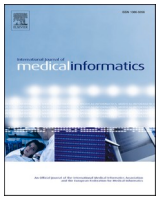
International Journal of Medical Informatics 161 (2022) 104738

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf



Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review

Anto Čartolovni [a,b,*], Ana Tomičić [a], Elvira Lazić Mosler [b,c]

[a] Digital Healthcare Ethics Laboratory (Digit-HeaL), Catholic University of Croatia, Ilica 242, 10 000 Zagreb, Croatia
[b] School of Medicine, Catholic University of Croatia, Ilica 242, 10 000 Zagreb, Croatia
[c] General Hospital Dr. Ivo Pedišić, Sisak, Croatia

ARTICLE INFO

Keywords:
Artificial intelligence
Medical ethics
Decision-making
Transparency
ELSI
Ethics by design
decision-making
medical AI
bioethics
digital health

ABSTRACT

Introduction: Recent developments in the field of Artificial Intelligence (AI) applied to healthcare promise to solve many of the existing global issues in advancing human health and managing global health challenges. This comprehensive review aims not only to surface the underlying ethical and legal but also social implications (ELSI) that have been overlooked in recent reviews while deserving equal attention in the development stage, and certainly ahead of implementation in healthcare. It is intended to guide various stakeholders (eg. designers, engineers, clinicians) in addressing the ELSI of AI at the design stage using the *Ethics by Design* (EbD) approach.
Methods: The authors followed a systematised scoping methodology and searched the following databases: Pubmed, Web of science, Ovid, Scopus, IEEE Xplore, EBSCO Search (Academic Search Premier, CINAHL, PSYCINFO, APA PsycArticles, ERIC) for the ELSI of AI in healthcare through January 2021. Data were charted and synthesised, and the authors conducted a descriptive and thematic analysis of the collected data.
Results: After reviewing 1108 papers, 94 were included in the final analysis. Our results show a growing interest in the academic community for ELSI in the field of AI. The main issues of concern identified in our analysis fall into four main clusters of impact: AI algorithms, physicians, patients, and healthcare in general. The most prevalent issues are patient safety, algorithmic transparency, lack of proper regulation, liability & accountability, impact on patient-physician relationship and governance of AI empowered healthcare.
Conclusions: The results of our review confirm the potential of AI to significantly improve patient care, but the drawbacks to its implementation relate to complex ELSI that have yet to be addressed. Most ELSI refer to the impact on and extension of the reciprocal and fiduciary patient-physician relationship. With the integration of AIbased decision making tools, a bilateral patient-physician relationship may shift into a trilateral one.

1. Introduction

1.1. Rationale

A global race in the field of Artificial Intelligence (AI) research and development is heating up across Europe, North America and Asia. Medicine and healthcare are not exempted from this race, as recent studies in the field of AI applied to medicine have shown promising results in dermatological diagnosis of skin cancer [1], in predicting the risk of emergency admissions [2], or in detecting breast cancer, outperforming radiology practitioners [3]. In a statement at the 70th General Assembly in 2019, the World Medical Association also acknowledged the vast potential of AI, suggesting, however, the term "augmented intelligence" as a narrow form of AI to underscore its supportive role. This variation in terms accurately depicts the real purpose of having AI coexist with human decision-making [1], in supporting and enhancing physicians' clinical decision-making [4]. However, in recent years, ethical and regulatory frameworks have proven to be a particularly daunting exercise in light of the increasing development of AI solutions in healthcare [5,6]. This is a major concern, as the current trend of reducing context-specific issues, namely healthcare, to abstract ethical principles or guidelines amounts to ethical whitewashing [7]. In the field of healthcare in particular, this can lead to physical or social harm [8,9] which impacts the public acceptability of algorithmic solutions beyond healthcare [10].

This scoping review focuses on the Ethical, Legal and Social issues

* Corresponding author.
E-mail addresses: anto.cartolovni@unicath.hr (A. Čartolovni), ana.tomicic@unicath.hr (A. Tomičić), elvira.lazic@unicath.hr (E. Lazić Mosler).

https://doi.org/10.1016/j.ijmedinf.2022.104738
Received 19 December 2021; Received in revised form 11 February 2022; Accepted 10 March 2022
Available online 14 March 2022

(ELSI) in the field of AI-based medical decision-support tools. It is worth noting that the purpose of a solely ethical reflection varies considerably from the ELSI approach in that their respective foci differ according to disciplinary, regulatory, and institutional contexts. An ELSI reflection aims to inform decision-makers and, more broadly, all stakeholders by providing the most comprehensive picture of the issues at stake. It should consider all aspects under evaluation, as well as the overall context in which the intervention (in this case - the implementation of AI in healthcare) is being carried out. This allows for a reflection aimed at highlighting the complexity of such issues, which in turn informs its conclusions. It also facilitates collaborative work between the different stakeholders tasked with evaluating these aspects, allowing each to refer to a common terminology and methodological grounds. On the basis of this, the need for a broader analysis, extending beyond mere normative considerations in ethical analysis to include social implications, had already been recognized: in the context of genomics, director of the Human Genome Project (HGP) James Watson coined the acronym ELSI to mean *ethical, legal, and social issues,* insisting that the project should also focus on societal aspects of the genetics research and its applications [11]. While it is difficult to clearly delineate between ethical, legal, and social issues, the latter should not be overlooked, as medicine is an important social practice affected by the social determinants of health (SDoH) [12]. Existing disparities and social inequalities in global health risk being replicated in AI-powered healthcare through biases in AI algorithms, causing further inequalities and potential discrimination. Earlier reviews have tended to focus solely on ethics [13–16], or only on regulatory perspectives [17], while a number have focused on specific fields such as neuroscience [18], nuclear medicine, gastroenterology, intensive care [19] and radiology [20]. The purpose of this review is to fill the gaps in the literature by bringing social issues into the discussion and to provide a coherent overview of ELSI from the broad spectrum of AI-based medical decision-making tools. This comprehensive review of ELSI is intended to guide various stakeholders (eg. designers, engineers, clinicians) in addressing these issues at the design stage using the Ethics by Design (EbD) approach [21].

1.2. Objectives

The objective of this scoping review is to collate and synthesise the existing literature on ELSI of AI-based medical decision-making tools, to outline the most common ELSI, and to identify the gaps in the literature. Our analysis will be driven by the following research question: *Which ELSI are most reported for AI-based medical decision support tools?* This review further aims to call the attention of healthcare policymakers to the more overlooked ELSI of AI-assisted medical decision-making tools.

2. Methods

To outline the breadth of ELSI in the emerging field of AI decision-making tools, we followed the method defined by Levac, Colquhoun, and O'Brien [22] and the Joanna Brigs Institute [23], and used the PRISMA-ScR checklist [24].

2.1. Eligibility criteria

We conducted a broad and comprehensive search, wherein all authors agreed on the following inclusion and exclusion criteria: included publications focused on ELSI of AI-based decision-making tools, and papers on Big data (i.e. data sets too large and complex for traditional data processing and management) and robotics were excluded. The corpus of analysis included original research published in peer-reviewed journals, books, book chapters, conference proceedings, perspectives or viewpoints written in English, while following systematized scoping methodology commentaries, editorials, business reports, dissertations, reviews were excluded. Theoretical, empirical/descriptive, empirical/experimental, empirical/quasi-experimental study designs were eligible for review. All publications published before January 1, 2021 were included in the review.

2.2. Information sources and search criteria

While developing a search strategy, all authors agreed to use the generic term artificial intelligence and its variants in the healthcare context, such as AI or augmented intelligence (ie, narrow AI), as it proved to be the most widely used term in our preliminary searches of ELSI literature. In these initial searches, we realized that the abbreviation AI yielded numerous results referring to apolipoprotein A-I, which we excluded from our Boolean search string. We conducted our search in the following online bibliographic databases: Pubmed, Web of science, Ovid, Scopus, IEEE Xplore, EBSCO Search (Academic Search Premier, CINAHL, PSYCINFO, APA PsycArticles, ERIC) on January 22nd 2021 using the following keywords: health*, medic*, clinic*, *ethic*, legal, social, "ELSI", "artificial intelligence", AI, "augmented intelligence", excluding "apolipoprotein". As we once again obtained a broad and unrefined spectrum of results by searching both title and abstracts, we opted to limit our search to titles only to obtain more focused results, which is recommended as a methodologically sound strategy [25,26]. We then filtered the databases by publication type (journal articles, books, chapters, and conference proceedings), by studies focusing on humans, and by publications written in English (See **Appendix A.**). We also performed a backward and forward snowballing search, drawing on the reference lists publications selected for analysis.

2.3. Selection of sources of evidence

After AČ, AT, ELM completed the search in the subsequent screening phases, the first and second author (AČ, AT) independently analysed the corpus of selected publications to avoid bias in the review results. Firstly, we inserted all publications from the primary literature search into a Mendeley reference manager and removed duplicates. Then, on the online platform Rayyan [27] both researchers double-blindly screened the titles and abstracts of the selected publications against the inclusion and exclusion criteria. Where there was ambiguity in the first screening phase, the publication was included for further evaluation during full-text analysis. The authors AČ and AT then proceeded to the full-texts analysis using Rayyan. In the third phase, the two researchers (AČ, AT) briefed frequently on indicator labelling to avoid coder's drift and resolved disagreement of screening decisions and full-texts analysis with the third researcher (ELM).

2.4. Data charting process and data items

In the preliminary searches and throughout the first screening phase, we followed the data charting process as suggested by the Joanna Briggs Institute's Manual for Evidence Synthesis for scoping reviews [23] and jointly outlined a *meta*-data table of relevant indicators to determine which variables to extract from our corpus (**Appendix A**). Relevant metadata were extracted during full-text analysis: year of publication, first author's country of affiliation, technology discussed, and category of ELSI, namely ethical, social, or legal.

2.5. Synthesis of results

We analyzed original theoretical and empirical work for descriptive and conceptual themes in the ELSI of AI-based decision-making tools. A descriptive analysis revealed the distribution of the lead authors' country of affiliation. For the thematic analysis, we collated descriptive codes based on the first 25 articles to label technological issues and ELSI. After reaching a consensus among authors on the appropriate codes, they were applied throughout the entire body of literature, with new codes added as necessary. A narrative form of the results is presented below, followed by figures which provide visual representation of key

findings. Full results, graphs, and tables are available in **Appendix A**.

3. Results

3.1. Selection of sources of evidence

Our initial search yielded 1108 publications from different databases based on title prevalence. The stages of the search and screening are shown in Fig. 1. After deduplication, 1060 publications remained. Once we filtered titles and abstracts against our inclusion criteria, 59 were included in the full-text analysis. We then conducted backward and forward snowballing based on the references of each publication and included 57 additional papers, resulting in 116 papers collated for full-text analysis. In full-text analysis, 22 publications were excluded, because they did not focus on ELSI (n = 12), did not address AI-based decision-making tools (n = 5), did not focus on AI in healthcare (n = 2), were not related to healthcare at all (n = 1), or because they were primarily concerned on Big data (n = 1) or robotics (n = 1).

3.2. Synthesis of results

3.2.1. Descriptive analytics

Our search yielded papers published between 2014 and 2020, predominantly published in 2019 (n = 32) and 2020 (n = 42). The majority of authors in our selected corpus of publications are affiliated with institutions in the United States (n = 37), the United Kingdom (n = 10) or Canada (n = 9) (Fig. 2).

Regarding the disciplinary field of the authors, as it might be presumed when it comes to ELSI we found that most studies were interdisciplinary (n = 49), followed by health sciences (n = 24) and law (n = 7). In terms of research design, 69% of the studies were theoretical, focusing exclusively on theoretically analyzing and discussing the various ELSIs. As for technology, in addition to the generic term AI (n = 40), the most frequently discussed were machine learning (n = 48) and deep learning (n = 20), along with several papers mentioning IBM's Watson (n = 11). Moreover, authors often associate various technologies with technical issues (e.g., algorithmic bias (n = 25), explainability (n = 22), accuracy (n = 22), and opacity (n = 20)), but also with ethical, legal

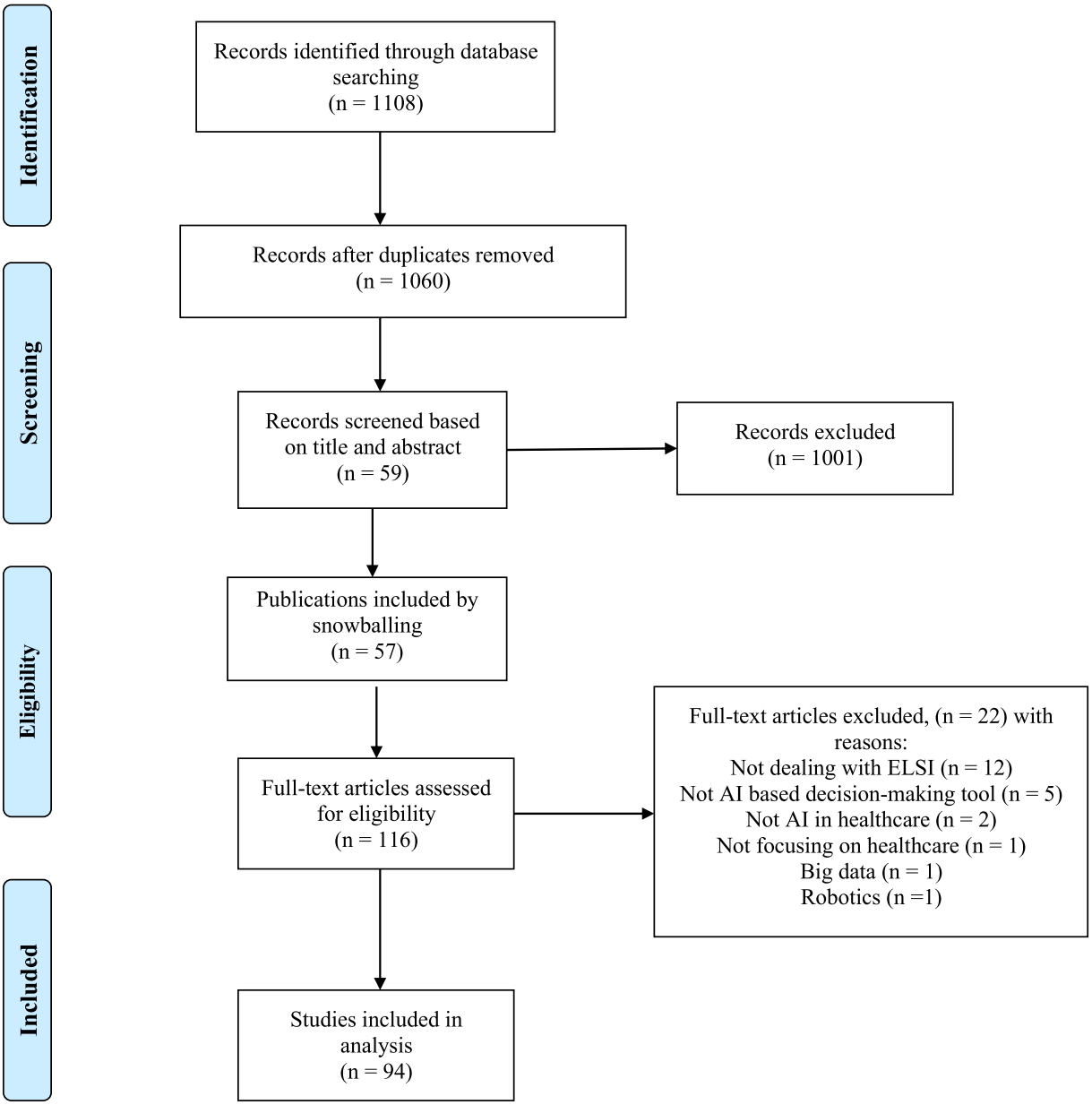


Fig. 1. PRISMA flow chart of the stages in the scoping literature review.

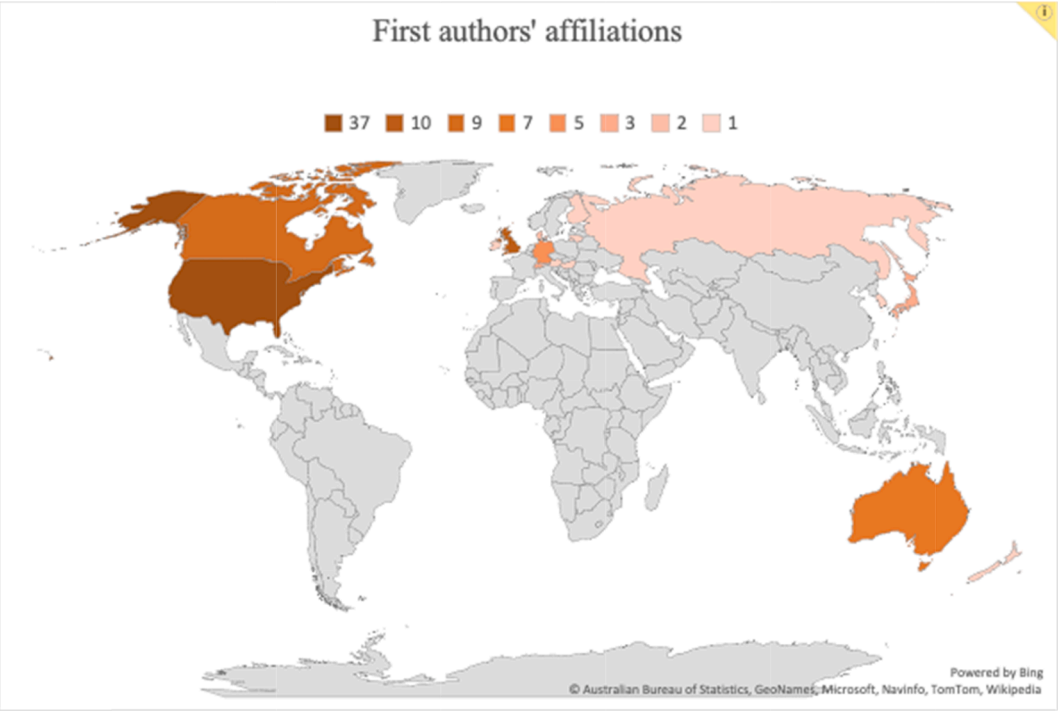


Fig. 2. Global distribution of papers according to primary author's affiliation.

or social implications. (See **Appendix A** for a more detailed description).

3.2.2. Overview of ethical, legal and social implications

For heuristic purposes we classified the issues in our analysis into three categories (ethical, legal, and social), although the vast majority could not be isolated and many of the reported ELSI were strongly intertwined. We therefore proceeded to present them visually as a coherent body of ELSI across four different clusters (AI algorithm, Physician, Patient and Healthcare) distributed according to their relevance and interrelation (Fig. 3). We schematized their distribution and interplay, keeping in mind their frequency in the literature and emphasizing it with font size (the largest being most represented and the smallest being least represented in the literature). Certain ELSI are spread across all four (safety and transparency, for instance), and others are at the intersection of two or more clusters, such as responsibility, accountability, and impact on patient-physician relationship creating an interdependent relationship between different clusters.

Given that our search focused on healthcare, a prevailing ethical issue identified in the reviewed literature was patient safety (n = 49), along with closely related ethical issues such as: transparency (n = 44), together with the related issues of biased decision-making (n = 43), explainability (n = 38), black-box (n = 30), trustworthiness (n = 29), opacity (n = 26), validity and reliability (n = 31), accountability (n = 22) and responsibility (n = 19) [28]. In the context of AI, the term "black box" refers to the lack of transparency [9,29,30], while unpredictability due to complexity or the inability to discern decision-making from within is termed opacity [31–34]. Not disclosing the use of an opaque AI-based decision-making support tool may be unethical [35] in that it may negatively impact patient-physician relationship [31], undermine patient autonomy and trust [36,37], and potentially compromise informed consent [33,38,39]. Furthermore, privacy emerged as a major ethical and legal concern (n = 42), often in the context of the misuse of electronic medical records (which has been publicly problematized) and of the issue of data ownership (n = 14).

The lack of appropriate regulation emerged as the most prevalent legal issue in our analysis (n = 50). As AI becomes increasingly ubiquitous, the current legislative and regulatory void will need to be addressed [10,40]. This is particularly so as it applies to healthcare, and more specifically to safety-effectiveness and liability within fault-based legal systems in case of self-learning algorithms embedded in AI-based decision-making tools [41,42]. However, when AI-based decision-making tools are designed to suggest medical recommendations and guidance [37,38,43] another prevailing issue in literature is liability and accountability for patient harm (n = 40). Who would be responsible and who should be held accountable? Who should patients or their relatives seek compensation from in the event of patient harm? [44,45] Most of the difficulties in establishing liability for AI-based tools stems from its opacity and black-box nature. The difficulty lies in demonstrating that the treatment was inappropriate, that the patient's rights were violated or that the harm caused by the algorithms and the chain of legal causality has been impaired [46].

Given that AI-based decision-making tools will clearly transform the future of healthcare, the most critical social issue identified by our literature review is the impact of AI on the patient-physician relationship (n = 43). AI-based decision-making tools are expected to have a major impact on the fiduciary relationship between physician and patient: the use of black-box algorithms may indeed undermine the level of trust that exists in the patient-physician relationship today. Other social issues that were highlighted by our analysis of literature focused on governance and policy (n = 35), public trust (n = 34) and acceptability (n = 31). Some studies emphasize and share Eric Topol's [47] enthusiastic views on the impact of healthcare quality (n = 6) whereby the implications of AI-based decision support tools in improving diagnostic and prognostic efficiency would unlock additional time for healthcare professionals to shift their focus toward the nurturing side of care and engage in trusting relationships with empathy and compassion [48,49]. Moreover, the benefit of this increased availability in time will not only result in more trustworthy and empathic care, but will also impact the workforce (n = 22) in terms of reduced stress and work overload for healthcare workers [48].

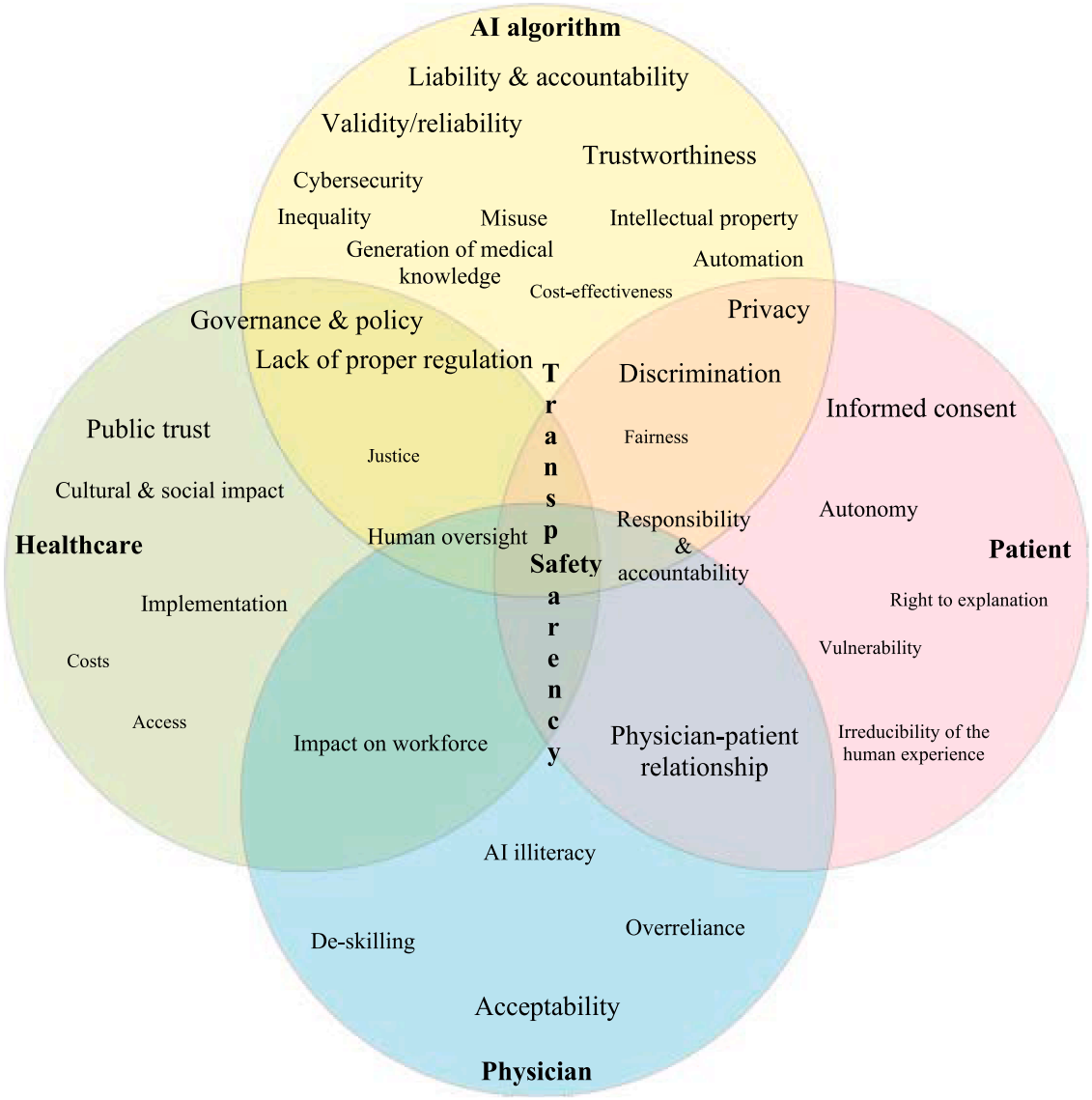


Fig. 3. Graphical representation of ELSI distribution and their interplay.

4. Discussion

4.1. Summary of evidence

A coherent overview of the ELSI of AI-based medical decision support tools presented in Fig. 3. constitutes the first phase of an *Ethics by Design (EbD)* approach in bringing to the attention of developers, designers, and engineers the underpinning issues which may vary throughout the development, deployment, and implementation phases [22]. The *EbD* approach entails reflecting on and addressing these issues through the implementation of ethical principles during the development of such systems. This approach is an additional tool for addressing ethical concerns and for demonstrating ethical compliance towards the responsible development and deployment of AI-based medical decision support tools [50].

4.1.1. Safety and transparency

As recognized by our literature corpus, safety and transparency constitute the two most pressing issues that pervade the entire AI-based medical decision-making process as an important and determining factor for trust in and acceptance of the technology [51]. As noted in the literature, safety has primarily focused on the reliability of the AI system [52] i.e. confidence in the proper functioning of the AI system, which is very often compromised by the algorithmic bias leading to unintended harm [53], and in the context of healthcare, against one of the key bioethical principles of non-maleficence [46,54]. This means that in a setting of daily clinical practice, if AI-based decision-making tools increase the risk of diagnostic error, either with false positives or false negatives [55] their use would then be unethical [56].

According to Sarah Gerke and colleagues [57], in achieving the safety and effectiveness of AI in healthcare two elements must be considered. Firstly, reliability and validity of the datasets - meaning refined datasets used for training, and secondly transparency. Second, a low level of transparency may hinder the trustworthiness of AI models in healthcare [29,51], or impair validation of the clinical recommendations for the model and identification of any errors or biases [41,58]. Therefore, some studies argue that non-explainable AI should be prohibited in healthcare [10] while at the same time representing one of the biggest regulatory challenges [59], notably from the accountability perspective [60] i.e. the moral responsibility for patient harm. The use of black-box algorithms in medicine is a novel epistemic challenge to current healthcare ideals and practices [5,28,37], which strives to be a person-centered medicine [37], especially in situations of disagreement between the physician's and algorithmic decision [32]. Opacity in black-

box medicine can manifest itself in two forms: literal opacity, when established relationships are completely hidden even if the algorithmic decision-making process is clear, and practical or approachable opacity, halfway between fully opaque black-box medicine and explicit personalized medicine [61]. However, although algorithmic opacity and algorithmic bias are related and yet independent from each other, this does not mean that greater transparency necessarily guarantees less bias [62]. Moreover, AI algorithms used in healthcare, besides being sufficiently transparent, also require monitoring over the reasoning and effects of AI-based decisions by providing a reasonable explanation to the patient [63,64].

In the medical context, transparency entails not only explainability but also a clear communication with patients about the limitations and performance gaps in protecting patient identity and mitigating harm [65]. Explainability is a key feature for physicians themselves and is a critical factor in their acceptance of the recommendations made by an AI decision-making tool [66], which may affect the ethical relationship between physician and patient. With this in mind, several questions pertaining to the degree of AI transparency arise [36]:

- Can it be achieved, what is considered as adequate explanation from the patient's perspective?
- How to address residual opacity, at which level could the decision-making procedure be explained?
- Should we prioritize input-oriented transparency focused on data quality, or the output related transparency to explain how the AI based decision making support tool arrived at a decision?

Even if transparency is achieved, a subsequent question is how to communicate AI decisions to the patient; in other words how to integrate it into the shared decision making process [28]. Therefore, AI-based decision-making support tools should incorporate a reasonable degree of transparency and mechanisms for responsible disclosure of their role in the decision-making process by enabling medical practitioners, patients, and their relatives to understand AI-based decisions [6].

4.1.2. *Lack of proper regulation - liability & accountability*

Despite attempts at regulation in the EU and the US, there are no globally harmonized standards or laws [6]. Comparing the approaches between Europe and other continents, Europe's regulatory framework reflects a more cautious stance: lack of harm does not warrantee safe deployment and use of the technology. Currently, there is no explicit European legislation on AI in healthcare and AI is defined as a medical device or as any software used for prevention, diagnosis, treatment or monitoring [63]. The recently enforced European Medical Device Regulation (MDR) 2017/745 classifies AI-based tools and decision-making software for diagnosis or therapeutic purposes as class IIa devices (in Rule 11), unless they result in death, irreversible deterioration of an individual's health (in that case, it is classified as class III), or a surgical intervention (class IIb) [35]. In the US, the Food and Drug Administration (FDA) took a step forward under the Digital Health Innovation Action Plan, launching a pilot pre-certification program evaluating medical software under development based on five criteria of excellence [59]. Until recently, current regulatory frameworks did not include the algorithmic self-learning capability of ML thereby failing to address specific risks of ML such as explainability or false positives/negatives [5]. But the FDA's new regulatory framework, Software as a medical device (SaMD), is inching towards addressing this issue as well as that of the algorithms' ability to update themselves through self-learning mechanisms (learning, adapting and optimizing in real time) [6,67,68].

The current legal realm for medical malpractice which determines the liability for patient harm, cannot fully address issues of accountability that AI-based decision-making tools bring to medicine, prompting Richman [69] to suggest a shift away from them toward corporate and product liability, i.e. the legal realm defining blameworthiness for patient harm in manufacturers, distributors, suppliers, and retailers. Others are more inclined toward tort law and medical malpractice regimes and suggest consulting corporate and product liability law only if AI-based decision-making support tools develop a level of sophistication such that their decisions are deemed a medical standard to follow. If AI-based treatments achieve greater accuracy than non-AI treatments, it will constitute the new "due standard of medical care" making healthcare professionals accountable for not exploiting their potential [5]. Results from a recent survey of physicians show that nearly half (49.3%) believe responsibility should lie with doctors, 31.2% with patients who consented to its use, and only 19.4% with the company which created the AI-based tool [70].

Price [71] argues that flexibility rather than rigid laws should be the path to a resilient legal framework. This should imply lighter premarket scrutiny, coupled with robust post-market scrutiny as these technologies enter the clinical care context, to avoid stifling innovation. New regulatory systems should promote meaningful human control by keeping humans in the loop [28], and their appraisal in regulatory impact assessments will strongly depend on the level of human control over AI, complemented by validation and certification mechanisms for algorithms [28]. Several countries, including EU states, announced new regulations and adapted general liability frameworks for AI to improve clarity, transparency and public trust in AI-driven solutions in healthcare [57].

4.1.3. *Impact on patient-physician relationship and governance of future healthcare*

AI's disruptive power may contribute to dismantling the conventional bilateral physician-patient relationship by involving various actors with AI systems: programmers, product manufacturers and AI-based tools, which will all challenge the existing medical ethics model of "shared decision-making" [46], downgrading medicine into a new kind of *paternalism*, namely *computernalism* [48,72]. Should clinicians embrace AI-based tools not only as supportive tools but as outright peers/actors, medical ethics will need to accommodate this inclusion of new morally relevant actors impacting the fiduciary relationship, making it less strained and prone to the physician's personal responsibility [73]. Admittedly, the redefinition of this bilateral relationship is already underway, and is transitioning to a trilateral "physician-computer-patient" relationship [74]. As a consequence, Triberti et al. [75] point out this new kind of relationship may lead to a "third wheel" effect, undermining the effectiveness of shared decision-making through the perception that the physician is only an intermediary between the patient and the algorithm. Humane relational practices should by no means be excluded from the emerging trilateral relationship [76]. Empathy and compassion constitute the human aspects of medical care, and recent study results among consumers confirm patients' and physicians' concerns that AI devices may reduce these human aspects, such as face-to-face cues and personal interactions with physicians, who might assume a more passive position in healthcare decision making in the future [42].

As elaborated above, the benefits AI would bring – in terms of increasing the time spent on the humanness in medicine, may however lead to precision and efficiency becoming more valued than empathy and professional judgement. This might potentially shift the focus of medicine away from human-specific skills [48], which might in turn contribute, as mentioned earlier, to the concerns that increasingly accurate AI algorithms might substitute physicians' judgement [43,77]. In this respect, a recent study of UK general practitioners' views suggests it is rather unlikely physicians will be substituted in the provision of empathic patient care [52]. Despite the epistemic superiority of AI-based decision-making tools, maintaining the clinician's professional integrity as the final decision-maker - "the human-in-the-loop", the one bearing responsibility for the decision-making - is arguably crucial [31,78]. Eschewing over-reliance and complacency by recognizing that AI cannot make moral judgments is key here [32]. However, negative

prejudices about AI's epistemic superiority also exist among physicians. An online survey of physicians by Oh et. al. [70] revealed that fewer than half of the survey participants agreed that "AI is superior to doctors' experience" (44%) and that "AI could replace doctors" (35.4%). Stripping physicians of epistemic authority and transferring it to an AI-based decision-making support tool would lead to shallow and deficient trust relationships in the practice of medicine that have already been seen in paternalistic medicine shifting the medicine and healthcare a few decades back [79,80].

Although current AI regulatory frameworks are still in their infancy, a possible solution to address the ELSI outlined in this analysis, and to avoid deploying AI models with potentially unsafe and morally questionable outcomes in healthcare, is to explore different models of AI governance in healthcare [41,45,81]. As outlined in the recent WHO guidance *Ethics and Governance of Artificial Intelligence for Health* [82], they should be based on policies governing the implementation of AI-based tools and comply with the ethical principles of reproducibility, transparency, fairness and human dignity. Various approaches to tackle these issues have been proposed, exemplified by the global ethical framework and governance system suggested by Guan [83]. However, they may pose a logistically unfeasible solution given they would entail regular audits and monitoring mechanisms that need to be enforced at the global level to mitigate risks, with ongoing assessment of the safety, quality, transparency, and ethical factors of AI-based services [42]. A global effort and solidarity spanning regulatory aspects in both high and low-income countries is currently lacking. Our results are also consistent with this shortcoming, wherein most authors discussing ELSI of AI in healthcare are from high-income-countries. In the wake of *EbD* with its main aim to design and create a technology for social good, the already outlined and mapped ELSI with an emphasis on currently overlooked social issues represent a significant challenge for the development and deployment of AI-based decision-making tools and warrant timely attention to fulfill the promise of an impending AI-empowered healthcare.

5. Limitations

We recognize several limitations to our study. Firstly, our search strategy was limited to the broad terms - *artificial intelligence* and *augmented intelligence*. While additional terms and concepts (machine learning, deep learning etc.) might have further complemented our search, the scope of our review was very broad, encompassing all ethical, legal and social implications in the sphere of AI-based decision-making tools. Due to time and resource limitations, we did not expand our search to include additional terms, and opted to proceed with umbrella terms only, given they are prevalent in the literature on the ELSI of AI. Furthermore, language was limited to English, so we may have overlooked influential articles published in other languages. Additionally, in reporting the geographical distribution of papers across the globe, the first author's country of affiliation is not necessarily their country of activity.

6. Concluding remarks

Our review results support the view that there is potential to improve healthcare systems and advance patient care provision through AI, but its implementation implies several drawbacks, or ELSI, that require consideration. The ELSI we identified in our review reflect how AI decision-making tools will impact the future of healthcare and can be categorized into four clusters: issues that concern AI algorithms, issues with high potential to impact patients, issues with high potential to impact physicians, and issues with high potential to impact healthcare overall. The most represented are patient safety, algorithmic transparency, lack of proper regulation, liability & accountability, and impact on patient-physician relationship and governance of AI-based healthcare. By far the most prevalent is the last cluster referring to the impact on the reciprocal and fiduciary relationship between patient and physician. With the integration of AI-based decision-making tools, that relationship will be extended from its bilateral foundation to a third party. In its current configuration, the physician-patient relationship is grounded in principles of medical ethics. Integrating a new moral actor challenges its applicability and requires both consideration and vigilance towards all identified ELSI leading us towards more responsible healthcare innovation. The findings of this review can guide further research in the form of systematic reviews and meta-analyses of ELSI, to deliver more detailed insights. Based on our findings, we encourage more empirical research using qualitative and quantitative methods, to inform practice and policy. We further consider that comparison with existing principles of medical ethics is a key step in the effort to obtain such information, as they have recently guided the patient-physician relationship. In addition to a continuous reviewing of the literature, and to theoretical analyses, qualitative and quantitative empirical studies are imperative to further investigate the ELSI of AI with a focus on overlooked ELSI among key stakeholders, as a number of ELSI have been less frequently reported in the literature but deserve equal attention.

Authors' contributions.

AČ, AT and ELM designed the study. ELM and AČ collected data. AT and AČ performed data extraction and analysis, and all authors interpreted the data. AČ drafted the manuscript, and all authors critically revised the manuscript and gave approval for the publication of the final version. All authors had full access to the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding

This work was supported by the Croatian Science Foundation (CSF), grant "(New) Ethical and Social Challenges of Digital Technologies in the Healthcare Domain", grant number UIP-2019-04-3212. The funder had no role in the design of this study and its execution, analyses, interpretation of the data, or decision to submit results.

8. Declarations

Declarations of interest: none.

Conflicts of interest

Authors declare no conflict of interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2022.104738.

References

[1] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.

[2] F. Rahimian, G. Salimi-Khorshidi, A.H. Payberah, J. Tran, R. Ayala Solares, F. Raimondi, et al., Predicting the risk of emergency admission with machine learning: development and validation using linked electronic health records, PLoS Med. 15 (2018) e1002695. https://doi.org/10.1371/journal.pmed.1002695.

[3] D. Killock, AI outperforms radiologists in mammographic screening, Nat. Rev. Clin. Oncol. 17 (3) (2020).

[4] S. Gaube, H. Suresh, M. Raue, A. Merritt, S.J. Berkowitz, E. Lermer, J.F. Coughlin, J.V. Guttag, E. Colak, M. Ghassemi, Do as AI say: susceptibility in deployment of clinical decision-aids, NPJ Digit. Med. 4 (1) (2021), https://doi.org/10.1038/s41746-021-00385-9.

[5] D. Schneeberger, K. Stöger, A. Holzinger, The European Legal Framework for Medical AI. 4th IFIP TC 5, TC 12, WG 84, WG 89, WG 129 Int Cross-Domain Conf Mach Learn Knowl Extr CD-MAKE 2020 2020;12279 LNCS:209–26. https://doi.org/10.1007/978-3-030-57321-8_12.

[6] T. Minssen, S. Gerke, M. Aboy, N. Price, G. Cohen, Regulatory responses to medical machine learning, J. Law Biosci. (2020), https://doi.org/10.1093/jlb/lsaa002.

[7] C.-M. Mörch, A. Gupta, B.L. Mishara, C.-M. Morch, A. Gupta, B.L. Mishara, Canada protocol: an ethical checklist for the use of artificial Intelligence in suicide prevention and mental health, Artif. Intell Med. 108 (2020), 101934, https://doi.org/10.1016/j.artmed.2020.101934.

[8] K. Ferryman, Addressing health disparities in the Food and Drug Administration's artificial intelligence and machine learning regulatory framework, J. Am. Med. Informatics Assoc. 27 (2020) 2016–2019, https://doi.org/10.1093/jamia/ocaa133.

[9] T. Matsuzaki, Ethical issues of artificial intelligence in medicine, Calif West Law Rev 55 (2018) 255–274.

[10] S.M. Carter, W. Rogers, K.T. Win, H. Frazer, B. Richards, N. Houssami, The ethical, legal and social implications of using artificial intelligence systems in breast cancer care, BREAST 49 (2020) 25–32, https://doi.org/10.1016/j.breast.2019.10.001.

[11] H. Zwart, A. Nelis, What is ELSA genomics? EMBO Rep. 10 (2009) 540–544, https://doi.org/10.1038/embor.2009.115.

[12] Magnan S. Social Determinants of Health 101 for Health Care: Five Plus Five. NAM Perspect 2017;7. https://doi.org/10.31478/201710C.

[13] J. Morley, C.C.V. Machado, C. Burr, J. Cowls, I. Joshi, M. Taddeo, et al., The ethics of AI in health care: a mapping review, Soc. Sci. Med. 260 (2020) 113172. https://doi.org/10.1016/j.socscimed.2020.113172.

[14] K. Murphy, E. Di Ruggiero, R. Upshur, D.J. Willison, N. Malhotra, J.C. Cai, N. Malhotra, V. Lui, J. Gibson, Artificial intelligence for good health: a scoping review of the ethics literature, BMC Med. Ethics 22 (1) (2021), https://doi.org/10.1186/s12910-021-00577-8.

[15] C. Nebeker, J. Torous, R.J. Bartlett Ellis, Building the case for actionable ethics in digital health research supported by artificial intelligence, BMC Med. 17 (2019) 137, https://doi.org/10.1186/s12916-019-1377-7.

[16] S.H. Park, Y.-H. Kim, J.Y. Lee, S. Yoo, C.J. Kim, Ethical challenges regarding artificial intelligence in medicine from the perspective of scientific editing and peer review, Sci. Ed. 6 (2019) 91–98, https://doi.org/10.6087/kcse.164.

[17] H.B. Harvey, V. Gowda, How the FDA regulates AI, Acad. Radiol. 27 (2020) 58–61, https://doi.org/10.1016/j.acra.2019.09.017.

[18] P. Kellmeyer, Artificial intelligence in basic and clinical neuroscience: opportunities and ethical challenges, Neuroforum 25 (2019) 241–250, https://doi.org/10.1515/nf-2019-0018.

[19] M. Beil, I. Proft, D. van Heerden, S. Sviri, P.V. van Heerden, Ethical considerations about artificial intelligence for prognostication in intensive care, Intensive Care Med. Exp. 7 (2019) 70, https://doi.org/10.1186/s40635-019-0286-6.

[20] F. Pesapane, C. Volonté, M. Codari, F. Sardanelli, Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States, Insights Imaging 9 (2018) 745–753, https://doi.org/10.1007/s13244-018-0645-y.

[21] World Economic Forum. Ethics by Design: An organizational approach to responsible use of technology | World Economic Forum. 2021.

[22] D. Levac, H. Colquhoun, K.K. O'Brien, Scoping studies: advancing the methodology, Implement Sci. 5 (2010) 69, https://doi.org/10.1186/1748-5908-5-69.

[23] M.D.J. Peters, C. Marnie, A.C. Tricco, D. Pollock, Z. Munn, L. Alexander, P. McInerney, C.M. Godfrey, H. Khalil, Updated methodological guidance for the conduct of scoping reviews, JBI Evid. Implement. 19 (1) (2021) 3–10.

[24] A.C. Tricco, E. Lillie, W. Zarin, K.K. O'Brien, H. Colquhoun, D. Levac, D. Moher, M. D.J. Peters, T. Horsley, L. Weeks, S. Hempel, E.A. Akl, C. Chang, J. McGowan, L. Stewart, L. Hartling, A. Aldcroft, M.G. Wilson, C. Garritty, S. Lewin, C. M. Godfrey, M.T. Macdonald, E.V. Langlois, K. Soares-Weiser, J.o. Moriarty, T. Clifford, Ö. Tunçalp, S.E. Straus, PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation, Ann. Int. Med. 169 (7) (2018) 467–473.

[25] F.J. Mateen, J. Oh, A.I. Tergas, N.H. Bhayani, B.B. Kamdar, Titles versus titles and abstracts for initial screening of articles for systematic reviews, Clin. Epidemiol. 5 (2013) 89–95, https://doi.org/10.2147/CLEP.S43118.

[26] A.C. Dowd, R.M. Johnson, Why Publish a Systematic Review: An Editor's and Reader's Perspective. In: O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond, K. Buntins, editors, Syst. Rev. Educ. Res., Wiesbaden: Springer Fachmedien Wiesbaden; 2020, p. 69–87. https://doi.org/10.1007/978-3-658-27602-7_5.

[27] M. Ouzzani, H. Hammady, Z. Fedorowicz, A. Elmagarmid, Rayyan-a web and mobile app for systematic reviews, Syst. Rev. 5 (2016) 210, https://doi.org/10.1186/s13643-016-0384-4.

[28] M. Braun, P. Hummel, S. Beck, P. Dabrock, Primer on an ethics of AI-based decision support systems in the clinic, J. Med. Ethics 47 (12) (2021).

[29] T.J. Kasperbauer, Conflicting roles for humans in learning health systems and AI-enabled healthcare, J. Eval. Clin. Pract. (2020), https://doi.org/10.1111/jep.13510.

[30] H.R. Sullivan, S.J. Schweikart, Are current tort liability doctrines adequate for addressing injury caused by AI? AMA J. Ethics 21 (2019) 160–166, https://doi.org/10.1001/amajethics.2019.160.

[31] A. Ferretti, M. Schneider, A. Blasimme, Machine learning in medicine: opening the new data protection black box, Eur. Data Prot. Law Rev. 2018;4.

[32] T. Lysaght, H.Y. Lim, V. Xafis, K.Y. Ngiam, AI-assisted decision-making in healthcare: the application of an ethics framework for big data in health and research, Asian Bioeth. Rev. 11 (2019) 299–314, https://doi.org/10.1007/s41649-019-00096-0.

[33] M. Kiener, Artificial intelligence in medicine and the disclosure of risks, AI Soc. 36 (3) (2021) 705–713.

[34] K. Begley, C. Begley, V. Smith, Shared decision-making and maternity care in the deep learning age: acknowledging and overcoming inherited defeaters, J. Eval. Clin. Pract. 27 (3) (2021) 497–503.

[35] D.D. Luxton, Recommendations for the ethical use and design of artificial intelligent care providers, Artif. Intell Med. 62 (2014) 1–10, https://doi.org/10.1016/j.artmed.2014.06.004.

[36] K. Astromskė, E. Peičius, P. Astromskis, Ethical and legal challenges of informed consent applying artificial intelligence in medical diagnostic consultations, AI Soc 36 (2) (2021) 509–520.

[37] J.C. Bjerring, J. Busch, Artificial intelligence and patient-centered decision-making, Philos. Technol. 34 (2) (2021) 349–371.

[38] J. Amann, A. Blasimme, E. Vayena, D. Frey, V.I. Madai, Explainability for artificial intelligence in healthcare: a multidisciplinary perspective, BMC Med. Inform. Decis. Mak. 20 (2020) 310, https://doi.org/10.1186/s12911-020-01332-6.

[39] I. de Miguel Beriain, Should we have a right to refuse diagnostics and treatment planning by artificial intelligence? Med. Heal Care Philos. 23 (2) (2020) 247–252.

[40] J. Powles, H. Hodson, Google DeepMind and healthcare in an age of algorithms, Health Technol. (Berl) 7 (2017) 351–367, https://doi.org/10.1007/s12553-017-0179-1.

[41] S. Reddy, S. Allan, S. Coghlan, P. Cooper, A governance model for the application of AI in health care, J. Am. Med. Informatics Assoc. 27 (2020) 491–497, https://doi.org/10.1093/jamia/ocz192.

[42] P. Esmaeilzadeh, Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives, BMC Med. Inform. Decis. Mak. 20 (2020) 170, https://doi.org/10.1186/s12911-020-01191-1.

[43] S. Gao, L. He, Y. Chen, D. Li, K. Lai, Public perception of artificial intelligence in medical care: content analysis of social media, J. Med. Internet Res. 22 (2020), e16649, https://doi.org/10.2196/16649.

[44] K. Chockley, E. Emanuel, The end of radiology? Three threats to the future practice of radiology, J. Am. Coll Radiol. 13 (2016) 1415–1420, https://doi.org/10.1016/j.jacr.2016.07.010.

[45] I. Bartoletti, AI in Healthcare: Ethical and Privacy Challenges. In: D. Riano, S. Wilk, S. TenTeije, edito, Artif. Intell. Med. AIME 2019, vol. 11526, 2019, p. 7–10. https://doi.org/10.1007/978-3-030-21642-9_2.

[46] F. Molnár-Gábor, Artificial intelligence in healthcare: Doctors, patients and liabilities. In: T. Wischmeyer, T. Rademacher, editors. Regul. Artif. Intell., Cham: Springer International Publishing; 2019, p. 337–60. https://doi.org/10.1007/978-3-030-32361-5_15.

[47] E. Topol, Deep Medicine - How Artificial Intelligence Can Make Healthcare Human Again, Basic books, New York, NY, USA, 2019.

[48] A. Kerasidou, Artificial intelligence and the ongoing need for empathy, compassion and trust in healthcare, Bull World Health Organ 98 (2020) 245–250, https://doi.org/10.2471/BLT.19.237198.

[49] R.D. Truog, Of slide rules and stethoscopes: AI and the future of doctoring, Hastings Cent Rep 49 (5) (2019).

[50] European Commission. Ethics By Design and Ethics of Use Approaches for Artificial Intelligence. 2021.

[51] M. Hengstler, E. Enkel, S. Duelli, Applied artificial intelligence and trust-The case of autonomous vehicles and medical assistance devices, Technol Forecast Soc Change 105 (2016) 105–120, https://doi.org/10.1016/j.techfore.2015.12.014.

[52] C. Blease, T.J. Kaptchuk, M.H. Bernstein, K.D. Mandl, J.D. Halamka, C. M. Desroches, Artificial intelligence and the future of primary care: exploratory qualitative study of UK general practitioners' views, J. Med. Internet Res. 21 (2019), e12802, https://doi.org/10.2196/12802.

[53] T. Nadarzynski, O. Miles, A. Cowie, D. Ridge, Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit Heal 2019;5: 205520761987180. https://doi.org/10.1177/2055207619871808.

[54] M.D. McCradden, A. Baba, A. Saha, S. Ahmad, K. Boparai, P. Fadaiefard, et al., Addressing health disparities in the Food and Drug Administration's artificial intelligence and machine learning regulatory framework, AI Soc. 27 (2020) 63–71, https://doi.org/10.1177/2055207619871808.

[55] S. Jabri, Artificial Intelligence and Healthcare: Products and Procedures, in: T. Wischmeyer, T. Rademacher, editors, Regul. Artif. Intell., Cham: Springer International Publishing; 2020, p. 307–35. https://doi.org/10.1007/978-3-030-32361-5_14.

[56] K.W. Goodman, Ethical and Legal Issues in Decision Support. In: E.S. Berner editor, Clin. Decis. Support Syst., New York, NY: Springer New York; 2007, p. 126–39. https://doi.org/10.1007/978-0-387-38319-4_6.

[57] S. Gerke, T. Minssen, G. Cohen, Ethical and legal challenges of artificial intelligence-driven healthcare, Artif. Intell Healthc. (2020) 295–336, https://doi.org/10.1016/b978-0-12-818438-7.00012-5.

[58] T. Panch, H. Mattie, R. Atun, Artificial intelligence and algorithmic bias: Implications for health systems, J. Glob. Health 9 (2019) 10318, https://doi.org/10.7189/jogh.09.020318.

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

[60] C.L. Thompson, H.M. Morgan, Ethical barriers to artificial intelligence in the national health service, United Kingdom of Great Britain and Northern Ireland, Bull World Health Organ 98 (2020) 293–295, https://doi.org/10.2471/BLT.19.237230.

[61] Price WNII. Black-Box Medicine. Harv J Law Technol 2015;28:419–68.

[62] A. Tupasela, E. Di Nucci, Concordance as evidence in the Watson for Oncology decision-support system, AI Soc. 35 (2020) 811–818, https://doi.org/10.1007/s00146-020-00945-9.

[63] M.P. Recht, M. Dewey, K. Dreyer, C. Langlotz, W. Niessen, B. Prainsack, J.J. Smith, Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations, Eur. Radiol. 30 (6) (2020) 3576–3584.

[64] G.S. Nelson, Bias in Artificial Intelligence, N C Med. J. 80 (2019) 220–222. https://doi.org/10.18043/ncm.80.4.220.
[65] M.D. McCradden, S. Joshi, J.A. Anderson, M. Mazwi, A. Goldenberg, R.Z. Shaul, Patient safety and quality improvement: ethical principles for a regulatory approach to bias in healthcare machine learning, J. Am. Med. Informatics Assoc. 27 (2020) 2024–2027, https://doi.org/10.1093/jamia/ocaa085.
[66] W.K. Diprose, N. Buist, N. Hua, Q. Thurier, G. Shand, R. Robinson, Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator, J. Am. Med. Inform. Assoc. 27 (2020) 592–600, https://doi.org/10.1093/jamia/ocz229.
[67] B. Meskó, M. Görög, A short guide for medical professionals in the era of artificial intelligence, NPJ. Digit Med. 3 (2020) 126, https://doi.org/10.1038/s41746-020-00333-z.
[68] J. Wiens, S. Saria, M. Sendak, M. Ghassemi, V.X. Liu, F. Doshi-Velez, K. Jung, K. Heller, D. Kale, M. Saeed, P.N. Ossorio, S. Thadaney-Israni, A. Goldenberg, Do no harm: a roadmap for responsible machine learning for health care, Nat. Med. 25 (9) (2019) 1337–1340.
[69] B. Richman, Health regulation for the digital age — correcting the mismatch, N Engl. J. Med. 379 (2018) 1694–1695, https://doi.org/10.1056/nejmp1806848.
[70] S. Oh, J.H. Kim, S.W. Choi, H.J. Lee, J. Hong, S.H. Kwon, Physician confidence in artificial intelligence: an online mobile survey, J. Med. Internet Res. 21 (2019), e12422, https://doi.org/10.2196/12422.
[71] W.N. Price II, Artificial Intelligence in Health Care: Applications and Legal Issues by W. Nicholson Price II. The SciTech Lawyer, vol. 14, 2017.
[72] R.J. McDougall, Computer knows best? The need for value-flexibility in medical AI, J. Med. Ethics 45 (2019) 156–160, https://doi.org/10.1136/medethics-2018-105118.
[73] D.S. Char, N.H. Shah, D. Magnus, Implementing machine learning in health care — addressing ethical challenges, N Engl. J. Med. 378 (2018) 981–983, https://doi.org/10.1056/nejmp1714229.
[74] I. de Miguel, B. Sanz, G. Lazcoz, Machine learning in the EU health care context: exploring the ethical, legal and social issues, Inf. Commun. Soc. 23 (2020) 1139–1153, https://doi.org/10.1080/1369118X.2020.1719185.
[75] S. Triberti, I. Durosini, G. Pravettoni, A "third wheel" effect in health decision making involving artificial entities: a psychological perspective. Front Public Heal 2020;8:117. https://doi.org/10.3389/fpubh.2020.00117.
[76] O.O. Varlamov, D.A. Chuvikov, L.E. Adamova, M.A. Petrov, I.K. Zabolotskaya, T. N. Zhilina, Logical, philosophical and ethical aspects of AI in medicine, Int. J. Mach. Learn Comput. 9 (2019) 868–873. https://doi.org/10.18178/ijmlc.2019.9.6.885.
[77] B. Chin-Yee, R. Upshur, Three problems with big data and artificial intelligence in medicine, Perspect. Biol. Med. 62 (2019) 237–256, https://doi.org/10.1353/pbm.2019.0012.
[78] T. Hoeren, M. Niehoff, Artificial intelligence in medical diagnoses and the right to explanation, Eur. Data Prot. Law Rev. 4 (2018) 308–319. https://doi.org/10.21552/edpl/2018/3/9.
[79] J.J. Hatherley, Limits of trust in medical AI, J. Med. Ethics 46 (2020) 478–481, https://doi.org/10.1136/medethics-2019-105935.
[80] S. Mahomed, Healthcare, artificial intelligence and the Fourth Industrial Revolution: Ethical, social and legal considerations, South African, J. Bioeth Law 11 (2018) 93, https://doi.org/10.7196/SAJBL.2018.v11i2.664.
[81] M.C. Mehta, I.T. Katz, A.K. Jha, Transforming global health with AI, N Engl. J. Med. 382 (2020) 791–793, https://doi.org/10.1056/NEJMp1912079.
[82] WHO. Ethics and Governance of Artificial Intelligence for Health: WHO guidance. 2021.
[83] J. Guan, Artificial intelligence in healthcare and medicine: promises, ethical challenges and governance, Chinese Med. Sci. J. 34 (2019) 76–83. https://doi.org/10.24920/003611.