MASTER THESIS | MASTER'S THESIS

Titel | Title

Racial discrimination in AI-based medical devices: the lack of diversity and representation of black women in training data.

verfasst von | submitted by

Paola Polignano BA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Laws (LL.M.)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 999 079
Universitätslehrgang lt. Studienblatt | Postgraduate programme as it appears on the student record sheet:	Human Rights (LL.M.)
Betreut von | Supervisor:	Univ.-Prof. Mag. Dr. Karl Stöger MJur

Table of Contents

1. **Introduction** 6
   1.1 **Background** 6
   1.2 **Purpose** 6
   1.3 **Research Questions** 7
   1.4 **Method** 7
   1.5 **Material** 8
   1.6 **Literature Review** 9
2. **Artificial Intelligence in medical devices** 9
   2.1 **Key Definitions and Concepts in Medical AI** 11
   2.2 **Advantages of AI in Healthcare** 14
   2.3 Level of adoption of AI in the healthcare sector in the EU 15
3. **What is discrimination?** 16
   3.1 **Algorithmic computer-based** 17
   3.2 **Stages of Algorithm Development** 19
      3.2.1 **Planning stage** 19
      3.2.2 **Development stage** 20
      3.2.3. Decision-making and use stage 20
4. **Ethical consideration that leads to algorithm bias** 20
   4.1 **Black women discrimination** 21
   4.2 **Algorithm discrimination: origin and consequences** 23
      4.2.1 **Cognitive biases** 24
      4.2.2 The Algorithmic Data Challenge 24
5. **Algorithmic bias in medical devices: insights from 'Gender Shades'** 25
   5.1 **Analysis of the Gender Shades Case** 26
   5.1.1 **The Pilot Parliaments Benchmark and comparative results** 28
   5.2 **Skin Tone and Accuracy: The Role of Data Representation in Dermatology AI Systems and Pulse Oximeters** 29
   5.3 Intersectional bias in medical diagnostics and AI: the dual impact of race and gender 31
6. **Legal analysis** 33
   6.1 **Intersectional Disparities in Healthcare Regulations** 34
      6.1.1 **The analysis of the limitations** 35
   6.2 **International legal framework** 36
      6.2.1 Limitations 37

6.3 **European legal framework** ... 39
6.3.1 **Analysis and Limitations** ... 40
6.4 **European non-discrimination regulation in medical devices** ... 42
6.4.1 **Limitations** ... 44
7. **Conclusion** ... 46
8. **Bibliography** ... 47
8.1 **Primary Sources** ... 47
8.2 **Secondary Sources** ... 49
8.3 **Cases** ... 57

List of Figures

Figure 1 – Relationship between artificial intelligence, machine learning and deep learning, of the European Parliament, Directorate-General for Parliamentary Research Services, Scientific Foresight Unit 2022→

Figure 2 – AI Users' statements regarding the use and impact of AI technologies in healthcare

Figure 3 – Self-reported unmet need for medical care, by sex, EU, 2010-2022

Figure 4 – Example images and average faces from the new Pilot Parliaments Benchmark (PPB) of the Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification (Buolamwini and Gebru, 2018)

List of Table

Table 1 – Main definitions and concepts in medical AI

Abbreviations:

AI - Artificial Intelligence

GDPR - General Data Protection Regulation

ICESCR - International Covenant on Economic, Social, and Cultural Rights

UDHR - Universal Declaration of Human Rights

CFR - Charter of Fundamental Rights of the European Union

ECHR - European Convention on Human Rights

NGOs - Non-Governmental Organizations

EU - European Union

AI Act - Artificial Intelligence Act

WHO - World Health Organization (if mentioned)

UN - United Nations

ECtHR - European Court of Human Rights (if mentioned)

ECJ - European Court of Justice (if mentioned)

1. Introduction

1.1 Background

The rapid integration of artificial intelligence (AI) into the healthcare sector carries a transformative potential, promising significant improvements in diagnostic accuracy, personalized treatment approaches, and overall patient outcomes. However, this technological advancement is not without its challenges. The increasing reliance on AI systems in medical decision-making processes raises critical ethical and legal concerns, particularly regarding racial discrimination and the safeguarding of fundamental rights. Since AI systems are influenced by the data on which they are trained, there is a risk of discriminatory outcomes that could disproportionately affect marginalized groups, especially racial minorities. An analysis of the cases of Buolamwini and Gebru (2018), as well as the studies by M. Groh et al. (2023) and Marie V. Plaisime (2023), it became clear that AI systems can inadvertently perpetuate existing biases present in training data, therefore leading to unequal healthcare outcomes. Despite the benefit of AI in healthcare, the current regulatory framework needs to adequately address the potential of AI racial discrimination, which leads to disparities in medical outcomes for vulnerable groups.

International and European regulations, such as the GDPR and the proposed AI Act, need to incorporate intersectional considerations to ensure that AI-driven medical devices do not disproportionately harm individuals, particularly those from vulnerable groups like Black women. By addressing the unique and intersecting forms of discrimination that these groups face, the regulations aim to promote fairness and equity in healthcare, to protect people from biases in AI systems.

1.2 Purpose

The purpose of this thesis is to explore the intersection between AI, racial discrimination, and the International, and European legal framework. It will evaluate compliance with the current regulations, identifying the areas needing improvement to address efficiently these challenges. The main focus, regarding legal analysis, will be on how fundamental rights are protected in the context of medical AI and on the identification of regulatory gaps that could eventually lead to fairness and effectiveness in AI-based medical devices.

1.3 Research Questions

The core objective of this thesis is to evaluate how European legal norms address the challenges provided by racial discrimination in AI medical devices. The research will be guided by the following key questions:

- To what extent do current European legal frameworks, including the European Convention on Human Rights (ECHR) and the Charter of Fundamental Rights (CFR), address racial discrimination in AI medical technologies?
- How do GDPR and the forthcoming Artificial Intelligence Act (AI Act) manage the integration of AI in healthcare while ensuring non-discrimination and data protection?
- What are the limitations of existing regulations in addressing intersectional discrimination, particularly in the context of AI medical devices?

1.4Method

This thesis explores the interdisciplinary field of legal research, combining legal analysis with ethical considerations, particularly regarding the issues of racial discrimination and fundamental rights in the context of artificial intelligence AI and medical devices.

The methodology of this thesis adopts an interdisciplinary approach, combining legal, ethical, and technical analysis to address the complex challenges posed by artificial intelligence AI in medical devices.

The second and third chapters utilize a descriptive approach to clarify what AI in medical devices entails, outlining its advantages and technical aspects throughout various stages of development. This foundational understanding is essential for contextualizing the subsequent analysis.

The fourth chapter shifts focus to the ethical dimension, drawing insights from various authors to explore how algorithmic bias can lead to discrimination, particularly against Black women. This chapter emphasizes the importance of recognizing and addressing such biases within AI systems to prevent existing inequalities.

The fifth chapter adopts a case study method, conducting in-depth examinations of three cases: Gender Shades, dermatology AI tools, and pulse oximeters. By analysing these cases, the thesis aims to highlight the real-world implications of AI biases and their potential to harm marginalized groups. The rationale behind selecting these cases was to develop a hypothesis underscoring the necessity of including intersectional biases in reports on AI and medical devices, with a specific focus on the underrepresentation of Black women.

The sixth chapter examines the existing international and European legal framework and outlines the limitations of compliance with human rights standards in the development and implementation of AI medical technologies. This chapter aims to identify gaps within the legal framework and propose recommendations for improving regulatory approaches to ensure that AI medical technologies uphold human rights and prevent discrimination.

1.5 Material

The thesis uses a diverse range of research materials to support its analysis of AI in medical devices and the associated ethical and legal challenges. These materials include books, academic articles, and research papers on artificial intelligence, ethics, and racial discrimination in healthcare. Additionally, reports and guidelines from the European Union, and the Council of Europe are incorporated to provide a regulatory perspective on the intersection of AI and human rights.

The thesis also draws upon relevant case studies such as the Gender Shades project, AI in dermatology, and pulse oximeters, to illustrate real-world implications of algorithmic bias. Furthermore, it reviews international and European legal frameworks, including the European Convention on Human Rights (ECHR), the Charter of Fundamental Rights of the European Union (CFR), and the General Data Protection Regulation (GDPR), as well as recent legislative proposals like the AI Act.

To ensure a comprehensive understanding, the thesis integrates insights from various disciplines, including law, ethics, artificial intelligence, healthcare, sociology, and human rights. Reliable internet sources, reports from human rights and AI-focused non-governmental organizations, and relevant opinion pieces also contribute to the breadth of the research materials used in this study.

1.6Literature Review

An in-depth literature review was conducted to support the ethical and legal analysis of the challenges posed by AI medical devices. This review included academic articles, policy reports, and relevant case studies. Keywords such as "intersectionality" were crucial to the analysis. As expressed in Kimberlé Crenshaw's 1989 work, intersectionality helps situate the analysis within an academic tradition that recognizes different types of discrimination, such as those based on gender and race. This review also discusses the potential for AI medical devices to perpetuate biases, emphasizing the need for specific and robust regulatory frameworks.

2. Artificial Intelligence in medical devices

The continuous development and evolution of artificial intelligence (AI), and consequently its algorithms, seems to be increasing very rapidly, introducing more and more new ethical and legal issues. For example, the UNESCO (2019) report, states that we interact with AI daily, both professionally and personal contexts, in areas such as job recruitment and being approved for loan approval, medical diagnoses, and beyond.

Artificial intelligence (AI) in medical devices represents a transformative integration of computational techniques into healthcare technology, aimed to increase diagnostic accuracy, treatment efficacy, and patient's outcome (Abbaoui et al., 2024). AI technologies utilize machine-learning algorithms and sophisticated data analysis techniques to extract necessary insights from medical data (ibid). These technologies help healthcare professionals to make informed decisions, increase patient outcomes, and optimize the efficiency of the healthcare delivery process (Bhati, 2023).

According to Thomas J. Fuchs, "Today, patients are dying not because of AI, but because of the lack of it," emphasizing the transformative potential of AI in healthcare. The U.S. Food and Drug Administration (FDA) has seen a significant rise in recent years in the development of Artificial Intelligence and Machine Learning-Enabled Medical Devices, particularly due to the availability of data, especially in radiology. This constantly evolving scenario, influenced by both the FDA in the United States and the European Commission, aims to ensure the safety and ethical use of AI in healthcare.

On one hand, the FDA provides the regulatory framework and guidelines for AI use in medical devices within the US (FDA 2024). On the other hand, the European Commission and the European Parliament set forth regulations and directives to protect patients and ensure ethical AI applications in healthcare across Europe (Lekadir, K. et al., AI Act European Parliamentary Research Service, 2022).

According to the University of Copenhagen, Project to Prevent Bias and Discrimination in Medical AI, the technique used in machine learning within AI in healthcare solutions involves programming medical devices, through algorithms, to recognize patterns in data and make predictions. For this to happen, it is necessary to 'feed' the machines with the necessary data, to provide accurate and effective feedback. When machines make decisions, they rely on past datasets they have interpreted. If the algorithms lack certain individual variables, they become biased, resulting in insufficient responses for those who may require it most.

Therefore, the AI decision-making algorithm consists of mechanisms by which machines make decisions based on past datasets they have analysed. Automatically, if individual variables are missing, this would lead to inaccuracies in developing specific feedback for each individual who may need it the most. Because, due to the complex applications of AI, these in particular must be monitored by humans to ensure a better outcome that aligns with human thought and decision-making, with AI serving to support decisions, diagnostic findings, or treatment proposals, rather than making autonomous decisions. Based on this premise, it is evident that human oversight is mandatorily required by law according to EU fundamental rights (Schneeberger, et al., 2020).

Artificial intelligence can be considered quite a complex term; to define AI, it is necessary to introduce its components, such as “data” and “algorithms”. Algorithms rely on a set of data to perform the necessary operations to achieve the desired goal and are trained to intervene in certain aspects (White Paper EU, 2020).

According to the definition proposed by the EU, the term refers to a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions […] Artificial intelligence systems use machine- and human-based inputs to perceive real and virtual environments Executive Order on AI, Sec. 3 (The White House, 2023). The complexity of the AI system is therefore caused by the multiple elements of which it is composed, and by their interaction to achieve specific scopes.

2.1 Key Definitions and Concepts in Medical AI

In the following chapters, some definitions offered by the European Parliament, Directorate-General for Parliamentary Research Services, Scientific Foresight Unit 2022, will be introduced in order to provide a comprehensive understanding of the key concepts and terminologies that will be discussed throughout the report.

Table 1 – Main definitions and concepts in medical AI

Term	Definitions
Artificial Intelligence	Here we will first use the historical definition of AI, i.e. when a machine is able to mimic human intelligence or even surpass it to perform a given task such as prediction or reasoning. However, in this thesis, we will mostly focus on one subfield of AI that is dominant in the healthcare area, namely machine learning (ML).
Machine learning	ML is a subfield of AI and concerns the methods that learn to perform given tasks, such as prediction or classification, based on existing data.
Deep learning	DL refers to NNs with more than three layers; in this case, the availability of big data is needed to estimate the optimal values of the parameters for this larger, more complex type of deep neural

	network (Goodfellow et al., 2016). Note that not all AI and ML tools are based on deep learning or NNs. Other techniques such as decision trees or support vector machines are widely used, especially when the data sample is not sufficiently large to build NNs or deep NNs
Medical AI or healthcare AI	This is a type of AI which is focused on specific applications in medicine or healthcare.

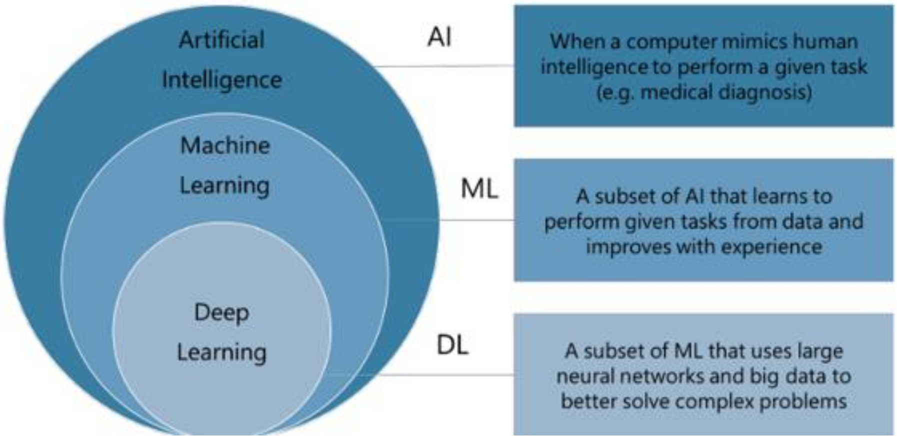
In regard to the understanding of the role of artificial intelligence in medical devices, it is essential to mention the subcategories of AI, such as Machine Learning (ML), which is one of the most popular approaches in the medical field. ML allows computational systems to improve their effectiveness by processing the data with which they have been specifically programmed.

A computational system in Machine Learning refers to a system designed to analyse cognitive processes using the same method. Its function is to address AI problems such as natural language understanding, planning, configuration, or diagnosis, which are computationally difficult. Identifying cognitive processes as computational processes applies the same limitations and capabilities of computational theory, such as computational complexity, in order to derive conclusions to solve problems (G. Brewka, ed.,1997).

Through this analysis, it is possible to understand the structure of problems and identify complexities inefficient clinical decision-making and diagnostic processes (N. Mehta, M.V. Devarakonda, 2018). Given the significant increase in cognitive load and information management, errors missed diagnoses, and the ordering of unnecessary tests can occur. Additionally, as the amount of data and medical literature continues to grow, this will further impact the system's processing capabilities.

Machine learning and natural language processing can support clinical and diagnostic decision-making systems by creating intelligent assistants for information retrieval (ibid.).

Figure 1 – Relationship between artificial intelligence, machine learning, and deep learning



As illustrated in Figure 1 of the European Parliament, Directorate-General for Parliamentary Research Services, Scientific Foresight Unit 2022, machine learning can be classified as a subset of artificial intelligence, while deep learning, as described, falls under machine learning. Deep learning utilizes artificial neural connections with multiple layers to detect patterns within databases, also significant are the ethical and legal challenges that arise when machine learning algorithms operate as "black boxes"(N. Mehta, M.V. Devarakonda, 2018, K.H. Yu, et al., 2018).

For instance, recent research in the medical context has utilized deep learning algorithms to predict molecules with potential antimicrobial activity. Therefore, the algorithm analysed over a billion molecules and virtually tested more than 107 million of them, identifying eight new antibacterial components that were different from known antibiotics (Stokes et al., 2020).

As explained in the article by Mehta and Devarakonda (2018) ML, first conceptualized in the 1950s, is essentially a form of applied statistics. Traditional ML requires cognitive engineers to identify input or predictive variables (known as features) that models use to predict outcomes or dependent variables (known as labels). In contrast, deep learning automatically determines which features predict the outcomes. Recent advancements in ML, especially in deep learning, have been driven by powerful processors like graphical processing units, new algorithms, and vast amounts of data. NLP, which focuses on enabling computers to "understand" spoken or written language, is another critical area.

This section explores the definitions of NLP and different types of ML, Real-case examples will be provided later in the chapters.

2.2 Advantages of AI in Healthcare

Published by the public health sector of the European Commission, this chapter investigates the significant advantages offered by artificial intelligence (AI) in healthcare. It explores how AI innovations are shaping current and future sustainable economic growth, improving societal well-being, and revolutionizing the value created by data within the healthcare sector. With over 500,000 different types of medical devices on the market, most people may need to use one at some point in their lives. These can range from simple items like contact lenses to more sophisticated devices such as pacemakers. The European medical devices sector prioritizes the protection and health of patients. It is distinguished by the competitiveness and innovation of small and medium-sized enterprises, supported by a regulatory framework, such as The Medical Devices Regulation (MDR) (Regulation (EU) 2017/745), aimed at ensuring the proper functioning of the internal market.

As mentioned by Badnjević et al., (2021), AI is revolutionizing healthcare by using technologies such as medical imaging (echocardiograms, computed tomography (CT), endoscopy, and skin photographs) and analysing tissue histology and physiological data (such as electrocardiograms (ECG)). These advancements span from medical applications to clinical engineering, showcasing their immense potential in transforming healthcare.

Artificial intelligence is generally a system that helps not only patients with advanced treatments but also offers the possibility of accessing alternative prescription medications (Basu, et al., 2020). Moreover, it enables medical access for people living in the most remote areas of the world (ibid.).

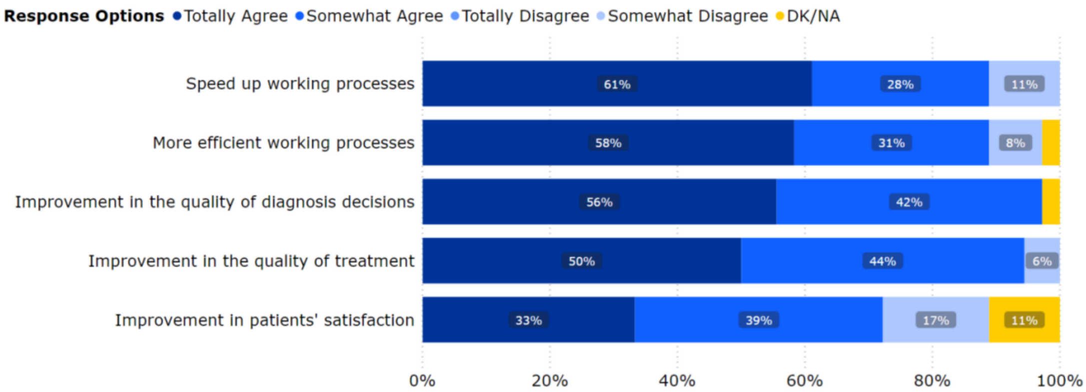
The foundational technology of AI, through its continuous monitoring and updates, ensures improved outcomes in diagnosing diseases, adapting treatments, and thereby enhancing performance to help people stay healthy (Bohr and Memarzadeh, 2020).

AI, through its ability to provide significant improvements across all healthcare domains from diagnostics to treatment, is already demonstrating substantial capabilities (ibid.). There is ample evidence indicating that AI algorithms perform as well as or even better than humans in tasks such as analysing medical images and correlating symptoms and biomarkers from electronic medical records (EMRs) to characterize and predict disease prognosis (ibid.).

2.3 Level of adoption of AI in the healthcare sector in the EU

According to the Study on eHealth, Interoperability of Health Data and Artificial Intelligence for Health and Care in the European Union report offered by the European Commission in 2021, based on various surveys, significant improvements have been identified in the use of artificial intelligence not only by professionals but also within the healthcare system.

Figure 2 AI Users' statements regarding the use and impact of AI technologies in healthcare.



Study on eHealth, Interoperability of Health Data and Artificial Intelligence for Health and Care in the European Union Lot 2: Artificial Intelligence for health and care in the EU Final Study Report, 2021

By linking diagnostic accuracy, timely detection of various diseases, and effective treatments for patients, it is evident how AI technologies in the healthcare sector directly influence the quality of patients' lives. On the other hand, these technologies can also positively impact healthcare professionals.

Consequently, doctors are not only able to provide better diagnoses and treatments for patients but also receive support that allows them to save time and focus on more productive tasks. The automation and improved efficiency of processes also provide social benefits, allowing patients to receive more attention from medical staff. Financially, costs will be naturally reduced, making healthcare more accessible to a wider variety of patients.

As shown in Figure 2, about 95% of participants agreed that the system led to improvements in the quality of diagnostic decisions. More than 90% of AI users fully agree that there has been an improvement in the quality of treatments in the healthcare system, and over 80% agree that the use of AI in the healthcare sector can make the work process faster and more efficient.

Despite the various benefits discussed in this chapter, there are notable limitations in the adoption of AI in the healthcare sector. According to the European Commission's 2021 report, the lack of high-quality data necessary for the operation and implementation of AI technologies in healthcare can serve as a major barrier. Consequently, this underscores the importance of enhanced national coordination to address these gaps and invest in the development of AI, fully realizing its benefits for patients.

In the following chapter, the potential limitations in the use of AI in the medical sector will be identified.

3. What is discrimination?

With a particular focus on the issue of racial discrimination in AI technologies used for patient diagnosis and treatment, this chapter discusses some of the notions of discrimination from a human rights perspective. By analysing the various stages of algorithm development, it will be possible to explore and understand the concept of algorithmic discrimination, identifying its origins, applications, and possible interconnections between gender, ethnicity, and artificial intelligence technologies.

Discrimination law is a multidisciplinary field, that is closely tied to constitutional and human rights law. It has its roots in employment law, and tort law is predominantly related to discrimination law (Khaitan,2015). Six of the main international human rights legal documents explicitly prohibit discrimination. Furthermore, the majority of countries around the world have constitutional or statutory measures in place to ban discriminatory practices (Stanford Encyclopedia of Philosophy, 2020).

Article 1 of the Convention on the Elimination of All Forms of Discrimination Against Women defines “discrimination against women” as any distinction, exclusion, or restriction based on sex that impairs or nullifies the recognition, enjoyment, or exercise of human rights and fundamental freedoms by women, on an equal basis with men, in political, economic, social, cultural, civil, or any other field, regardless of their marital status.

Amnesty International's general concept of discrimination undermines the essence of human dignity by infringing upon individuals' rights based only on their identity or beliefs. It inflicts harm and perpetuates inequality. Furthermore, depriving a person of their human rights or preventing them from enjoying equal legal rights can lead to unjustified distinctions made in policy or law.

The Human Rights Committee, in its 1989 examination of the International Covenant on Civil and Political Rights, observed that the Covenant neither defines "discrimination" nor specifies what constitutes it. However, Article 1 of the International Convention on the Elimination of All Forms of Racial Discrimination denifes "racial discrimination."

It describes it as any distinction, exclusion, restriction, or preference based on race, color, descent, or national or ethnic origin that aims to nullify or impair the recognition, enjoyment, or exercise, on an equal footing, of human rights and fundamental freedoms in political, economic, social, cultural, or any other field of public life.

Additionally, discrimination based on gender under the EU non-discrimination law is established in Directive 2006/54/EC of the European Parliament and Council on the 5th July 2006. It covers the implementation of equal opportunities and treatment of men and women in matters of employment and occupation.

The directive highlights the differences between direct and indirect discrimination. Direct discrimination, as outlined in Article 2(2)(a) of the Racial Equality Directive 2000/43/EC, is when a person is treated less favorably than others in the same situation, due to their racial or ethnic origin. According to the European Court of Human Rights, indirect discrimination is difference in treatment in the form of “disproportionately prejudicial effects of a general policy or measure”. Indirect discrimination does not require a discriminatory intent” (ECtHR, Biao v. Denmark, 2016)

3.1 Algorithmic computer-based

Algorithms can be defined as sequences of computer instructions that, using input data, generate specific values or sets of values as output (Kulk, et al., 2020). These examples illustrate the diversity of algorithms and their varied functions. Some algorithms directly influence decisions, such as determining eligibility for social security benefits or issuing fines for speeding violations (ibid.). Others primarily calculate probabilities, such as the likelihood that abnormal human cell deviations indicate cancer or the suitability of an individual for a specific job based on qualifications. While these algorithms do not make decisions themselves, they can assist human decision-makers. For instance, a doctor might consider the probability calculations from an algorithm when diagnosing cancer, or police might decide to increase patrols in a specific neighborhood based on crime probability assessments (ibid.).

It is possible to distinguish various types of algorithms, each capable of performing different functions such as decision-making, pattern detection, profiling, classification, clustering, and probability calculations (Vetzo, et al., 2018).

As previously discussed in chapter 2.1, machine learning serves as a process that supports decision-making systems. It functions as an algorithm distinguished by its capacity to "learn" autonomously, adapting, evolving, and enhancing to optimize specific outcomes based on input data, without requiring explicit programming for each step of improvement. In particular, leverage various analytical tools to identify correlations and patterns in large datasets (Gerards and Xenidis, 2020).

Algorithms in general, need to be specifically developed and "trained" to analyse data in particular ways. In the analysis of (big) data of medical devices, both supervised and unsupervised learning can be employed, depending on specific tasks. The method of supervised learning involves introducing new data to the system during the 'training' phase. The algorithm is tasked with identifying the same or similar patterns in this data. If the algorithm correctly recognizes the relevant patterns in the labelled data (as determined by data scientists), it receives positive feedback.

Conversely, it receives negative feedback if it misclassifies the data. Through this process, the algorithm gradually adjusts and improves its accuracy until it can reliably identify the relevant categories. Once it reaches this level of accuracy, the algorithm can be validated, allowing it to function outside the controlled environment of a data lab and assist in real-world decision-making processes without the need for pre-labelled data (Gerards and Xenidis, 2020).

Other algorithms may rely on unsupervised learning and these algorithms are equipped with instructions and extensive training data, enabling them to autonomously identify correlations and patterns, primarily through clustering and regression techniques (Huq, 2020). The distinction from supervised learning lies in that these algorithms are not trained using annotated data. Consequently, it is not feasible to consistently verify whether they identify patterns and correlations in the same manner as humans would (Hamon, et al, 2020). By supervising this machine learning, it is only possible to relate to the algorithm's output, which could or not adhere to the human idea (Hacker, 2018).

3.2Stages of Algorithm Development

As claimed by Kulk, Van Deursen, and others 2020, there are generally three stages that can be distinguished in the process of algorithmic decision-making. The following subsections are based on the differentiation of the various stages of algorithm development, as outlined by Gerards and Xenidis, Algorithmic discrimination in Europe 2020.

3.2.1 Planning stage

The planning stage is the first part, where a company or public body decides to define the scope of the use of the algorithm. Therefore, once the objective has been defined, it is possible to choose the type of algorithm that should fit the best for the purpose of the work.

Part of the planning phase, for a decision-making process, is to decide the output of the algorithm that should be used. As soon as the algorithm planning stage comes to an end, that should be no surprises. Because an algorithm exhibits considerable rigidity: It generally lacks the ability to autonomously adapt to evolving contextual circumstances, like new ideas. However, algorithms can be designed to adapt if they are trained appropriately, allowing them to respond to changes in their environment over time.

As introduced in chapter 3.1, supervised learning algorithms are trained using labelled data, which is typically derived from a dataset representing a specific moment in time. This implies that the data they are fed is heavily influenced by the external context. Consequently, regular updates are necessary for them to function effectively in a highly dynamic environment.

Regarding unsupervised learning and deep learning algorithms, they can be considered more flexible as they have the capacity to train themselves in the face of social changes and shifts in the world. They can automatically adapt to new contexts, thereby creating new pathways with new data (Vetzo, et al., 2018).

In summary, during the planning stage, it is among others, decided whether (it will continue to learn independently to create a new algorithm or use an existing one.

3.2.2 Development stage

During the development phase, experts write the computer code and determine how to integrate the existing technical system with the algorithm to meet the predefined objectives set by the user. It is essential to determine which data analysis technologies are most suitable for achieving the established objectives. The final part of this phase involves training, receiving feedback, and potentially testing and validating or certifying the algorithm.

3.2.3. Decision-making and use stage

This stage defines, the final step before the use of the algorithm, after have been tested and validated. This implies that once provided with new input data, the algorithm can generate output to fulfil the objectives established during the planning stage. The utilization of algorithmic output can vary depending on the type of algorithm and the user's objectives.

4. Ethical consideration that leads to algorithm bias

In this chapter, we will present some of the ethical factors that form the foundation of one of this study's research objectives, which is the intersectionality between race gender, and AI technologies.

To understand the aspects leading to algorithmic discriminatory outcomes, it is essential to rely on the concept of intersectionality.

As stated in the Oxford English Dictionary 2015, the sociological definition of the word "intersectionality" means:

> “The interconnected nature of social categorizations such as race, class, and gender, regarded as creating overlapping and interdependent systems of discrimination or disadvantage; a theoretical approach based on such a premise.”

Oxford English Dictionary

The concept deals with different social factors such as gender, race, ethnicity, and socio-economic background. All these social identities need to be analysed to uncover the origins of the various inequalities that result in the marginalization and disadvantage of underprivileged groups (Crenshaw, 1991; Verloo, 2006).

Kimberlé Crenshaw (1989) originally coined the term intersectionality, employing the analogy of intersecting paths to illustrate how various forms of oppression based on gender and race intertwine to generate a distinct experience of marginalization and discrimination.

As expressed by Bauer and Lizotte in their 2021 publication in the American Journal of Public Health, intersectionality, heterogeneity, and public health ethics are deeply interconnected. It is crucial to note that individuals representing minorities in society are often not given adequate importance in overall evaluations. Intersectionality and ethics necessitate addressing these under-representations of minority groups to enhance overall well-being. Therefore, an intersectional application of AI enables the promotion of ethical principles aligned with human rights values and core ethical principles.

It's very interesting and essential to consider the concept of intersectionality that Crenshaw introduced to challenge the reduction and simplified analysis of oppression. Originating from the advocacy of black women seeking recognition within the broader feminist movement of the 1960s and 1970s, intersectionality theory expanded to emphasize that the feminist struggles of black women are intricately intertwined with their efforts against racism, classism, and other barriers to achieving equal opportunities and social justice (Samuels and Ross-Sheriff, 2008).

Starting from the premise that this thesis will analyse cases of discrimination against black women through artificial intelligence systems and medical devices, as Crenshaw 1989 states, “any analysis that does not take intersectionality into account cannot sufficiently address the particular manner in which Black women are subordinated”.

By discussing diversity and its importance in the implementation of AI technologies, the following chapter will then address how to overcome algorithmic discrimination, particularly due to the lack of representation of women and, specifically, black women.

4.1 Black women discrimination

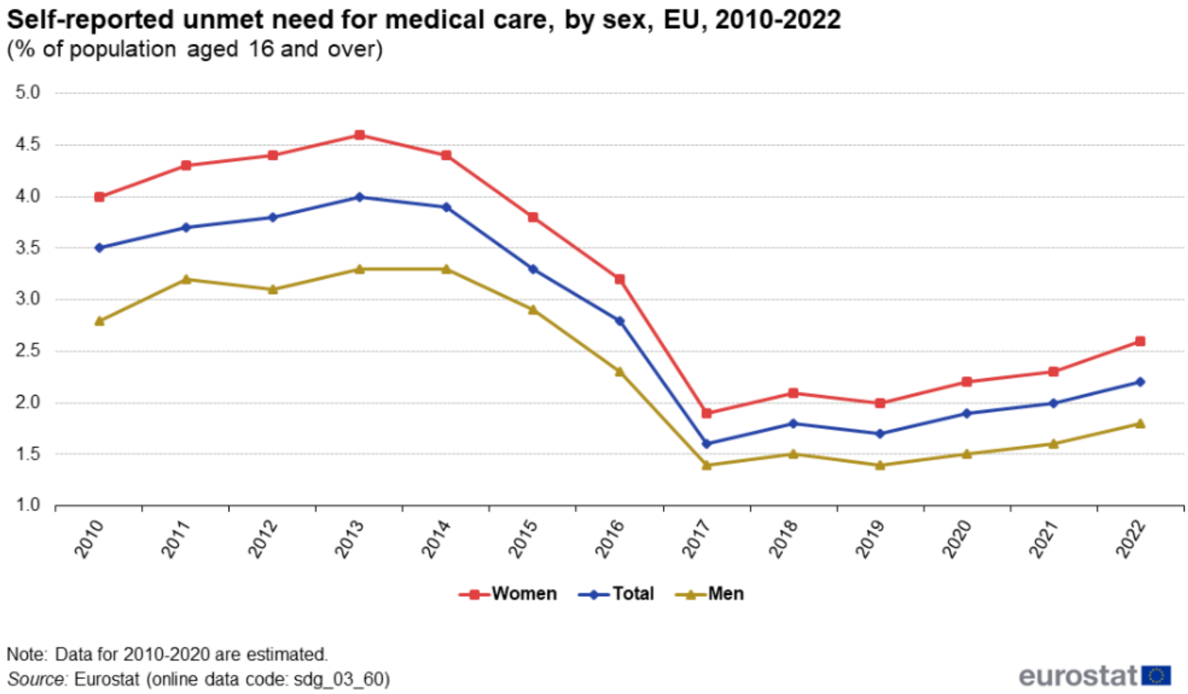
Taking into consideration the report published by the European Union Agency for Fundamental Rights in 2023 regarding "Being Black in the EU: Experiences of people of African descent," it is evident that black people continue to experience discrimination within the 13 EU Member States.

Through this overview, not only will the data collected by the FRA regarding experiences of discrimination based on skin color be presented, but it will also provide a general view of the discriminatory experiences that occur not only in the medical AI sector but also reflect the reality of the society we live in.

The survey data indicate that, in general, women of African descent tend to report incidents of discrimination more frequently than men (12% compared to 6%). Regrettably, Black women may face suboptimal patient-provider relationships due to perceived discrimination (Washington, and Randall, 2023) EU. People of African descent are more likely to report having unmet medical needs in most EU Member States according to a study of the Fundamental Rights Agency, based on survey data collected in 15 EU Member States in 2022.

As shown in figure 3, I utilize this statistic from SDG monitoring, which measures the proportion of individuals aged 16 and above who report unmet medical care needs because of financial limitations, waiting lists, or geographical barriers. This data is derived from the EU Statistics on Income and Living Conditions (EU-SILC) and represents individuals' self-reported assessments of necessary medical care that they did not receive or pursue, with dental care excluded from the analysis.

Figure 3



As shown in this image, only women experience the highest levels of inaccessibility to basic care due to various factors ranging from financial to geographical and others.

Discrimination within medical settings can inflict harm. For Black women, perceiving discrimination serves as a persistent stressor that detrimentally affects their physical and mental well-being, as well as their overall quality of life (Pascoe and Smart, 2009).

Through this analysis, the aim was to demonstrate that gender and race are extremely high-risk factors affecting accessibility to healthcare systems and common practices, such as basic healthcare services. Black people, particularly black women, face some of the highest chances of experiencing discriminatory acts, as evidenced by the data developed from various reports mentioned in this chapter. Because this inequality against Black women exists within society, it is almost inevitable that it will be reflected in artificial intelligence systems. As discussed previously, AI systems are programmed, developed, and monitored by humans, who create the data, thereby inevitably reflecting existing societal biases.

The following chapter will therefore introduce the concept of algorithmic discrimination, examining the origins and developments of these biases in AI medical device systems.

4.2 Algorithm discrimination: origin and consequences

After defining the algorithm and its various development stages in chapters 3.1 and 3.2, and emphasizing the importance of the concept of intersectionality in chapter 4, the following chapter will address the necessity of this concept to explain the phenomenon of algorithmic discrimination. This chapter aims to understand the origins of discrimination in the programming stages of the algorithm, presenting evidence of the underrepresentation of black women. Consequently, due to their gender and ethnicity, these individuals face unforeseen inequalities within AI medical technologies.

As mentioned before, humans are almost always involved and responsible before, during, and after the preparation of medical computer algorithms. This means that the algorithm can support human decision-making, operating through a "teamwork" dynamic between humans and machines. Humans are therefore the most important element during the development stages, as they are responsible for preparing and categorizing the data and eventually evaluating the algorithm during its learning phase (Katyal, 2019).

As can be understood, the algorithm depends on humans and the inputs planned by them to develop an accurate output. Specifically, medical specialists are an example of how humans practically collaborate with algorithms. Indeed, medical specialists use these technologies to consult the algorithms, to ensure, for example, a more accurate diagnosis (Gerards and Xenidis, 2020).

4.2.1 Cognitive biases

At what stage in the algorithm development process can the emergence of biases, that lead to discrimination, be identified? The emergence of discriminatory algorithms is connected to various stages in their development. Cognitive biases, for example, can pose significant challenges during the use of artificial intelligence, leading to discriminatory outcomes.

According to the seminal work by Tversky and Kahneman 1974, cognitive biases represent a systematic pattern of deviation from rational judgment, wherein individuals create their own "subjective reality" based on their perception of input. This subjective perception, rather than the objective input, can influence their behaviour, resulting in distorted and inaccurate judgments. While cognitive biases and their effects on decision-making are well-known and extensively studied, AI-assisted decision-making introduces a new paradigm. It is crucial to analyse and empirically study the role of cognitive biases within this new decision-making context (ibid.). This demonstrates how cognitive biases can lead to discriminatory outcomes during the human decision-making process. Since this occurs when humans are assisted by algorithms, it implies a collaborative decision-making setting.

Recently, as AI systems become integral to high-stakes human decisions, comprehending human behaviour and dependence on technology has become crucial. Ineffective collaboration between humans and automation will lead to increasingly costly and catastrophic outcomes (Lee and Katrina, 2004)

4.2.2 The Algorithmic Data Challenge

The challenge of training data is a major concern today. As discussed in Chapter 4.1, it can be reflected in the algorithm, potentially resulting in discriminatory patterns. In other words, human discriminatory attitudes risk being embedded and mirrored in the algorithms they create.

Data origins like electronic health records (EHR) or insurance claims often contain inaccuracies, systemic prejudices, and data deficiencies, which could disproportionately affect minorities receiving suboptimal care (Parikh et al.,2021).

As previously discussed, cognitive biases can manifest in various ways, such as "automation bias." This occurs when humans trust the outcome of the algorithm regardless of whether it may produce results that lead to discrimination.

Specifically, in this context, medical specialists may rely on certain medical devices without considering that they might make errors, resulting in discriminatory outcomes.

5. Algorithmic bias in medical devices: insights from 'Gender Shades'

In this chapter, the case study Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification by Joy Buolamwini and Timnit Gebru's 2018 will be analysed. This research highlights significant disparities in the accuracy of commercial gender classification algorithms across different demographic groups, particularly based on race and gender. Moreover, it will also analyse two more case studies conducted by M. Groh et al. 2023, and another one by Marie V Plaisime, 2023, unlike the first case, these cases analyse specific medical examples such as dermatology AI systems, and pulse oximeters. These examples illustrate how biases in AI algorithms when lacking sufficient representation of diverse skin colours, particularly in individuals like Black women, can result in disparities in various blood values.

For instance, inaccuracies in pulse oximeters may affect the measurement of blood oxygen levels in individuals with darker skin tones due to historical biases in data used for calibration. Similarly, dermatology AI systems may misdiagnose skin conditions in patients of different ethnicities if not adequately trained on diverse datasets.

By drawing parallels between the findings of "Gender Shades" and the biases observed in medical devices, this chapter aims to underscore the critical importance of addressing algorithmic biases and ensuring diversity in training data.

5.1 Analysis of the Gender Shades Case

To give a general overview of the case study and highlight the research by J. Buolamwini and T. Gebru (2018), this study evaluates the biases found in automated facial analysis algorithms and datasets related to different phenotypic subgroups. Through the use of the dermatologist-endorsed Fitzpatrick Skin Type classification system, this study examines the gender and skin type distribution within two facial analysis benchmarks, IJB-A and Adience. The Fitzpatrick Skin Type classification system, introduced by Thomas B. Fitzpatrick in 1972, categorizes individual skin tones and their tendency to burn or tan under sunlight exposure (Sharma and Patel, 2023).

Fitzpatrick skin types, typically determined by a physician, have been subject to numerous attempts to create an objective classification of skin type. In the medical field, these classifications are used to estimate dosages, assess cancer risk, or provide guidance during counselling.

However, this system is acknowledged to be ineffective in certain populations, particularly those with darker, non-Caucasian ethnicities, who are at higher risk of experiencing post-procedural pigmentary and scarring-related adverse effects (ibid).
As detailed in the 'Gender Shades' study, the IJB-A dataset, introduced by the National Institute of Standards and Technology (NIST) in 2015, comprises 500 unique public figures with images categorized by Fitzpatrick skin types and aims to be geographically representative.
The Adience dataset, a more recent gender classification benchmark from 2014, includes 2,284 distinct individuals, with 2,194 of their images labelled for both skin type and gender. Each dataset provides a single annotated image per individual for assessment purposes.

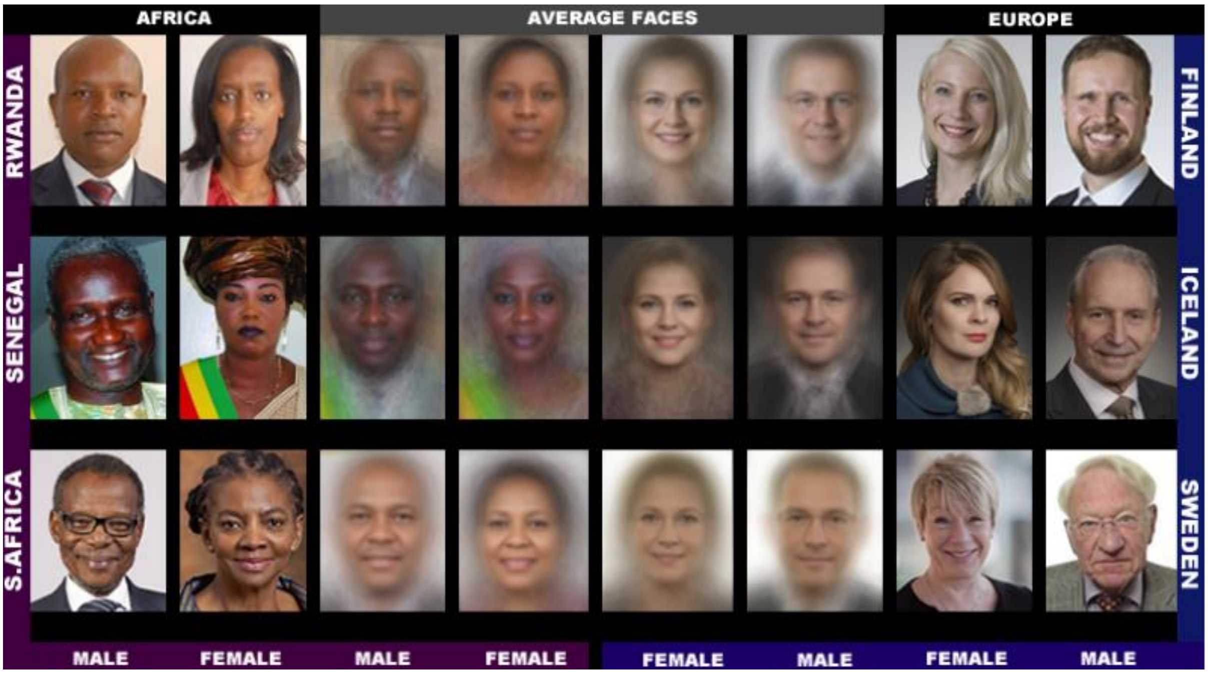
The focus of this thesis is not on the commercial gender classification systems evaluated in the 'Gender Shades' research, but rather on the datasets themselves. These datasets, used in facial recognition systems and AI medical devices, can exhibit similar biases, particularly when medical technologies depend on non-representative data. The findings from the 'Gender Shades' research illustrate that such data can be particularly unrepresentative of Black women.
While the data itself may not be incorrect, it is incomplete when it comes to the "non-Caucasian" population. This research highlights the underrepresentation of individuals of African descent by examining the gender classification of African faces and the impact of phenotypic traits on gender classification.

Many AI systems, such as facial recognition tools, rely on machine learning algorithms trained with labelled data (Sharma and Patel, 2023). As discussed in Section 4.2, several factors can contribute to algorithmic biases. For example, biased outcomes in facial recognition algorithms can stem from three main factors: biased datasets, ingrained social injustices, and individual choices. First, if training data is incomplete or biased, AI may perform well in tests but fail in real-world scenarios, such as misidentifying underrepresented groups in facial recognition. Second, even with diverse data, AI can still produce biased outcomes if societal practices are deeply embedded, as seen in Amazon's gender-biased HR algorithm, which favoured male candidates due to historical data. Lastly, when programmers select or exclude parameters, they may unintentionally embed their biases, leading to biased outcomes in sensitive situations like loan approvals, where individual choices greatly influence the algorithm's behaviour (Dhunnoo, 2023). Moreover, Esteva et al. showed that without datasets that include labels for various skin characteristics, such as colour, thickness, and hair density, it is challenging to assess the accuracy of automated skin cancer detection systems across different skin types (Esteva et al., 2017).

5.1.1 The Pilot Parliaments Benchmark and comparative results

The Pilot Parliaments Benchmark (PPB), was created to achieve a better inclusiveness of gender and skin types. The dataset used in the research analyses in this chapter, presents a very unique gender classification benchmark, based on the Fitzpatrick skin type categorization, ensuring a more varied and inclusive sample to address biases in current benchmarks. It is composed of 1270 individuals from three main African countries (Rwanda Senegal and South Africa) and three European Countries (Iceland, Finland, and Sweden).

Figure 4



Example images and average faces from the new Pilot Parliaments Benchmark (PPB) of the Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification (Buolamwini and Gebru, 2018)

Figure 4 depicts subjects comprising male and female parliamentarians from six nations included in the dataset, ranked according to their gender parity as evaluated by the Inter-Parliamentary Union.

This dataset was created to ensure a better balance between race and gender by promoting an intersectional approach, addressing the phenotypic imbalance present in the benchmarks.

The accuracy of the evaluation is determined by four main subgroups: darker females, darker males, lighter females, and lighter males.

The Kinmethodology used in this research consists in the use of the Fitzpatrick labelling system, which categorizes skin from Type I to Type VI, specifically defining lighter subjects with skin type I, II, III, and darker subjects with Fitzpatrick skin type IV, V, VI.

Based on the PPB decomposition by gender and skin type, the dataset comprises 1,270 individuals with a distribution of 44.6% female and 55.4% male. The differentiation between darker and lighter skin is 46.4% and 53.6%, respectively, offering a much more detailed representation compared to the Adience dataset, which reveals a distribution of 86.2% lighter-skinned individuals. Moreover, it appears that, compared to the datasets created in the presented research, the IJB-A (4.4%) and Adience (6.4%) datasets show that darker-skinned females are the least represented in their respective benchmarks, which instead reveal an overrepresentation of lighter-skinned males.

It is interesting to notice how overall, the PPB dataset provides a more equitable representation of both lighter and darker-skinned individuals compared to the IJB-A and Adience datasets.

Furthermore, as a result of the 'Gender Shades' research analysis, it is evident that all three commercial gender classifiers (Microsoft, IBM, Face ++) used to evaluate the accuracy in classifying female subjects generally perform worse with images of Black women, showing error rates ranging from 20.8% to 34.7%.

This not only reinforces the core argument of this thesis, that there is still a significant underrepresentation of black women in the datasets used but also highlights that, similar to facial recognition systems which could perform worse for people of black women, AI medical devices can present similar biases if trained on non-representative data.

5.2 Skin Tone and Accuracy: The Role of Data Representation in Dermatology AI Systems and Pulse Oximeters

The issue of the dataset bias could extend beyond facial recognition systems and infects could affect various domains of AI included for instance medical technologies.

Beginning with the experiment conducted by Groh et al. (2023), on deep learning-aided decision support for diagnosing skin diseases across various skin tones, this chosen study was designed as a comprehensive digital experiment to evaluate physicians' accuracy in diagnosing inflammatory skin diseases from images. This approach replicates the store-and-forward teledermatology process, where physicians receive patient images with limited clinical context. They were provided with a set of 264 images representing 46 skin diseases, ensuring near uniform distribution across various skin tones as measured by previously mentioned Fitzpatrick skin types.

The experiment aimed to assess the diagnostic accuracy of both dermatology specialists (BCDs) and primary care physicians (PCPs). Participants were shown images of skin diseases and asked to provide differential diagnoses. The images included eight prevalent diseases, atopic dermatitis, and Lyme disease, distributed over light and dark skin types. In order to measure the effect of the “decision support systems”, two different diagnostic support interfaces were compared, such as baseline neural network (with modest accuracy) and an enhanced version (designed to predict future advancements).

The results revealed a significant disparity in diagnostic accuracy, as physicians identified diseases in dark skin images with lower accuracy compared to light skin images. Specifically, BCDs and PCPs were 10% and 22% less accurate for dark skin diagnoses. These disparities underscore the influence of data representation on diagnostic performance and emphasize the need for inclusive datasets in medical research. This issue is particularly pronounced in many European countries, where it is easier to train AI on images of lighter skin tones due to predominantly white populations, further exacerbating the problem of underrepresentation in AI training data.

By analysing the diagnostic practices explored in the presented case, it is possible to observe through a critical examination how biases are perpetuated and the necessity to reinforce the representation in both training datasets and medical devices.

Based on the article by Marie V. Plaisime, titled "Invited Commentary: Undiagnosed and Undertreated—the Suffocating Consequences of the Use of Racially Biased Medical Devices During the COVID-19 Pandemic," published in 2023, this final case study in the thesis explores how pulse oximeters have contributed to racial disparities in healthcare outcomes during the pandemic.

During the peak of the COVID-19 pandemic, the Centers for Disease Control and Prevention (CDC) established clinical protocols to manage therapeutic medical interventions, leading to a significant increase in the use of medical technology in both home and hospital settings. Medical staff heavily relied on patient-reported symptoms and data from medical tools like pulse oximeters and thermometers to make critical decisions regarding hospital admissions and the allocation of supplemental oxygen therapies, including mechanical ventilation.

Even though medical technology is often seen as neutral, many devices incorporate racially biased algorithms that tend to prioritize care for White patients over Black patients, who may require more urgent medical attention. An article by Sudat et al. (2023), highlights substantial inaccuracies in skin color blood type readings among Black patients. The study found that these devices overestimated oxygen saturation levels in Black patients, impacting hospital admissions, treatment delays, healthcare delivery, and hospital readmission rates. This overestimation can lead to unrecognized or delayed recognition of a patient's need for supplemental oxygen, disproportionately penalizing Black patients during the pandemic.

Pulse oximeters, which measure peripheral oxygen saturation in the blood, were initially developed during World War II.

These devices often use spectrophotometry, which is used by scientists to figure out the concentration or amount of a particular substance in a sample and are calibrated using the Fitzpatrick Scale, which has limited and rigid options that contribute to biased outputs. Research shows that patients with darker skin tones are more likely to experience errors in pulse oximeter measurements, leading to undertreatment and underdiagnosis.

Considering these findings of the case, it becomes clear how crucial it is to observe how medical technology perpetuates health inequities and address the racial biases in devices like pulse oximeter function tests. Improving the accuracy of these tools is a vital step towards creating a more equitable healthcare system.

5.3 Intersectional bias in medical diagnostics and AI: the dual impact of race and gender

By examining the reasoning behind the selected cases, the scope was to develop a hypothesis highlighting the necessity of including intersectional biases in the reports on AI and medical devices, specifically addressing the underrepresentation of Black women. These examples demonstrate that AI is not merely a neutral instrument but is co-created with society. Consequently, it has significant political and social ramifications in perpetuating existing power dynamics, discrimination, and structural inequities (Benjamin, 2019; Noble, 2018; Ulnicane & Aden, 2023).

The selection of the first case of J. Buolamwini and T. Gebru's (2018), involving the PPB dataset, illustrates the significant underrepresentation of Black women in datasets used for gender classification. This case was chosen to emphasize the poor representation of Black women and the resulting bias in AI systems. Despite not being a medical case, it demonstrates how non-representative datasets can lead to significant disparities in algorithmic performance.

The second case of M. Groh et al. (2023), involving dermatology specialists and primary care physicians, was selected to show how similar biases manifest in the medical field. The experiment conducted in the dermatology context reveals race bias in diagnostic accuracy, with darker skin tones being diagnosed less accurately than lighter skin tones. This supports the hypothesis that the same type of race discrimination seen in facial recognition datasets can occur in medical diagnosis due to non-representative data.

The third case based on the publication of Marie V Plaisime (2023), focusing on pulse oximeters, further underscores the intersectionality of race and gender in medical technology biases. Pulse oximeters, which are crucial for measuring peripheral oxygen saturation, have been shown to overestimate oxygen levels in Black patients, particularly impacting marginalized Black women.

This case was selected to highlight how intersectional biases are often underestimated and underreported, not only in general medical diagnostics but specifically in the context of medical devices like pulse oximeters. This reinforces the thesis's objective to shed light on how the compounded effects of race and gender bias continue to impact the accuracy and fairness of medical diagnostics and AI systems.

Furthermore, it is important to underscore the intersectionality of race and gender in these biases. While the dermatology case does not explicitly focus on gender disparities, the PPB dataset case emphasizes that Black women are particularly underrepresented. This intersectional bias remains underreported in medical diagnostics, including areas such as pulse oximetry, where devices may perform inadequately for Black women.

The fact that this intersectionality is not adequately addressed in both the dermatology case and the medical device case is a crucial aspect of this thesis. By highlighting these intersectional disparities, this work aims to reveal how the combination of racial and gender biases can impact the accuracy and fairness of AI in both facial recognition and medical diagnostics. This reinforces the necessity for more inclusive and representative datasets to ensure fair and accurate AI outcomes for all demographic groups, especially those who are doubly marginalized.

Underestimating the disparities not only by race but also by gender underscores the unspoken challenges Black women face in medical diagnostics. Even in revelations about biases in medical devices or diagnostics, the unique experiences of Black women are often overlooked. This thesis, as previously mentioned, not only aims to draw attention to these biases but also seeks a deeper understanding of how intersectional factors can influence medical outcomes.

In light of these findings, the next chapter will address the legal questions arising from these intersectional biases in AI and medical diagnostics. The purpose is also to raise awareness of this issue but also to detect possible inconsistencies between AI medical technologies and European human rights standards.

The forthcoming analysis will research several critical aspects related to the intersection of EU regulations and algorithmic discrimination. It will explore the complexities and limitations inherent in the EU framework for gender equality and non-discrimination, particularly in the context of algorithmic bias. Additionally, the discussion will address how current data limitations within the healthcare sector might affect the reliability of predictive diagnostic systems for diverse patient populations. Lastly, the examination will consider the defining features and challenges of algorithmic discrimination and evaluate potential strategies for mitigating these issues within the scope of existing EU legal standards.

6. Legal analysis

This part of the research seeks to evaluate the degree to which current practices in the development and implementation of AI medical technology adhere to European human rights standards. By providing a brief overview of the equality and anti-discrimination norms in EU law, including the European Convention on Human Rights ECHR, and the Charter of Fundamental Rights (CFR).

What are the obligations of states under international human rights law addressing discrimination in medical technologies? To answer this question, an overview of the Universal Declaration of Human Rights (UDHR) and the International Covenant on Economic, Social, and Cultural Rights (ICESCR) will be provided, and then highlighting some of the state’s obligations which include non-discrimination.

Through the analysis of existing legal frameworks, sources of primary law in the European and International context, including secondary law, such as EU directives, regulations, and non-binding soft-low, it aims firstly, to assess the extent to which current practices in AI medical technology development and implementation comply with European human rights standards. Secondly to prove the gap in the material scope complying with EU non-discrimination law.

Human rights constitute a key element of the highest level of European Union legislation, often referred to as "primary law," whose rights establish an essential legal foundation for the development and application of medical AI, aside from ethical considerations. Secondary and tertiary EU law, on the other hand, must align with this foundational framework. Unlike these latter layers, it is expected that the primary legal principles will remain stable over time (D. Schneeberger et al.,2020). For the purpose of this thesis, the most relevant types of primary law instruments chosen are the European Convention on Human Rights (ECHR) and the Charter of Fundamental Rights (CFR),

6.1 Intersectional Disparities in Healthcare Regulations

By examining the general regulations and limitations at the European level and international level concerning gender equality, and then delving into the specific laws against discrimination in the healthcare system, we can already identify initial gaps in representation, particularly regarding the intersection of gender and racial disparities in healthcare.

Equality and non-discrimination concerns are covered by both primary and secondary EU legal sources.

The EU legal framework on gender equality and non-discrimination is comprehensive, enshrined in primary law through **Article 2 TEU** and **Article 3(3) TEU**, which establish the promotion of gender equality as a fundamental EU value and mission. **Article 8 TFEU** incorporates gender equality into all EU activities, while **Article 10 TFEU** supports efforts against sex discrimination. **Directive 2004/113/EC** ensures equal treatment in the access to and supply of goods and services. Furthermore, **Directive 92/85/EEC** regarding occupational safety and health for pregnant, breastfeeding women, and those who have recently given birth.

Furthermore, addressing discrimination specifically related to women in healthcare on an international level, **Article 12** of the Convention on the Elimination of All Forms of

Discrimination Against Women (**CEDAW**), highlights the responsibility of States Parties to guarantee non-discriminatory access for women in healthcare services.

Aiming to ensure equitable access to healthcare services for women. As it is possible to reelevate in this brief presentation of some of the primary sources that engage with women's equality in healthcare.

6.1.1 The analysis of the limitations

By analysing the EU Non-Discrimination Law in the context of healthcare there are several limitations that can be identified, particularly regarding the hierarchy of protection and its implications.

The **Directive 2004/113/EC** by covering equal treatment and the access and supply of good and services between women and men, faces several limitations and gaps in its practical application. Despite the extended scope of the Directive its application at the national level, is inconsistent and leads to unequal protection and enforcement of rights. This is due to the persistent gaps in how it is applied over Member States.

Furthermore, as outlined in the European Parliament (2017), report on the proposal for a directive on certain aspects concerning contracts for the sales of goods, Committee on the Internal Market and Consumer Protection, the Directive, which primarily focuses on gender discrimination and does not specifically nor adequately address the intersection of discrimination for instance race, but also age or disabilities or other vulnerable groups. The lack of an intersectional approach shows the necessity to tackle the complexity of gender inequality in connection with other forms of discrimination, but also the lack of awareness among service providers, users, and legal bodies about the rights and obligations established by the Directive. This emphasizes the need for the practical enforcement of equal treatment principles.

Moreover, the role of equality bodies, plays a crucial role in the monitoring and implementation of the Directive. The Commission should support and encourage these bodies, in order to ensure the enforcement of the Directive at the national level. Finally, the Directive lacks clear sectorial gender-mainstreaming recommendations, which are essential for ensuring equal treatment in various sectors. Such recommendations could help in addressing specific gender disparities within different areas of goods and services provision.

While **Directive 92/85/EEC** addresses women, who have recently given birth and are breastfeeding, it does not extend to other healthcare discrimination issues, it narrowly focuses on workplace safety rather than general access and treatment to healthcare. The intersection of gender equality and non-discrimination laws may not be particularly covered within the current legal framework, and therefore may fail to address biases, especially for marginalized groups like black women.

6.2 International legal framework

To provide a complete analysis of the legal aspects concerning the obligation of states under international human rights law to address discrimination in AI medical technologies, this work will first examine the international legal framework. The two most relevant international primary law sources chosen are the Universal Declaration of Human Rights (UDHR) and the International Covenant on Economic, Social and Cultural Rights (ICESCR).

The Universal Declaration of Human Rights (UDHR), proclaimed in 1948, is a landmark document in human rights history and it established fundamental human rights to be universally protected. The UDHR has a significant influence on the creation of several human rights treaties, globally and regionally applicable (United Nations, UDHR). Furthermore, the International Covenant on Economic, Social and Cultural Rights (ICESCR) in accordance with the UDHR, with the idea of free individuals enjoying freedom from fear and want can only be realized if conditions are established where everyone can enjoy their economic, social, and cultural rights, alongside their civil and political rights (Office of the High Commissioner for Human Rights).

The obligation of states under international human rights law to address discrimination in AI medical technologies is enshrined in **Article 2 UDHR** and **Article 2 ICESCR**. This first article mandates that everyone is entitled to all the rights and freedoms set forth in the UDHR without distinction of any kind, such as race, sex, language, religion, political or other opinion, national or social origin, property, birth, or other status. This article requires states to guarantee that the rights enunciated in the ICESCR will be exercised without discrimination of any kind. While UDHR underlines the obligation of states to ensure that no discriminatory practices exist in the provision of medical technologies, the ICESCR emphasizes the obligation of states to eliminate any discrimination in access to medical technologies and ensure equal treatment in healthcare.

Furthermore, concerning the right to health, **Article 12** of the **ICESCR** recognizes the right of everyone to the highest attainable standard of physical and mental health. It obligates states to take steps necessary for the prevention, treatment, and control of diseases and to create conditions that would assure all medical services and medical attention in the event of sickness.

In regards to the right to information and participation **Article 19** and **Article 25** of **UDHR**, firstly guarantee the right to freedom of opinion and expression, which includes the freedom to seek, receive, and impart information and ideas through any media. This implies the obligation of states to ensure transparency and access to information regarding media (such as medical devices), **Article 25** declares that everyone has the right to a standard of living adequate for the health and well-being of themselves and their family, which includes medical care.

Additionally, **Article 12 ICESCR** alongside the right to health, encompasses the right to access information and participation in public health decisions. States are required to provide accessible information and participation in public health decisions. States are required to provide accessible information about medical technologies and ensure that affected communities, including marginalized groups, can participate in decision-making processes.

6.2.1 Limitations

Article 2 of UDHR assert that all individuals are entitled to rights and freedoms without any distinction, however, the lack of specific guidelines for the implementation of the rights within the context of AI medical technologies explicit the absence of monitoring mechanisms and enforcement of non-discrimination in AI, leave considerable space for interpretation and inconsistent between states.

Similarly, **Article 2 ICESCR** obliges states to take steps to fully realize non-discrimination mandates, also does not provide detailed or instruction in addressing the unique challenges of AI in medical technologies. Both articles highlight the need for more precise directives and robust enforcement to effectively address the issues posed by AI-related discrimination in healthcare.

Article 12 ICESCR Although the right to the highest attainable standards of physical and mental health can be recognized, certain inconsistencies can be identified. These inconsistencies are due to the absence of data representation requirements, leading to inadequate treatment or diagnoses among a diverse group of people, and thus not ensuring the standard of health ensured by this provision. This provision, does not account for potential biases and lacks measures to address such issues.

Article 12 of the ICESCR ensures the highest attainable standard of physical and mental health can be recognized. However, it was drafted long before the presence of AI in health care, resulting in the provision not addressing the unique challenges imposed by AI to achieve the highest attainable standard of health. For instance, the biased training of data might lead to a consistent underdiagnose or false diagnostic among specific ethnic groups. While this clearly impacts the right to health, Article 12 does not provide guidelines on addressing such technological biases.

Article 12 also does not mandate the use of representative data in health decision-making processes. This might lead to the failure of accurate diagnostics. Lastly, Article 12 does not ensure algorithmic transparency, the right to health does not include provisions for understanding how health-related decisions are taken (black box phenomenon, mentioned in chapter 2.1). This means there is no explicit right for the patient to understand how the diagnosis was made, therefore this does not ensure the highest level of attainable health.

Article 19 of UDHR guarantees the right to information, including the right to seek and receive information, but does not address the need for clear accessible, and unbiased information regarding AI medical devices. The lack of provisions for algorithmic transparency especially in automated decision-making systems reduces the legitimacy behind decisions made for eligibility to some medical treatments. The absence of provisions for correcting AI-generated misinformation might generate or propagate health-related misinformation. For example, if an AI medical devices program provides inaccurate health information, article 19 does not provide a framework for correcting this information nor does it hold developers accountable for spreading this misinformation. Lastly, this provision does not address the ethical standard in AI development systems. For instance, AI systems can be programmed to provide health information and might prioritize treatments or medications based on commercial interests rather than medical efficacy.

To conclude, while both articles provide important human rights protections, they were not designed with AI technologies in mind. Therefore, they lack specific provisions to address unique challenges posed by AI healthcare including biases, transparency, data representation, and the need for ongoing updates to keep pace with technological advancements. Highlighting the need for legal frameworks that explicitly mention AI in regards to health issues.

The right to information guaranteed by **Article 19 UDHR** through its transparency, does not explicitly address the need for clear, accessible, and unbiased information about AI medical technologies.

In the end **Article 25 of UDHR** to the right to participation, explains the right to participate in government and public affairs, but it does not directly integrate the affected communities in the development and regulation of AI medical technologies.

Overall, the analysed articles which aim for non-discrimination and therefore right to health, and rights to information and participation, show the lack of detailed implementation guidelines, enforcement mechanisms, and provisions to ensure fully transparency and participation in regard to the AI sector.

Given these gaps, states are left with insufficient efforts to combat discrimination in AI-based healthcare technologies. The absence of specific directives and effective monitoring could therefore lead to the continued perpetuation of these biases in AI medical technologies.

6.3 European legal framework

This first part of the legal analysis will cover the European framework by providing a brief overview of the equality and anti-discrimination norms in EU law, including the European Convention on Human Rights (ECHR), and Charter of Fundamental Rights (CFR).

Despite diverse national laws in EU Member States, these instruments ensure uniform protection of fundamental rights, crucial for medical AI due to its significant impact on physical and mental integrity. Fundamental rights not only protect individuals from state intervention but also require the state to safeguard freedoms from third-party interference (D. Schneeberger et al.,2020).

Human rights constitute a key element of the highest level of European Union legislation, often referred to as "primary law," whose rights establish an essential legal foundation for the development and application of medical AI, aside from ethical considerations. Secondary and tertiary EU law, on the other hand, must align with this foundational framework. Unlike these latter layers, it is expected that the primary legal principles will remain stable over time (ibid.).

The compliance and intensity to which current practices address racial discrimination in AI medical devices are enshrined through **Article 14 ECHR and Article 21 CFR**, these mandates prohibit discrimination on various grounds including race, color, ethnic origin, and genetic features. Furthermore, **Article 8 ECHR** and **Article 7 CFR**, guarantee the right to respect for private and family life.

Additionally, **Article 13 ECHR** and **Article 47 CFR,** are the right to an effective remedy and fair trial guarantee the remedies for violations of rights and freedoms under the ECHR.

6.3.1 Analysis and Limitations

Both articles, 14 ECHR and **21 of the CFR**, despite their non-discrimination mandate, do not provide specific guidelines or mechanisms for implementation in AI medical technologies. These deficiencies within various jurisdictions may consequentially lead to discrimination.

However, even though **Article 8 ECHR** and **Article 7 CFR** include the protection of data privacy, in the context of AI medical devices they do not address specific challenges posed by racial biases in training data. The lack of explicit directives to ensure diverse representative data sets prevent the effectiveness of the AI system in general but also in medical devices.

In the end **Article 13 ECHR** and **Article 47 CFR** include the right for effective remedies, showing a difficulty in the procedural complexities due to the limited evidence.

By answering the question of how legal mandates of the ECHR and CFR comply with the posed challenges and to which degree of intensity is the intersection of racial discrimination in AI medical devices, it highlights the lack of specificity in primary laws, which rely heavily on implementation and enforcement for their effectiveness.

Based on these analyses, additional requirements for the development and operation of medical AI in Europe can be deduced. As previously mentioned, continuous monitoring of AI devices is as necessary as the thorough verification of training data to identify and correct possible biases.

Moreover, these limitations reveal themselves to be closely related to the problem of being able to verify the mere existence of AI-automated discrimination. In an attempt to establish a discriminatory outcome of medical devices due to the algorithm's automatic decision-making process, AI in this regard can be challenging, just verifying and proving the existence of discrimination.

According to Wachter et al. (2020), AI is more abstract, intangible, and elusive compared to traditional forms of discrimination, which consequently poses significant challenges in recognizing it. Therefore, individuals such as black women, who may seek to assert their rights under EU non-discrimination law, face challenges in proving a link between discriminatory treatment and resulting harm, such as receiving a misdiagnosis due to AI-driven discrimination.

These missing elements are essential to making a so-called prima facie case of discrimination (Amdiju, 2023). It seems that an algorithmic decision-making system could violate the prohibition of indirect discrimination, leading to more complexities for marginalised groups to claim their rights.

This occurs when practice that initially appears neutral, but in the end results in discrimination against individuals based on their ethnic origin or other protected characteristic (Tobler, 2005).

As described by the European Court of Human Rights, indirect discrimination:

a difference in treatment may take the form of disproportionately prejudicial effects of a general policy or measure which, though couched in neutral terms, discriminates against a group. Such a situation may amount to 'indirect discrimination', which does not necessarily require a discriminatory intent.

ECtHR, Biao v. Denmark (Grand Chamber), No. 38590/10, 24 May 2016, para. 103.

When considering the impact of indirect discrimination, the focus should not be on whether the alleged discriminator intended to discriminate or acted deliberately. Instead, the emphasis should be on the effects of the practice itself. (ECtHR, Biao v. Denmark 2016).

On the other hand, as further elaborated by Wachter et al. (2020) AI poses challenges to the scope of non-discrimination law itself. The lack of specificity, as seen in the mandates analysed in Chapter 6.3, highlights the need for collaboration between legislators and society to revisit the scope of non-discrimination laws. This collaboration is necessary to adapt these laws to contemporary needs, thereby incorporating not only human beings and organizations but also "machines."

Non-discrimination law must then help people against algorithmic discrimination, even though, the law shows several weaknesses because of the lack of transparency but also firstly because of the lack of evidence to prove the existence of a discriminatory act if self.

Despite various legislative advances made in recent times and in specific sectors, there remains an inadequate analysis and understanding of the limitations posed by specific complex algorithmic systems. This issue will be further discussed in the limitations section of chapter 6.3.1.

In conclusion, the limitations and gaps in the previously mentioned articles, reveal that while these provisions prohibit discrimination and protect privacy, they lack the necessary specificity for addressing the complexity of AI medical technologies. The absence of clear guidelines on AI’s potential biases, particularly in training data and automated decision-making processes, presents challenges in preventing racial discrimination and ensuring adequate solutions.

6.4 European non-discrimination regulation in medical devices

Based on the various discussions conducted so far regarding the analysis of potential methodologies and applications during the training phase of algorithms in medical devices, the concept of algorithmic discrimination was introduced. Despite its advantages, integrating data and AI into healthcare presents numerous challenges and ethical complexities that require thorough consideration. These challenges can manifest at different stages of algorithm development, including the planning stage, development stage, and decision-making stage, and use of each stage potentially leading to various discriminatory outcomes.

According to K. Stöger et al. (2021), European medical artificial intelligence should not rely only on machine decisions but rather on the cooperation between AI systems and humans. Human monitoring and oversight are essential to ensure AI-supported decisions, diagnostic findings, or treatment proposals, to avoid misdiagnoses.

As a result, through the analysis of some regulations from *European Parliament legislative resolution of 13 March 2024 on the Artificial Intelligence Act*, this chapter begins with the key definitions of the General Data Protection Regulation (GDPR) and then concludes with *Regulation (EU) 2017/746 on in vitro diagnostic medical devices*.

Since the fast development of multiple technologies, such as software as a medical device (SaMD), and the General Data Protection Regulation GDPR they replace the previous Data Protection Directive (DPD) of 1995 and aim to protect the rights and personal data of all individuals in the European Union (B. Wolford, 2020).

The GDPR, as secondary legislation, is grounded in the Treaties and the Charter of Fundamental Rights of the European Union. It cannot expand, alter, or restrict the scope of primary law and must therefore align with the primary legal framework. Consequently, the GDPR has a delicate role, as it must balance data protection with other objectives, such as the freedom of business and the general interests of technological development (Granmar, 2021).

The GDPR includes different provisions that could be relevant for this analysis, and therefore essential to analyse the existing limitations in the data of the healthcare sector that can lead to unreliable predictive systems of diagnosis in diverse patients, in particular on black women.

With respect to the planning phase (3.2.1), development phase (3.2.2), and decision-making (3.2.3) algorithmic phase, **Article 25 GDPR** mandates that data protection needs to integrate the necessary safeguards into the processing in order to meet the requirements included in this Regulation and ensure the rights of data subjects are upheld, essential for addressing initial biases.

Article 5 of GDPR states how personal data shall be collected and processed for specified legitimate purposes and only limited to that. To ensure, for instance, that during the development phase, the data processing is aligned with the prefix scope. Additionally, this Article needs accuracy and constant update, to ensure the inclusion of new developments and findings.

Article 22 GDPR is, relevant for the decision-making phase, **and** allows individuals to avoid being subject to a decision based only on automated processing, which affects them in case certain conditions are met, patients are allowed the right to challenge biases in AI medical devices decisions ensuring their proper treatment. Moreover, this article highlights the necessity of human intervention on medical devices before the contest to the decision.

According to the European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act), this secondary legal instrument chosen for this thesis and amending Union Legislative that adhere to high standards of safety and fairness.

Regulation (50) of Position of the European Parliament adopted at first reading on 13 March 2024 with a view to the adoption of Regulation (EU) 2024/ (Artificial Intelligence Act), underlines the importance of incorporate this provision for the so classified high-risk medical devices, and therefore the requirement for a conformity assessment bodies from a third-party in order to identify and mitigate eventual biases. This regulation accentuates the safety and compliance with the Union harmonization legislation so that AI could adhere to strict standards.

Regulation (58) of Position of the European Parliament adopted at first reading on 13 March 2024 with a view to the adoption of Regulation (EU) 2024/ (Artificial Intelligence Act), describes the importance of considering the impact of AI used to determine access to essential public assistance and services, such as healthcare, which could impact individuals' lives.

An interesting acknowledgment for this thesis is highlighted in the regulation in regards of the potential of AI systems to infringe such a fundamental right, the right to social protection, non-discrimination, and human dignity. Therefore, the text explicitly declares the possibility of AI to perpetuate historical patterns of discrimination subsequentially going to create new forms of discriminatory factors such as racial or ethnic origins.

The last legal instrument that will be analysed in this chapter focuses on regulations that ensure the safety and health of the regulatory framework for medical devices.

(1) Revision of Directives 90/385/EEC and 93/42/EEC states the importance to address the contemporary needs, with the establishment of a robust, transparent, and sustainable regulatory framework.
(2) Regulation on Medical Devices, aimed to design and support the proper function of internal market of medical devices to ensure the protection of patient and user health.

Both Regulations aim to guarantee certain standards of safety and efficacy of medical devices due to protection, while also fostering innovation.

6.4.1 Limitations

Article 25 GDPR, could be related to the planning phase of algorithmic development (*3.2.1*) means that data is represented including diversity, since the existing healthcare data could lack sufficient representation, specifically for black women. Therefore, the underrepresentation could be connected to the reliability and fairness of automated systems from the beginning of their stages.

Article 5 is necessary for the development of AI system without any biases, but the update stage of the data could lead to inaccuracies.

Additionally**, Article 22** essential for the decision-making phase, could result in challenges and therefore the requirement of human monitoring highlights the necessity of potential biases arising during the process.

As can be noted through the analyzed directives and regulations, the focus has mostly been on addressing and referring to the protection of patients under the EU non-discrimination directives.

Although Regulation (58), addresses the possibility of AI perpetuating historical patterns of discrimination and consequently creating new forms of discriminatory factors such as racial or ethnic origins, none of the other mentioned secondary law instruments refer to the specific discrimination that particularly vulnerable groups may face compared to other individuals.

The so-called multi-dimensional discrimination, as explained by Ellis & Watson (2012), refers to specific groups of people who may face discrimination based on more than one protected characteristic. For instance, instead of being treated separately as "black" and "woman," an individual may be discriminated against for being black and a woman, and therefore it explains the intrinsic concept of intersectionality.

The lack of specificity regarding the composition of disadvantaged groups within the EU non-discrimination directives leads to a specific level of generality in their formulation, potentially overlooking the needs of those most vulnerable to intersectional discrimination.

Therefore, while European non-discrimination regulations provide a framework for safeguarding patients' rights and the efficiency of medical devices, numerous challenges still persist. The integration of AI in healthcare highlights the necessity for specific legal instruments in order to address the rights of those who face algorithmic discrimination, for instance in regards to black women who face intersectional discrimination.

The limitations identified in this chapter, particularly concerning GDPR and the underrepresentation of diverse data, show the demand for constant revisions and monitoring of the regulations. It is crucial that future legislation, not only ensures the highest standards of safety and advancement but also to preserve the protection against discrimination by taking into consideration the complexity of the dynamics that AI involves in healthcare technologies.

7. Conclusion

This thesis has examined the intersection of artificial intelligence (AI) and human rights within the context of medical technologies, focusing on non-discrimination and the right to health.

The aim was to evaluate the effectiveness of international and European legal frameworks in addressing the particular challenges posed by AI in healthcare, including biases and transparency issues.

The case studies of algorithmic bias in facial recognition, dermatology AI, and pulse oximeters underscore the urgent need for more inclusive and representative datasets in the development of medical AI. The pervasive underrepresentation of black women in data training has led to disparities in diagnostic accuracy and treatment outcomes, perpetuating existing healthcare inequities

As a result of the inadequate representation in training datasets, biases could reflect in underdiagnoses or incorrect diagnoses for specific ethnic groups. Additionally, the fact that many of the AI systems, due to their programmed structure of a "black box", makes the decision-making process quite unclear, it might therefore impede patients' ability to understand and contest AI. This highlights the necessity to a more urgent need for transparency and an inclusive approach in the AI system. In reference to the legal analysis, the investigation.

By examining the legal framework, while the European Convention on Human Rights, the EU Charter of Fundamental Rights, and regulations such as GDPR and the proposed AI Act provide important safeguards, they lack the necessary specificity and enforcement mechanisms to effectively address intersectional discrimination in AI-driven medical devices. The gaps in the legal framework, especially by ensuring diverse data representation and algorithmic transparency, highlight the pressing need for a more comprehensive and intersectional approach to regulating the integration of AI in healthcare.

To promote fairness and equity in the application of AI technologies, policymakers and regulators must prioritize addressing the unique and compounded forms of discrimination faced by marginalized groups. By incorporating an intersectional perspective into the development of future regulations, the European Union can better protect the fundamental rights of all individuals, regardless of their race, gender, or other intersecting identities, in the face of the rapid advancements in medical AI.

8. Bibliography

8.1 Primary Sources

Council of the European Union. 1992. "EUR-Lex - 31992L0085 - EN - EUR-Lex." Eur-Lex.europa.eu. November 28, 1992. https://eur-lex.europa.eu/eli/dir/1992/85/oj.

COUNCIL DIRECTIVE 2004/113/EC. 2004. "Implementing the Principle of Equal Treatment between Men and Women in the Access to and Supply of Goods and Services." *Official Journal of the European Union*.

European Commission. 2012. "EUR-Lex - 52012SC0273 - EN Commission Staff Working Document Impact Assessment on the Revision of the Regulatory Framework for Medical Devices. (1) Revision of Directives 90/385/EEC and 93/42/EEC ." *Europa.eu*. Vol. SWD/2012/0273. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=CELEX:52012SC0273:EN:HTML.

European Parliament. 2024. "Texts Adopted - Artificial Intelligence Act - Wednesday, 13 March 2024." Www.europarl.europa.eu. June 19, 2024. https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html#title2.

European Parliament and the Council of the European Union. 2017. "EUR-Lex - 32017R0746 - EN - EUR-Lex." Eur-Lex.europa.eu. April 5, 2017. https://eur-lex.europa.eu/eli/reg/2017/746/oj.

European Parliament. Texts Adopted - Artificial Intelligence Act - Wednesday, 13 March 2024. 2024a. "Regulation (50) of Position of the European Parliament Adopted at First Reading on 13 March 2024 with a View to the Adoption of Regulation (EU) 2024/ (Artificial Intelligence Act),." Www.europarl.europa.eu. March 13, 2024. https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html#title2.

European Parliament. 2024b. "Regulation (58) of Position of the European Parliament Adopted at First Reading on 13 March 2024 with a View to the Adoption of Regulation (EU) 2024/ (Artificial Intelligence Act),." Www.europarl.europa.eu.

March 13, 2024. https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html#title2.

European Union. 2000. “Charter of Fundamental Rights of the European Union .” https://www.europarl.europa.eu/charter/pdf/text_en.pdf.

European Union. 2017. “EUR-Lex - 32017R0745 - EN - EUR-Lex, REGULATION (EU) 2017/745 of the EUROPEAN PARLIAMENT and of the COUNCIL.” Europa.eu. Official Journal of the European Union. April 5, 2017. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745.

European, Commission. 2020. “On Artificial Intelligence -A European Approach to Excellence and Trust White Paper on Artificial Intelligence a European Approach to Excellence and Trust.” https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_en?filename=commission-white-paper-artificial-intelligence-feb2020_en.pdf..

GDPR. 2018. “General Data Protection Regulation (GDPR).” General Data Protection Regulation (GDPR). 2018. https://gdpr-info.eu/.

Official Journal of the European Union. 2006. “DIRECTIVE 2006/54/EC of the EUROPEAN PARLIAMENT and of the COUNCIL of 5 July 2006 on the Implementation of the Principle of Equal Opportunities and Equal Treatment of Men and Women in Matters of Employment and Occupation (Recast).” https://eige.europa.eu/sites/default/files/ep_council_eu_directive_2006-54-ec.pdf.

United Nations. 1948. “Universal Declaration of Human Rights.” United Nations. 1948. https://www.un.org/en/about-us/universal-declaration-of-human-rights.

United Nations. 1965. “International Convention on the Elimination of All Forms of Racial Discrimination.” OHCHR. United Nations. December 21, 1965. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-convention-elimination-all-forms-racial.

United Nations. 1966. "International Covenant on Economic, Social and Cultural Rights." United Nations. General Assembly resolution 2200A (XXI). December 16, 1966. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights.

United Nations General Assembly. 1979. "Convention on the Elimination of All Forms of Discrimination against Women New York, 18 December 1979." OHCHR. United Nations. December 18, 1979. https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-elimination-all-forms-discrimination-against-women.

United Nations. Convention on the Elimination of All Forms of Discrimination against Women (CEDAW). 2003. "UNITED NATIONS INTER-PARLIAMENTARY UNION the Convention on the Elimination of All Forms of Discrimination against Women and Its Optional Protocol Handbook for Parliamentarians." http://archive.ipu.org/pdf/publications/cedaw_en.pdf.

8.2 Secondary Sources

Abbaoui, Wafae, Sara Retal, Brahim El Bhiri, Nassim Kharmoum, and Soumia Ziti. 2024. "Towards Revolutionizing Precision Healthcare: A Systematic Literature Review of Artificial Intelligence Methods in Precision Medicine." *Informatics in Medicine Unlocked* 46 (101475): 101475. https://doi.org/10.1016/j.imu.2024.101475.

Agnieszka KOZŁOWSKA. 2017. "REPORT on the Application of Council Directive 2004/113/EC Implementing the Principle of Equal Treatment between Men and Women in the Access to and Supply of Goods and Services | A8-0043/2017 | European Parliament." Europa.eu. 2017. https://www.europarl.europa.eu/doceo/document/A-8-2017-0043_EN.html#_section2.

Altman, Andrew. 2011. "Discrimination (Stanford Encyclopedia of Philosophy)." Stanford.edu. February 1, 2011. https://plato.stanford.edu/entries/discrimination/.

Amdiju, Natasha, Zhaneta Independent consultan PhD Poposka, and OSCE Mission to Skopje . 2023. "GUIDELINES on SHIFTING the BURDEN of PROOF of the COMMISSION for PREVENTION and PROTECTION against DISCRIMINATION." https://www.osce.org/files/f/documents/f/5/548245.pdf.

Amnesty International. 2019. "Discrimination." Amnesty.org. 2019. https://www.amnesty.org/en/what-we-do/discrimination.

Badnjević, Almir, Halida Avdihodžić, and Lejla Gurbeta Pokvić. 2022. "Artificial Intelligence in Medical Devices: Past, Present and Future." *Science, Art and Religion* 1 (1-2): 101–6. https://doi.org/10.5005/sar-1-1-2-101.

Basu, Kanadpriya, Ritwik Sinha, Aihui Ong, and Treena Basu. 2020. "Artificial Intelligence: How Is It Changing Medical Sciences and Its Future?" *Indian Journal of Dermatology* 65 (5): 365–70. https://doi.org/10.4103/ijd.IJD_421_20.

Bauer, Greta R., and Daniel J. Lizotte. 2021. "Artificial Intelligence, Intersectionality, and the Future of Public Health." *American Journal of Public Health* 111 (1): 98–100. https://doi.org/10.2105/ajph.2020.306006.

Bekbolatova, Molly, Jonathan Mayer, Chi Wei Ong, and Milan Toma. 2024. "Transformative Potential of AI in Healthcare: Definitions, Applications, and Navigating the Ethical Landscape and Public Perspectives." *Healthcare* 12 (2): 125–25. https://doi.org/10.3390/healthcare12020125.

Benjamin, R. 2019. *Race after Technology: Abolitionist Tools for the New Jim Code*. Polity.

Bhati, Deepak. 2023. "Improving Patient Outcomes through Effective Hospital Administration: A Comprehensive Review." *Cureus* 15 (10): 1–12. https://doi.org/10.7759/cureus.47731.

Bohr, Adam, and Kaveh Memarzadeh. 2020. "The Rise of Artificial Intelligence in Healthcare Applications." *Artificial Intelligence in Healthcare* 1 (1): 25–60. https://doi.org/10.1016/B978-0-12-818438-7.00002-2.

Commissioner, Office of the. 2019. "FDA Rules and Regulations." FDA. May 7, 2019. https://www.fda.gov/regulatory-information/fda-rules-and-regulations.

Crenshaw, Kimberlé. 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 (1): 139–67.

Crenshaw, Kimberle . 1991. "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review* 43 (6): 1241–99.

Dhunnoo, Pranavsingh. 2019. "Your Guide to Facial Recognition Technology in Healthcare." The Medical Futurist. November 19, 2019. https://medicalfuturist.com/your-guide-to-facial-recognition-technology-in-healthcare/.

Directorate-General for Communications Networks, Content and Technology (European Commission), and PwC. 2021. *Study on EHealth, Interoperability of Health Data and Artificial Intelligence for Health and Care in the European Union: Final Study Report. Lot 2, Artificial Intelligence for Health and Care in the EU*. *Publications Office of the European Union*. LU: Publications Office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/fb8d8ec2-55a0-11ed-92ed-01aa75ed71a1.

Ellis, Evelyn, and Philippa Watson. 2012. "Key Concepts in EU Anti-Discrimination Law." *Oxford University Press EBooks, 2nd Edn, Oxford European Union Law Library*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199698462.003.0004.

Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. 2017. "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks." *Nature* 542 (7639): 115–18. https://doi.org/10.1038/nature21056.

European Commission. 2021. “Artificial Intelligence in Healthcare Report | Shaping Europe’s Digital Future.” Digital-Strategy.ec.europa.eu. November 11, 2021. https://digital-strategy.ec.europa.eu/en/library/artificial-intelligence-healthcare-report.

European Commission. n.d. “Overview.” Health.ec.europa.eu. Public Health. Accessed August 23, 2024. https://health.ec.europa.eu/medical-devices-sector/overview_en.

European Commission, Eurostat. n.d. “Self-Reported Unmet Need for Medical Care, by Sex, EU, 2010-2022.” Https://Ec.europa.eu/Eurostat/Statistics-Explained/Index.php?Title=File:Self-Reported_unmet_need_for_medical_care,_by_sex,_EU,_2010-2022.Png. Accessed August 24, 2024.

European Court of Human Rights. 2016. “HUDOC - European Court of Human Rights: GRAND CHAMBER CASE of BIAO v. DENMARK (Application No. 38590/10).” Coe.int. May 24, 2016. https://hudoc.echr.coe.int/fre#.

European Court of Human Rights. 2024. “Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention Prohibition of Discrimination.” Council of Europe/European Court of Human Rights.

European Medicines Agency. 2019. “Medical Devices | European Medicines Agency.” Www.ema.europa.eu. 2019. https://www.ema.europa.eu/en/human-regulatory-overview/medical-devices.

European Parliament. 2010. “Equality between Men and Women | Fact Sheets on the European Union | European Parliament.” Europa.eu. 2010. https://www.europarl.europa.eu/factsheets/en/sheet/59/equality-between-men-and-women.

European Union Agency for Fundamental Rights, 2023. 2023. “Being Black in the EU - Experiences of People of African Descent EU Survey on Immigrants and Descendants of Immigrants.” https://fra.europa.eu/sites/default/files/fra_uploads/fra-2023-being-black_in_the_eu_en.pdf.

Garson, G. David . 1998. “G. David Garson: Neural Networks: An Introductory Guide for Social Scientists.” Jasss. London: Sage Publications. 1998. https://www.jasss.org/4/3/reviews/garson.html.

Gerhard, Brewka, Jürgen Dix, and Kurt Konolige. 1997. “Nonmonotonic Reasoning: An Overview. .” *Stanford: CSLI Publications* 73.

Goodfellow. 2016. “Goodfellow, I., et Al. (2016) Deep Learning. MIT Press, Cambridge, MA. - References - Scientific Research Publishing.” Www.scirp.org. 2016. https://www.scirp.org/reference/referencespapers?referenceid=2791883.

Granmar, Claes G. 2021. “Global Applicability of the GDPR in Context.” *Oup.com* 11 (3): 225–44. https://doi.org/10.1093.

Hacker, P. 2018. “Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Decision-Making under EU Law. 55 Common Market Law Review 1143 .”

Hamon, R., H. Junklewitz, and I. Sanchez. 2020. “Robustness and Explainability of Artificial Intelligence.” European Commission, Brussels: JRC Technical Report, EU Science Hub.

Hug, AZ. 2020. “A Right to a Human Decision 105 Virginia Law Review .”

“Intersectionality (N.), .” 2023. In *Oxford English Dictionary*, sense 2.

Intersoft Consulting. 2013. “General Data Protection Regulation (GDPR) – Final Text Neatly Arranged.” General Data Protection Regulation (GDPR). 2013. https://gdpr-info.eu/art-25-gdpr/.

Katyal, S. K. 2019. “Private Accountability in the Age of Artificial Intelligence.” *UCLA Law Review*.

Khaitan, Tarunabh. 2015. “A Theory of Discrimination Law.” Ssrn.com. May 15, 2015. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2628112.

Kristensen, Tina Virenfeldt. 2020. "New Project Will Help to Prevent Bias and Discrimination in Medical AI." Edited by Katarzyna Wac. Di.ku.dk. Department of Computer Science University of Copenhagen Universitetsparken 1 DK-2100 Copenhagen . May 22, 2020. https://di.ku.dk/english/news/2020/new-project-will-help-to-prevent-bias-and-discrimination-in-medical-ai/.

Kulk, S., S. Van Deursen, and and Others. 2020. "Juridische Aspecten van Algoritmen Die Besluiten Nemen." Een Verkennend Onderzoek WODC.

Lee, J. D., and Katrina A. See. 2004. "Trust in Automation: Designing for Appropriate Reliance." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46 (1): 50–80. https://doi.org/10.1518/hfes.46.1.50_30392.

Lekadir, Karim , Gianluca Quaglio, and Anna Tselioudis Garmendia. 2022. "Artificial Intelligence in Healthcare Applications, Risks, and Ethical and Societal Impacts." EPRS | European Parliamentary Research Service: European Parliament .

Lekadir, Karim, Gianluca Quaglio, Anna Tselioudis Garmendia, and Chaterine Gallin. 2022. "Artificial Intelligence in Healthcare: Applications, Risks, and Ethical and Societal Impacts." Study by the Panel for the Future of Science and Technology (STOA), European Parliamentary Research Service (EPRS), PE 729.512.

Mehta, Neil, and Murthy V. Devarakonda. 2018. "Machine Learning, Natural Language Programming, and Electronic Health Records: The next Step in the Artificial Intelligence Journey?" *Journal of Allergy and Clinical Immunology* 141 (6): 2019-2021.e1. https://doi.org/10.1016/j.jaci.2018.02.025.

Noble, S. U.,. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.

Office of the High Commissioner for human rights . 1989. "CCPR General Comment No. 18: Non-Discrimination." Refworld. November 10, 1989. https://www.refworld.org/legal/general/hrc/1989/en/6268.

Parikh, Aparna R., Emily E. Van Seventer, Giulia Siravegna, Anna V. Hartwig, Ariel Jaimovich, Yupeng He, Katie Kanter, et al. 2021. "Minimal Residual Disease Detection Using a Plasma-Only Circulating Tumor DNA Assay in Patients with Colorectal Cancer." *Clinical Cancer Research* 27 (20): 5586–94. https://doi.org/10.1158/1078-0432.CCR-21-0410.

Pascoe, Elizabeth A., and Laura Smart Richman. 2009. "Perceived Discrimination and Health: A Meta-Analytic Review." *Psychological Bulletin* 135 (4): 531–54. https://doi.org/10.1037/a0016059.

Samuels, Gina E. Miranda, and Fariyal Ross-Sheriff . 2008. "Identity, Oppression, and Power: Feminisms and Intersectionality Theory." *Journal of Womenand Social Work* 23 (1). https://doi.org/10.1177/0886109907310475.

Schneeberger, David, Karl Stöger, and Andreas Holzinger. 2020. "The European Legal Framework for Medical AI." *Lecture Notes in Computer Science* 12279: 209–26. https://doi.org/10.1007/978-3-030-57321-8_12.

Sharma, Ajay N., and Bhupendra C. Patel. 2021. "Laser Fitzpatrick Skin Type Recommendations." PubMed. Treasure Island (FL): StatPearls Publishing. 2021. https://www.ncbi.nlm.nih.gov/books/NBK557626/.

Stöger, Karl, David Schneeberger, and Andreas Holzinger. 2021. "Medical Artificial Intelligence." *Communications of the ACM* 64 (11): 34–36. https://doi.org/10.1145/3458652.

Stokes, Jonathan M., Kevin Yang, Kyle Swanson, Wengong Jin, Andres Cubillos-Ruiz, Nina M. Donghia, Craig R. MacNair, et al. 2020. "A Deep Learning Approach to Antibiotic Discovery." *Cell* 180 (4): 688-702.e13. https://doi.org/10.1016/j.cell.2020.01.021.

Sudat, Sylvia E K, Paul Wesson, Kim F Rhoads, Stephanie Brown, Noha Aboelata, Alice R Pressman, Aravind Mani, and Kristen M J Azar. 2022. "Racial Disparities in Pulse Oximeter Device Inaccuracy and Estimated Clinical Impact on COVID-19 Treatment Course." *American Journal of Epidemiology* 192 (5). https://doi.org/10.1093/aje/kwac164.

The White House. 2023. "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence.". October 30, 2023. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

Tobler, Christa . 2008. "Limits and Potential of the Concept of Indirect Discrimination. European Network of Legal Experts in the Non-Discrimination Field." Luxembourg: Office for Official Publications of the European Communities.

Treffinger, Stephen . 2023. "The New Wave of AI in Healthcare - NYAS." The New York Academy of Sciences. May 25, 2023. https://www.nyas.org/ideas-insights/blog/the-new-wave-of-ai-in-healthcare/.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124–31.

Ulnicane, Inga, and Aini Aden. 2023. "Power and Politics in Framing Bias in Artificial Intelligence Policy." *Review of Policy Research* 40 (5). https://doi.org/10.1111/ropr.12567.

UNESCO. 2022. "Ethics of Artificial Intelligence." Www.unesco.org. UNESCO. 2022. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics.

Vetzo, MJ, JH Gerards, and R Nehmelman . 2018. "Algoritmes En Grondrechten (Algorithms and Fundamental Rights) ." Den Haag, Boom Juridisch.

Wachter, Sandra, Brent Mittelstadt, and Chris Russell. 2020. "Why Fairness Cannot Be Automated: Bridging the Gap between EU Non-Discrimination Law and AI." *SSRN Electronic Journal* 105567. https://doi.org/10.2139/ssrn.3547922.

Washington, Ariel, and Jill Randall. 2022. "'We're Not Taken Seriously': Describing the Experiences of Perceived Discrimination in Medical Settings for Black Women."

Journal of Racial and Ethnic Health Disparities 10 (2). https://doi.org/10.1007/s40615-022-01276-9.

Wolford, Ben. 2024. “What Is GDPR, the EU’s New Data Protection Law?” GDPR.eu. 2024. https://gdpr.eu/what-is-gdpr/.

World Commission on the Ethics of Scientific Knowledge and Technology. 2019. “Preliminary Study on the Ethics of Artificial Intelligence.” Unesco.org. 2019. https://unesdoc.unesco.org/ark:/48223/pf0000367823.

Yu , Kun-Hsing, Andrew L Beam, and Isaac S Kohane. 2018. “Artificial Intelligence in Healthcare.” *Nature Biomedical Engineering* 10 (PMID: 31015651.): 719–31. https://doi.org/10.1038/s41551-018-0305-z..

ZÁBORSKÁ, Anna. 2015. “REPORT on the Application of Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the Implementation of the Principle of Equal Opportunities and Equal Treatment of Men and Women in Matters of Employment and Occupation | A8-0213/2015 | European Parliament.” Www.europarl.europa.eu. June 25, 2015. https://www.europarl.europa.eu/doceo/document/A-8-2015-0213_EN.html.

8.3 Cases

Buolamwini, Joy, Timnit Gebru, Sorelle Friedler, and Christo Wilson. 2018. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification *.” *Proceedings of Machine Learning Research* 81: 1–15.

Gerards (Utrecht University) , Janneke , and Raphaële Xenidis (University of Edinburgh and University of Copenhagen) . 2021. “Algorithmic Discrimination in Europe: Challenges and Opportunities for Gender Equality and Non-Discrimination Law.” European Commission.

Plaisime, Marie V. 2023. “Invited Commentary: Undiagnosed and Undertreated—the Suffocating Consequences of the Use of Racially Biased Medical Devices during the

COVID-19 Pandemic." *American Journal of Epidemiology* 192 (5).
https://doi.org/10.1093/aje/kwad019.

Abstract

The integration of artificial intelligence (AI) into the healthcare sector is rapidly advancing, promising significant improvements in diagnostic accuracy, personalized treatment approaches, and overall patient outcomes. However, this progress brings critical challenges, particularly in regards to racial and gender discrimination. This study examines the legal obligations of states under international and European human rights law to address discrimination in medical technologies. By identifying gaps in the representation of Black women in AI medical devices and exploring limitations in data that lead to unreliable diagnostic systems, the research highlights disparities in medical treatments. It aims to detect inconsistencies between AI-based medical technologies and European human rights standards and offers recommendations to promote equality in healthcare.

Abstrakt

Die Integration von künstlicher Intelligenz (KI) in den Gesundheitssektor schreitet schnell voran und verspricht erhebliche Verbesserungen in der diagnostischen Genauigkeit, personalisierten Behandlungsmethoden und den allgemeinen Patientenergebnissen. Dieser Fortschritt bringt jedoch kritische Herausforderungen mit sich, insbesondere in Bezug auf Rassen- und Geschlechterdiskriminierung. Diese Studie untersucht die rechtlichen Verpflichtungen der Staaten gemäß internationalem und europäischem Menschenrecht, um Diskriminierung in medizinischen Technologien zu bekämpfen. Durch die Identifizierung von Lücken in der Darstellung von Frauen mit dunkler Hautfarbe in KI-basierten medizinischen Geräten und die Untersuchung von Datenbeschränkungen, die zu unzuverlässigen Diagnosesystemen führen, hebt die Forschung Ungleichheiten in der medizinischen Behandlung hervor. Ziel ist es, Inkonsistenzen zwischen KI-basierten medizinischen Technologien und europäischen Menschenrechtsstandards zu erkennen und Empfehlungen zur Förderung der Gleichstellung im Gesundheitswesen zu geben.