Cuadernos de Bioética. 2026; 37(119): 23-50
DOI: 10.30444/CB.201

A QUALITATIVE EVIDENCE SYNTHESIS ON THE BIOETHICAL ISSUES OF THE USE OF ARTIFICIAL INTELLIGENCE IN HEALTHCARE AND PROPOSED SOLUTIONS

SÍNTESIS DE EVIDENCIA CUALITATIVA SOBRE LOS PROBLEMAS BIOÉTICOS DEL USO DE INTELIGENCIA ARTIFICIAL EN SALUD Y SOLUCIONES PROPUESTAS

RAMÓN J. PALMER-COLL[1] *https://orcid.org/0009-0008-3342-1975*
VERÓNICA A. BROKKE[2] *https://orcid.org/0009-0002-9094-3047*
PATRICI CALVO[3] *https://orcid.org/0000-0002-3228-9019*
1. Universidad Jaume I de Castelló. Kennedy Institute of Ethics, Georgetown University. al384209@uji.es
2. Universidad CEU Cardenal Herrera.
3. Universitat Jaume I de Castelló.

RESUMEN:

Palabras clave:
Bioética, Revisión Sistemática, Inteligencia Artificial, Big Data, Ética Clínica

Recibido: 03/03/2025
Aceptado: 10/11/2025

Antecedentes: El aumento exponencial del uso y desarrollo de inteligencia artificial (IA) en salud ha traído varios problemas éticos, pero no hay una guía clara sobre como garantizar un uso ético de IA médica. **Método**: Se llevó a cabo una revisión sistemática cualitativa de estudios sobre el panorama bioético con relación a problemas éticos, principios y soluciones del uso de IA médica. Se realizó una búsqueda en MEDLINE, Scopus y otras tres bases de datos, junto con otros documentos relevantes. Se usó la herramienta GRADECERQual para evaluar la confianza en nuestros hallazgos. Se usó un análisis temático para analizar los resultados. **Resultados**: Se seleccionaron 180 estudios. Se clasificaron los resultados en: problemas éticos, principios éticos y soluciones. **Discusión**: Se desarrolló un modelo teórico del uso de IA en salud basado en nuestros resultados. Se definió y analizó la necesidad de incluir *explicabilidad* como quinto principio bioético y se desarrolló un *checklist* para el uso clínico de IA.

ABSTRACT:

Keywords:
Bioethics, Systematic Review, Artificial Intelligence, Big Data, Clinical Ethics

Background: The exponential increase of the use and development of artificial intelligence (AI) in healthcare has brought several ethical concerns, but there is no clear guidance on how to guarantee an ethical use of Medical-AI. **Method**: We carried out a qualitative systematic review of studies about the bioethical landscape regarding ethical issues, principles and solutions of using Medical-AI. We searched MEDLINE, Scopus, and three other databases, together with relevant guidelines and frameworks. The GRACE-CERQual tool was used for confidence assessment of findings. Thematic analysis was used to analyze results. **Results**: 180 studies were selected. We developed a framework to classify our findings into: Ethical Issues, Ethical Principles and Solutions. **Discussion**: We used our findings to develop a theoretical logic model for an ethical medical-AI. We defined and discussed the need for explicability as the fifth bioethical principle and developed a checklist for an ethical use of medical-AI by clinicians.

1. Introduction

In 1950 Artificial Intelligence (AI) was defined as the branch of computer science in which machine-based approaches were used to mimic what human intelligence might do(1). In recent years, there has been an increase in the use and design of medical-AI. The Federal Drug Administration (FDA) has approved several AI systems(2). More recently, the European Union AI Act defines AI as:

"A machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments." (Article 3, paragraph 1)

And classifies Medical-AI devices as high risk[3].

AI is being used to analyze medical images like mammograms[4] and dermatological images[5], to recognize patterns in electrocardiograms[6], ultrasound[7] and patients voices to diagnose psychiatric disorders[8]; to develop prosthetic limbs[9], to sequence DNA and RNA to evaluate drug efficacy[10], to quantify disease severity[11], in epidemiology[12,13], to assist decision making[14], implement informed consent (IC) by representation[15], to discover new drugs and decrease costs[16].

However, Medical-AI has raised several ethical concerns. AI could reduce collective and individual freedoms, life-quality, affect the economy and clash with fundamental rights[8,14,15], but results are heterogeneous. There is no clear guidance on how to guarantee an ethical medical-AI. Furthermore, the implications of the European Union AI Act[3] for digital medicine are unclear[17]. Ethical issues, solutions and principles vary. Concerns exist on the effect these problems might have on the four bioethical principles[18], and if these principles need expanding, but with no consensus. The great number of documents published and lack of qualitative synthesis reviews, make it difficult to create a clear framework and a robust theoretical logic model for AI-bioethics. There is need for better understanding the bioethical problems of using Medical-AI, their effect on bioethical principles and what solutions exist to address them.

We asked what the bioethical implications of using AI in healthcare are and what proposals, including ethical principles, exist to address these problems, using the PerSPecTIF Question formulation[19]:

- **Perspective:** clinical bioethics.
- **Setting:** global.
- **Phenomenon of interest:** use of AI.
- **Environment:** healthcare.
- **Comparison:** to not using AI.
- **Time:** Up to 18/11/2024.
- **Findings:** ethical issues, ethical principles and solutions proposed.

2. Objectives

Overall aim: carry out a qualitative review of the bioethical landscape regarding the use of AI in healthcare.

Specific objectives:

- Identify, appraise and synthesize qualitative research on the ethical problems for the implementation of AI in healthcare.
- Develop a framework to classify ethical issues regarding the use of medical-AI.
- Identify, appraise and synthesize qualitative research evidence on bioethical principles proposals to guarantee an ethical use of AI in healthcare.
- Identify and synthesize qualitative research on solutions for an AI technology that meets bioethical standards.
- Develop a framework to classify proposed solutions for bioethical use of medical-AI.
- Establish a classification of ethical principles for medical-AI.
- Integrate our findings to develop a theoretical-logic model for a bioethical medical-AI.
- Integrate our findings to develop a checklist for an ethical use of AI in medical practice.
- Use our findings to define the bioethical principle of Explicability and establish the need/or not for this principle as the fifth bioethical principle.

3. Method

We carried out a systematic qualitative evidence synthesis following an adapted PRISMA protocol[20], as proposed by Flemming et al (2018)[21]. We used the Cochrane Protocol and Review Template for Qualitative Evidence Synthesis as a guide to develop this review[22].

3.1. Criteria for study consideration

3.1.1. Types of studies

We included primary studies with qualitative study designs such as ethnography, phenomenology, case studies, grounded theory studies and qualitative process evaluations. We included studies that use qualitative methods for data collection and analysis. Based on Noyes et al (2019)[23], we included: non-research-policy and regulatory frameworks, qualitative studies on views and experiences, and discourse analysis.

We included mixed methods studies where it was possible to extract and analyze data qualitatively. We included studies regardless of whether they were conducted alongside studies of the effectiveness of using AI in medicine or not.

We excluded studies that collect data using qualitative methods but do not use qualitative analysis methods.

3.1.2. Topics of interest

Inclusion criteria: studies about ethical problems of use of AI in healthcare, studies about AI and bioethical principles, studies about bioethical principles proposals for an ethical AI in healthcare, studies proposing solutions to ethical problems of using AI in healthcare, studies in English or Spanish language only.

Studies not meeting inclusion criteria were excluded, as well as: papers on Big Data or AI not mentioning Bioethics, studies dealing only with AI development and not with clinical practice, lack of integration between AI and ethics (i.e. articles whose main focus is not a qualitative analysis of bioethics and AI and merely mention AI or bioethics), articles not related to human health, previous review articles, and duplicates.

3.2. Sources of information and search strategy

We carried out a systematic review using MEDLINE PubMed Ovid, Scopus Elservier, TRIP, Epistemonikos and LILACS. The reference list of selected articles was reviewed, as well as books and other relevant documents from national and international agencies. For these, the same selection criteria were used. Last search date was November 18, 2024.

Table 1. Search Strategies

Data Base	Search Strategy
PubMed	(((((artificial intelligence[Title/Abstract])) OR (AI[Title/Abstract])) OR (artificial intelligence[MeSH Terms]) AND (ffrft[Filter])) AND ((((((((((bioethics[Title/Abstract]) OR (bioethics[MeSH Terms])) OR (medical ethics[MeSH Terms])) OR (medical ethics[Title/Abstract])) OR (morals[Title/ Abstract])) OR (morals[MeSH Terms])) OR (clinical ethics[MeSH Terms])) OR (clinical ethics[Title/Abstract])) OR (ethics[Title/Abstract])) OR (ethics[MeSH Terms]) AND (ffrft[Filter]))) AND ((((Principles based ethics[MeSH Terms]) OR (principles based ethics[Title/Abstract])) OR (Bioethics principles[Title/ Abstract])) OR (principles[Title/Abstract]) AND (ffrft[Filter]))
Scopus	(TITLE-ABS (“artificial intelligence”) OR INDEXTERMS (“artificial intelligence”)) AND (TITLE-ABS (“bioethics”) OR INDEXTERMS (“bioethics”) OR TITLE-ABS (“medical ethics”) OR INDEXTERMS (“medical ethics”) OR TITLE-ABS (“morals”) OR INDEXTERMS (“morals”) OR TITLE-ABS (“clinical ethics”) OR INDEXTERMS (“clinical ethics”) OR TITLE-ABS (“ethics”) OR INDEXTERMS (“ethics”)) AND (TITLE-ABS (“bioethics principles”) OR INDEXTERMS (“bioethics principles”) OR TITLE-ABS (“principles based ethics”) OR INDEXTERMS (“principles based ethics”) OR TITLE-ABS (“principles”) OR INDEXTERMS (“principles”)) SUBJAREA (medi OR nurs OR heal OR bioc OR immu OR neur OR phar)
TRIP	(bioethics OR clinical ethics OR ethics) AND (artificial intelligence) AND (bioethics principles OR principles based ethics OR principles)
LILACS	inteligencia artificial Y bioetica
Epistemnikos	bioethics AND artificial intelligence

Our search strategy can be seen in **Table 1**. Terms used for search strategy where: Artificial intelligence, AI, Ethics, bioethics, medical ethics, clinical ethics, morals, principles-based ethics, bioethics principles, principles. A first search was carried out in PubMed. This search strategy was modified to fit each database. We did not apply limits on language, publication date or other criteria.

3.3. Sampling of studies

A first selection was carried out using paper title and abstract, to identify those meeting selection criteria. Eligibility of each paper was evaluated by reading full text. To avoid duplicates, RefWorks was used. 2 authors (RJPC, VABG) reviewed the data.

Qualitative evidence synthesis aims for variation in concepts[24]. Deductive and inductive saturation approaches were discarded as we did not have a priori categories, but rather developed categories based on results. Furthermore, there is no consensus on how inductive saturation should be achieved. This way 'fill gaps in the data'[25] following saturation can guarantee a more complete theory.

3.4. Data extraction

Data extracted from each article was: study design, study participants (authors if there were no participants), institutional affiliation of the first author, country, publication year and method. These were introduced into a **Characteristics of Study Table (Appendix Table 1)**. We classified our findings into three categories: ethical issues, ethical principles, solution proposals. We developed a framework to classify ethical issues and solutions and integrated it into a summary of findings table.

3.5. Risk and methodological limitations assessment

As this is a synthesis of qualitative studies, no risk of bias assessment was carried out. We assessed study quality as follows.

We used CASP (2018) to assess methodological limitations. For Research value, we classified articles as low, moderate, or high, based on[26]:

- Discussing contributions to existing knowledge or understanding.
- Identifying areas where research is necessary.
- Discussing whether or how findings can be transferred to populations or ways to use findings.

Articles were classified as high value and downgraded if did not fulfil criteria.

We reported our assessments in a **Methodological Limitations Table (Appendix Table 2)**. We included studies meeting eligibility criteria regardless of quality. Poorer quality studies contribute less to synthesis[27]. Synthesis becomes "weighted" towards findings of better-quality studies[24].

3.6. Data management, analysis and synthesis

A thematic synthesis of ethical problems from using medical-AI, bioethical principles proposals and other solutions was carried out, as well as a quantitative synthesis of ethical issues and principles in included studies. We examined findings and developed prompts for future implementers. We used a logic model approach[28,29] to develop a logical flow of theoretical connections through which AI characteristics relate to ethical concerns and how these relate to proposed solutions. We developed a checklist (see Implications for Practice Section), phrased as questions, not recommendations. We sent this to different medical colleges and associations to gather feedback about relevance and way they were phrased and presented.

3.7. Assessing confidence in review findings

GRADE-CERQual was used to assess confidence in each finding[30]. After assessing each component, we judged overall confidence in evidence supporting each finding. We judged confidence as high, moderate, low, or very low. All findings started as high and were downgraded if there were concerns regarding any GRADE-CERQual component.

We present summaries on our confidence assessments in our **Summary of Qualitative Findings**, and detailed descriptions of our confidence assessment in an

Evidence Profile (**Appendix Table 3)**. The summary of qualitative findings table (**Table 2**) and appendix tables can be found at the repository at SSRN at https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=7552022

4. Results

180 sources published between 1989 and 2024 were analyzed. A PRISMA flow diagram was created **(Figure 1)**[31]. Articles were in English or Spanish. Most were thematic or framework synthesis, and critical analysis. For detailed descriptions of included studies see **Characteristics of Included Studies (Appendix Table 1)**.

4.1. Methodological evaluation

Most studies (n=120) were classified as moderate. Most failed to address methodological limitations or justify their data collection strategy, especially hermeneutical and critical analysis. For individual methodological limitations assessments, see **Appendix Table 2**.

4.2. Confidence evaluation

Most findings were moderate confidence. Findings where frequently downgraded due to methodological limitations, and low adequacy (quantity) of data. For individual confidence assessments see **Appendix Table 3**.

4.3. Findings

Findings were classified into three categories: ethical concerns, ethical principles and solutions. Frameworks for bioethical concerns and solutions were developed (**Figure 2**). Findings were classified using these frameworks. Extracted data was introduced into **Summary of Qualitative Findings Table.**

4.3.1. Ethical problems of the use of medical-AI (Findings 1-47)

Ethical concerns findings were organized into 12 overarching framework categories (**Figure 2**): Autonomy; Technological Imperialism; Maleficence (uncertain-

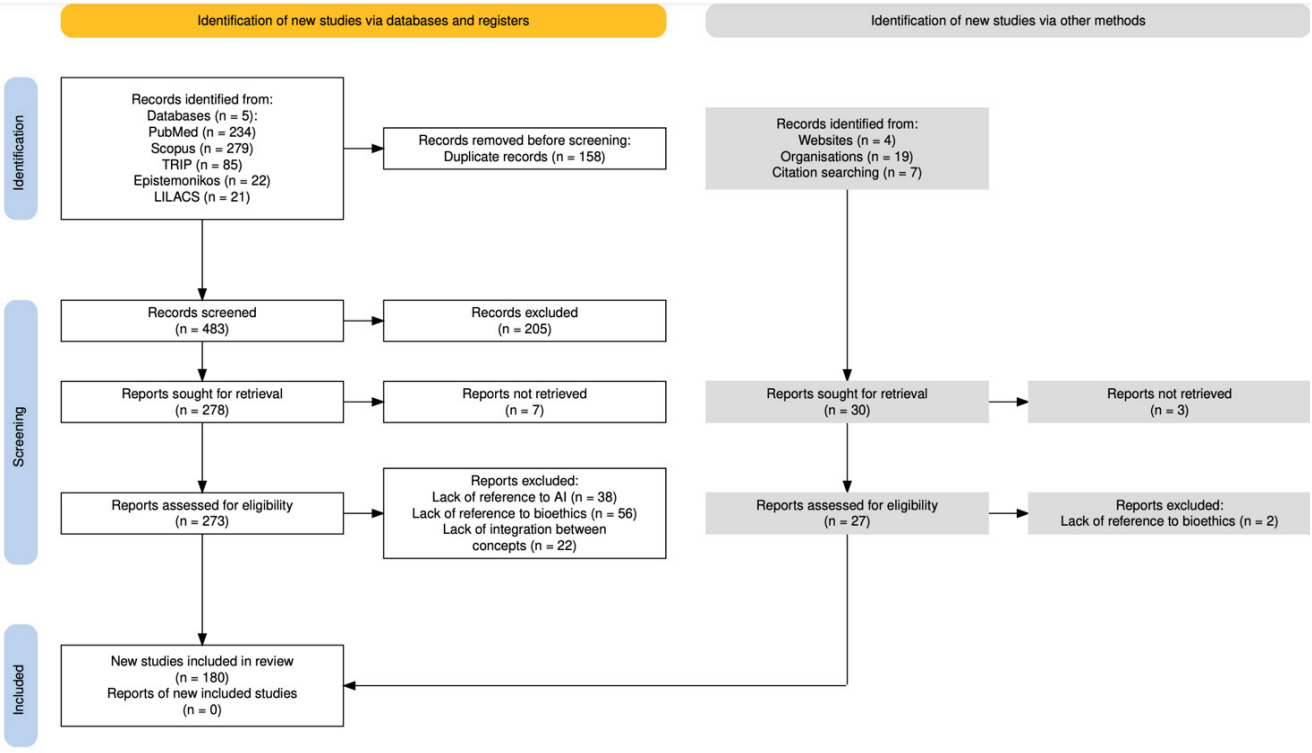

Figure 1. PRISMA flow diagram of included studies.

ty in AI behavior and caution), Clinical Relationships; Drug Development; Responsibility, Decision-Making and Transparency; Justice and Discrimination; Security; Human Rights; Employment; Sustainability; Self-Regulation. Most mentioned ethical issues included: lack of evidence to use AI safely (n=42), increased discrimination (n=40), lack of privacy (n=36), unclear responsibility (n=31), and using biomedical data (n=30).

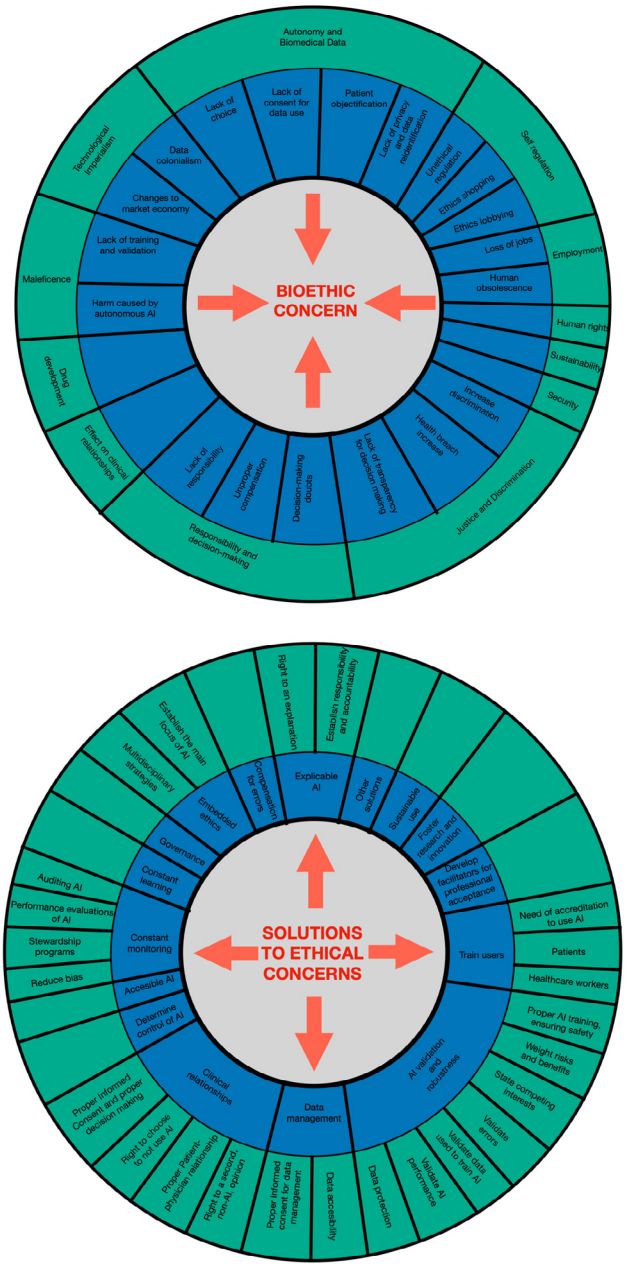

Figure 2. Framework of bioethical concerns and solutions for Medical-AI

4.3.1.1. Autonomy

Using AI, regardless of its benefits, might hinder autonomy[32]. Patients might lose the right to choose their treatment[33,34]. AI might be their only option, leading to lack of choice[35-42]. Insurance companies may force people to use AI, without alternative[33]. Language processing of documents might limit patient autonomy and their ability to choose their treatments[43]. (*Finding 1)*.

Due to the amount of data needed to train AI, patient records must be accessed[46]. This could lead to patients being treated as a source of data (objects) rather than human subjects, violating autonomy[14,47-49,142]. Patient's value would depend on their data production capacity[33,52-56,78]. (*Finding 2)*.

Data anonymization might deny someone's right to control their information[38]. Even using anonymization (*Findings 184-187*), reidentification is sometimes possible[57]. (*Finding 3)*. This can cause breaches in autonomy and privacy[52]. *(Finding 4).*

It is unclear what external parties might use patient data for[33,40,51,52,58-68], it could be used for means different to those initially intended[38,41,55], which is difficult to predict[37,43]. Possibly leading to data mercantilism[32], where personal information is used to amass wealth, with no ethical considerations; or harming people[69-73]. Data mercantilism may be due to a behavioral data surplus, which increases chances of data been used for other means[32]. *(Finding 5)*.

Sensible information might be required to develop AI, hindering patient privacy[32,35,36,38,40,42,46-48,50-53,56,67,72-74,78-90,104,175]. Lack of privacy may lead to increase stigmatization[150]. (*Finding 6)*. Furthermore, a breach in AI data, could expose biometric information, increasing discrimination further[52]. (*Finding 7)*. Data breaches could identify a person's political opinions, ethnicity, sexual orientation, and overall health: «Keyboard typing patterns may be used to infer emotional states"[52]. (*Finding 8)*.

High-processing power algorithms designed to ensure privacy might give good result when tested on small datasets but could underperform with larger datasets[52]. *(Finding 9)*.

Despite accreditation (*Finding 216*), "If an (accredited) adversary gets access to information. Consequences could be severe", as it is challenging to identify such intruders[52]. (*Finding 10)*.

AI can lead to patient infantilization, where patients are treated as non-independed humans, incapable of taking health decisions[47]. (*Finding 11)*. Lack of a proper IC can lead to patient information been used for purposes patients might not approve[32,33,35-38,43]. (*Finding 12)*. For these problems, several proposals exist, including proper informed consent has been proposed [38,40,43,46,49,52,56,58,68,82,89,94,106,163,165,172]. *(Finding 176).*

4.3.1.2. Technological imperialism

AI development by few companies, means AI will be in hands of a small group[33]. Creating a monopoly. Leading to fewer innovation that could potentially improve lives[33,50]. This way, AI could decrease life quality. *(Finding 13)*.

The use of patients as data producers *(Finding 2)*, can be accentuated with data colonialism. This "combines predatory practices used in historical colonialism with abstract quantification of computational methods" to extract data as a valuable resource[96]. Data is turned into products that might bring no benefit to populations of origin[41,78,88,96-98]. Without a deep understanding of context and significance of extracted data, even experts with good intentions might make errors[97]. *(Finding 14)*.

4.3.1.3. Maleficence (uncertainty in behavior of ai and caution)

Autonomous Robot-Caregivers could increase feelings of solitude and sadness, worsening vulnerability and social exclusion of elderly patients[99-102]. *(Finding 15)*.

A low trustworthiness of AI´s predictability models, and lack of scientific evidence to back up its safety could harm patients[39,49,56,63,65,66,68,75,80,87,98,102,104-110,112-129]. There is a risk of AI been implemented before subject to rigorous security and efficacy controls[46]. This could be due to "technological solutionism", where advanced technologies are used as "magic wands" to dissolve social, cultural, economic, and institutional barriers; or due to overestimating the system's capacity; increasing background noise and decreasing efficacy[33].

Some predictability models can represent an unnecessary expense[33]: an algorithm designed to predict mortality gave a higher risk to people that had visited a specific city, when this had no effect on life expectancy[111]. *(Finding 16)*.

4.3.1.4. Clinical relationships

AI could improve patient-physician relationships as AI might do bureaucratic work, allowing clinicians to spend more time with patients[33,101,131]. In this regard, not using AI could be unethical. *(Finding 17).*

Human emotions and human contact are an essential part of patient-physician relationship, which is a key element for proper healing, especially in psychiatry. AI could damage this relationship, decreasing quality of care[37,49,54,57,63,68,78,83,86,101,112,121,125,128,133-140]. Finding 18. Furthermore, Ai could decrease human contact[48]. *(Finding 18)*. Trust, both in physicians and healthcare systems, could be affected if patients found out AI was used without IC. If this is not managed, healthcare workers could feel less fulfilled with their jobs[33]. *(Finding 19).* For these reasons, inclusive communication[63,173] and the right to a second non-AI opinion[34], among other proposals were identified. *(Findings 180 and 183).*

4.3.1.5. Drug development

There are already AI-developed drugs tested[177]. Even though use of AI could reduce development and commercialization times and costs, pharmaceuticals could use IA to jump security and quality controls by performing less clinical trials or using smaller sample sizes[33,59,63,66,97,141]. AI can also expose patients to excessive pharmacological marketing[33]. *(Finding 20)*.

4.3.1.6. Responsibility, decision-making and transparency

If an AI system is involved, in an adverse event, it is unclear if AI can be held accountable, and what level of responsibility will be attributed to humans, regardless of the human capacity to overrun the mis-

take[37,39,43,47,56,61,65,71,78,81,82,86,102,105,106,110,122,127,128,131,133,134,139,143-149].

It is difficult to make one person responsible. Designers, healthcare workers, and institutions might all play a role, while no one is completely responsible[149], leading to dilution of responsibility[142]. On occasions responsibility to determine liability relies on institutions to decide who is responsible internally, possibly leading to scapegoat situations, arguing an event was human error, when humans had little to no control over AI[133]. *(Finding 21).*

In order to solve this issue, many studies mention the importance of establishing responsibility, but are unclear on who or what should be held accountable (*Finding 121*)[33,51,53,54,81,82,105,143], some articles establish humans are always responsible for AI actions (*Finding 164*)[37,43,50,136,148], some argue physicians are to be held accountable (*Finding 124*)[61,144,160,189], while others designers should be responsible (*Finding 124*)[122,133], and others established shared responsibility between all parties (*Finding 125*)[122,133].

Dilution of responsibility can lead to patients not being properly compensated, or to patients not fully understanding how harm was brought upon them[65,75,124,133]. *(Finding 22)*. Motives to look for compensation can be hindered, as many AIs are patented, privately owned, making them difficult to inspect, lacking transparency (*Finding 23*) and explicability[81,144]. Lack of transparency and disinformation about criteria and reasoning of AI (specially in Black Box algorithms)[33,49,113,114], could decreases AI-trust by patients and physicians and hinder clinical decision-making[33,39,43,49,54,56,66,67,75,80,86,88,112,118,120,122,129,137,146,149-155].

Overreliance on AI could harm patients and develop a sense of loss of control by professionals[33,35,42,50,61,65,69,70,74,87,109,113,114,116,118,121,127,129,138,140,149,158-162]. *(Finding 24)*. Factors that could influence this are:

- Physicians might not be able to know if an AI decision is biased/incorrect or if AI knows something physician does not[149,163,164]. Furthermore, we need to put this into a clinical context, where decisions are taken under imperfect conditions, hindering a proper AI evaluation[163]. *(Finding 25)*.
- Computer knows best bias closes the door to dialogue under the premise that algorithms know more than physicians[165]. *(Finding 26)*.
- Automatization bias: when it is believed something automatized is trustworthy and safe, regardless of if it really is[49,88,158]. *(Finding 27)*.
- AI might not be capable of returning control to humans, and even if it does, humans might not have enough time to react[80,118]. *(Finding 28)*.

Several explicable AI mechanisms have been proposed:

- There must be mechanisms to determine accountability, including professional codes of ethics and conduct to determine individual accountability, institutional policies to determine organizational accountability, transparency to stakeholders, and regulation by governments and professional entities to determine regulatory accountability[51]. *(Finding 127)*.
- Medical-AI must include mechanisms to make it understandable. Physicians must know why an AI took a certain decision and how it works. This could help humans trust AI decisions[51,56,119,127,147]. *(Finding 128)*.
- Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOT) model. This 22-item checklist consideres important points for transparent reporting of predictive models[147,154]. *(Finding 129)*.
- Individuals are currently not able to properly challenge results provided by algorithms. Individuals must have the right to an explanation of how and why an AI took a decision[113]. *(Finding 130)*.

4.3.1.7. Justice and discrimination

AI can discriminate people due to religion, race, gender, or other reasons[41]. Biased AI can deny certain populations a better and fair healthcare access[35,39,48,50,51,56,60,62,64-66,69,71,82,85-88,97,98,102,106,109,113-115,123,126,131,136,144,147,150,155,170]. Potential discrimination from insurance companies is also concerning[32,60,71]. *(Find-*

ing 29). This could increase health inequality[33,46,50,56,57,59,64,70,72,76,78,88,114,155,159,166-169]. *(Finding 30)*. The epistemic status of AI outputs has also been questioned[171].

Reasons for a biased AI include:

- Human intrinsic biasses could be introduced into algorithms. These could cause greater discrimination and harm (faster and affecting more people) than a single biased individual[33,69,74,107,135]. *(Finding 31)*.
- Non-representative or insufficient data sets could discriminate underrepresented grou ps[33,56,66,69,85,109,136,152,174]. *(Finding 32)*.
- If AI is trained with data from the place it is being developed, but sold globally, without taking into consideration that data used to train AI might not be ideal nor representative of all societies where it is indented to be sold at, IA could be biased towards them[33,69]. *(Finding 33)*.
- Models used to train and develop AI might not account for bias[69]. *(Finding 34)*.
- Lack of transparency and de-biasing techniques[69]. *(Finding 35)*.
- Lack of model validation[69]. (Further explored in *Findings 78 and 191-200). (Finding 36)*.
- Implementation of AI might differ across regions, leading to different AI outputs and potential biases[69]. *(Finding 37)*.

AI could not only improve clinical relationships (*Finding 17*), but potentially improve healthcare, life quality and expectancy. In this regard, not using AI would be unethical[33,122]. *(Finding 38)*.

It is possible that cheaper AI, more widely used and easily accessed, could come at the price of privacy, as more information could have to be shared to access IA. Certain individuals might have to give up autonomy and privacy to access AI[72]. *(Finding 39)*.

4.3.1.8. Security

AI could become a target for cyber-attacks, hindering patient rights, harming patients and workers[32,33,47,50,51,56,64,71,83,87,88,104,111,113,123,147,175]. *(Finding 40)*.

4.3.1.9. Human rights

With Human Rights being placed at the epicenter of AI regulatory frameworks[33,56,182,183,185], some concerns exist regarding the negative effect AI could have on human rights[33,37,53,78,88,142]. *(Finding 41).*

4.3.1.10. Employment

AI could increase unemployment by automatization of many human tasks[33,50,141,176]. *(Finding 42)*. This could increase feelings of human obsolescence against machines, leading to a further decrease in employment[177]. *(Finding 43)*.

4.3.1.11. Sustainability

The World Health Organization (WHO) considers climate change a global health issue[33]. There are concerns that AI development and use might be unsustainable[33,50,56,62,88,115,178]. *(Finding 44)*.

4.3.1.12. Self-Regulation

Companies developing internal AI ethical standards and guidelines, rises these concerns: lack of public involvement in development (not all stakeholders might be involved in developing these frameworks)[33,180], prospective application of standards (not accepting causal or retrospective responsibility), internal auditing and lack of transparency in their application[33], lack of external coercion for non-compliance with ethical standards[33,98,142,180]. *(Finding 45)*.

Companies might choose to use only ethical principles that allow them to keep their unethical practices; while arguing they follow ethical standards, creating a market for ethical principles[178]. *(Finding 46)*. Some groups might lobby to avoid certain ethical standards and regulations from affecting them and their practices[178]. *(Finding 47)*.

4.3.2. Ethics principles proposals (Findings 48-120)

73 different ethical principles proposals were identified. Most cited principles where: Transparency (n=35), Explicability (n=34), Justice (n=34) and Privacy (n=29). All principles cited can be seen in **Table 2**.

Table 2. Bioethical Principles Proposed for an Ethical Use of AI in Healthcare

Principle	Citations	Principle	Citations
Transparency	35	Trust	2
Justice	34	Avoidance of Bias	2
Explicability	34	Robustness	2
Autonomy	29	Inclusion	2
Privacy	29	Vigilance	1
Accountability	21	Body and Mental Integrity	1
Non-maleficence	19	Social well-being	1
Beneficence	18	Iterativity	1
Responsibility	18	Usability	1
Governance	18	Contestability	1
Trustworthiness	16	Decision-making	1
Sustainability	11	Lawfulness	1
Human Dignity	11	Application of Human Values	1
Safety	7	Liability	1
Solidarity	7	Equality	1
Diversity/Pluralism	6	Controllability	1
Data Protection	5	Validation	1
Security	5	Feasibility	1
Fairness	5	Consent of Data Use	1
Democracy	5	Generalizability	1
Scientific Inquiry	4	Sharing of Benefits	1
Compassion	4	Governability	1
Equity	4	Humanity	1
Human well-being	4	Quality Control and Supervision	1
Inclusivity	3	Teamwork	1
Non-discrimination	3	Relevance	1
Accessibility	3	Education of AI	1
Confidentiality	3	Reliability	1
Interventions	2	Updateability	1
Interpretability	2	Caution	1
Communication	2	Utility	1
Accuracy	2	Honesty	1
Academic Integrity	2	Validity	1
Prevention of Harm	2	Rule of Law	1
Informed Consent	2	Vulnerability	1
Cooperation	2	Awareness of Misuse	1
Trust	2	Human Rights	1

4.3.3. Solution proposals (Findings 121 – 229)

4.3.3.1. Responsibility (Findings 121 – 126)

Articles mention the importance of establishing responsibility but are unclear who or what should be held accountable for medical actions involving AI[33,51,53,54,81,82,105,143]. Some articles establish humans are always responsible for AI actions[37,43,50,136,148], others establish physicians specifically are to be held accountable for actions involving AI in healthcare[61,144,160,189], others state that developers must be accountable for the performance of their models[157]. Shared responsibility between all involved parties[122,133] has also being proposed. Fostering self-accountability can help achieve a more ethical AI in healthcare[192], for this reason a studie argued there must be mechanisms in place to determine accountability[51].

4.3.3.2. Explicable AI (Findings 127 – 130)

Medical-AI must have the proper mechanisms in place to be understandable by its users. Physicians must know why an AI took a certain decision and how the AI works. This could help humans trust an AI decision. For this purpose, XAI (explicable AI)[54,56,119,127,199] and the TRIPOT model (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis model)[54,154] have being proposed. The TRIPOT model is a checklist of 22 items considered important for transparent reporting of predictive models including model specification, performance and validation. Furthermore, individuals are currently not able to properly challenge the results provided by algorithms, so individuals must have the right to an explanation of how and why an AI takes a decision[113,172].

4.3.3.3. Compensation for errors (Findings 131 – 132)

We must go beyond establishing responsibility and establish proper systems to make sure patients are compensated for any harm caused by AI[33]. For this purpose, a framework of reimbursement in AI-healthcare has been proposed[65].

4.3.3.4. Embedding ethics into AI design (Findings 133 – 152)

To guarantee an ethical use of Medical-AI, many articles mention the need to embed ethics into the design of Medical-AI. Several proposals exist:

- AI focus: AI must have humans at the center and be design around them. AI must always serve human interests. Patients must always go first for the algorithm[89,100,116,124,126,156,160,200,201]. This implies a Human centered AI[53,56,71,80,131,142,144,159,167,169,182,183,185,199,202,203].
- Multidisciplinary strategies and ethical deliberation: there must be a multidisciplinary ethical deliberation about AI during its development and implementation[102]. This should include patients [69], healthcare workers [49,144], and other stake holders [35,54,68,73,88,102,114,128,133,148,156,196].
- Social responsibility strategies: to guarantee clear accountability and liability for any harm caused by AI [126].
- Value sensitive design: a value sensitive design could mitigate some ethical concerns regarding AI[200].
- Ethical alignment: this is a principles-driven approach to guarantee AI behaves in an ethically acceptable way[157,200].

However, other authors argue: "for an AI to perform "ethically", we cannot expect to "implement ethics into the system". Rather, we must design systems to be functional in context, including contexts of oversight and review"[133]. For this reason, constant learning[40,53,69,129,148,155,183], and constant monitoring[33,39,53,56,66,88,105,111,114,128,133,148,153,155,175,183], have been proposed. Similar proposals were also identified (*Findings 152-154*).

4.3.3.5. Governance (Findings 153 – 156)

There must be proper regulation to guarantee an ethical use of AI. Several regulatory frameworks already exist[3,33,46,50,53,56,61,80,81,104,113,128,130,166-168,182,183,185,197]. There must be cooperation and integration between agencies[148], and clinicians should incorporate guidelines of ethical use of AI into their practice[128]. For example, the GMAIH model of governance can improve fairness, transparency, accountability and trustworthiness[54,147].

4.3.3.6. Constant learning and monitoring (Findings 157 – 165)

AI should constantly receive new input data to improve its performance[40,53,69,129,148,155,183,199]. The ongoing monitoring can ensure AI outputs remain acceptable as clinical practice, disease prevalence and science evolve. Another option would be checking AI systems to make sure they maintain safety and performance standards, [3,33,39,53,56,66,88,105,111,114,128,130,133,148,153,155,175,183], this includes, stewardship programs, performance evaluations and AI audits.

4.3.3.7. Accessibility (Findings 166 – 169)

AI should be accessible to all[33,50,53,69,76,155,160,177]. Accessibility could help decrease the existing health breach, already worsen by the technological health breach.

4.3.3.8. Control (Findings 170 – 175)

There is controversy on who must be in control of AI. Some articles state AI should always be controlled by humans[80,124], others argue humans should always be able to overrun an AI decision[56,144], and others that AI should only assist clinicians[42].

4.3.3.9. Information, clinical relationships and autonomy (Findings 176 – 184)

Humans should always know when AI is been used on their health, the benefits and disadvantages of using it, and given the choice to use it or not. Patients must give prior authorization to use AI on their health. Patients must be given proper IC[38,40,43,46,49,52,56,58,68,82,89,94

,106,163,165,172) and compassion(84), ensuring patients never confuse AI with humans(151,194). For this purpose, inclusive communication between physicians and the public(63,173) and shared decision making(137) have been proposed.

4.3.3.10. Data management (Findings 185 – 209)

Several proposals exist to protect patient data and guarantee ethical data management strategies:

- Patient should be involved in data collection(35,74,114,155).
- Implement proper cybersecurity mechanisms(159).
- Patients must have the right to access their own personal data(38).

4.3.3.11. Validating AI (Findings 200 – 209)

AI itself should be validated scientifically (52,56,59,63,68,108,116,122,133,150,157,158,164,183). Data sets used to train AI must be accurate and scientifically proven (validated)(49,52-54,56,62,68,75,86,114,116,125,127,133,135,141,153,155,159,170,174,183,205,206). This way AI can avoid bias and harm. This implies, using only the best clinical knowledge available and empirical data to develop AI.

4.3.3.12. Training and education (Findings 210 – 216)

Healthcare workers must be trained to understand how AI works and be able to take proper decisions using AI (49,92,137,150,158,189,190,205,207). Training will also help clinicians adopt these technologies. Medical faculties must include AI courses into their curriculum(49). Patients must also be educated on AI(148,158). Furthermore, some articles argue for the need of accreditation to use Medical-AI(155).

4.3.3.13. Other proposals (Findings 217 – 229)

Other more general proposals exist to guarantee and ethical use and development of Medical-AI. This include:

- Defying and addressing clinicians' expectations when designing AI. AI must be adapted to fulfill performance expectations by healthcare workers (188).
- AI should foster research and innovation(56).
- Adapt Medical-AI to the social context(112).
- Follow Isaac Asimov's three laws for robots(35).
- A list of Do and Do not to use AI has been proposed(140).
- A revised Hippocratic oath has been proposed that incorporates the ethical issues posed by AI(79).
- A Rawlsian ethics model has been proposed(210).
- An ethics algorithm for the general and selective deployment of unrepresentative AI in healthcare has been proposed(211).

5. Discussion

5.1. Interpretation

We included 180 studies. We classified our findings into 3 categories: ethical concerns, ethical principles and solutions to ethical problems. We developed a framework to integrate our findings regarding ethical concerns and solutions.

We identified 12 areas that cause ethical concerns. Findings were consistent through studies, with most of them focusing on one or several specific issue and failing to address all.

We identified 73 ethical principles proposals and several other solutions. Many ethical principles proposed where synonyms or overlapped, leading to lack of clarity to establish a proper AI ethics framework. Most Principles identified can be englobed under one of four ethical principles proposed by Beauchamp and Childress (Georgetown Mantra)(18). Nevertheless, many of the new principles proposed overlap over two or more of the Georgetown Mantra. Furthermore, it is important to understand that the *new principles* debate can be framed in different bioethical paradigms. There are other ways we could classify the ethical principles identified. Following Sgreccia's personalistic approach to bioethical principles, and bioethics in general, we find an alternative classification. Our findings could be classified into personalistic principles (defense of physical life, totality or therapeutic, freedom and responsibility, and sociability and subsidiarity). It is also important to understand that, unlike Beauchamp and Childress, Sgreccia does establish a hierarchy

between principles, were the defense of physical life is the most basic principle, followed by freedom and responsibility, totality and lastly, sociability and subsidiarity. This is important as it avoids overlaps when classifying new principles into the personalistic principles, as classifying one new principle into a personalistic principle does not imply it does not involve another personalistic principle. For example, as totality is under defense of physical life, classifying a new principle into totality does not exclude it from being related to defense of physical life. We classified ethical principles proposed in **Figure 3,** comparing both principalistic and personalistic classifications. On the principalistic side, we left Explicability, and its related principles, separate, as there is disagreement on whether Explicability can be englobed inside these Principles[78,191] or it stands as a fifth principle of bioethics[64]. On the personalistic side, Explicability could fit into the principle of Responsibility. This alternative view of principles could provide further insight, expand the current debate and drive future research.

Solution proposals were consistent with our findings regarding ethical issues, and addressed ethical challenges previously identified. The number of proposals identified was greater than the number of issues they intent to solve, which could be problematic, as there is lack of consistency among findings on what specific solutions should be implemented. There is disagreement on who or what should be held accountable for errors where AI was involved, on who should receive education on AI technology and what standards and guidelines should be followed. Furthermore, it is unclear what is the best approach to guarantee privacy, safety and transparency, with some articles stating AI should be completely transparent and others stating it would be enough if doctors have evidence of AI being validated to certain standards, without total transparency[122].

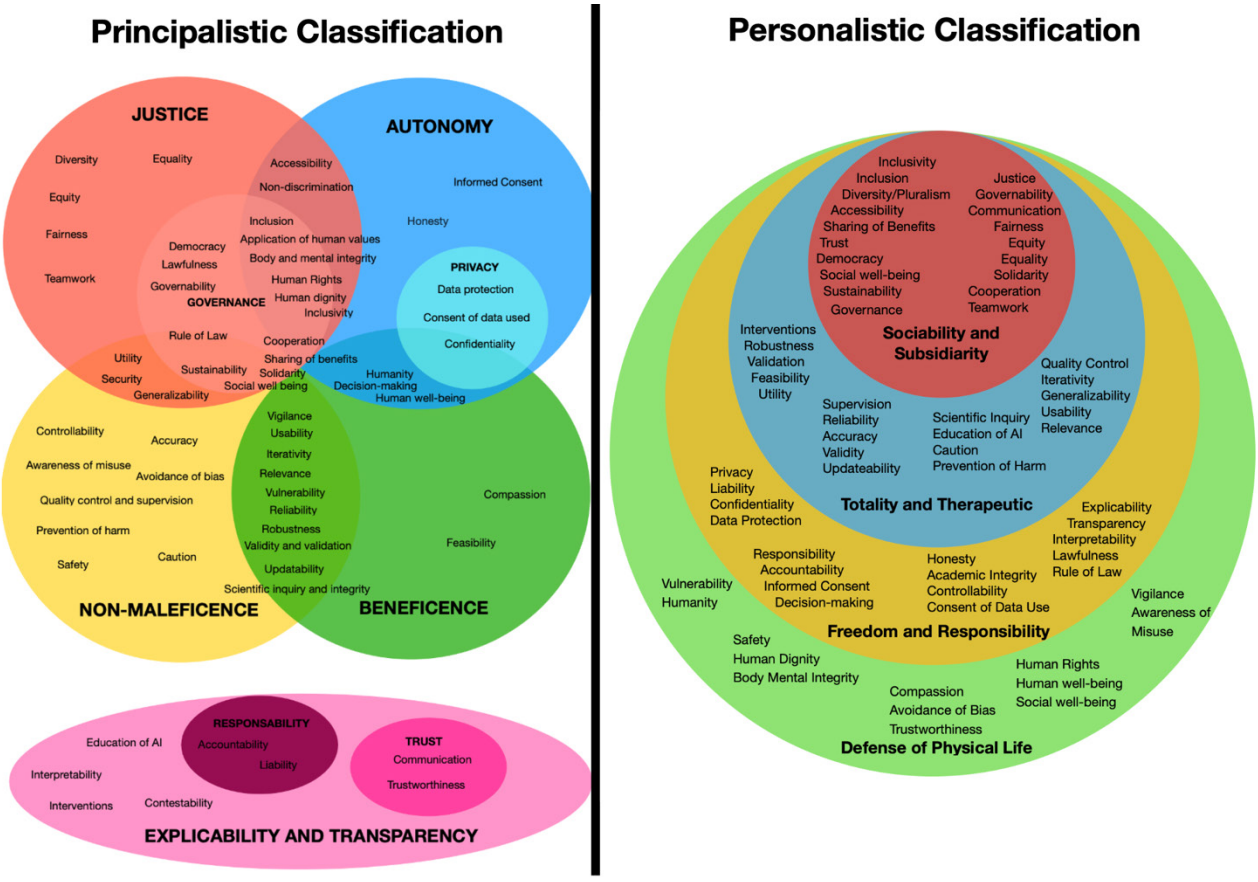

Figure 3. Classification of ethical principles proposals

5.2. Implications for practice

The new medical revolution, powered by AI, posses new important ethical challenges for the clinical practice. This includes the de-humanization of the medical practice, technological solutionism, patient safety, data management and privacy. Another big area of ethical concern are the possible biases integrated into a Medical-AI, as this could increase discrimination, the health breach and prevent underrepresented populations and social groups from a proper and ethical access to healthcare. Even though the ethical issues of the use of AI are widely accepted, the shear number of solution proposals and lack of scientific evidence to back up many of them, makes it hard to understand and address, the proper ethical path to follow.

However, Medical-AI also posses the possibility for a greater healthcare landscape, as the use of AI in the increasing prevalence of chronic disease and longer life span, in a world already affected by pandemics, can help alleviate the strain of healthcare systems and advance the ongoing healthcare transformation.

One of the greater risks of the use of AI in the medical practice is the avolition of the current medical-ethical landscape, this implies the abolition of a medicine centered around the Doctor-Patient Relationship, to stablish an AI-patient model. In this case, we could even ask ourselves the question if an AI-patient centered medicine could be considered medicine *per se*, or if this would be something else. In other words, would an AI-patient medicine maintain the essence of the medical practice? We argue, this would not be possible. In this regard, we can mention Edmund Pellegrino, who stablished a medicine based around the doctor-patient relationship. Furthermore, Laín Entralgo defined the clinical relationship as a particular and unique relationship, based on trust: trust in the technique to heal, trust in the professional's knowledge, and trust in the doctor's ethical values. As a current report from the Spanish Bioethics Committee argues: "the new healthcare revolution, that lives in the paradigm of the digital society, has the challenge to provide trust, propinquity, quality, warmth, continuity of care, time for the actors, and reduce the social breach, which is the first determinant for health" [213]. For these reasons, and in accordance with the World Medical Association, we argue the face-to-face medical (human) consult between patient and physician must be the gold standard for the medical practice.

A human gold standard thus implies AI could not fully substitute physicians, as even if both had a similar level of knowledge, the virtues and qualities of a human-to-human interaction, could not be substituted by an AI. In other words, the level of trusts and confidence between a patients and physician can not be match by an AI, as there is a human dimension to this relationship that can not be substituted by the cold, calculating mind of a Medical-AI. If the medical practice is based on a doctor-patient relationship (a human relationship), AI cannot fully take over the medical practice. AI can revolutionize the medical practice, but it does not pose a paradigm change to medicine, as the current human-to-human medical paradigm remains the gold standard for the medical act.

By this, we do not mean to say, AI will not revolutionize medicine, indeed, it is already doing so. However, the medical-AI revolution, does not (at least for now) have the potential to completely alter what we conceive as medicine. The new medical technologies, the automatization of medical processes, and diagnostic tools provided by AI, cannot compensate for a lack of human-to-human interaction to the point that medicine is completely striped of its human core. There will probably be a point were not using certain AI applications will be considered unethical, as it will be impossible to overlook the great benefits that they bring. However, even if an AI model acts autonomously or the use of AI is required by future medical guidelines, clinicians must always be guided by the basic ethical principles that stablish the core of medicine and guarantee the future of the medical practice.

The medical practice in its essence makes clinicians take ethical decisions every day. The ethical deliberation is present in every step of the clinical practice, even though it is sometimes left aside or forgotten. Even the smallest medical act, requires a practical ethical deliber-

ation to determine is it is good or bad to give someone a certain medicine, for example. Therefore, as long as AI and machines are uncapable of ethical though and action, which many argue they will never be, AI cannot fully substitute the patient-doctor relationship.

Another important point to address is the lack of scientific validation of Medical-AI and the solutions proposed to guarantee and ethical use of AI. As the Spanish Bioethics Committee says: "we lack data results comparing telemedicine to the classical in-person model" [213]. All we can be certain at this point is that Medical-AI brings the ability to revolutionize healthcare, however independent scientific research must be carried out to examine the real extend of these claims. Furthermore, different AI systems, that use different mechanisms and are designed for different purposes, will need to be individually tested. One general scientific test, to test to real practical use of Medical-AI is not appropriate. Individual scientific tests must be developed and carried out for different AI systems, with the intention to examine the real, practical validity and usability of these technologies. Until this happens, all we can talk about, is science-fiction. Once individual AI systems are developed and tested, we recommend an ethical deliberation to be carried out at all levels (development, management, clinical...), to provide individualized ethical solutions with true practical importance that go beyond scientific and philosophical-fiction and address real ethical issues that can really impact the clinical practice. As there is little certainty of what the future of Medical-AI will bring, in the next section we propose a theoretical logic model and checklist to help guide clinicians through the ethical uncertainty of future AI technologies.

5.2.1. Theoretical logic model

To test utility of findings for use in practice, we used our framework to develop a theoretical logic model (**Figure 4**). Each input box was populated by statements based directly on our findings. We show different aspects of AI technology and ethical concerns related to each aspect of AI (grey background), as well as proposals to solve these issues (white background). Arrow thick-

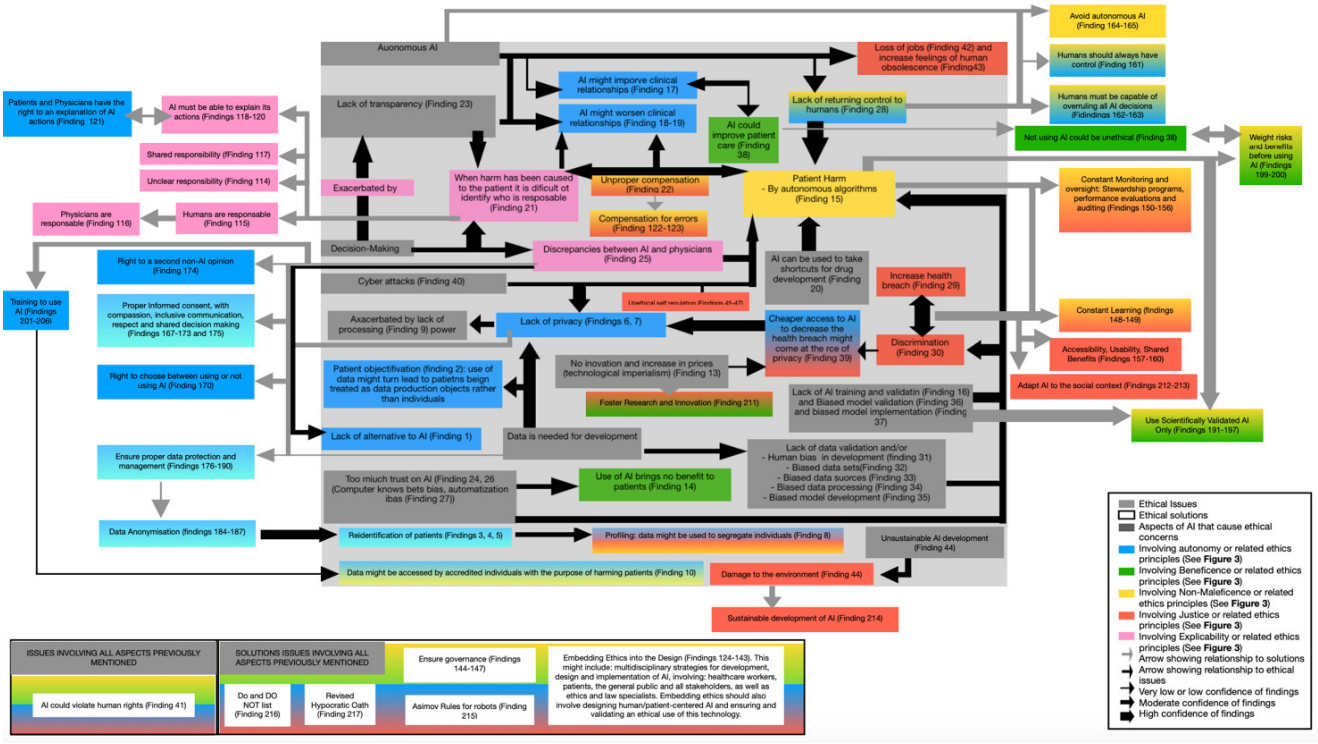

Figure 4. Theoretical Logic model of a bioethical medical-AI

ness is proportional to degree of confidence of findings in the box to which they point. To integrate our ethical principles findings, we used our ethical principles classification and painted each textbox with a color representing each principle: blue for autonomy, red for justice, green for beneficence, yellow for non-maleficence and pink for explicability. Several boxes have more than one color, as there is more than one ethical principle involved. Boxes regarding characteristics of AI only, and not involving ethical principles, where painted grey. Concerns regarding Human Rights (*Finding 41*) and several solution proposals, including embedded ethics (*Findings 124-143*) are represented separately, as they involve most of AI aspects and address most ethical concerns. These boxes were painted with several colors as they involve all ethical principles.

This logic model could help synthesize our findings and allow identification of different factors that lead to ethical issues and how these relate to solutions described in this review.

5.2.2. Checklist for clinicians

We developed a checklist for an ethical use of Medical-AI (**Figure 5**), which could help clinicians undertake a critical analysis of the ethical implications of AI and implement an ethics-based medicine. We sent this to medical colleges and patient associations, for feedback. We got no response.

5.2.3. Explicability as an ethical principle

We used our ethical principles classification and logic model to provide a definition of Explicability as an ethical principle, identity how Explicability relates to the four classic ethical principles and if should be considered as a fifth ethical principle or it fits under the four existing principles[18].

Explicability (*Finding 104*)/Transparency (*Finding 105*) refers to: how AI communicates (*Contestability, Finding 106*), how humans are able to understand reasoning behind AI outputs (*Interpretability, Finding 87*) and provide medical guidance based on them (*Interventions, Finding 110*), as well as use this information to establish responsibility (*Finding 100*), accountability (*Finding 99*) and liability (*Finding 101*) for actions where AI is involved; to establish trust on AI (*Findings 90-91*) and use it to improve patient health and healthcare.

Using this definition and our theoretical logic model, we deduce Explicability relates mainly to Autonomy,

Ethical considerations:
1. Does the decision taken with AI impact Human Right?
2. Am I doing an appropriate use of resources by using AI?
3. What benefit do I expect from using AI?
4. Has the patient been informed and consented the use of AI on his/her health?
5. Do I have sufficient control over this technology?
6. Where ethical issues addressed during the development process?
7. Has the technology being sufficiently validated to be used safely?
8. Have you been informed of the possible biases this AI has?
9. Is there a way to know when the AI is behaving in an unexpected or dangerous way?
10. Do I know the risks of using this AI?
11. Is the AI maintained, supervised and updated regularly?
12. Can I properly interpret AI results?
13. Can I trust this AI?
14. Do I agree with the AI decision?
15. Do I know and understand what am I accountable for, when using AI?

Figure 5. Checklist for a bioethical use of Medical-AI

as explaining why and establishing responsibility for medical actions is part of a proper IC, which is mainly based on the principle of autonomy[18]. In this sense, explicability is not a new principle, as establishing responsibility and explaining health-related actions to patients was already established. Traditionally, it was physicians who had to explain its actions to patients. Now it is AI who has to explain its actions. This will encompass AI as part of an expanded patient-physician-AI clinical relationship. Explicability relates to Justice and non-maleficence, as understanding AI outputs can reduce patient harm, discrimination and bias[33]. IC and traditional explicability (between patient and physician) shared links to justice and non-maleficence too. Nevertheless, even though explicability already existed, it is because of how AI can exacerbate lack of explicability, why Explicability is needed as a new principle[64,138]. We can conclude that, Explicability does not cover an area of bioethics that was previously unaccounted. AI increases bioethical concerns regarding lack of understanding and responsibility, which might explain why Explicability has been proposed as a new bioethical principle. Further research must be done to establish if AI possess a significant lack of responsibility and understanding to definitely determine whether Explicability should be included as the fifth bioethical principle.

For the moment, not enough evidence exists to give Explicability the same ontological level as the four bioethical principles have, as its use can be understood mostly as an extension of the principle of autonomy, with no generalizability beyond AI, not being able to be used in the rest of the clinical practice were AI is not used, therefore lacking the necessary generalizability, proper of bioethical principles. However, the claim for explicability can be understood in a context greater than ethical principles, it can be used to improve patient-doctor relationships. This hints at the problem we previously faced, the substitution of the patient-doctor relationship by an AI-patient relationship. In this sense, explicability is an appropriate value to incorporate to the clinical practice, not a principle. In fact, in this sense, explicability is already being used today (and has been used for a long time). This way, we can foster and improve trust and the human aspects necessary for proper healing and proper medicine. Nevertheless, we must not understand or use explicability to abolish the human aspect of the patient-doctor relationship, as fostering explicability to increase trust between AI and patients but at the same time diminishing the human aspects of care, is a contradiction. Finally, special care and necessary steps must be taken, beyond the addition, or not, of an ethical principle to guarantee an ethical use of Medical-AI.

5.3. Completeness and applicability

Our study could help synthesize existing work on bioethical medical-AI, guide future research and guidelines to ensure an ethical-AI. Further research on identifying best proposals that guarantee an ethical use of AI needs to be done. As well as reaching consensus on the need to include explicability as a fifth bioethical principles and establishing a common framework for medical-AI bioethics. Further research could be done on the practical-clinical aspects of AI-healthcare, as data is scarce.

5.4. Limitations

The same authors that reviewed studies for inclusion carried out the methodological and confidence assessments. Dissemination bias, potential weaknesses with search strategy could have ruled out the most relevant studies. There is a limitation on data bases accessed to retrieve data as well as the language of articles included. While logic models provide a guiding theory of how interventions are intended to work, critiques exist around their use, including their potential to oversimplify complex interventions[212].

Given this is a qualitative review, no bias assessment was conducted, we assessed confidence. In our confidence analysis, many articles were downgraded due to methodological limitations, as they were critical analysis, with little to no assessment of methods used for data collection and analysis.

6. Conclusion

Our overall aim was to carry out a qualitative review of the bioethical landscape regarding AI use in healthcare. To do this, we identified, appraised and synthesized qualitative research evidence on ethical problems for implementation of AI in healthcare and developed a framework for bioethical issues of medical-AI to classify our findings. We identified, appraised and synthesized qualitative research evidence on bioethical principles proposals, and classified them according to ethical principles of Autonomy, Justice, Beneficence, Non-Maleficence and Explicability. We identified and synthesized qualitative research evidence on solutions for AI to meet bioethical standards, and developed a framework for solutions for an ethical medical-AI to classify our findings. We integrate our findings into a theoretical logic model for AI-bioethics. We used our ethical principles classification and logic model to define and establish if Explicability should be included as a fifth bioethical principle, concluding it depends on the extent to which AI can compromise responsibility and transparency. Finally, we proposed practical applications for our findings developing a checklist to guarantee an ethical use of AI in clinical practice.

Acknowledgements

Consuelo M. García

Author Contributions

Ramón J. Palmer: conceptualization, design, literature search and analysis, manuscript preparation and review.

Verónica A. Brokke: literature search and analysis, manuscript review.

Patrici Calvo: conceptualization, design, manuscript review.

Disclosure of Competing Interests and Funding

All authors state they have no financial or professional relationships that may pose competing interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

All authors have read and approved the final version send for publication.

Funding

This publication is part of the project PID2022-139000OB-C22, funded by MCIU/AEI/10.13039/501100011033/FEDER, EU, the project PID2019-106420GA-I00, funded by MCIN/AEI/10.13039/501100011033.

References

1. McCarthy, J., Minsky, ML., Rochester, N., Shannon, C. «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. Agust 31, 1955». *AI Magazine*. 1955; 27(4): 12-14.
2. U.S Food and Drug Administration. [Publicación en línea] «Artificial Intelligence and machine learning (AI/ml)-enabled medical Devices». https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices> [Consulta: 18/11/2024]
3. European Union. «Artificial Intelligence Act (Regulation (EU) 2024/1689)». *Official Journal version of 13 June 2024*. 2024.
4. Langlotz, CP., Allen, B., Erickson, BJ., et al. «A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop». *Radiology*. 2019; 291(3): 781–791.
5. Jones, OT., Matin, RN., van der Schaar, M., et al. «Artificial intelligence and machine learning algorithms for early detection of skin cancer in community and primary care settings: a systematic review». *Lancet Digital Health*. 2022; 4(6): e466–e476.
6. Johnson, KW., Torres Soto, J., Glicksberg, BS., et al. «Artificial Intelligence in Cardiology». *Journal of the American College of Cardiology*. 2018; 71(23): 2668.

7. Dilsizian, S., Siegel, E. «Artificial Intelligence in Medicine and Cardiac Imaging: Harnessing Big Data and Advanced Computing to Provide Personalized Medical Diagnosis and Treatment». *Current Cardiology Reports*. 2013; 16(1): 12-13.
8. Monteith, S., Glenn, T., Geddes, J., et al. «Expectations for Artificial Intelligence (AI) in Psychiatry». *Current Psychiatry Reports*. 2022; 24(11): 709.
9. Bouton, C., Shaikhouni, A., Annetta, N., et al. «Restoring cortical control of functional movement in a human with quadriplegia». *Nature*. 2016; 533(7602): 247–250.
10. Xie, J., Luo, X., Deng, X., et al. «Advances in artificial intelligence to predict cancer immunotherapy efficacy». *Frontiers in Immunology*. 2023; 4(13): 1076883.
11. Phakhounthong, K., Chaovalit, P., Jittamala, P., et al. «Predicting the severity of dengue fever in children on admission based on clinical features and laboratory indicators: application of classification tree analysis». *BMC Pediatrics*. 2018; 18(1): 109.
12. Lazer, D., Kennedy, R., King, G., Vespignani, A. «The Parable of Google Flu: Traps in Big Data Analysis». *Science*. 2014; 343(6176): 1203–1205.
13. Jiang, D., Hao, M., Ding, F., et al. «Mapping the transmission risk of Zika virus using machine learning models». *Acta Tropica*. 2018; 185: 391–399.
14. Garcia-Vidal, C., Sanjuan, G., Puerta-Alcalde, P., et al. «Artificial intelligence to support clinical decision-making processes». *EBioMedicine*. 2019; 46: 27–29.
15. Ameneiros-Lago, E., Carballada-Rico, C., Garrido-Sanjuán, J., García Martínez, A. «Adecuación de las intervenciones clínicas en pacientes con enfermedad avanzada y compleja. Propuesta de algoritmo para la toma de decisiones». *Revista de calidad asistencial*. 2015; 30: 38–43.
16. Fleming, N. «How artificial intelligence is changing drug discovery». *Nature*. 2018; 557(7707): 55.
17. Gilbert, S. «The EU passes the AI Act and its implications for digital medicine are unclear». *npj Digital Medicine*. 2024; 7: 135.
18. Beauchamp, T., Childress, J. *Principles of biomedical ethics*, Oxford University Press, Oxford, 2001.
19. Booth, A., Noyes, J., Flemming, K., et al. «Formulating questions to explore complex interventions within qualitative evidence synthesis». *BMJ Global Health*. 2019; 4(Suppl 1): e001107.
20. Page, M., McKenzie, J., Bossuyt, P., et al. «The PRISMA 2020 statement: an updated guideline for reporting systematic reviews». *BMJ*. 2021; 372:n71.
21. Flemming, K., Booth, A., Hannes, K., et al. «Cochrane Qualitative and Implementation Methods Group guidance series—paper 6: reporting guidelines for qualitative, implementation, and process evaluation evidence syntheses». *Journal of Clinical Epidemiology*. 2018; 97: 79–85.
22. Glenton, C., Lewin, S., Downe, S., et al. «Qualitative Evidence Syntheses Within Cochrane Effective Practice and Organisation of Care: Developing a Template and Guidance». 2021.
23. Noyes, J., Booth, A., Moore, G., et al. «Synthesizing quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods». *BMJ Global Health*. 2019; 4(Suppl 1): e000893.
24. Glenton, C., Colvin, C., Carlsen, B., et al. «Barriers and facilitators to the implementation of lay health worker programmes to improve access to maternal and child health: a qualitative evidence synthesis». *Cochrane Database of Systematic Reviews*. 2013; 2019(3): CD010414.
25. Poletti, P., Platon, A., Rutschmann, O., et al. «Low-dose versus standard-dose CT protocol in patients with clinically suspected renal colic». *American Journal of Roentgenology*. 2007; 188(4): 927–933.
26. Critical Appraisal Skills Programme. «CASP Qualitative Checklist». 2018.
27. Atkins, S., Lewin, S., Smith, H., et al. «Conducting a meta-ethnography of qualitative literature: Lessons learnt». *BMC Medical Research Methodology*. 2008; 8(21).
28. Kneale, D., Thomas, J., Harris, K. «Developing and Optimising the Use of Logic Models in Systematic Reviews: Exploring Practice and Good Practice in the Use of Programme Theory in Reviews». *PloS One*. 2015; 10(11): e0142187.

29. Anderson, LM., Petticrew, M., Rehfuess, E., et al. «Using logic models to capture complexity in systematic reviews». *Research Synthesis Methods*. 2011; 2(1): 33–42.
30. Lewin, S., Booth, A., Glenton, C., et al. «Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series». *Implementation Science*. 2018; 13(S1): 2.
31. Haddaway, NR., Page, MJ., Pritchard, CC., McGuinness, LA. «PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis». *Campbell Systematic Review*. 2022; 18(2): e1230.
32. Price, NW., Cohen, GI. «Privacy in the age of medical big data». *Nature Medicine*. 2019; 25: 37–43.
33. World Health Organization. «Ethics and governance of artificial intelligence for health: guidance on large multi-modal models». 2024.
34. Ploug, T., Holm, S. «The right to a second opinion on Artificial Intelligence diagnosis—Remedying the inadequacy of a risk-based regulation». *Bioethics*. 2023; 37(3): 303–311.
35. Evans, BJ. «Rules for robots, and why medical AI breaks them». *Journal of Law and the Biosciences*. 2023; 10: lsad001.
36. Mirbabaie, M., Hofeditz, L., Frick, NRJ., Stieglitz, S. «Artificial intelligence in hospitals: providing a status quo of ethical considerations in academia to guide future research». *AI & Society*. 2022; 37(4): 1361–1382.
37. Luxton, DD. «Recommendations for the ethical use and design of artificial intelligent care providers». *Artificial Intelligence in Medicine*. 2014; 62(1): 1–10.
38. Vayena, E., Blasimme, A. «Biomedical Big Data: New Models of Control Over Access, Use and Governance». *Bioethical Inquiry*. 2017; 14(4): 501–513.
39. Redrup Hill, E., Mitchell, C., Brigden, T., Hall, A. «Ethical and legal considerations influencing human involvement in the implementation of artificial intelligence in a clinical pathway: a multi-stakeholder perspective». *Frontiers in Digital Health*. 2023; 5: 1139210.
40. Okun, S., Hanger, M., Browne-James, L., et al. «Commitments for Ethically Responsible Sourcing, Use, and Reuse of Patient Data in the Digital Age: Cocreation Process». *Journal of Medical Internet Research*. 2023; 25(6): e41095.
41. Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S. «Dissecting racial bias in an algorithm used to manage the health of populations». *Science*. 2019; 366(6464): 447–453.
42. Mouchabac, S., Adrien, V., Falala-Séchet, C., et al. «Psychiatric Advance Directives and Artificial Intelligence: A Conceptual Framework for Theoretical and Ethical Principles». *Frontiers in Psychiatry*. 2021; 11: 622506.
43. Suominen, H., Lehtikunnas, T., Back, B., et al. «Applying language technology to nursing documents: pros and cons with a focus on ethics». *International Journal of Medical Informatics*. 2007; 76: S293–S301.
44. Tiribelli, S., Calvaresi, D. «Rethinking Health Recommender Systems for Active Aging: An Autonomy-Based Ethical Analysis». *Science and Engineering Ethics*. 2024; 30(3): 22.
45. Khan, M., Ewuoso, C. «Epistemic (in)justice, social identity and the Black Box problem in patient care». *Medicine, Health Care and Philosophy*. 2024; 27(2): 227–240.
46. International Bioethics Committee, UNESCO. [Publicación en línea] «Preliminary Draft Report of the IBC on Big Data and Health» 2016. <https://unesdoc.unesco.org/ark:/48223/pf0000245256> [Consulta: 18/11/2024]
47. Sharkey, A., Sharkey, N. «Granny and the robots: ethical issues in robot care for the elderly». *Ethics and Information Technology*. 2010; 14: 27-40.
48. O'Neil, C. «Life in the age of the algorithm». *Science*. 2017; 355(6321): 137.
49. Quinn, T., Senadeera, M., Jacobs, S., et al. «Trust and medical AI: the challenges we face and the expertise needed to overcome them». *Journal of the American Medical Informatics Association*. 2021; 28(4): 890–894.
50. Montréal Declaration for a Responsible Development of Artificial Intelligence. [Publicación en línea] «Toolkit». 2024. <https://montrealdeclaration-responsibleai.com/toolkit/> [Consulta: 18/11/2024]

51. Jackson, BR., Ye, Y., Crawford, JM., et al. «The Ethics of Artificial Intelligence in Pathology and Laboratory Medicine: Principles and Practice». *Academic pathology*. 2021; 8: 2374289521990784.
52. Khalid, N., Qayyum, A., Bilal, M., et al. «Privacy-preserving artificial intelligence in healthcare: Techniques and applications». *Computers in biology and medicine*. 2023; 158: 106848.
53. Asamblea General de las Naciones Unidas. «A/HRC/47/54». 2021.
54. Reddy, S., Allan, S., Coghlan, S., Cooper, P. «A governance model for the application of AI in health care». *Journal of the American Medical Informatics Association*. 2020; 27(3): 491–497.
55. Wangmo, T., Lipps, M., Kressig, R., Ienca, M. «Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders». *BMC medical ethics*. 2019; 20(98).
56. High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI*, European Commission, 2019.
57. de Montjoye, Y., Hidalgo, CA., Verleysen, M., Blondel, VD. «Unique in the Crowd: The privacy bounds of human mobility». *Scientific Reports*. 2013; 3(1376).
58. Martinez-Martin, N., Kreitmair, K. «Ethical Issues for Direct-to-Consumer Digital Psychotherapy Apps: Addressing Accountability, Data Protection, and Consent». *JMIR mental health*. 2018; 5(2): e32.
59. Hosny, A., Aerts, HJWL. «Artificial intelligence for global health». *Science*. 2019; 366(6468): 955–956.
60. Ienca, M., Ignatiadis, K. «Artificial Intelligence in Clinical Neuroscience: Methodological and Ethical Challenges». *AJOB neuroscience*. 2020; 11(2): 77–87.
61. Organización Médica Colegial de España, *Código de Deontología Médica*, 2022.
62. Chiang, S., Picard, R., Chiong, W., et al. «Guidelines for Conducting Ethical Artificial Intelligence Research in Neurology: A Systematic Approach for Clinicians and Researchers». *Neurology*. 2021; 97(13): 632–640.
63. Parums, DV. «Editorial: Artificial Intelligence (AI) in Clinical Medicine and the 2020 CONSORT-AI Study Guidelines». *Medical science monitor*. 2021; 27: e933675.
64. Floridi, L., Cowls, J. «A Unified Framework of Five Principles for AI in Society». *Harvard Data Science Review*. 2019; 1.1.
65. Abràmoff, MD., Roehrenbeck, C., Trujillo, S., et al. «A reimbursement framework for artificial intelligence in healthcare». *npj Digital Medicine*. 2022; 5(72).
66. Leo, CG., Tumolo, MR., Sabina, S., et al. «Health Technology Assessment for In Silico Medicine: Social, Ethical and Legal Aspects». *International journal of environmental research and public health*. 2022; 19(3): 1510.
67. Garani-Papadatos, T., Natsiavas, P., Meyerheim, M., et al. «Ethical Principles in Digital Palliative Care for Children: The MyPal Project and Experiences Made in Designing a Trustworthy Approach». *Frontiers in digital health*. 2022; 4: 730430.
68. Coghlan, S., Leins, K., Sheldrick, S., et al. «To chat or bot to chat: Ethical issues with using chatbots in mental health». *Digital health*. 2023; 9: 20552076231183542.
69. Nazer, LH., Zatarah, R., Waldrip, S., et al. «Bias in artificial intelligence algorithms and recommendations for mitigation». *PLOS digital health*. 2023; 2(6): e0000278.
70. Ahmed, Z., Mohamed, K., Zeeshan, S., Dong, XQ. «Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine». *The Journal of Biological Databases and Curation*. 2020; 2020: baaa010.
71. Markowetz, A., Błaszkiewicz, K., Montag, C., et al. «Psycho-Informatics: Big Data shaping modern psychometrics». *Medical hypotheses.* 2014; 82(4): 405–411.
72. Kummer, M., Schulte, P. «When Private Information Settles the Bill: Money and Privacy in Google's Market for Smartphone Applications». *Management science*. 2019; 65(8): 3470–3494.
73. Goodman, K. «Ethics in Health Informatics». *Yearbook of medical informatics*. 2020; 29(1): 26-31.
74. Monteith, S., Glenn, T. «Automated Decision-Making and Big Data: Concerns for People with Mental Illness». *Current Psychiatry Reports*. 2016; 18(12): 112.
75. Thompson, CL., Morgan, HM. «Ethical barriers to artificial intelligence in the national health service, Unit-

ed Kingdom of Great Britain and Northern Ireland». *Bulletin of the World Health Organization*. 2020; 98(4): 293–295.

76. Couture, V., Roy, M., Dez, E., et al. «Ethical Implications of Artificial Intelligence in Population Health and the Public's Role in Its Governance: Perspectives from a Citizen and Expert Panel». *Journal of medical Internet research*. 2023; 25: e44357.
77. Agarwal, R., Dugas, M., Gao, G., et al. «Emerging technologies and analytics for a new era of value-centered marketing in healthcare». *Journal of the Academy of Marketing Science*. 2020; 48(1): 9–23.
78. Fournier-Tombs, E., McHardy, J. «A Medical Ethics Framework for Conversational Artificial Intelligence». *Journal of medical Internet research.* 2023; 25: e43068.
79. Meskó, B., Spiegel, B. «A Revised Hippocratic Oath for the Era of Digital Health». *Journal of medical Internet research*. 2022; 24(9): e39177.
80. Future of Life Institute. [Publicación en línea] «Asilomar AI principles». <https://futureoflife.org/open-letter/ai-principles/> [Consulta: 18/11/2024].
81. House of Lords. [Publicación en línea] «AI in the UK: ready, willing, able?». 2018. <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf.> [Consulta: 18/11/2024].
82. Yuste, R., Goering, S., Arcas, BAY., et al. «Four ethical priorities for neurotechnologies and AI». *Nature*. 2017; 551(7679): 159–163.
83. Bak, MAR., Ploem, MC., Tan, HL., et al. «Towards trust-based governance of health data research». *Medicine, Health Care and Philosophy*. 2023; 26(2): 185–200.
84. Macri, R., Roberts, SL. «The Use of Artificial Intelligence in Clinical Care: A Values-Based Guide for Shared Decision Making». *Current oncology*. 2023; 30(2): 2178–2186.
85. Car, J., Sheikh, A., Wicks, P., Williams, MS. «Beyond the hype of big data and artificial intelligence: building foundations for knowledge and wisdom». *BMC Medicine*. 2019; 17(143).
86. Vayena, E., Blasimme, A., Cohen, I. «Machine learning in medicine: Addressing ethical challenges». *PLOS Medicine*. 2018; 15(11): e1002689.
87. Pedersen, M., Verspoor, K., Jenkinson, M., et al. «Artificial intelligence for clinical decision support in neurology». *Brain communications*. 2020; 2(2): fcaa096.
88. Hasani, N., Morris, MA., Rahmim, A., et al. «Trustworthy Artificial Intelligence in Medical Imaging». *PET clinics*. 2022; 17(1): 1–12.
89. Kenny, LM., Nevin, M., Fitzpatrick, K. «Ethics and standards in the use of artificial intelligence in medicine on behalf of the Royal Australian and New Zealand College of Radiologists». *Journal of medical imaging and radiation oncology.* 2021; 65(5): 486–494.
90. Baric-Parker, J., Anderson, EE. «Patient Data-Sharing for AI: Ethical Challenges, Catholic Solutions». *The Linacre quarterly*. 2020; 87(4): 471–481.
91. Vandemeulebroucke, T. «The ethics of artificial intelligence systems in healthcare and medicine: from a local to a global perspective, and back». *Pflügers Archiv – European Journal of Physiology*. 2024; 477(4):591-601.
92. Lu, H., Alhaskawi, A., Dong, Y., et al. «Patient Autonomy in Medical Education: Navigating Ethical Challenges in the Age of Artificial Intelligence». *Inquiry; The journal of health care organization, provision, and financing*. 2024; 61: 469580241266364.
93. Ong, JCL., Chang, SY., William, W., et al. «Ethical and regulatory challenges of large language models in medicine». *The Lancet Digital Health*. 2024; 6(6): e428–e432.
94. Chen, NL., Sng, BL. «Ethical Principles of Clinical and Translational Research». *Anesthesiology clinics*. 2024; 42(4): 581–592.
95. Wang, Y., Ma, Z. «Ethical and legal challenges of medical AI on informed consent: China as an example». *Developing world bioethics*. 2024; 25: 46-54.
96. Couldry, N., Mejias, UA. «Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject». *Television & new media*. 2019; 20(4): 336–349.
97. Ferryman, K. «The Dangers of Data Colonialism in Precision Public Health». *Global policy*. 2021; 12(S6): 90–92.
98. Bærøe, K., Miyata-Sturm, A., Henden, E. «How to achieve trustworthy artificial intelligence for health». *Bulletin of the World Health Organization*. 2020; 98(4): 257–262.

99. Hoorn, JF., Winter, SD. «Here Comes the Bad News: Doctor Robot Taking Over». *International Journal of Social Robotics*. 2018; 10(4): 519–535.
100. Pirni, A., Balistreri, M., Capasso, M., et al. «Robot Care Ethics Between Autonomy and Vulnerability: Coupling Principles and Practices in Autonomous Systems for Care». *Frontiers in robotics and AI*. 2021; 8: 654298.
101. Yew, GCK. «Trust in and Ethical Design of Carebots: The Case for Ethics of Care». *International Journal of Social Robotics*. 2021; 13(4): 629–645.
102. Bhattacharya, S., Hossain, M., Juyal, R., et al. «Role of public health ethics for responsible use of artificial intelligence technologies». *Indian journal of community medicine*. 2021; 46(2): 178–181.
103. Lee, WT. «Artificial Intelligence in Medicine: A Caution About Good Intentions and Where It May Lead». *Otolaryngology-head and neck surgery*. 2024; 170(6): 1605–1606.
104. European Group on Ethics in Ethics and New Technologies. [Publicación en línea] «Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems». 2018. <https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1> [Consulta: 18/11/2024]
105. Hafner, AW., Filipowicz, AB., Whitely, WP. «Computers in medicine: Liability issues for physicians». *International journal of clinical monitoring and computing*. 1989; 6(3): 185–194.
106. Balthazar, P., Harri, P., Prater, A., Safdar, NM. «Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics». *Journal of the American College of Radiology*. 2018; 15(3): 580–586.
107. Benjamin, R. «Assessing risk, automating racism». *Science*. 2019; 366(6464): 421–422.
108. Eaneff, S., Obermeyer, Z., Butte, AJ. «The Case for Algorithmic Stewardship for Artificial Intelligence and Machine Learning Technologies». *The Journal of the American Medical Association*. 2020; 324(14): 1397–1398.
109. McCradden, MD., Joshi, S., Mazwi, M., Anderson, JA. «Ethical limitations of algorithmic fairness solutions in health care machine learning». *Lancet Digital Health*. 2020; 2(5): e221–e223.
110. O'Brolcháin, F. «Robots and people with dementia: Unintended consequences and moral hazard». *Nursing ethics*. 2019; 26(4): 962–972.
111. Chen, J. «Potential Trade-Offs and Unintended Consequences of Artificial Intelligence». En: *Artificial Intelligence in Health Care*. National Academies Press, United States, 2023, 99–130.
112. Lysaght, T., Lim, HY., Xafis, V., Ngiam, KY. «AI-Assisted Decision-making in Healthcare». *Asian Bioethics Review*. 2019; 11(3): 299–314.
113. House of Commons Science and Technology Committee. [Publicación en línea] « Algorithms in decision making: Government Response to the Committee's Fourth Report». 2018. <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/1544/1544.pdf> [Consulta: 18/11/2024]
114. Rajkomar, A., Hardt, M., Howell, MD., et al. «Ensuring Fairness in Machine Learning to Advance Health Equity». *Annals of Internal Medicine*. 2018; 169(12): 866–872.
115. Richie, C. «Environmentally sustainable development and use of artificial intelligence in health care». *Bioethics*. 2022; 36(5): 547–555.
116. Cresswell, K., Rigby, M., Magrabi, F., et al. «The need to strengthen the evaluation of the impact of Artificial Intelligence-based decision support systems on healthcare provision». *Health policy*. 2023; 136: 104889.
117. Petersson, L., Vincent, K., Svedberg, P., et al. «Ethical considerations in implementing AI for mortality prediction in the emergency department: Linking theory and practice». *Digital health*. 2023; 9: 20552076231206588.
118. Parikh, RB., Obermeyer, Z., Navathe, AS. «Regulation of predictive analytics in medicine». *Science*. 2019; 363(6429): 810–812.
119. Malinverno, L., Barros, V., Ghisoni, F., et al. «A historical perspective of biomedical explainable AI research». *Patterns*. 2023; 4(9): 100830.
120. Alvarez-Melis, D., Jaakkola, T. «Towards Robust Interpretability with Self-Explaining Neural Net-

works». 32nd *Conference on Neural Information Processing Systems (NeurIPS 2018)*, Montréal, Canada. 2018.

121. London, AJ. «Groundhog Day for Medical Artificial Intelligence». *The Hastings Center report*. 2018; 48(3): 2-4.
122. Paladino, MS. «Artificial Intelligence in Medicine. Ethical reflections from the thought of Edmund Pellegrino». *Cuadernos de Bioética*. 2023; 34(110): 25–35.
123. Pasquale, F. *The black box society: the secret algorithms that control money and information*, Harvard University Press, Cambridge, 2015.
124. Braun, M., Hummel, P., Beck, S., Dabrock, P. «Primer on an ethics of AI-based decision support systems in the clinic». *Journal of Medical Ethics*. 2020; 47(12): e3.
125. Cheng, M., Li, X., Xu, J. «Promoting Healthcare Workers' Adoption Intention of Artificial-Intelligence-Assisted Diagnosis and Treatment: The Chain Mediation of Social Influence and Human–Computer Trust». *International journal of environmental research and public health*. 2022; 19(20): 13311.
126. McLennan, S., Fiske, A., Tigard, D., et al. «Embedded ethics: a proposal for integrating ethics into the development of medical AI». *BMC medical ethics*. 2022; 23(1): 6.
127. Holm, S. «On the Justified Use of AI Decision Support in Evidence-Based Medicine: Validity, Explainability, and Responsibility». *Cambridge quarterly of healthcare ethics*. 2023: 1-7.
128. Liao, F., Adelaine, S., Afshar, M., Patterson, BW. «Governance of Clinical AI applications to facilitate safe and equitable deployment in a large health system: Key elements and early successes». *Frontiers in digital health*. 2022; 4: 931439.
129. Yu, K., Kohane, IS. «Framing the challenges of artificial intelligence in medicine». *BMJ Quality & Safety*. 2019; 28(3): 238–241.
130. Ho, CW., Caals, K. «How the EU AI Act Seeks to Establish an Epistemic Environment of Trust». *Asian Bioethics Review*. 2024; 16(3): 345–372.
131. Szalai, J. «The potential use of artificial intelligence in the therapy of borderline personality disorder». *Journal of evaluation in clinical practice*. 2021; 27(3): 491–496.
132. Cuzzolin, F., Morelli, A., Cîrstea, B., Sahakian, BJ. «Knowing me, knowing you: theory of mind in AI». *Psychological Medicine*. 2020; 50(7): 1057–1061.
133. Elish, MC. «Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction». *Engaging science, technology, and society*. 2019; 5: 40–60.
134. Siqueira-Batista, R., Souza, CR., Maia, PM., Siquiera, SL. «Robotic Surgery: Bioethical Aspects». *Arquivos Brasileiros de Cirurgia Digestiva*. 2016; 29(4): 287–290.
135. Verghese, A., Shah, NH., Harrington, RA. «What This Computer Needs Is a Physician». *The Journal of the Americal Medical Association*. 2018; 319(1): 19-20.
136. Char, DS., Shah, NH., Magnus, D. «Implementing Machine Learning in Health Care - Addressing Ethical Challenges». *New England Journal of Medicine*. 2018; 378(11): 981–983.
137. Sauerbrei, A., Kerasidou, A., Lucivero, F., Hallowell, N. «The impact of artificial intelligence on the person-centred, doctor-patient relationship: some problems and solutions». *BMC medical informatics and decision making*. 2023; 23(73).
138. Adams, J. «Defending explicability as a principle for the ethics of artificial intelligence in medicine». *Medicine, Health Care and Philosophy*. 2023; 26(4): 615–623.
139. Felzmann, H., Fosch-Villaronga, E., Lutz, C., Tamo-Larrieux, A. «Towards Transparency by Design for Artificial Intelligence». *Science and Engineering Ethics*. 2020; 26(6): 3333–3361.
140. Beil, M., Proft, I., van Heerden, D., et al. «Ethical considerations about artificial intelligence for prognostication in intensive care». *Intensive Care Medicine Experimental*. 2019; 7(70).
141. Blanco-González, A., Cabezón, A., Seco-González, A., et al. «The Role of AI in Drug Discovery: Challenges, Opportunities, and Strategies». *Pharmaceuticals*. 2023; 16(6): 891.
142. Expert Committee on human rights dimensions of automanted data processing and different forms of artificial intelligence (MSI-AUT), Rapporteur: Yeung,

K. [Publicación en línea] «A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework». 2018. <https://edoc.coe.int/en/artificial-intelligence/8026-responsibility-and-ai.html> [Consulta: 18/11/2024]

143. Hao, K. [Publicación en línea] «When algorithms mess up, the nearest human gets the blame». MIT Technology Review. 2019. <https://www.technologyreview.com> [Consulta: 18/11/2024]
144. Fanni, R., Steinkogler, VE., Zampedri, G., Pierson, J. «Enhancing human agency through redress in Artificial Intelligence Systems». *AI & Society*. 2023; 38(2): 537–547.
145. Fulmer, R., Joerin, A., Gentile, B., et al. «Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial». *JMIR Mental Health*. 2018; 5(4): e64.
146. Amann, J., Blasimme, A., Vayena, E., et al. «Explainability for artificial intelligence in healthcare: a multidisciplinary perspective». *BMC Medical Informatics and Decision Making*. 2020; 20(310).
147. Halligan, A. «Implementing clinical governance: turning vision into reality». *British Medical Journal*. 2001; 322(7299): 1413–1417.
148. Zhang, J., Zhang, Z. «Ethics and governance of trustworthy medical artificial intelligence». *BMC medical informatics and decision making*. 2023; 23(7).
149. Grote, T., Berens, P. «On the ethics of algorithmic decision-making in healthcare». *Journal of Medical Ethics*. 2020; 46(3): 205–211.
150. Briganti, G., Le Moine, O. «Artificial Intelligence in Medicine: Today and Tomorrow». *Frontiers in Medicine*. 2020; 7: 27.
151. Di Nucci, E. «Should we be afraid of medical AI?». *Journal of Medical Ethics*. 2019; 45(8): 556–558.
152. Ching, T., Himmelstein, D., Beaulieu-Jones, B., et al. «Opportunities and obstacles for deep learning in biology and medicine». *Journal of The Royal Society Interface*. 2018; 15(141): 20170387.
153. Lee, C., Houssami, N., Elmore, J., Buist, DSM. «Pathways to breast cancer screening artificial intelligence algorithm validation». *The Breast*. 2020; 52: 146–149.
154. Collins, GS., Reitsma, JB., Altman, DG., Moons, KGM. «Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement». *European Urology*. 2015; 67(6): 1142–1151.
155. Vollmer, S., Mateen, BA., Bohner, G., et al. «Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness». *British Medical Journal*. 2020; 368: l6927.
156. Uygun Ilikhan, S., Özer, M., Tanberkan, H., Bozkurt, V. «How to mitigate the risks of deployment of artificial intelligence in medicine?». *Turkish journal of medical sciences*. 2024; 54(3): 483–492.
157. Daneshvar, N., Pandita, D., Erickson, S., Snyder Sulmasy, L. «Artificial Intelligence in the Provision of Health Care: An American College of Physicians Policy Position Paper». *Annals of internal medicine*. 2024; 177(7): 964–967.
158. Goddard, K., Roudsari, A., Wyatt, JC. «Automation bias: a systematic review of frequency, effect mediators, and mitigators». *Journal of the American Medical Informatics Association*. 2012; 19(1): 121–127.
159. Alami, H., Rivard, L., Lehoux, P., et al. «Artificial intelligence in health care: laying the Foundation for Responsible, sustainable, and inclusive innovation in low- and middle-income countries». *Globalization and health*. 2020; 16(52).
160. Gundersen, T., Bærøe, K. «The Future Ethics of Artificial Intelligence in Medicine: Making Sense of Collaborative Models». *Science and Engineering Ethics*. 2022; 28(17).
161. Skuban-Eiseler, T., Orzechowski, M., Denkinger, M., et al. «Artificial Intelligence–based Clinical Decision Support Systems in Geriatrics: An Ethical Analysis». *Journal of the American Medical Directors Association*. 2023; 24(9): 1271–1276.
162. Fletcher, RR., Nakeshimana, A., Olubeko, O. «Addressing Fairness, Bias, and Appropriate Use of Artificial Intelligence and Machine Learning in Global Health». *Frontiers in artificial intelligence*. 2021; 3: 561802.

163. Elwyn, G., Frosch, D., Thomson, R., et al. «Shared Decision Making: A Model for Clinical Practice». *Journal of General Internal Medicine*. 2012; 27(10): 1361–1367.
164. Masic, I., Miokovic, M., Muhamedagic, B. «Evidence Based Medicine - New Approaches and Challenges». *Acta Informatica Medica*. 2008; 16(4): 219-225.
165. McDougall, RJ. «Computer knows best? The need for value-flexibility in medical AI». *Journal of Medical Ethics*. 2019; 45(3): 156–160.
166. UN High-Level Panel on Digital Cooperation. [Publicación en línea] «The age of digital interdependence: report of the UN Secretary-General's High-Level Panel on Digital Cooperation». 2019 <https://digitallibrary.un.org/record/3865925?v=pdf> [Consulta: 18/11/2024]
167. Horgan, D., Romao, M., Morré, SA., Kalra, D. «Artificial Intelligence». *Public health genomics*. 2019; 22(5-6): 145–161.
168. Dubrovsky, E. [Publicación en línea] «The age of digital interdependence: report of the UN Secretary-General's High-Level Panel on Digital Cooperation». 2019 <https://news.un.org/en/story/2019/03/1034831> [Consulta: 18/11/2024]
169. Mehta, MC., Katz, IT., Jha, AK. «Transforming Global Health with AI». *The New England journal of medicine*. 2020; 382(9): 791–793.
170. Senders, JT., Zaki, MM., Karhade, AV., et al. «An introduction and overview of machine learning in neurosurgical care». *Acta Neurochirurgica*. 2018; 160(1): 29–38.
171. Baeva, AV. «Epistemic status of artificial intelligence in medical practice: Ethical challenges». *Digital diagnostics*. 2024; 5(1): 120–132.
172. Hurley, ME., Lang, BH., Kostick-Quenet, KM., et al. «Patient Consent and The Right to Notice and Explanation of AI Systems Used in Health Care». *American journal of bioethics*. 2024; 25(3): 102-114.
173. Hendricks-Sturrup, R., Simmons, M., Anders, S., et al. «Developing Ethics and Equity Principles, Terms, and Engagement Tools to Advance Health Equity and Researcher Diversity in AI and Machine Learning: Modified Delphi Approach». *JMIR AI*. 2023; 2: e52888.
174. Williams, AM., Liu, Y., Regner, KR., et al. «Artificial intelligence, physiological genomics, and precision medicine». *Physiological genomics*. 2018; 50(4): 237–243.
175. Nicholson, W., Price, I. «Health Care Artificial Intelligence: Law, Regulation, and Policy». En: *Artificial Intelligence in Health Care: the hope, the hype, the promise, the peril*. Whicher, D., Ahmed, M., Israni, ST., Michael Matheny. National Academy Press, Washington DC, 2024.
176. Davenport, T., Kalakota, R. «The potential for artificial intelligence in healthcare». *Future Healthcare Journal*. 2019; 6(2): 94–98.
177. Federspiel, F., Mitchell, R., Asokan, A., et al. «Threats by artificial intelligence to human health and human existence». *BMJ Global Health*. 2023; 8(5): e010435.
178. Li, P., Yang, J., Islam, MA., Ren, S. [Publicación en línea] «Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models». Communications of the ACM. 2023. <https://www.semanticscholar.org/paper/Making-AI-Less-%27Thirsty%27-Li-Yang/05fae03311d24e3f4a82e8021153f3d68d3bdf63> [Consulta: 18/11/2024]
179. Talasila, S., Neelima, N., Kumar, AS., et al. «The Eco-Ethical Nexus». En: *Artificial Intelligence Techniques for Sustainable Development*. *CRC Press*, 2024, 26–43.
180. Floridi, L. «Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical». *Philosophy & Technology*. 2019; 32(2): 185–193.
181. United Nations. [Publicación en línea] «SECRETARY-GENERAL'S REMARKS TO INFORMAL MEETING OF THE GENERAL ASSEMBLY ON THE INDEPENDENT HIGH-LEVEL PANEL ON DIGITAL COOPERATION». 2019. <https://www.un.org/en/sg-digital-cooperation-panel> [Consulta: 18/11/2024]
182. Amnesty International, Access Now. [Publicación en línea] «The Toronto Declaration». 2018. <https://www.torontodeclaration.org> [Consulta: 18/11/2024]
183. Committee of Experts on Internet Intermediaries (MSI-NET). [Publicación en línea] «Address-

ing the impacts of Algorithms on Human Rights». Council of Europe, 2018. <https://edoc.coe.int/en/internet/7589-algorithms-and-human-rights-study-on-the-human-rights-dimensions-of-automated-data-processing-techniques-and-possible-regulatory-implications.html> [Consulta: 18/11/2024]

184. Oniani, D., Hilsman, J., Peng, Y., et al. «Adopting and expanding ethical principles for generative artificial intelligence from military to healthcare». *npj Digital Medicine*. 2023; 6(225).
185. United Nations. [Publicación en línea] «UN SECRETARY-GENERAL ANTONIO GUTERRES CALLS FOR A GLOBAL RESET IN 2021». 2021. <https://unric.org/en/un-secretary-general-antonio-guterres-calls-for-a-global-reset-in-2021/> [Consulta: 18/11/2024]
186. Kahraman, F., Aktas, A., Bayrakceken, S., et al. «Physicians' ethical concerns about artificial intelligence in medicine: a qualitative study: "The final decision should rest with a human"». *Frontiers in public health*. 2024; 7: 1616827.
187. Frischhut, M., Werner-Felmayer, G. «Medical ethical principles against the background of 'European values'». *Health Care Ethics and Communication*. 2024; 52(7): 409–412.
188. Peng, ML., Wickersham, JA., Altice, FL., et al. «Formative Evaluation of the Acceptance of HIV Prevention Artificial Intelligence Chatbots By Men Who Have Sex With Men in Malaysia: Focus Group Study». *JMIR formative research*. 2022; 6(10): e42055.
189. Predel, C., Steger, F. «Ethical Challenges With Smartwatch-Based Screening for Atrial Fibrillation: Putting Users at Risk for Marketing Purposes?». *Frontiers in cardiovascular medicine*. 2021; 7: 615927.
190. De Gagne, JC., Hwang, H., Jung, D. «Cyberethics in nursing education: Ethical implications of artificial intelligence». *Nursing ethics*. 2024; 31(6): 1021–1030.
191. Segers, S., De Proost, M. «Take five? A coherentist argument why medical AI does not require a new ethical principle». *Theoretical Medicine and Bioethics*. 2024; 45(5): 387–400.
192. Wang, W., Wang, Y., Chen, L., et al. «Justice at the Forefront: Cultivating felt accountability towards Artificial Intelligence among healthcare professionals». *Social science & medicine*. 2024; 347: 116717.
193. Zhui, L., Fenghe, L., Xuehu, W., et al. «Ethical Considerations and Fundamental Principles of Large Language Models in Medical Education: Viewpoint». *Journal of medical Internet research*. 2024; 26: e60083.
194. Decker, M. «Caregiving robots and ethical reflection: the perspective of interdisciplinary technology assessment». *AI & Society*. 2008; 22(3): 315–330.
195. Makridis, C., Hurley, S., Klote, M., Alterovitz, G. «Ethical Applications of Artificial Intelligence: Evidence From Health Research on Veterans». *JMIR medical informatics*. 2021; 9(6): e28921.
196. Frez Pulgar, GA. «Bioética e Integración de la Inteligencia Artificial en la Medicina». *Aesthethika*. 2023; 19(2): 41–47.
197. National Health System. [Publicación en línea] «The Topol Review Interim Report». 2018. <https://www.hee.nhs.uk/sites/default/files/documents/Topol%20Review%20interim%20report_0.pdf> [Consulta: 18/11/2024]
198. Donnelly, D. «First Do No Harm: Legal Principles Regulating the Future of Artificial Intelligence in Health Care in South Africa». *Potchefstroom electronic law journal*. 2022; 25: 1–43.
199. Lukkien, DRM., Stolwijk, NE., Ipakchian Askari, S., et al. «AI-Assisted Decision-Making in Long-Term Care: Qualitative Study on Prerequisites for Responsible Innovation». *JMIR nursing*. 2024; 7: e55962.
200. Bleher, H., Braun, M. «Reflections on Putting AI Ethics into Practice: How Three AI Ethics Approaches Conceptualize Theory and Practice». *Science and engineering ethics*. 2023; 29(3).
201. Economou-Zavlanos, NJ., Bessias, S., Cary, MP., et al. «Translating ethical and quality principles for the effective, safe and fair development, deployment and use of artificial intelligence technologies in healthcare». *Journal of the American Medical Informatics Association*. 2024; 31(3): 705–713.
202. Grover, V., Balusamy, B., Rajasekaran, VA., Milanova, MG. *Approaches to human-centered AI in healthcare*, IGI Global, Hershey PA, 2024.

203. Lyons, JB., Hobbs, K., Rogers, S., Clouse, SH. «Responsible (use of) AI». *Frontiers in neuroergonomics*. 2023; 4: 1201777.
204. Smith, JA., Abhari, RE., Hussain, Z., et al. «Industry ties and evidence in public comments on the FDA framework for modifications to artificial intelligence/machine learning-based medical devices: a cross sectional study». *British Medical Journal*. 2020; 10(10): e039969.
205. McCoy, LG., Nagaraj, S., Morgado, F., et al. «What do medical students actually need to know about artificial intelligence?». *npj Digital Medicine*. 2020; 3(86).
206. Refolo, P., Duthie, K., Hofmann, B., et al. «Ethical challenges for Health Technology Assessment (HTA) in the evolving evidence landscape». *International journal of technology assessment in health care*. 2024; 40(1): e39.
207. Gray, K., Slavotinek, J., Dimaguila, GL., Choo, D. «Artificial Intelligence Education for the Health Workforce: Expert Survey of Approaches and Needs». *JMIR medical education*. 2022; 8(2): e35223.
208. Weidener, L., Fischer, M. «Proposing a Principle-Based Approach for Teaching AI Ethics in Medical Education». *JMIR medical education*. 2024; 10: e55368.
209. Shameer, K., Johnson, KW., Glicksberg, BS., et al. «Machine learning in cardiovascular medicine: are we there yet?». *Heart*. 2018; 104(14): 1156–1164.
210. Westerstrand, S. «Reconstructing AI Ethics Principles: Rawlsian Ethics of Artificial Intelligence». *Science and Engineering Ethics*. 2024; 30(46).
211. Vandersluis, R., Savulescu, J. «The selective deployment of AI in healthcare: An ethical algorithm for algorithms». *Bioethics*. 2024; 38(5): 391–400.
212. Rohwer, A., Pfadenhauer, L., Burns, J., et al. «Series: Clinical Epidemiology in South Africa. Paper 3: Logic models help make sense of complexity in systematic reviews and health technology assessments». *Journal of clinical epidemiology*. 2017; 83: 37–47.
213. Comité de Bioética de España. [Publicación en línea] «Informe del Comité de Bioética de España sobre Aspectos Bioéticos de la Telemedicina en el Contexto de la Relación Clínica». 2022. <https://comitedebioetica.isciii.es/wp-content/uploads/2023/10/CBE_Informe-sobre-aspectos-bioeticos-de-la-telemedicina-en-el-contexto-de-la-relacion-clinica.pdf> [Consulta: 18/11/2024]