*Sociology and Social Work Review*
*Volume 9 (Issue 2)/2025* | pp. 7-26



DOI: 10.58179/SSWR9201
https://globalresearchpublishing.com/sswr/

*Original Article*

# Artificial intelligence as a social phenomenon: A bibliometric reading through the lenses of structural violence, intersectionality, and surveillance

Radu-Mihai Dumitrescu[a*], Adrian-Nicolae Dan[b]

[a] *University of Bucharest, Romania*
[b] *University of Bucharest, Romania*

**Abstract**
This study investigates how social vulnerabilities associated with the use of artificial intelligence (AI) in medicine are reflected in recent biomedical literature and how these patterns correlate with central theoretical directions in sociology and social work. Through a bibliometric analysis of 2,589 meta-analyses and systematic reviews published between 2020 and 2025 in PubMed, the research maps the conceptual structure of the field using co-occurrence networks at two thresholds (5 and 20). The results show a concentration of discussions on the technical and clinical aspects of AI (diagnosis, predictive modelling, electronic health records, large language models), while terms expressing social and ethical concerns (equity, algorithmic bias, privacy, ethics, health disparities, clinical competence) occupy semi-peripheral positions in the network. Interpreting these structures through theoretical lenses such as structural violence, social determinants of health, intersectionality, algorithmic oppression, surveillance capitalism, and care ethics reveals that AI risks reproducing and intensifying pre-existing inequalities. The analysis emphasises that algorithmic bias, unequal data infrastructures, model opacity, and changes in the distribution of clinical work are not isolated phenomena, but manifestations of broader social processes that shape vulnerability and exclusion. Therefore, the study argues for the need to integrate sociological and social work perspectives into the development and evaluation of medical AI and advocates for interdisciplinary approaches that place equity, transparency, and the experiences of marginalised populations at the centre. Such an orientation is essential for AI in medicine to contribute to reducing — rather than amplifying — social inequalities in health.

***Keywords***: *artificial intelligence, health equity, algorithmic bias, structural violence, social determinants of health, intersectionality, surveillance capitalism, ethics of care, digital health, bibliometric analysis, medical sociology, social work and health disparities.*

[*]Corresponding author. Radu-Mihai Dumitrescu. Tel.:004-0722-776-783; *E-mail address: dum_mihu@yahoo.com*

## Introduction

Artificial intelligence promises efficiency, faster diagnosis and expanded access to medical services. However, if implemented without caution, AI can exacerbate inequalities, create risks to patient safety and erode public trust. The World Health Organisation (WHO) and key regulators (EU, FDA) explicitly call for governance, transparency and rigorous assessments throughout the life cycle of systems, precisely to prevent adverse social effects (World Health Organisation, 2025).

The social vulnerabilities identified in the application of AI in medicine can be directly anchored in several major theoretical traditions in sociology, social work, and the humanities. The idea that AI models amplify health disparities and structural disadvantages refers to the concept of structural violence, whereby political and economic structures located "far" from the clinic systematically produce illness and avoidable death for certain groups (Farmer, et al., 2006b), including through insufficient or selective medical infrastructure. This ties in with the framework of social determinants of health, which shows that the distribution of disease follows the distribution of social resources (income, education, work, housing), and that technological policies (including AI) can either reduce or deepen these inequalities (Marmot, 2005; Serban, 2025). The fact that AI systems perform differently on the basis of race, gender, age or class can be understood through the lens of intersectionality (Kimberlé W. Crenshaw, 1991), which explains why the effects of a technology cannot be understood separately on isolated "axes" (race or gender), but at their intersection, where disadvantages accumulate (Kimberlé W. Crenshaw, 1989). Works on algorithmic discrimination and the "digital dragnet" (Eubanks, 2018; Noble, 2019) show that data and automatic scoring infrastructures tend to monitor, profile and penalise poor, racialised or already marginalised people in particular, continuing old logics of social control in the form of a "New Jim Code" (algorithms that are apparently neutral but anchored in histories of racism and poverty) (Benjamin, 2019). AI systems tend, at least at this point, to use stereotypes because of biases in training data and algorithms. These biases manifest themselves in various personnel recruitment tools, image generation and decision-making processes, perpetuating pre-existing stereotypes in the real world. Real-world cases highlight the seriousness of the problem and the ongoing legal challenges, some of which are even more sensitive in the case of medical practice. In the same way, gender gaps can also be perpetuated. The dimensions of confidentiality, consent and digital surveillance in medical AI resonate with the analysis of surveillance capitalism, in which human experience is treated as raw material for data extraction and commercial predictions, with risks of expropriation of autonomy and exploitation of vulnerable groups (Zuboff, 2019). Concerns about the doctor-patient relationship, care work and the deskilling of professionals can be read through the lens of care ethics, which insists that care is a deeply relational activity, unevenly distributed across social and gender groups. introducing AI without paying attention to "who bears the burden of caring for and supervising algorithms" risks reproducing the same devaluation of care work that critical ethics criticises (Tronto, 1993). In this way, bibliometric maps can be read not only as keyword structures, but as meeting points between medical AI and the major theories of inequality, power, surveillance and care in the social sciences.

Recent academic literature indexed in PubMed on the application of artificial intelligence in medicine highlights a coherent set of social vulnerabilities that recur in meta-analyses and systematic reviews published in recent years. The first are vulnerabilities generated by algorithmic inequalities, which include the perpetuation of demographic biases and the differentiated performance of models across social, racial, age or gender groups. A second category concerns decision dependency and "automation bias", whereby AI models can reorient clinical decisions in a way that is difficult to challenge or oversee, with disproportionate effects on vulnerable patients. At the same time, the literature points to systemic risks related to confidentiality, data reuse and digital surveillance, especially with the expansion of generative models and access to massive sets of sensitive data. Other vulnerabilities stem from the opacity of AI models and the phenomenon of hallucinations, which complicate professional accountability and can produce errors that are difficult to detect in practice. A significant body of work analyses the impact of AI on clinical work, on the autonomy of professionals and on the quality of the doctor-patient relationship. These types of vulnerabilities appear consistently in the research synthesised in PubMed and constitute the analytical framework for the further interpretation of bibliometric results.

**Algorithmic inequality and the perpetuation of structural biases**

Models trained on historical clinical data or general texts can learn and amplify existing inequalities, affecting underrepresented patient groups in particular. Recent literature shows demographic biases in both clinical decision support systems and uses of LLMs for mental health: lower accuracy for patients of colour, different recommendations based on racial criteria, and the perpetuation of racist medical myths in general chatbot responses (Cross, Choma, & Onofrey, 2024). The social consequences are significant: unequal access to diagnosis, suboptimal triage, and erosion of trust in the system for already disadvantaged communities. The WHO emphasises that multimodal models „can improve health only if risks are identified and managed to overcome persistent inequities" (World Health Organization, 2024).

**Overinvestment in AI and the risk of "automation bias"**

When an AI system makes a recommendation, clinicians tend to follow it even when it contradicts clinical evidence, an effect called automation bias. Studies from 2024–2025 show that non-specialist doctors are more vulnerable, and AI assistance in chest pain triage can alter decisions in a way that accentuates demographic differences. Socially, this means unequal distribution of time and resources in emergency departments (Kücking et al., 2024). Cases of detection failures (e.g., sepsis models) show that poor performance, unrecognised in time, can persist in practice until independent evaluations, with high social costs due to delays in treatment (Papareddy et al., 2025).

**Security, confidentiality and consent for data reuse**

Medical AI relies on vast sets of sensitive data. Risks include re-identification, data triangulation, secondary sharing, and lack of patient control over future uses of data (including for training LLMs). Recent studies map global challenges (such as GDPR, CCPA) and highlight real barriers to obtaining informed consent for secondary uses of data in AI. Socially, perceptions of "digital surveillance" can decrease healthcare attendance and accentuate distrust (Conduah, Ofoe, & Siaw-Marfo, 2025).

**Opaque design, "hallucinations" and diffuse responsibility**

LLMs can produce plausible, but erroneous ("hallucinations"), including under adversarial attacks (a false or misleading output generated by an AI model, caused by an adversarial input, i.e. a question or image specifically designed to exploit weaknesses in the model and cause it to produce incorrect information); in clinical settings, these translate into direct risks for patients and diffuse responsibility between the provider, hospital, and developer. Research from 2025 proposes safety assessment frameworks and shows multimodel vulnerabilities to adversarial hallucinations, highlighting the need for human verification and traceability. Socially, patients may be disproportionately affected where resources for a "second pair of eyes" are lacking (Asgari et al., 2025).

**Impact on the doctor–patient relationship and clinical work**

The integration of AI may change the roles of professionals, increase oversight and audit work and risk long-term "deskilling" (less practice for rare skills). The WHO calls for governance that protects clinician autonomy and ensures context-appropriate disclosure; otherwise, professional agency and the quality of patient interaction may be eroded (World Health Organisation, 2025).

**Governance and regulation: what the "latest" frameworks say**

In the European Union, the AI Act considers AI systems that are part of medical devices to be high-risk, with strict requirements for risk management, data quality, human oversight, and post-market monitoring; recent technical documents clarify the interaction with the medical device regime. Socially, this may reduce discrepancies between hospitals through common minimum standards (Aboy, Minssen, & Vayena, 2024). In the US, from 2024–2025, the FDA issued (and updated) guidance on transparency, predetermined change plans (PCCPs), and full life cycle management for AI/ML software devices. The emphasis on transparency and controlled updates is crucial for maintaining public trust (Food and Drug Administration, 2025). The WHO, with regard to LMM in health, recommends pre- and post-implementation assessments, documentation of limitations, bias audits, and the involvement of affected parties (including vulnerable communities) in design (World Health Organization, 2024).

**Recommendations for mitigating social risks**

A series of recommendations to mitigate the risks of including AI in medicine have begun to appear in academic medical literature and in various documents from international organisations. Patient-centred data governance is envisaged, through clear policies on consent for reuse, opt-out options and the introduction of data usage logs and re-identification audits (Conduah et al., 2025). Equity by design is a trend that requires diverse data sets, standardised performance reporting by subgroups (gender, age, ethnicity, socio-economic status) and continuous real-life monitoring (Cross et al., 2024). Automation bias can be controlled through interfaces that display uncertainty, the requirement for independent clinical justification, and regular staff training (Kücking et al., 2024). Safety assessment for LLM will need to include hallucination testing, red teaming, and double-check usage policies for high-risk tasks (Asgari et al., 2025). When we talk about transparency and traceability, we refer to compliance with AI Act/FDA requirements for documentation, PCCP, and post-market surveillance, but also to public reporting of incidents (Aboy et al., 2024). Last but not least, when we refer to inclusion and co-design, we must move towards involving affected communities in defining the objectives of models and success metrics, in order to prevent solutions that "work" technically but produce injustices (World Health Organization, 2024).

AI in medicine can be an accelerator of equity or, conversely, a multiplier of inequities. Recent data show concrete risks, from bias and automation bias to confidentiality and hallucinations, but also an emerging framework of solutions: strict governance (WHO, EU, FDA), robust evaluations, and equity-centred design. The social direction depends on how institutions and developers translate these standards into clinical practice. Observing how these vulnerabilities are reflected and constructed in academic medical language can provide valuable information, and by repeating the questions, we can identify emerging issues or trends, observe the inclusion of AI systems in routine practice, and see how a whole range of inequalities are managed, resolved, or even accentuated.

**Objectives**

The aim of this research is to explore how the social vulnerabilities of applying artificial intelligence in medicine are articulated in recent academic medical literature, through a bibliometric analysis of meta-analyses and systematic reviews from 2020–2025. We aim to map the thematic landscape of AI applications in medicine by analysing the co-occurrence of keywords and identifying major clusters of concepts (clinical, technical, social, and ethical) in the synthesis literature published in the last five years. By identifying and describing how social vulnerabilities (such as inequality and health disparities, algorithmic bias, data privacy and security, impact on clinical work and the doctor-patient relationship) we want to see how these are reflected in thematic clusters and connected to technical nodes (artificial intelligence, machine learning, large language models, etc.). Comparing conceptual structures at two co-occurrence thresholds (5 and 20 occurrences) serves to distinguish between emerging or niche themes and the stable conceptual core of the field, and to assess the extent to which social-ethical terms (such as *"equity"*, *"bias"*, *"privacy"*, *"ethics"*, *"governance"*) are integrated into the mainstream of medical AI research. The analysis of the positioning and connectivity of social-ethical terms in the network aims to assess whether the discourse on social vulnerabilities (equity, transparency, trust, governance) is organically integrated into applied research, or remains relatively peripheral and segmented from the technical-clinical core.

**Methodology**

To investigate the social vulnerabilities of applying artificial intelligence in medicine, we conducted a PubMed database query; the query combined terms for artificial intelligence, terms for medicine and healthcare, terms for social/ethical/equity dimensions, and was limited to recent publications (from the last 5 years).

The query formula used is shown below:
(("artificial intelligence"[MeSH Terms] OR "artificial intelligence"[Title/Abstract] OR "machine learning"[Title/Abstract] OR "large language model*"[Title/Abstract] OR "deep learning"[Title/Abstract]) AND ("medicine"[MeSH Terms] OR "healthcare"[Title/Abstract] OR "clinical practice"[Title/Abstract] OR "medical applications"[Title/Abstract]) AND ("social vulnerability"[Title/Abstract] OR "health equity"[MeSH Terms] OR "inequity"[Title/Abstract] OR "bias"[Title/Abstract] OR "ethics"[MeSH Terms] OR "ethical"[Title/Abstract] OR "governance"[Title/Abstract] OR "privacy"[Title/Abstract] OR "consent"[Title/Abstract] OR "automation bias"[Title/Abstract] OR "trust"[Title/Abstract] OR "public perception"[Title/Abstract])) AND ("2019/01/01"[Date - Publication] : "3000"[Date - Publication])

The query yielded a total of 2,701 results; for the period 2020-2025, 2,589 results were filtered by meta-analysis, review, scoping/systematic review. In these papers, a total of 5,287 keywords (full count) were identified, 563 at a co-occurrence threshold of 5 and 120 at a threshold of 20. For these two thresholds, bibliometric maps (VOSviewer) of keyword co-occurrence were created.

## Results

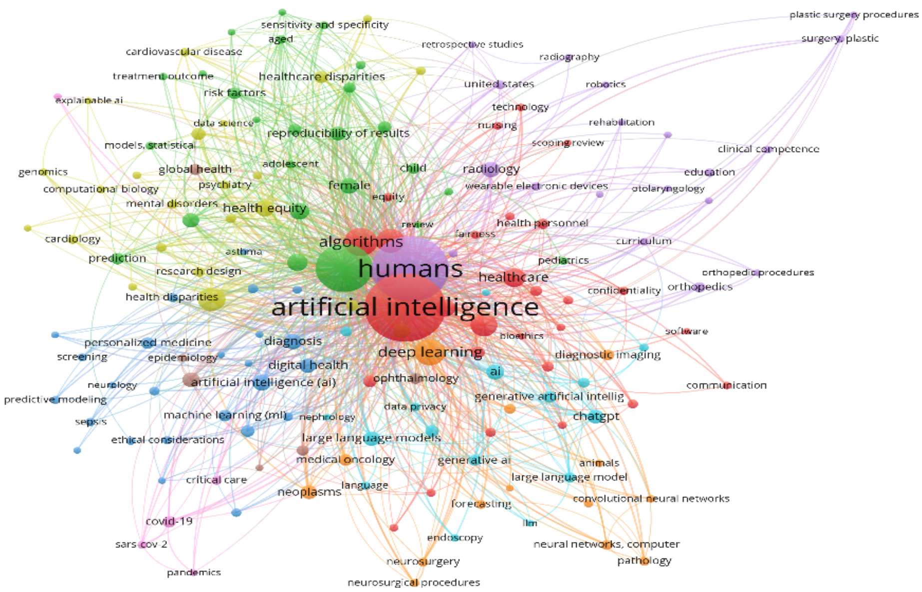
### Bibliometric analysis of the literature on AI in medicine (2020–2025) - Context and method of analysis

This analysis is based on bibliometric maps generated with VOSviewer, using the co-occurrence of keywords from 2,589 articles published between 2020 and 2025. The dataset exclusively includes meta-analyses, systematic reviews, and scoping reviews focused on the application of artificial intelligence (AI) in medicine, with an emphasis on social and ethical dimensions.

The PubMed query included terms such as *"artificial intelligence"*, *"healthcare"*, *"ethics"*, *"equity"*, *"bias"*, *"trust"*, *"privacy",* etc., to highlight concerns related to the social aspects of AI in health. A total of 5,287 terms (full count) were identified, of which 563 have at least 5 occurrences, and 120 keywords exceed the threshold of 20 occurrences. The co-occurrence analysis of these terms revealed the thematic structure of recent literature, highlighting clusters of topics and the links between them.

1. **Thematic clusters identified in the co-occurrence map at a co-occurrence threshold of 5**

The bibliometric map at the threshold of 5 (Figure 1) suggests the existence of several distinct thematic clusters, reflecting the main directions in which AI has been applied in medicine in recent years. Each cluster groups closely related terms, indicated by the same colour on the VOSviewer map. Based on similar literature and frequent keywords in the sample, the composition of the dominant clusters and the links between them can be interpreted.




• *Figure 1: The expanded landscape of artificial intelligence in medicine: themes, conceptual networks, and emerging vulnerabilities*

- **General structure.** The largest nodes, "*humans"* and "*artificial intelligence"*, followed by *"deep learning"*, *"algorithms"*, *"healthcare"* and *"health equity"*, show that the discussion about social vulnerabilities is anchored simultaneously in people, technology and equity. From the centre, links extend to specialised clusters: *"health disparities"*, *"health equity"*, *"digital health"*, *"personalised medicine"*, *"predictive modelling"*, *"covid-19"*, *"critical care"*, *"large language models"*, *"chatgpt"*, *"data privacy"*, *"education"*, *"clinical competence",* and *"nursing"*.
- **Red cluster: clinical AI core and governance.** The main (large/central) nodes are *"artificial intelligence"*, *"deep learning"*, *"algorithms"*, *"healthcare"*, *"humans"* (partially green, but strongly connected here); social-normative terms are directly linked: *"fairness", "equity", "bioethics", "confidentiality"*. Organisational and implementation terms are *"nursing", "health personnel", "technology", "software", "diagnostic imaging", "wearable electronic devices",* and *"scoping review"*. This is the cluster in which AI is explicitly anchored in *healthcare* and in the actual work *of health personnel* and *nursing*. The simultaneous presence of *"fairness", "equity", "bioethics",* and *"confidentiality"* in the same cluster as *"algorithms"* and *"software"* suggests that the literature treats social vulnerabilities as part of the design and implementation of systems. We can identify links to social vulnerabilities through the presence of *"fairness", "equity",* and *"healthcare",* which reflect concerns that algorithms may generate different treatment for different groups; here, they are directly linked to *"algorithms"* and specialties such as *"radiology"* and *"diagnostic imaging"*. *"Confidentiality"* and its connections to *"software", "wearable electronic devices"*, and *"digital health"* (through edges that cross over to the blue cluster) show the dimension related to confidentiality and data security. The nodes *"nursing", "health personnel",* and *"scoping review"* suggest that the literature discusses the impact on clinical work (task redistribution, need for supervision).
- **Green cluster*:* equity, populations and outcomes.** The main nodes are *"health equity"* (large, highly connected node), *"healthcare disparities"* and *"health disparities"*. The demographic variables are *"female"*, *"child"*, *"adolescent"* and *"aged"*. We find terms related to methods and outcomes, such as *"treatment outcome"*, *"risk factors", "reproducibility of results", "prediction", "sensitivity and specificity"* and *"research design"*; diseases and clinical areas are represented by *"cardiovascular disease", "asthma", "psychiatry", "mental disorders",* and *"global health"*. The cluster groups dimensions of *"health equity"* and *"disparities"* with demographic variables and method terms. The strong connection to *"humans"* and *"artificial intelligence"* indicates that equity studies are central, not peripheral. Social vulnerabilities are also reflected in various forms. The combination *of "health equity" – "health disparities" – "healthcare disparities"* with *"female"*, *"child'*, *"adolescent"* and *"aged"* suggests a focus on the differentiated performance of algorithms between age groups and genders. Links to *"cardiology"*, *"cardiovascular disease"*, *"asthma"*, *"psychiatry",* and *"mental disorders"* show that equity is being discussed in specific pathologies, where a model may triage or diagnose differently. *"Reproducibility of results"* and *"research design"* related to *"health equity"* suggest concern about the lack of reproducibility that can accentuate *"healthcare disparities"*.

- **Yellow cluster: data science, genomics and disparities.** The nodes are represented by *"data science"*, *"models"* and *"statistical"*, which are related to scientific fields such as *"genomics"* and *"computational biology"* and to socio-clinical terms such as *"global health", "health disparities", "mental disorders", "psychiatry"* and *"cardiology"*. In terms of content, the cluster links the area of *"genomics"* – *"computational biology"* and *"data science"* to social terms such as *"global health"* and *"health disparities*". It is placed between the green cluster (equity) and the blue cluster (clinical applications), acting as a transition area. The link to social vulnerabilities is found in the combination of *"genomics"* – *"health disparities"* – *"global health"*, which suggests a concern for how molecular data sets (often dominated by certain populations) can perpetuate global disparities. Links to *"mental disorders"* and *"psychiatry"* indicate that social vulnerabilities also arise in the field of mental health, where *data science* can be used to predict or classify patients.
- **The blue cluster: screening, predictive modelling and clinical specialties.** The main nodes are *"digital health"*, *"diagnosis"*, *"artificial intelligence (AI)"* and *"machine learning (ML)"*, which are related to *"personalised medicine"*, *"predictive modelling"* and *"screening*". Several medical specialties are evident, namely *"neurology"*, *"nephrology"*, *"sepsis"*, *"critical care",* and *"epidemiology"*, in connection with a social-normative term: *"ethical considerations"*. This is a cluster oriented towards concrete clinical applications of AI: *"screening"*, *"diagnosis"*, *"predictive modelling"* for *"sepsis"*, *"critical care"*, *"neurology"* and *"nephrology"*. *"Digital health"* and *"personalized medicine"* connect these applications to broader digital infrastructures. Social vulnerabilities are highlighted by the presence of *"ethical considerations"* in the same cluster as *"predictive modelling"*, *"screening",* and *"critical care"*, suggesting discussions about the effects of automated decisions in critical situations, about who receives treatment or screening. The strong links between *"personalised medicine"*, *"predictive modelling",* and equity nodes (*"health equity"*, through connections to the green cluster) indicate concerns that personalisation based on *"machine learning (ML)"* may amplify *"health disparities"*.
- **Purple cluster: pandemics, COVID-19 and critical care.** The main nodes, *"COVID-19"*, *"SARS-CoV-2"* and *"pandemics"*, are strongly connected to *"critical care"*, *"sepsis"*, *"ethical considerations"* and *"digital health*". This is where AI intersects with *"pandemics"* and *"critical care"*, suggesting the use of algorithms for triage, prognosis, or resource allocation during COVID-19. The link to *"ethical considerations"* reinforces the idea that, in the context of *"pandemics"*, social vulnerabilities related to limited resources and *"health disparities"* are centrally discussed.
- **The turquoise cluster: large language models, data privacy, and chatgpt.** The main nodes, *"large language models"*, *"language models"*, *"language"*, *"generative ai"*, *"ai"*, and *"chatgpt"*, are connected to social protection nodes, such as *"data privacy"*, and have clinical links: *"endoscopy"*, *"neurosurgery"*, *"forecasting"*, *"animals"*, *"neural networks"*, *"computer"* and *"pathology"*. The cluster is centred on *"large language models"* and *"chatgpt"*, connected to *"data privacy"* and clinical terms (e.g. *"endoscopy"*, *"pathology"* and *"forecasting"*). It is closely linked to the red cluster through *"artificial intelligence"*, *"deep*

*learning"*, and *"healthcare",* and to the brown cluster through *"medical oncology"* and *"ophthalmology"*. Social vulnerabilities are revealed by the fact that *"data privacy"* is within the same cluster as *"chatgpt"*, *"generative ai"*, and *"language models",* showing recognition of the tension between the use of LLM and data protection. Links to *"endoscopy"*, *"neurosurgery"*, and *"pathology"* suggest concerns about the use of LLMs and *"generative AI"* in high-risk specialties, where language errors or „hallucinations" can directly affect patients. The connection to *"forecasting"* suggests the role of generative AI in clinical predictions, which can influence resource allocation and, implicitly, *"equity"*.

- **Brown cluster: oncology and ophthalmology specialties in the LLM era.** The nodes *"medical oncology"*, *"neoplasms"*, *"ophthalmology"*, *"neurosurgery"* and *"neurosurgical procedures"* are linked to IA terms: *"large language models"*, *"deep learning"* and *"forecasting"*. The cluster shows a focus on AI applications (including large language models) in various medical specialties. Social vulnerabilities here relate to unequal access to advanced technologies; these specialties are often concentrated in large centres. Links to *"health equity"* and *"global health"* (through edges originating from central nodes) suggest discussions about *"healthcare disparities"* in the treatment of *"neoplasms"*.
- **The education–competence–surgical specialties cluster (purple to the right).** The nodes *"education"*, *"curriculum",* and *"clinical competence"* are connected to various specialties: *"orthopaedics"*, *"orthopaedic procedures"*, *"otolaryngology"*, *"plastic surgery procedures"*, *"plastic surgery"*, *"rehabilitation"*, *"radiology"*, *"nursing"*, *"health personnel"*, *"technology"* and *"wearable electronic devices"*. The cluster explicitly links *"education"* and *"curriculum"* to *"clinical competence"* in various surgical specialties and to *"technology" – "wearable electronic devices"*. It is connected to the central nodes *"artificial intelligence"*, *"healthcare"* and *"algorithms"*. In terms of social vulnerabilities, the link between *"education" – "curriculum" – "clinical competence" – "technology"* suggests concern about the unequal training of *"health personnel"* and *"nursing"* in the use of AI. This leads to differences in *"clinical competence"*, and therefore to *"healthcare disparities"* between hospitals or regions. The presence of procedural specialities (*"orthopaedic procedures"*, *"plastic surgery procedures"*) indicates the risk of deskilling or dependence on AI in surgical decisions.
- **Links between clusters and the overall picture of social vulnerabilities. Equity and disparities.** The terms *"health equity"*, *"equity"*, *"fairness"*, *"health disparities"*, *"healthcare disparities",* and *"global health"* are widespread among the green, yellow, and red clusters. Their strong connections to *"artificial intelligence"*, *"deep learning"*, *"digital health",* and *"personalised medicine"* show that inequalities are discussed in direct relation to models and clinical applications. **Confidentiality and data.** *"Confidentiality"* (red) is strongly connected to *"healthcare"*, *"diagnostic imaging"*, *"software",* and the LLM cluster with *"data privacy"*. From there, it links to *"digital health"*, *"wearable electronic devices"*, and *"large language models"*, suggesting vulnerabilities related to extensive data collection and its reuse in *"language models"* and *"chatgpt"*. **Ethics and bioethics.** *"Bioethics"* (red) is close to *"deep learning"* and *"diagnostic imaging"*; *"ethical considerations"* (blue) appears next *to*

*"sepsis"*, *"critical care"*, and *"predictive modelling"*. Ethics is placed precisely in high-risk areas: critical care, pandemics, diagnostic imaging. **Clinical work and competence.** The terms *"nursing"*, *"health personnel", "education"*, *"curriculum"*, and *"clinical competence"* are positioned at the intersection of the red and purple clusters. Social vulnerabilities also refer to how AI is changing the roles of professionals and the skills required, with the risk of differences between groups of *"health personnel"*. **Crises and special contexts.** The cluster *"covid-19"* / *"pandemics"* / *"critical care"* / *"sepsis"* is connected to *"ethical considerations"* and *"digital health"*, and the central nodes show that crisis situations highlight inequalities and ethical tensions in the use of AI. **Emerging technologies – LLM and generative AI.** The cluster *"large language models"* / *"generative AI"* / *"chatgpt"* / *"data privacy"* is closely linked to *"deep learning"*, *"healthcare"*, but also to specialties such as *"endoscopy"*, *"ophthalmology"*, *"medical oncology"*. Recent social vulnerabilities (hallucinations, confidentiality, impact on communication with *"humans"*) are incorporated into the network, not separated.

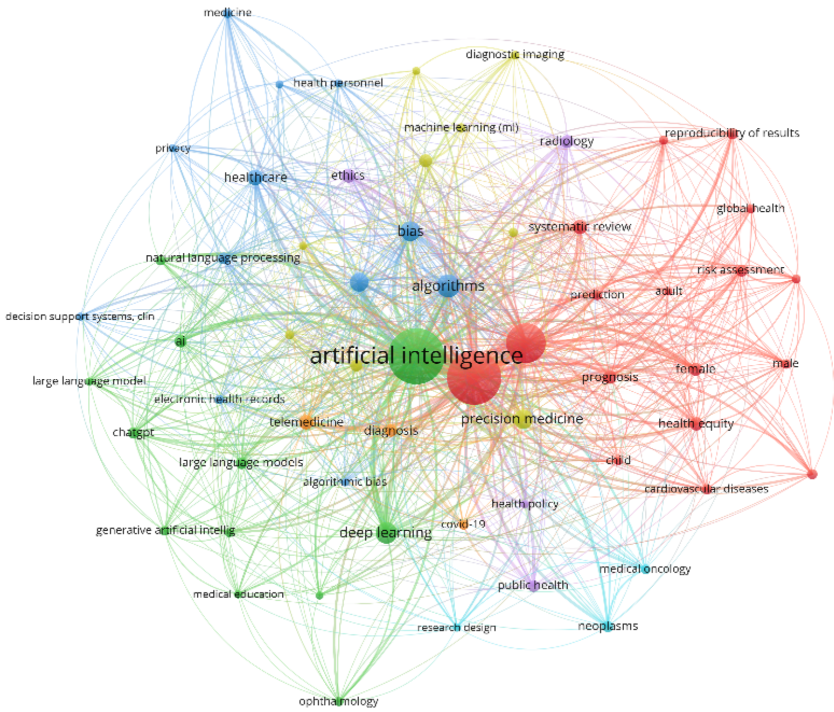
The detailed map of terms shows that social vulnerabilities (equity, *"health disparities"*, "*confidentiality"*, *"data privacy"*, *"bioethics"*, *"ethical considerations"*, *"clinical competence"*) are intertwined and scattered around technical nodes (*"artificial intelligence"*, *"deep learning"*, *"large language models"*, *"algorithms"*) and clinical specialties. *"Health equity"*, *"fairness",* and *"equity"* are directly connected either to demographic variables (*"female"*, *"child"*, *"adolescent"*, *"aged"*) or to domains (*"global health"*, *"healthcare disparities"*, *"cardiology"*, *"mental disorders"*), which supports the idea of algorithmic inequalities and *"automation bias"*. *"Confidentiality"* and *"data privacy"* are linked to *"digital health"*, *"wearable electronic devices"*, *"large language models"* and *"chatgpt"*, which supports the discussion on consent, digital surveillance and data reuse. The nodes related to *"education"*, *"curriculum"*, *"clinical competence"*, *"nursing"*, and *"health personnel"* show concern for how AI is changing the doctor-patient relationship and professional work. In other words, just by looking at the words in the network, the size of the nodes and the connections, it is clear that the literature on AI in medicine articulates social vulnerabilities along three main axes: equity/disparities, confidentiality/data privacy, and clinical work/competence, all in direct contact with core technologies (*"deep learning"*, *"large language models"*, *"chatgpt"*, *"digital health"*, *"personalised medicine"*).

## 2. Thematic clusters identified in the co-occurrence map at a co-occurrence threshold of 20

The co-occurrence map at threshold 20 (Figure 2) represents the „hard skeleton" of the field, i.e. those concepts that not only appear frequently in the literature on artificial intelligence in medicine, but are sufficiently central that they tend to be associated with each other repeatedly and systematically. While the threshold 5 map provides a very rich picture with many ramifications, including emerging or niche topics, the threshold 20 map reduces complexity and brings to the fore the stable conceptual structure of the domain. In practice, we move from a granular, highly detailed picture to an 'epistemic core' of the literature, where it becomes clearer which themes are truly dominant and how academic discourse is organised around them.

Compared to the threshold 5 map, where social vulnerabilities – *„bias"*, *„equity"*, *„privacy"*, *„disparities"* – were scattered across several peripheral and intermediate clusters, the threshold 20 map allows for a more accurate identification of how these concepts are integrated into the mainstream of the discussion on clinical AI. The remaining nodes have passed the frequency and relevance filter, so the relationships between them have stronger conceptual significance: if two ideas constantly co-appear in the analysis at a high threshold, it means that the literature considers them structurally related, not incidental.

In addition, the threshold 20 map helps to understand conceptual „polarisation": which themes cluster around artificial intelligence, which themes are located on the periphery, which relationships are robust enough to pass the strict co-occurrence filter. In the context of social vulnerabilities, this map shows not only where terms such as *„bias"*, *„health equity"*, or *„privacy"* appear, but also how central their role is in the conceptual architecture of AI in medicine. Thus, the threshold 20 analysis not only simplifies the map, but also reveals where the „heavy nodes" of the discussion on social risks are and, by absence, which topics are not yet sufficiently consolidated in the current literature.




- *Figure 2: The conceptual core of artificial intelligence in medicine: robust relationships and dominant themes*

- **Red cluster: equity, demographics, and clinical outcomes.** The main nodes are *"female"*, *"male"*, *"adult"* and *"child"*, connected to *"health equity"*, *"cardiovascular diseases"*, *"global health"*, *"risk assessment"*, *"prediction"*, *"prognosis"*, *"reproducibility of results"* and *"systematic review"*. Structurally, the cluster is very dense, with close links between demographic variables (*"female"*, *"male"*, *"adult"*, *"child"*) and terms related to model performance and validation (*"prediction"*, *"prognosis"*, *"reproducibility of results"*). *"Health equity"* appears integrated into this core, not marginal, indicating that the literature treats equity as a constituent part of AI evaluation. Strong connections to *"cardiovascular diseases"* and *"global health"* show that demographic differences are discussed in pathologies and contexts with a high population impact. Implications for social vulnerabilities become evident; the relationships between *"female", "male", "adult", "child",* and *"health equity"* show that algorithmic inequalities are conceptually anchored in demographic differences. *"Reproducibility of results"* is connected to *"prediction"* and *"risk assessment"*, suggesting that lack of reproducibility is perceived as a structural factor of inequity. The link to *"global health"* indicates concern for differences between health systems and populations, not just between individuals.

- **Green cluster: LLM, generative AI, clinical data, and digital infrastructure.** The main nodes are: *"large language models"*, *"natural language processing"*, *"chatgpt",* connected to *"generative artificial intelligence"*, *"AI"*, *"deep learning", "electronic health records"*, *"telemedicine"*, *"diagnosis", "algorithmic bias"*, *"decision support systems"* and *"clinical"*. Structural observations may be related to the fact that this cluster is centred on recent technologies (*"LLM"*, *"chatgpt"*, *"generative AI"*), which shows the maturation of technological discourse in medical literature. The connection of these nodes to *"electronic health records"*, *"diagnosis"* and *"decision support systems"* suggests the integration of LLMs into concrete clinical processes. *"Algorithmic bias"* appears within the cluster, indicating that bias is discussed directly in the context of digital infrastructure and automated clinical systems. *"Algorithmic bias"* connected to *"decision support systems"* and *"diagnosis"* indicates concern about automated errors affecting patient triage and assessment. The strong connection between *"electronic health records"* and LLMs suggests concerns about data reuse and the potential for amplifying existing bias in EHRs. *"Telemedicine"* linked to LLMs and AI shows the discussion about unequal access to technology and digital health services.
- **Blue cluster: clinical medicine, ethics, and healthcare professionals.** The main nodes are: *"healthcare"*, *"medicine"*, *"health personnel",* connected to *"ethics"*, *"privacy", "diagnostic imaging"*, *"radiology"*, and *"machine learning (ml)".* This is the cluster where terms describing the broader clinical environment, medical professionals and ethical tensions appear. *"Health personnel"* is simultaneously connected to *"medicine"*, *"healthcare"* and *"ethics"*, suggesting that the literature analyses the impact of AI on clinical work and professional responsibility. *"Privacy"* is positioned in the cluster but oriented towards technical nodes, indicating that data protection is perceived as a structural problem of digital clinical systems. From a social implications' perspective, the connections between *"privacy"*, *"health personnel",* and *"medicine"* show that data vulnerabilities are understood as part of everyday clinical work, not as a purely technical element. The presence *of "ethics"* alongside *"machine learning (ml)"* and *"diagnostic imaging"* suggests concern about the use of AI in high-risk procedures. The link to *"radiology"* reflects the specialty in which AI is already widely implemented, and therefore where tensions related to responsibility and quality are most evident.

- **The purple–turquoise cluster (intermediate): oncology, public health and pandemics.** The visible nodes are represented by *"medical oncology"*, *"neoplasms"*, *"public health"*, *"health policy"*, *"covid-19"* and *"research design"*. This cluster connects two sub-themes: cancer (where AI is very active) and public health (where AI is used in policy, prediction, and epidemiological surveillance). *"Covid-19"* and *"research design"* are transition nodes, suggesting the role of the pandemic in accelerating the use of AI and the methodological re-evaluation of its tools. The presence of *"health policy"* and *"public health"* indicates the systemic discussion about the impact of AI at the population level. *"Covid-19"* remains a central example of a context in which AI can create or amplify inequalities, depending on the quality of the models.
- **Yellow cluster: technical validation and predictive performance.** The visible nodes are *"algorithms"*, *"machine learning (ml)", "bias", "systematic review"* (connected to the red cluster), *"reproducibility of results", "diagnosis"*, *"prediction"*. Although small in number of words, the cluster brings together terms that are essential to the discussion of vulnerabilities: *"bias"*, reproducibility, algorithmic performance. The simultaneous connection to the red cluster (*"fairness"*), the green cluster (*"LLM"*), and the blue cluster (*"ethics"*) shows that *„bias"* is a bridge node between technical, social, and clinical topics. *"Bias"* is not marginal but positioned almost centrally, indicating that the issue of algorithmic inequities is recognised as fundamental, not secondary. *"Prediction"* and *"diagnosis"* are simultaneously linked to *"algorithms"* and demographics, clearly suggesting that the literature discusses the differentiated performance of AI for different groups.

*"Health equity"*, *"bias",* and *"privacy"* are integrated into the central clusters of the map, not on the periphery; social vulnerabilities are considered a structural part of the discussion about AI in medicine. The red cluster shows a focus on demographics and clinical outcomes, the green cluster on emerging technologies and digital infrastructure, blue on clinical practice and ethics, and yellow on technical validation and reproducibility. The links between these clusters outline the chain of social vulnerabilities: data → models → equity → clinical practice → population outcomes. The threshold 20 map confirms what the threshold 5 map showed diffusely: social vulnerabilities are omnipresent and articulated in the fundamental thematic cores of AI research in medicine.

## Discussions

### The position and connectivity of social-ethical terms in the network

Analysing the positions of nodes representing social and ethical terms (such as *"ethics", "bias", "equity", "trust", "privacy",* etc.), it can be observed that these concepts are often on the edge of the network or grouped in a dedicated cluster, rather than scattered centrally among technical terms. Terms such as *"privacy"* and *"security"* form a well-connected internal cluster (focused on data security), but their links to clinical or technical clusters are limited to a few connections (e.g., *"privacy"* with *"data/big data"*, *"security"* with *"IoT"*). Similarly, *"ethics"* and *"governance"* appear connected to each other and to terms such as *"policy"*, but less so to *"deep learning"* or *"radiology"*, suggesting that the ethical discussion takes place in a somewhat parallel framework to applied research. This peripheral nature is confirmed by the lower weight of these nodes: in bibliometric examples, *"ethics"* had a much lower link score than the central AI nodes (Torun, 2022), and *"political economics"* or *"governance"* appear as isolated or secondary nodes. *"Equity"* and algorithmic bias are terms that are present but not among the 20 most frequent; however, their appearance above the threshold of 20 suggests growing attention. *"Equity"* is often discussed in the context of fairness in access to technology and the impact on health disparities, but in the network, it may be relatively far from the technological core, closer to terms such as *"disparities"* or *"public health"*. Bias, on the other hand, appears to be more closely connected to technical language; for example, the phrase *"algorithmic bias"* links the concept of bias to *"machine learning"*, indicating awareness in the technical community of the problem of algorithmic bias. Thus, *"bias"* acts partly as a bridge between the technological and ethical clusters; it is a technical subject (mitigating bias in models) with social implications (inequity).

Other social terms have specific connections with clinical topics. *"Trust"*, for example, links discussions about the acceptance of AI by medical staff and patients with the need for transparency and explainability of models. The word *"transparency"* connects with both *"ethics/accountability"* and *"explainable AI (XAI)"* in the technology cluster. The presence of these links suggests that, although social-ethical terms are largely grouped separately, there are interactions between technical and ethical discourse. However, the intensity of these interactions is low; the maps show that elements of ethics and social responsibility have, on the whole, a lower degree of connectivity, indicating partial rather than full integration into mainstream medical AI research. In the literature as a whole, it has been noted that most research has focused on AI performance, and aspects of fairness, trustworthiness, legality, and ethics have received attention but remain secondary to (Steerling et al., 2023). This reality is also reflected on the map: ethics and fairness issues do not appear as central themes, but as complementary components.

An indication of the peripheral position of these themes is also given by the language used in the articles analysed. Many ethical terms (e.g. *"autonomy", "beneficence", "justice"*, *"accountability"*) appear in titles or conceptual discussions rather than as dominant keywords. Similarly, terms such as *"guidelines"*, *"regulation",* or *"governance framework"* rarely appear in the main network, suggesting that the governance dimension of AI in health is still emerging and not strongly integrated into practical discussions. There are exceptions: the concept of „AI governance" is addressed in some health policy studies, but these works do not constitute a critical mass in the body of reviews analysed, so they do not form central nodes on the map. Thus, the discourse on AI governance and implementation policies remains marginal in our network, signalling a possible gap.

**Integration vs. marginalisation of discourse on social vulnerabilities**

The connectivity assessment shows that discourse on the social vulnerabilities and implications of AI is present but partially marginalised in the literature. Topics such as fairness, bias and transparency are well represented as subjects of interest (especially in the ethics cluster), but are not centralised in the overall map. In other words, the scientific community recognises the importance of these topics, but they are often treated in dedicated sections (e.g. sections on „Ethical considerations" in reviews) or in articles specifically focused on ethics, rather than being organically integrated into most applied studies. For example, ethical considerations related to data confidentiality and bias are explicitly mentioned as challenges in research (Abdulsalam et al., 2025), which shows awareness of the issues. However, these considerations usually appear at the end of the papers (in the form of ethical discussions) and are not part of the main objective of many technically oriented studies.

On the bibliometric map, social and ethical nodes tend to be peripheral, indicating a degree of insularity in the discourse on social vulnerabilities in relation to the technological and clinical core. Terms such as *"accountability"* or *"bias"* are connected to few other concepts, a sign that only a subset of the literature addresses them directly. For example, *"bias"* could be connected to *"algorithm"* and *"data"*, but it does not appear in connection with *"oncology"* or *"diagnostic accuracy"*, suggesting that not all clinical studies take the issue of bias into account. Equity is also treated more theoretically; the idea that AI should be equitable is promoted, but practical implementation (e.g., studies evaluating the impact of AI on health inequalities) is rare, which explains the peripheral position of the term. In addition, the fact that *"governance"* and *"regulation"* are weak nodes indicates that concrete approaches to public policy and regulation of medical AI have not yet been widely discussed in the articles in our sample.

This relative marginalisation does not mean that social issues are completely neglected, but that they are still a specialised discourse carried out by a segment of the scientific community. In fact, a review of the literature highlights that ethical dilemmas related to confidentiality, trust and transparency are major obstacles to the implementation of AI in the healthcare system (Ahmed et al., 2023). The fact that they are perceived as practical barriers indicates the need for their closer integration: for AI to be widely adopted, these vulnerabilities must be addressed (e.g., lack of transparency creates mistrust among clinicians and patients (Ahmed et al., 2023)). In our map, this idea is reflected in the modest connection between the ethical cluster and the others: interaction exists (through terms such as *"trust"* or *"bias"* that partially link the clusters), but it is not strong enough, suggesting that the discourse on social vulnerabilities is still in an early stage of integration.

## Connections between bibliometric findings and sociological and humanistic theoretical frameworks

Bibliometric maps reveal a conceptual landscape in which technical terms related to artificial intelligence (*"artificial intelligence"*, *"deep learning"*, *"algorithms"*, *"large language models"*, *"predictive modelling"*) are linked to socio-ethical terms (*health equity*, *bias*, *privacy*, *ethics*). These relationships visible in the network can be read through the lens of solid theoretical traditions in sociology, social work, and the humanities, which provide a framework for interpreting social vulnerabilities. The green cluster (equity, demographics, clinical outcomes), identified at both thresholds 5 and 20, resonates with the theory of social determinants of health, which states that health status is structurally shaped by the social, economic, and political factors of the population. The dense connection between *"health equity"*, demographic variables (*"female"*, *"child"*, *"aged"*), major diseases (*"cardiovascular diseases"*) and algorithmic performance indicators (*"prediction"*, *"prognosis"*, *"reproducibility of results"*) suggests that algorithmic bias not just a technical flaw, but a form of reproduction of existing social inequalities, a phenomenon anticipated in the theory of structural violence (Farmer et al., 2006a), which shows that health systems and public policies can, through seemingly neutral mechanisms, cause systematic harm to vulnerable groups.

The cluster dedicated to the terms *"algorithmic bias"*, *"decision support systems"*, *"electronic health records"*, *"LLM"*, *"telemedicine"* aligns with the literature in the humanities on algorithmic discrimination (Noble, 2019; Benjamin, 2019; Eubanks, 2018). The map shows that *"bias"* functions as a bridge node between technical and social clusters, confirming the central thesis of these works: digital technologies can operate as extensions of power structures and historical prejudices. For example, the connectivity between *"algorithmic bias"* and *"diagnosis"*, visible especially in threshold map 20, is consistent with Ruha Benjamin's analysis of the „New Jim Code," technological mechanisms that produce exclusion under the guise of objectivity. At the same time, the association between *"electronic health* records*"*, *"LLM",* and *"data privacy"* reflects central themes in „surveillance capitalism" (Zuboff, 2019), where massive data aggregation becomes an infrastructure that can disadvantage the underrepresented, either through disproportionate surveillance or through the unauthorised reuse of their data.

Bibliometric maps also highlight an intersectional structure of vulnerabilities: demographic nodes (*"female"*, *"male"*, *"child"*, *"aged"*) are positioned in proximity to terms of equity and algorithmic performance. This proximity is explained theoretically by intersectionality (Kimberlé Williams Crenshaw, 1991), which shows that the effects of discrimination are not one-dimensional, but manifest themselves at the intersection of race, gender, age and social class. In the network, the connections between demographics, *"health disparities"*, *"prediction",* and *"global health"* signal that AI models can disproportionately affect people at the intersection of multiple forms of vulnerability (e.g., elderly women with comorbidities), which transforms technical analysis into a social one par excellence.

Terms related to professionals (*"health personnel"*, *"nursing"*, *"clinical competence"*, *"education"*) are connected to *"ethics"* and *"technology"*, a pattern that can be interpreted through the ethics of care (Tronto, 1993). This theory emphasises the relational and distributed nature of care. On the map, the presence of clusters linking AI to education, curriculum and professional competence suggests a tension: AI can redistribute tasks, intensify monitoring and generate „deskilling", disproportionately affecting already overburdened professionals — often women in nursing or caregiving roles. From the perspective of care ethics, this is not just an organisational problem, but a moral vulnerability, as it erodes the quality of the doctor-patient relationship and affects subjects who have little power in the design of technology.

The cluster associated with pandemics (*"COVID-19"*, *"critical care"*, *"ethical considerations"*) can be linked to sociological literature on inequalities in crisis conditions, where technologies tend to amplify pre-existing vulnerabilities. The connections visible on the map show that predictive models and automated triage systems are frequently associated with the context of the pandemic, a situation in which algorithmic decisions can have acute consequences, and in which the theory of structural violence becomes relevant for understanding differences in access to treatment.

Overall, bibliometric maps not only chart the dominant themes in medical AI research, but also reflect, through their topology, the tensions identified by major sociological theories: the unequal distribution of risks (structural violence), the social determinants of health, the intersection of vulnerabilities, systemic bias in technological infrastructures, and the precarious nature of care work. This dialogue between bibliometric data and social theory shows that AI vulnerabilities are not add-ons, but are closely interconnected with the dynamics of power, social stratification and surveillance, making the integration of social dimensions an essential condition for the responsible implementation of AI in medicine.

## Conclusions

The bibliometric analysis of the literature on artificial intelligence in medicine (2020–2025), integrated with sociological and humanistic theoretical perspectives, highlights a complex conceptual landscape in which technological advances coexist with subtle but persistent forms of social vulnerability. Co-occurrence maps show that the discourse on AI in health is dominated by technical and clinical themes such as *"deep learning"*, *"diagnosis"*, *"predictive modelling"* and *"large language models"* — while dimensions related to equity, bias, ethics, privacy and care work appear integrated, but often on the margins of the technological core. This distribution is not random; it reproduces social structures well documented by health sociology theories. At the centre of the network are nodes associated with algorithmic performance and model validation, while social terms are connected pointwise or function as bridges between technical and clinical clusters, suggesting that social reflection is present but insufficiently absorbed into dominant practice.

Rereading the maps through the lens of structural violence theory (Farmer) clarifies that AI, in the form in which it is implemented today, can act as a medium for amplifying existing inequalities. The consistent presence of the terms *"health equity"*, *"health disparities"*, *"female"*, *"child"*, *"aged"* in proximity to the concepts of *"prediction"*, *"risk assessment"* and *"prognosis"* shows that algorithmic performance is not uniformly distributed, but follows the lines of vulnerability of the social system. Thus, AI not only reflects but can also intensify the social determinants of health, confirming the position of the „social determinants of health" theory that the risks of disease and, in this case, the risks generated by technological tools are structurally shaped.

The cluster structure also supports the link with the literature on algorithmic discrimination and algorithmic oppression (Benjamin, Noble, Eubanks). The *"algorithmic bias"* node, connected to both *"diagnosis"* and *"electronic health records"* and *"LLM"*, indicates that bias is not a marginal defect, but an emergent property of data infrastructures and the way models are designed and trained. The technical bibliography only partially captures these effects, but when recontextualised sociologically, they become expressions of broader structures of exclusion, in which populations differentiated on the basis of race, the elderly, patients with comorbidities, or those from disadvantaged backgrounds are the most vulnerable.

The phenomenon of *"data privacy"* and its proximity to *"LLM"* and *"generative AI"* reflects another fundamental theoretical dimension: surveillance capitalism (Zuboff). The structure of the network indicates the permanent tension between the clinical need for data and its exploitation as a resource for prediction, optimisation or the development of generative models. From this perspective, vulnerabilities related to privacy, consent, and data reuse are not anomalies, but structural elements of a digital economy in which the patient becomes an involuntary supplier of informational raw material. Bibliometrics confirms this interpretation: confidentiality terms are linked to emerging technologies, not to solid ethical structures, signalling the insufficient integration of the regulatory framework into technological development.

Another important tension concerns the changing distribution of clinical work, reflected in the education–competence–professionals cluster (*"education"*, *"curriculum"*, *"clinical competence"*, *"nursing"*). From the perspective of care ethics (Tronto, 1993), this network suggests a tacit redistribution of responsibility, in which medical staff become guardians of AI, bearing the burden of supervising and verifying models. The phenomenon of „deskilling" identified in the literature is thus part of a broader ethical issue: technology risks undermining the very care relationships on which medicine is fundamentally based.

Overall, the bibliometric conclusions show that the social vulnerabilities of AI in medicine are not conceptual accidents, but manifestations of structural dynamics described by:

- structural violence (unequal distribution of technological risks);
- social determinants of health (differentiated AI performance across groups);
- intersectionality (accumulation of vulnerabilities at the intersection of social identities);
- algorithmic discrimination (systemic bias in data infrastructures);
- surveillance capitalism (exploitation of patient data as an economic resource);
- the ethics of care (erosion of the therapeutic relationship and clinical skills).

On this basis, the paper shows that the discussion about AI vulnerabilities needs to shift from the technical realm (where it is treated as an „add-on") to a structural, interdisciplinary approach capable of explaining not only how vulnerabilities arise, but why certain groups are disproportionately affected. A real integration of social dimensions into AI development requires a shift from models focused on technical performance to models oriented towards equity, transparency, redistribution of responsibility and protection of patient autonomy. Only in this way can AI become, not a multiplier of inequalities, but a tool for reducing them and strengthening a health system that works for everyone.

**Authors contributions**

R.M.D. was involved in research design, the literature review, data collection, analysis and interpretation, and drafting conclusions. A.N.D. was involved in the literature review, data analysis and interpretation, and drafting of conclusions.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**ORCID ID**

Radu Mihai Dumitrescu https://orcid.org/0009-0008-4481-394X
Adrian Nicolae Dan https://orcid.org/0000-0002-4033-3358

**References:**

Abdulsalam, K. O., Torera, M. N., Dawuda, A. A., Musa, S., & Olumide, O. O. (2025). Emerging Trends and Insights: A Comprehensive Bibliometric Analysis of Artificial Intelligence Applications in Healthcare and Psychology. Journal of Advances in Mathematics and Computer Science, 40(1), 54–71. https://doi.org/10.9734/JAMCS/2025/V40I11962

Aboy, M., Minssen, T., & Vayena, E. (2024). Navigating the EU AI Act: implications for regulated digital medical products. NPJ Digital Medicine, 7(1), 237. https://doi.org/10.1038/S41746-024-01232-3

Ahmed, M. I., Spooner, B., Isherwood, J., Lane, M., Orrock, E., & Dennison, A. (2023). A Systematic Review of the Barriers to the Implementation of Artificial Intelligence in Healthcare. Cureus, 15(10), e46454. https://doi.org/10.7759/CUREUS.46454

Asgari, E., Montaña-Brown, N., Dubois, M., Khalil, S., Balloch, J., Yeung, J. A., & Pimenta, D. (2025). A framework to assess clinical safety and hallucination rates of LLMs for medical text summarisation. NPJ Digital Medicine, 8(1), 274. https://doi.org/10.1038/s41746-025-01670-7

Benjamin, R. (2019). Race After Technology Abolitionist Tools for the New Jim Code. London: Polity Press.

Conduah, A. K., Ofoe, S., & Siaw-Marfo, D. (2025). Data privacy in healthcare: Global challenges and solutions. Digital Health, 11, 20552076251343960. https://doi.org/10.1177/20552076251343959

Crenshaw, Kimberlé W. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. University of Chicago Legal Forum, 1989, 139. Retrieved from https://scholarship.law.columbia.edu/faculty_scholarship/3007

Crenshaw, Kimberlé Williams. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. Stanford Law Review, 43(6), 1241–1299. https://doi.org/10.2307/1229039

Cross, J. L., Choma, M. A., & Onofrey, J. A. (2024). Bias in medical AI: Implications for clinical decision-making. PLOS Digital Health, 3(11), e0000651. https://doi.org/10.1371/JOURNAL.PDIG.0000651

Eubanks, V. (2018). Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor. New York: Picador, St. Martin's Press.

Farmer, P. E., Nizeye, B., Stulac, S., & Keshavjee, S. (2006a). Structural violence and clinical medicine. PLoS Medicine, 3(10), 1686–1691. https://doi.org/10.1371/JOURNAL.PMED.0030449

Farmer, P. E., Nizeye, B., Stulac, S., & Keshavjee, S. (2006b). Structural Violence and Clinical Medicine. PLoS Medicine, 3(10), e449. https://doi.org/10.1371/JOURNAL.PMED.0030449

Food and Drug Administration. (2025). Artificial Intelligence in Software as a Medical Device | FDA. Retrieved December 8, 2025, from https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device

Kücking, F., Hübner, U., Przysucha, M., Hannemann, N., Kutza, J. O., Moelleken, M., … Busch, D. (2024). Automation Bias in AI-Decision Support: Results from an Empirical Study. Studies in Health Technology and Informatics, 317, 298–304. https://doi.org/10.3233/SHTI240871

Marmot, M. (2005). Social determinants of health inequalities. Lancet, 365(9464), 1099–1104. https://doi.org/10.1016/S0140-6736(05)71146-6

Noble, S. U. (2019). Algorithms of Oppression: How Search Engines Reinforce Racism. New York University Press.

Papareddy, P., Lobo, T. J., Holub, M., Bouma, H., Maca, J., Strodthoff, N., & Herwald, H. (2025). Transforming sepsis management: AI-driven innovations in early detection and tailored therapies. Critical Care, 29(1), 366. https://doi.org/10.1186/S13054-025-05588-0

Serban, I.V. (2025). Neuroscience, Genetics, Education, and AI: Charting New Frontiers in Understanding Human Behaviour and Criminal Responsibility. BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 16 (Sup1), 399-414. https://doi.org/10.70594/brain/16.S1/31

Steerling, E., Siira, E., Nilsen, P., Svedberg, P., & Nygren, J. (2023). Implementing AI in healthcare—the relevance of trust: a scoping review. Frontiers in Health Services, 3, 1211150. https://doi.org/10.3389/FRHS.2023.1211150/BIBTEX

Torun, M. (2022). VOS Viewer analysis of artificial intelligence in Healthcare and Economics. In E. Demir & M. Torun (Eds.), Artificial Intelligence Applications and their economic effects on the field of Health Care (pp. 109–122). Istanbul: EFE Academy Publishing.

Tronto, J. C. (1993). Moral Boundaries: A Political Argument for an Ethic of Care. Moral Boundaries (1st Edition). New York: Routledge. https://doi.org/10.4324/9781003070672

World Health Organization. (2024). WHO releases AI ethics and governance guidance for large multi-modal models. Retrieved December 8, 2025, from https://www.who.int/news/item/18-01-2024-who-releases-ai-ethics-and-governance-guidance-for-large-multi-modal-models

World Health Organization. (2025). Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models. Retrieved December 8, 2025, from https://www.who.int/publications/i/item/9789240084759

Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: Public Affairs. Retrieved from https://www.hbs.edu/faculty/Pages/item.aspx?num=56791

**Authors biographies**

Radu-Mihai Dumitrescu is an Associate Professor at the Faculty of Sociology and Social Work, at the University of Bucharest, Romania.

Adrian-Nicolae Dan is an Associate Professor at the Faculty of Sociology and Social Work, at the University of Bucharest, Romania.

Received: July 14, 2025
Reviewed date: July 29, 2025
Accepted for Publication: December 12, 2025