in the UK and Europe have recently begun to take steps to examine medical software, including offering standards and recommendations for the use of AI in healthcare. Therefore, in order to transparently regulate and govern AI-powered healthcare ecosystems, a thorough and rigorous approach must be developed **(Baric-Parker and Anderson, 2020).**

Research agenda – future potentials of responsible AI

Continuous discussions are needed to comprehensively understand responsible AI use in healthcare. There are two primary future potentials on appropriate AI use in medical fields. Prior to executing the SHIFT of AI, we must comprehend the societal, organisational, and individual barriers. Previous research on medical AI has focused on analysing the financial advantages of AI applications and the state of technology for doing so. Further research is needed to fully understand the practises, frameworks, and environments that support the ethical application of AI in healthcare. Second, despite the potential for AI-enabled tools to enhance clinical practises, there are certain barriers that prevent medical professionals from actually utilising them. To address the challenges associated with the use of AI, AI solution providers and developers should aim at designing and putting into practise moral, open, and responsible AI solutions.

Healthcare organisations could reduce risks and preserve trust by taking AI ethics into account. Future studies are therefore required to

comprehend the role that ethical AI use plays from a SHIFT perspective in order to add value and lower potential risks in the healthcare industry. Thus, more empirical research examining and evaluating the efficacy of AI-based digital health interventions in enhancing healthcare decision- making and quality is needed to support the suggested advantages of these technologies **(Bukowski et al., 2020).** It is important to anticipate the societal ramifications (at the system level) of scientific analysis and evaluation rather than merely highlighting the ethical standards in health care services research. In the field of medicine, responsible AI is a promising strategy. However, additional study is required to address the broader ethical and societal concerns of these technologies, which are propelling the evolution of digital healthcare.

Discussion and conclusion

Academic and practical implications

The promising future of reducing health disparities by incorporating AI into clinical procedures is accompanied by an unforeseen difficulty: healthcare institutions are now more likely than ever to violate moral or ethical standards. Governments, regulators, and other stakeholders have put pressure on healthcare service providers and medical algorithm designers to act responsibly. To create a responsible AI framework, 253 articles from the past 20 years in a variety of domains (such as bioethics, information systems, healthcare management, and medicine) were methodically arranged and reviewed in the current review. The framework outlined responsible AI initiatives across five main themes and fourteen sub-themes, all of which were supported by virtue ethics theory. Previously mentioned, virtue ethics theory has been applied to investigate what constitutes a moral action in the context of business management by focusing on the ethical aspects of everyday business

Research agenda – future potentials of responsible AI

Continuous discussions are needed to comprehensively understand responsible AI use in healthcare. There are two primary future potentials on appropriate AI use in medical fields. Prior to executing the SHIFT of AI, we must comprehend the societal, organisational, and individual barriers. Previous research on medical AI has focused on analysing the financial advantages of AI applications and the state of technology for doing so. Further research is needed to fully understand the practises, frameworks, and environments that support the ethical application of AI in healthcare. Second, despite the potential for AI-enabled tools to enhance clinical practises, there are certain barriers that prevent medical professionals from actually utilising them.

To address the challenges associated with the use of AI, AI solution providers and developers should aim at designing and putting into practise moral, open, and responsible AI solutions.

Healthcare organisations could reduce risks and preserve trust by taking AI ethics into account. Future studies are therefore required to comprehend the role that ethical AI use plays from a SHIFT perspective in order to add value and lower potential risks in the healthcare industry. Thus, more empirical research examining and evaluating the efficacy of AI-based digital health interventions in enhancing healthcare decision- making and quality is needed to support the suggested advantages of these technologies **(Bukowski et al., 2020).** It is important to anticipate the societal ramifications (at the system level) of scientific analysis and evaluation rather than merely highlighting the ethical standards in health care services research. In the field of medicine, responsible AI is a promising strategy. However, additional study is required to address the broader ethical and societal concerns of these technologies, which are propelling the evolution of digital healthcare.

Discussion and conclusion

Academic and practical implications

The promising future of reducing health disparities by incorporating AI into clinical procedures is accompanied by an unforeseen difficulty: healthcare institutions are now more likely than ever to violate moral or ethical standards. Governments, regulators, and other stakeholders have put pressure on healthcare service providers and medical algorithm designers to act responsibly. To create a responsible AI framework, 253 articles from the past 20 years in a variety of domains (such as bioethics, information systems, healthcare

management, and medicine) were methodically arranged and reviewed in the current review. The framework outlined responsible AI initiatives across five main themes and fourteen sub-themes, all of which were supported by virtue ethics theory. Previously mentioned, virtue ethics theory has been applied to investigate what constitutes a moral action in the context of business management by focusing on the ethical aspects of everyday business operations **(Audi, 2012).**

Through the application of this theory to a context that has not received much attention - the use and application of AI in healthcare - our review may help researchers and practitioners move past the symbolic advocacy of AI ethics and towards healthcare practises and responsible governance of AI.

Table 3
A research agenda for future research on responsible AI from a to z.

Research themes	Sample research questions pertinent to each theme
Sustainable AI	a. What effective policies and actions can be taken by healthcare organizations or governments to leverage AI for the purpose of social sustainability? b. What kind of AI ethics training should be integrated into medical school curriculums? c. How can AI tools improve social impacts by reinforcing the regulations in tackling irresponsible medical practices in the digital world? d. What coordination mechanisms can mitigate ethical concerns regarding AI commercialization and ensure a more sustainabl
Human-centric	healthcare ecosystem in the long run?
How can AI systems	incorporate and translate human
AI	judgements to generate accurate medical knowledge and insights? e. How should ethics education be designed and integrated into the training of AI solution developers? f. What are the effective coordination mechanisms for multi-actor decision making that involves amongst others, AI agents and healthcare professionals? g. How do AI technologies enhance co-ordination among health policy makers, healthcare organizations, and patients? h. Does AI and human coordination impede the effectiveness of AIand clinical efficacy? i. What defines the paradoXical nature of bias from the sociomateriality perspective of algorithmic decision-making?
Inclusive AI	k. What are the implications of inclusive communication for AI-based digital health management? l. What type of individual roles and disciplines in an ethics panel contribute most to a better understanding of inclusive communication in AI-based digital healthcare? m. How can AI-enabled health systems operate in a coordinated manner to deliver inclusive care to patients?
Fair AI	n. To what extent do AI algorithms affect healthcare practitioners' efficiency and quality of care? o. What algorithmic attributes and characteristics are required for reducing biases in the data and prediction model? p. What safeguards measures can healthcare organizations take to ensure that patients are treated fairly when a medical decision is delegated to an AI-based health system? q. In what ways can minorities and marginalized groups be involved in consultations to mitigate biases and structural inequalities? r. How to empower local actors and foster local collaboration between stakeholders to develop equitable AI solutions ? s. How can AI systems and applications be designed in way to enhance patients' perceptions of fairness and trust?
Transparent AI	t. What kind of low-cost AI solutions can be deployed to address health disparity in low resource settings (e.g., LMICs)? u. How can the risks of cybersecurity, data loss and patient identity theft be reduced or mitigated through health data governance? v. To what extent should AI healthcare system be transparent in terms of data use and algorithm management? w. How should AI generate results that are discernible and lucid to users and health practitioners? x. What factors drive patients or healthcare professionals to share data in AI-driven digital health environments? y. How can data quality assurance and programming norm be cultivated in the AI development stages? z. How can health data be utilized evidence-based medicine using AI approaches?

One of the main goals of responsible AI, for example, is to develop comprehensive education and training programmes, as our review of the sustainable AI theme demonstrates. According to our review, the most suitable function for AI-powered health systems in the human-centric AI theme is that of an assistant, helping human practitioners make clinical care decisions. Regardless of the application, these systems will require careful testing and evaluation; however, there is a philosophical debate between AI and humans regarding whether or not some degree of imperfection can be advantageous for medical intervention and treatment (LuXton, 2014).

Secondly, we highlight the most urgent issues with responsible AI.

(refer to Section 5) and urge resolving these issues by responding to

In order to enable the provision of more effective and responsible AI- powered healthcare services, the main focus of this review is to emphasise to patients, healthcare providers, legislators, and designers of medical algorithms how important it is to have a thorough understanding of responsible AI initiatives. As such, we offer three real-world applications. First, by describing in great detail how AI is applied responsibly in healthcare settings, our analysis goes beyond simply investigating AI ethics.

One of the main goals of responsible AI, for example, is to develop comprehensive education and training programmes, as our review of the sustainable AI theme demonstrates. According to our review, the most suitable function for AI-powered health systems in the human- centric AI theme is that of an assistant, helping human practitioners make clinical care decisions. Regardless of the application, these systems will require careful testing and evaluation; however, there is a philosophical debate between AI and humans regarding whether or not some degree of imperfection can be advantageous for medical intervention and treatment **(LuXton, 2014).**

Secondly, we highlight the most urgent issues with responsible AI. (refer to Section 5) and urge resolving these issues by responding to our suggested research questions. For instance, our review reveals that in order to create fair medical AI solutions, a varied set of stakeholders and experts is required. It has been proposed that in order to develop common rules specific to AI that are grounded in virtue ethics, the federal government, healthcare practitioners and providers, researchers, and providers of AI health technology should collaborate. The question of how minorities and marginalised groups can engage in consultations to lessen biases and structural inequalities, however, remains unanswered in the body of existing literature.

Third, a thorough examination of the ethical strategies for integrating AI into healthcare has been supplied by this review. By establishing a framework of openness and shared accountability that holds all parties involved in the supply chain of AI algorithms accountable, this perspective can promote the creation of reasonable ethical policies and regulatory actions. In particular, to develop medical AI responsibly, five initiatives must be included: sustainability, human-centeredness, inclusivity, fairness, and transparency. These can be evaluated by legislators and policymakers to see if biases or inherent risks are adequately mitigated for better AI adoption

Limitations

It is crucial to recognise that, even though our systematic review is supported by a sizable number of trustworthy sources, the majority of the papers it reviewed have a Western perspective (they are primarily from West Europe and North America). Therefore, in order to gain a more comprehensive or nuanced understanding of what ethical AI in healthcare entails, future research should take into account looking up and analysing studies from different languages or continents.

There is a main reason why we decided against performing bibliometric analyses. By tracking back the beginnings of topic, authorship, and citation over time and presenting the results in a descriptive way, the bibliometric approach concentrates on an author- centric review. Given that our main goal is to identify ethical AI initiatives so that medical professionals can benefit from our findings, it doesn't seem appropriate to conduct this kind of review. We do, however, advise future researchers to employ bibliometric analyses to illustrate descriptive findings, such as the creation of topics for initiatives involving responsible AI.

The inclusion and exclusion criteria not included in this meta-analysis Because it was better to applied or included in experimental studies.

References:

Abra`moff, M.D., Leng, T., Ting, D.S.W., Rhee, K., Horton, M.B., Brady, C.J., Chiang, M.F., 2020. Automated and computer-assisted detection, classification, and diagnosis of diabetic retinopathy. Telemed. e-Health 26 (4), 544–550.

Agarwal, R., Dugas, M., Gao, G.G., Kannan, P.K., 2020. Emerging technologies and providedUniver- by the Sheffield University Management School, The sity of Sheffield, United Kingdom. We would like to thank the

anony- mous reviewers for their constructive comments that led to the improvement of the paper.

Ahmed, Z., Mohamed, K., Zeeshan, S., Dong, X., 2020. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database. https://doi.org/10.1093/database/baaa010.

Aitken, M., Tully, M.P., Porteous, C., Denegri, S., Cunningham-Burley, S., Banner, N., et al., 2019. Consensus statement on public involvement and engagement with data intensive health research. Int. J. Popul. Data Sci. 4 (1), 586.

Alami, H., Gagnon, M.-P., Fortin, J.-P., 2017. Digital health and the challenge of health systems transformation. mHealth 3, 31.

Alami, H., Gagnon, M.-P., Fortin, J.-P., 2019. Some multidimensional unintended consequences of telehealth utilization: a multi-project evaluation synthesis. Int. J. Health Pol. Manag. 8 (6), 337–352.

Alami, H., Rivard, L., LehouX, P., Hoffman, S.J., Cadeddu, S.B.M., Savoldelli, M., Samri, M.A., Ahmed, M.A.A., Fleet, R., Fortin, J.-P., 2020. Artificial intelligence in health care: laying the Foundation for Responsible, sustainable, and inclusive innovation in low- and middle-income countries. Glob. Health 16 (1), 52.

All of Us Research Program Investigators, 2019. The "All of Us" research program. N. Engl. J. Med. 381 (7), 668–676.

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al., 2020. EXplainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion 58, 82–115.

Audi, R., 2012. Virtue ethics as a resource in business. Bus. Ethics Q. 22 (2), 273–291. Baig, M.A., Almuhaizea, M.A., Alshehri, J., Bazarbashi, M.S., Al-Shagathrh, F., 2020.

Baric-Parker, J., Anderson, E.E., 2020. Patient data sharing for AI: ethical challenges, Catholic solutions. Linacre Q. 87 (4), 471–481.

Benke, K., Benke, G., 2018. Artificial intelligence and big data in public health. Int. J. Environ. Res. Publ. Health 15 (12), 2796.

Blease, C., Kaptchuk, T.J., Bernstein, M.H., Mandl, K.D., Halamka, J.D., DesRoches, C.M., 2019. Artificial Intelligence and the future of primary care: exploratory qualitative study of UK general practitioners' views. J. Med. Internet Res. 21 (3), e12802. Blobel, B., Ruotsalainen, P., Brochhausen, M., Oemig, F., Uribe, G.A., 2020. Autonomous systems and artificial intelligence in healthcare transformation to 5P medicine – ethical challenges. Stud. Health Technol. Inf. 270, 1089–1093.

Boell, S.K., Cecez-Kecmanovic, D., 2014. A hermeneutic approach for conducting literature reviews and literature searches. Commun. Assoc. Inf. Syst. 34 (1), 12.

Briganti, G., Le Moine, O., 2020. Artificial Intelligence in medicine: today and tomorrow. Front. Med. 7, 27.

Bukowski, M., Farkas, R., Beyan, O., Moll, L., Hahn, H., Kiessling, F., Schmitz-Rode, T., 2020. Implementation of eHealth and AI integrated diagnostics with multidisciplinary digitized data: are we ready from an international perspective?Eur. Radiol. 30 (10), 5510–5524.

Cabitza, F., Rasoini, R., Gensini, G.F., 2017. Unintended consequences of machine learning in medicine. JAMA 318 (6), 517–518.

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.

Chadwick, R., Levitt, M., Shickle, D., 2014. The Right to Know and the Right Not to Know: Genetic Privacy and Responsibility. Cambridge University Press. Chakrabarty, S., Bass, A.E., 2015. Comparing virtue, consequentialist, and deontological ethics-based corporate social responsibility: mitigating microfinance risk in institutional voids. J. Bus. Ethics 126 (3), 487–512.

Char, D.S., Abra`moff, M.D., Feudtner, C., 2020. Identifying ethical considerations for machine learning healthcare applications. Am. J. Bioeth. 20 (11), 7–17.

Char, D.S., Shah, N.H., Magnus, D., 2018. Implementing machine learning in health care – addressing ethical challenges. N. Engl. J. Med. 378 (11), 981–983.

Chatterjee, S., Sarker, S., Fuller, M., 2009. A deontological approach to designing ethical collaboration. J. Assoc. Inf. Syst. Online 10 (3), 138–169.

Chen, I.Y., Pierson, E., Rose, S., Joshi, S., Ferryman, K., Ghassemi, M., 2020. Ethical machine learning in healthcare. Annu. Rev. Biomed. Data Sci. 4. —beyond the medicine. JAMA 315 (6), 551–552.

Chen, J.H., Asch, S.M., 2017. Machine learning and prediction in medicine peak of inflated expectations. N. Engl. J. Med. 376 (26), 2507–2509.

Christie, G., 2018. Progressing the health agenda: responsibly innovating in health technology. J. Resp. Innovat. 5 (1), 143–148.

Chun, R., 2005. Ethical character and virtue of organizations: an empirical assessment and strategic implications. J. Bus. Ethics 57 (3), 269–284.

Cipriani, A., Geddes, J., 2003. Comparison of systematic and narrative reviews: the example of the atypical antipsychotics. Epidemiol. Psychiatr. Sci. 12 (3), 146–153. Clarke, V., Braun, V., 2013. Teaching thematic analysis: overcoming challenges and developing strategies for effective learning. Psychol. 26 (2), 120–123.

Combs, C.D., Combs, P.F., 2019. Emerging roles of virtual patients in the age of AI. AMA J. Ethics 21 (2), E153–E159.

Cotton, E., 2021. Mental health services in England are being 'Uberised – and that s bad for patients and therapists. Available at: https://theconversation.com/mental-health - services-in-england-are-being-uberised-and-thats-bad-for-patients-and-therapists -167065.

Culnan, M.J., Williams, C.C., 2009. How ethics can enhance organizational privacy: lessons from the choicepoint and TJX data breaches. MIS Q. 33 (4), 673–687. Dalton-Brown, S., 2020. The ethics of medical AI and the physician-patient relationship. Camb. Q. Healthc. Ethics 29 (1), 115–121.

Darcy, A.M., Louie, A.K., Roberts, L.W., 2016. Machine learning and the profession of analytics for a new era of value-centered marketing in healthcare. J. Acad. Market. Sci. 48 (1), 9–23.

Davenport, T., Kalakota, R., 2019. The potential for artificial intelligence in healthcare. Futur. Healthc. J. 6 (2), 94–98.

Deo, R.C., 2015. Machine learning in medicine. Circulation 132 (20), 1920–1930.

Di Ieva, A., 2019. AI-augmented multidisciplinary teams: hype or hope? Lancet 394 (10211), 1801.

Downey, A., 2019. NHS bosses meet with tech giants to discuss commercial patient database. Digital Health available at: https:// www. digitalhealth.net/2019/12/nhs- bosses-meet-with-tech-giants-to-discuss-commercial-patient-database/.

Drew, L., 2019. The ethics of brain-computer interfaces. Nature 571, S19.

Eaneff, S., Obermeyer, Z., Butte, A.J., 2020. The case for algorithmic stewardship for artificial intelligence and machine learning technologies. JAMA 324 (14), 1397–1398.

Eduard, F.-V., Jordi, A.-C., 2019. I'll take care of you,' said the robot. Paladyn. J. Behav. Rob. 10 (1), 77–93.

European Medicines Agency, 2020. Guidance on the management of clinical trials during the COVID-19 (Coronavirus) pandemic. available at: https://ec.europa.eu/health/ sites/ health/files/files/eudralex/vol-10/ guidanceclinicaltrials_covid19_en.pdf. Faraj, S., Pachidi, S., Sayegh, K., 2018. Working and organizing in the age of the learning algorithm. Inf. Organ. 28 (1), 62–70.

Fenech, M., Strukelj, N., Buston, O., 2018. Ethical, social and political challenges of artificial intelligence in health. Avaiable at: https://wellcome.ac.uk/sites/default /files/ai-in-health-ethical-social-political-challenges.pdf.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R.,

Chazerand, P., Dignum, V., et al., 2018. AI4People - an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds Mach. 28 (4), 689–707.

Floridi, L., Luetge, C., Pagallo, U., Schafer, B., Valcke, P., Vayena, E., et al., 2019. Key ethical challenges in the European medical information framework. Minds Mach. 29 (3), 355–371.

Flyverbom, M., Deibert, R., Matten, D., 2019. The governance of digital technology, big data, and the internet: new roles and responsibilities for business. Bus. Soc. 58 (1), 3–19.

Galetsi, P., Katsaliaki, K., Kumar, S., 2019. Values, challenges and future directions of big data analytics in healthcare: a systematic review. Soc. Sci. Med. 241, 112533. Holub, P., Kozera, L., Florindi, F., van Enckevort, E., Swertz, M., Reihs, R., Wutte, A., Valik, D., Mayrhofer, M. Th., 2020. BBMRI–ERIC's contributions to research and knowledge exchange on COVID-19. Eur. J. Hum. Genet. 28 (6), 728–731. Hoorn, J.F., Winter, S.D., 2018. Here comes the bad news: doctor robot taking over. Int. J. Soc. Robot. 10 (4), 519–535.

Horgan, D., Romao, M., Morre,´ S.A., Kalra, D., 2019. Artificial Intelligence: power for civilisation – and for better healthcare. Publ. Health Gen. 22 (5–6), 145–161. Hosny, A., Aerts, H.J., 2019. Artificial intelligence for global health. Science 366 (6468), 955–956.

Ienca, M., Ignatiadis, K., 2020. Artificial Intelligence in clinical neuroscience: methodological and ethical challenges. AJOB Neurosci. 11 (2), 77–87.

ISO, 2014. Robots and robotic devices – safety requirements for personal care robots. Avaiable at. https:// www. iso. org/ standard/ 53820. html.

Jeste, D.V., Graham, S.A., Nguyen, T.T., Depp, C.A., Lee, E.E., Kim, H.-C., 2020. Beyond artificial intelligence: exploring artificial wisdom. Int. Psychogeriatr. 32 (8),993– 1001.

Jimenez´ -Luna, J., Grisoni, F., Schneider, G., 2020. Drug discovery with explainable artificial intelligence. Nat. Mach. Intel. 2 (10), 573–584.

Jobin, A., Lenca, M., Vayena, E., 2019. The global landscape of AI ethics guidelines. Nat. Mach. Intel. 1, 389–399.

Jung, H.H., Pfister, F.M., 2020. Blockchain-enabled clinical study consent management. Technol. Innovat. Manag. Rev. 10 (2), 14–24.

Kandalaft, M.R., Didehbani, N., Krawczyk, D.C., Allen, T.T., Chapman, S.B., 2013. Virtual reality social cognition training for young adults with high-functioning autism. J. Autism Dev. Disord. 43 (1), 34–44.

Kayaalp, M., 2018. Patient privacy in the era of big data. Balkan Med. J. 35 (1), 8–17. Larson, D.B., Magnus, D.C., Lungren, M.P., Shah, N.H., Langlotz, C.P., 2020. Ethics ofusing and sharing clinical imaging data for artificial intelligence: a proposed framework. Radiology 295 (3), 192536.

Littmann, M., Selig, K., Cohen-Lavi, L., Frank, Y., Ho¨nigschmid, P., Kataka, E., et al., 2020. Validity of machine learning in biology and medicine increased through collaborations across fields of expertise. Nat. Mach. Intel. 2 (1), 18–24. Lombard, M., Snyder-Duch, J., Bracken, C.C., 2002. Content analysis in mass communication: assessment and reporting of intercoder reliability. Hum. Commun. Res. 28 (4), 587–604.

Lupton, M., 2018. Some ethical and legal consequences of the application of artificial intelligence in the field of medicine. Trends Med. 18 (4), 1–7.

LuXton, D.D., 2014. Recommendations for the ethical use and design of artificial intelligent care providers. Artif. Intell. Med. 62 (1), 1–10.

Ma, H., Guo, X., Ping, Y., Wang, B., Yang, Y., Zhang, Z., Zhou, J., 2019. PPCD: privacy-preserving clinical decision with cloud support. PLoS One 14 (5), e0217349. Manrique de Lara, A., Pela´ez-Ballestas, I., 2020. Big data and data processing in rheumatology: bioethical perspectives. Clin. Rheumatol. 39 (4), 1007–1014.

Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. BMJ 368. https:// doi.org/10.1136/bmj.l6927.

Machine learning in cardiovascular medicine: are we there yet? Heart 104 (14), 1156–1164.

Martinez-Martin, N., Kreitmair, K., 2018. Ethical issues for direct-to-consumer digital psychotherapy apps: addressing accountability, data protection, and consent. JMIR Ment. Health 5 (2), e32.

McCall, B., 2020. COVID-19 and artificial intelligence: protecting health-care workers and curbing the spread. Lancet Dig. Health 2 (4), e166–e167.

McCoy, L.G., Nagaraj, S., Morgado, F., Harish, V., Das, S., Celi, L.A., 2020. What do medical students actually need to know about artificial intelligence? NPJ Dig. Med. 3 (1), 1–3.

McCradden, M.D., Joshi, S., Anderson, J.A., Mazwi, M., Goldenberg, A., Zlotnik Shaul, R., 2020a. Patient safety and quality improvement: ethical principles for a regulatory approach to bias in healthcare machine learning. J. Am. Med. Inf. Assoc. 27 (12), 2024–2027.

McCradden, M.D., Joshi, S., Mazwi, M., Anderson, J.A., 2020b. Ethical limitations of algorithmic fairness solutions in health care machine learning. Lancet Dig. Health 2 (5), e221–e223.

McGraw, D., Petersen, C., 2020. From commercialization to accountability: responsible health data collection, use, and disclosure for the 21st century. Appl. Clin. Inf. 11, 366–373, 02.

McNair, D., Price, W.N., Health care AI: Law, regulation, and policy, 2019. Artificial intelligence in health care: the hope, the hype, the promise, the peril. Washington DC: Nat. Acad. Med.

Mehta, M.C., Katz, I.T., Jha, A.K., 2020. Transforming global health with AI. N. Engl. J. Med. 382 (9), 791–793.

Mesko´, B., Hetenyi,´ G., Gyo˝rffy, Z., 2018. Will artificial intelligence solve the human resource crisis in healthcare? BMC Health Serv. Res. 18 (1), 545. https://doi.org/ 10.1186/s12913-018-3359-4.

Miller, D.D., 2020. Machine intelligence in cardiovascular medicine. Cardiol. Rev. 28 (2), 53–64.

Miller, E., Polson, D., 2019. Apps, avatars, and robots: the future of mental healthcare. Issues Ment. Health Nurs. 40 (3), 208–214.

Mittelstadt, B., 2017. Ethics of the health-related internet of things: a narrative review. Ethics Inf. Technol. 19 (3), 157–175.

Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., Floridi, L., 2016. The ethics of algorithms: mapping the debate. Big Data Soc. 3 (2), 1–21.

Mollura, D.J., Culp, M.P., Pollack, E., Battino, G., Scheel, J.R., Mango, V.L., Dako, F., 2020. Artificial intelligence in low-and middle-income countries: innovating global health radiology. Radiology 297 (3), 513–520.

Monegain, B., 2016. IBM Watson takes analytics prowess overseas: supercomputer to work on big data and genomics in Italy. Available at: https:// www.healthcareitnews. com/ news/ ibm- watson- takes-analytics-prowess-overseas- supercomputer-work-big- data-and-genomics-italy.

Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M., Floridi, L., 2021. Operationalising AI ethics: barriers, enablers and next steps. AI Soc. https:// doi. org/ 10. 1007/ s00146-021-01308-8.

Morley, J., Machado, C.C., Burr, C., Cowls, J., Joshi, I., Taddeo, M., Floridi, L., 2020. The ethics of AI in health care: a mapping review. Soc. Sci. Med. 260, 113172.

Newell, S., Marabelli, M., 2015. Strategic opportunities (and challenges) of algorithmic decision-making: a call for action on the long-term societal effects of 'datification. J. Strat. Inf. Syst. 24 (1), 3–14.

Noorbakhsh-Sabet, N., Zand, R., Zhang, Y., Abedi, V., 2019. Artificial intelligence transforms the future of health care. Am. J. Med. 132, 795–801.

Norgeot, B., Quer, G., Beaulieu-Jones, B.K., Torkamani, A., Dias, R., Gianfrancesco, M., et al., 2020. Minimum information about clinical artificial intelligence modeling: the ' MI-CLAIM checklist. Nat. Med. 26 (9), 1320–1324.

O Neil, C., 2016. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Penguin.

O Sullivan, S., Nevejans, N., Allen, C., Blyth, A., Leonard, S., Pagallo, U., et al., 2019. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. Int. J. Med. Robot. Comput. Assist. Surg. 15 (1), e1968.

Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S., 2019. Dissecting racial bias in an ' algorithm used to manage the health of populations. Science 366 (6464), 447–453.

Panch, T., Mattie, H., Celi, L.A., 2019. The "inconvenient truth" about AI in healthcare. NPJ Dig. Med. 2 (1), 1–3.

Panchmatia, J.R., Visenio, M.R., Panch, T., 2018. The role of artificial intelligence in orthopaedic surgery. Br. J. Hosp. Med. 79 (12), 676–681.

Park, S.H., Do, K.H., Kim, S., Park, J.H., Lim, Y.S., 2019. What should medical students know about artificial intelligence in medicine? J. Educ. Eval. Health Prof. 16, 18.

Peters, D., Vold, K., Robinson, D., Calvo, R.A., 2020. Responsible AI—two frameworks for ethical design practice. IEEE Trans. Technol. Soc. 1 (1), 34–47.

Pfohl, S.R., Foryciarz, A., Shah, N.H., 2020. An empirical characterization of fair machine learning for clinical risk prediction. J. Biomed. Inf. https://doi.org/ 10.1016/j.jbi.2020.103621.

Ploug, T., Holm, S., 2016. Meta consent–a flexible solution to the problem of secondary use of health data. Bioethics 30 (9), 721–732.

Poulsen, A., Fosch-Villaronga, E., Burmeister, O.K., 2020. Cybersecurity, value sensing robots for LGBTIQ elderly, and the need for revised codes of conduct. Austr. J. Inf. Syst. 24.

Powell, J., 2019. Trust me, i'm a chatbot: how artificial intelligence in health care fails the Turing Test. J. Med. Internet Res. 21 (10), e16222.

Powell, J., Fitton, R., Fitton, C., 2006. Sharing electronic health records: the patient view. J. Innovat. Health Inf. 14 (1), 55–57.

Powles, J., Hodson, H., 2017. Google DeepMind and healthcare in an age of algorithms. Health Technol. 7 (4), 351–367.

Price, W.N., Cohen, I.G., 2019. Privacy in the age of medical big data. Nat. Med. 25 (1), 37–43.

Quinn, T.P., Senadeera, M., Jacobs, S., Coghlan, S., Le, V., 2021. Trust and medical AI: the challenges we face and the expertise needed to overcome them. J. Am. Med. Inf. Assoc. 28 (4), 890–894.

Rai, A., 2020. EXplainable AI: from black boX to glass boX. J. Acad. Market. Sci. 48 (1), 137–141.

Rajkomar, A., Dean, J., Kohane, I., 2019. Machine learning in medicine. N. Engl. J. Med. 380 (14), 1347–1358.

Rajkomar, A., Hardt, M., Howell, M.D., Corrado, G., Chin, M.H., 2018. Ensuring fairness in machine learning to advance health equity. Ann. Intern. Med. 169 (12), 866–872.

Reddy, S., Allan, S., Coghlan, S., Cooper, P., 2020. A governance model for the application of AI in health care. J. Am. Med. Inf. Assoc. 27 (3), 491–497.

Rickert, J., 2020.'On patient safety: the lure of artificial intelligence—are we jeopardizing our patients privacy? Clin. Orthop. Relat. Res. 478 (4), 712–714.

Rigby, M.J., 2019. Ethical dimensions of using artificial intelligence in health care. AMAJ. Ethics 21 (2), 121–124. '

Ross, C., Swetlitz, I., 2018. IBM s Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show. Stat. News available at: https:// www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorre ct-treatments/.

Samuel, G., Derrick, G., 2020. Defining ethical standards for the application of digital tools to population health research. Bull. World Health Organ. 98 (4), 239–244.

Schwalbe, N., Wahl, B., 2020. Artificial intelligence and the future of global health. Lancet 395 (10236), 1579–1586.

Shameer, K., Johnson, K.W., Glicksberg, B.S., Dudley, J.T., Sengupta, P.P., 2018.

Shaw, J., Rudzicz, F., Jamieson, T., Goldfarb, A., 2019. Artificial intelligence and the implementation challenge. J. Med. Internet Res. 21 (7), e13659.

Siddaway, A., Wood, A., Hedges, L., 2019. How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. Annu. Rev. Psychol. 70, 747–770.

Song, S.Y., Kim, Y.-K., 2018. Theory of virtue ethics: do consumers' good traits predict their socially responsible consumption? J. Bus. Ethics 152 (4), 1159–1175.

Strydom, S.K., Strydom, M., 2018. Big data governance and perspectives in knowledge management. IGI Global.

Toh, T.S., Dondelinger, F., Wang, D., 2019. Looking beyond the hype: applied AI and machine learning in translational medicine. EBioMedicine 47, 607–615.

Trocin, C., Mikalef, P., Papamitsiou, Z., Conboy, K., 2021. Responsible AI for digital health: a synthesis and a research agenda. Inf. Syst. Front 1–19. https:// doi. org/ 10.1007/s10796-021-10146-4.

Urgent need for developing a framework for the governance of AI in healthcare. Stud. Health Technol. Inf. 272, 253–256.

Vayena, E., Blasimme, A., Cohen, I.G., 2018. Machine learning in medicine: addressing ethical challenges. PLoS Med. 15 (11), e1002689.

Vayena, E., Salathe,´ M., Madoff, L.C., Brownstein, J.S., 2015. Ethical challenges of big data in public health. PLoS Comput. Biol. 11 (2), e1003904.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., et al., 2020. The role of artificial intelligence in achieving the Sustainable Development Goals. Nat. Commun. 11 (1), 1–10.

Vollmer Dahlke, D., Ory, M.G., 2020. Emerging issues of intelligent assistive technology use among people with dementia and their caregivers: a US perspective. Front. Public Health 8 (191). https://doi. org/ 10. 3389/ fpubh. 2020.00191.

Vollmer, S., Mateen, B.A., Bohner, G., Kira´ly, F.J., Ghani, R., Jonsson, P., et al., 2020.

Wamba, S.F., Queiroz, M.M., 2021. Responsible artificial intelligence as a secret ingredient for digital health: bibliometric analysis, insights, and research directions. Inf. Syst. Front 1–16. https://doi.org/10.1007/s10796-021-10142-8.

Wang, Y., Kosinski, M., 2017. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. J. Pers. Soc. Psychol. 114 (2), 246–257.

Wang, Y., Kung, L., Byrd, T.A., 2018. Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. Technol. Forecast. Soc. Change 126, 3–13.

Wangmo, T., Lipps, M., Kressig, R.W., Ienca, M., 2019. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. BMC Med. Ethics 20 (1), 98.

Wearn, O.R., Freeman, R., Jacoby, D.M., 2019. Responsible AI for conservation. Nat. Mach. Intel. 1 (2), 72–73.

World Health Organization, 2021. Ethics and governance of artificial intelligence for health. Available at: https:// www. who. int/ publications/ i/item/9789240029200. Wright,

S.A., Schultz, A.E., 2018. The rising tide of artificial intelligence and business automation: developing an ethical framework. Bus. Horiz. 61 (6), 823–832.

Xafis, V., Schaefer, G.O., Labude, M.K., Brassington, I., Ballantyne, A., Lim, H.Y., Lipworth, W., Lysaght, T., Stewart, C., Sun, S., Laurie, G.T., Tai, E.S., 2019. An ethics framework for big data in health and research. Asian Bioethics Rev. 11 (3), 227–254.

Yew, G.C.K., 2020. Trust in and ethical design of carebots: the case for ethics of care. Int. J. Soc. Robot. https://doi.org/10.1007/s12369¨ ¨ -020-00653-w. Yüksel, B., Küpçü, A., Ozkasap, O., 2017. Research issues for privacy and security of electronic health services. Future Generat. Comput. Syst. 68, 1–13.

Zhang, J., Hon, H.W., 2020. Towards responsible digital transformation. Calif. Manag. Rev. 62 (3).

Zitnik, M., Nguyen, F., Wang, B., Leskovec, J., Goldenberg, A., Hoffman, M.M., 2019. Machine learning for integrating data in biology and medicine: principles, practice, and opportunities. Inf. Fusion 50, 71–91.