A Survey of Bias in Healthcare: Pitfalls of Using Biased Datasets and Applications

Bojana Velichkovska, Daniel Denkovski, Hristijan Gjoreski, Marija Kalendar, Venet Osmani

[1] Faculty of Electrical Engineering and Information Technologies, Ss. Cyril and Methodius University in Skopje, N. Macedonia
[2] Fondazione Bruno Kessler, Trento, Italy

{bojanav, danield, hristijang, marijaka}@feit.ukim.edu.mk, vosmani@fbk.eu

Abstract. Artificial intelligence (AI) is widely used in medical applications to support outcome prediction and treatment optimisation based on collected patient data. With the increasing use of AI in medical applications, there is a need to identify and address potential sources of bias that may lead to unfair decisions. There have been many reported cases of bias in healthcare professionals, medical equipment, medical datasets, and actively used medical applications. These cases have severely impacted the quality of patients' healthcare, and despite awareness campaigns, bias has persisted or in certain cases even exacerbated. In this paper, we survey reported cases of different forms of bias in medical practice, medical technology, medical datasets, and medical applications, and analyse the impact these reports have in the access and quality of care provided for certain patient groups. In the end, we discuss possible pitfalls of using biased datasets and applications, and thus, provide the reasoning behind the need for robust and equitable medical technologies.

Keywords: healthcare bias, attitudes of healthcare professionals, biased datasets, medical applications

1 Introduction

With the rapid development of artificial intelligence (AI) and the efficiency which AI offers, there are entire industries that rely on its application in solving everyday challenges. Advantageously, AI is an important part of a vast range of actively-used applications, spanning from social media and personalized recommender systems, all the way to smart homes, smart cars, surveillance, and so on. With that, the benefits offered by AI are manifold. However, there are risks that social media and recommender systems can use information to influence user's opinions, especially for high-stakes events, such as elections.

These issues then raise the question, “With AI having a high error margin, what happens when that same AI is employed to sensitive areas such as medicine?” There is a plethora of actively-used medical applications based on artificial intelligence [1]. Namely, AI is widely used in medical setting, with applications in everything from patient care and maintaining medical records to billing. Therefore, the use of AI in medicine directly, and sometimes indirectly, determines the treatments given to patients, and consequently, patients’ outcomes. One of the many applications of AI in medicine nowadays is the identification and diagnosis of different medical conditions [2]. A subbranch of AI in diagnostic capacity is medical imaging diagnosis, where AI algorithms are taught to recognize complex patterns in imaging data and provide health assessments of the patients’ conditions, which as a diagnostic tool has had large success in past years [3]. Other applications include, but are not limited to, personalised medicine [4], smart health records [5], clinical trial and research [6], drug discovery and manufacturing [7], etc.

However, technology is prone to different forms of malfunction, so it is expected that issues can arise with actively-used applications. Of all potential issues, there are some which happen as an unintentional and unexpected by-product of AI, and these issues are therefore more serious than others, because they are “silent” or “hidden”. With this, the use of AI in medicine and healthcare is full of potential for clinical, social, and ethical conflicts. Namely, there is a risk of patient harm due to prevailing errors in AI models, influenced by biased inequities in the health system, exacerbated by lack of transparency in patient selection when creating medical datasets, as well as the evident lack of transparency in development of AI-based medical applications [8].

With these risks in mind, it becomes important to understand healthcare biases in medical practice, and address the potential pitfalls of using biased datasets and biased AI applications. Furthermore, increasing awareness would lead to improving application of AI in clinical practices and consequently better outcomes.

This paper is organised into four sections. Section two gives the definition of bias and the different types of bias possible in medical AI. The next section gives a survey of reported cases of bias in medical datasets and medical AI-based applications. Section four gives a summary of the different types of bias and the lessons learned. The paper concludes in section five.

2 Bias and different types of bias

According to the Merriam Webster dictionary, bias is defined as:

- a personal and often unreasoned judgment for or against one side;
- an unreasoned and unfair distortion of judgment in favour of or against a person or thing;
- a settled and predictable leaning in one direction and connotes unfair prejudice;
- a systematic error introduced into sampling or testing by selecting or encouraging one outcome or answer over others.

Therefore, the meaning of the word bias can be surmised into a propensity to show partiality towards certain persons or groups in favour of others, which often negatively impacts the marginalised group. From this definition, we can infer that the presence of bias in medical setting would mean overlooking a group of patients discriminatively, which can result in significantly poorer health services to certain patient groups that might lead to long-term complications, impairments, and in worst case scenarios even deaths which could have been prevented.

According to a recent presentation on AI in healthcare delivered by Marzyeh Ghassemi, bias is already present in the clinical landscape [9]. The bias present in the healthcare system can: one, exist consciously and be exhibited through prejudiced ideas such as racism and sexism; two, be unconscious, but cemented thoughts based on learned stereotypes; or, three, can happen mistakenly by being based on conclusions drawn only from working with a uniform portion of the population. With either of the listed, there are different forms of bias which can occur: racial bias, bias based on sex, gender or sexual identity, socioeconomical bias, educational bias, as well as bias arising from geographical location, overweight and obesity, or age (see Fig. 1). Some forms of bias are more dominant than other, but each have manifold consequences that can negatively impact certain patient groups.

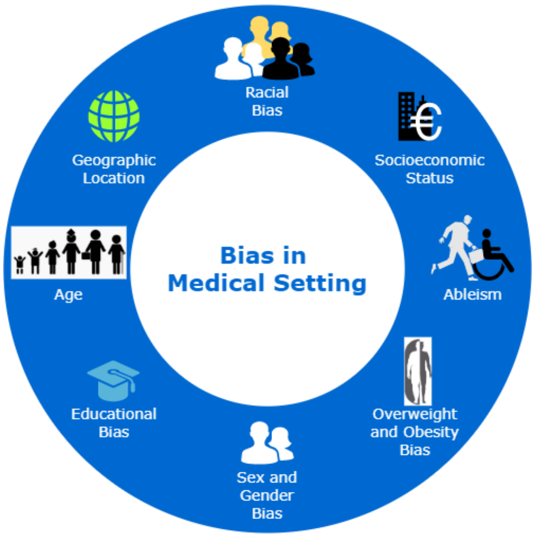

Fig. 1. Bias in Medical Setting: Different types of bias present.

3 Bias in Medicine

AI-related bias in medical setting can be dissected from four distinct aspects: data-driven, algorithmic, technological, and human [10]. Essentially, all these aspects are, directly or indirectly, linked to the human aspect in medicine, wherein biased medical decisions, attitudes, and behaviour are a daily occurrence, and unfortunately also get propagated and taught to medical students [11].

3.1 Human Aspect of Bias

The first direction we decided to investigate was the human aspect of bias. According to the key findings presented in the U.S. National Healthcare and Disparities Report [12] conducted in 2019, it was observed that in spite of efforts disparities in medicine have persisted and some have even worsened, mostly for poorer and uninsured populations. The report also showed disparities based on residence location. Additionally, there were racial and ethnic disparities, i.e., White patients were found to have received better care compared to other racial and ethnic groups: Blacks, American Indians, Alaska Natives, Hispanics, Asians, and Native Hawaiians (Pacific Islanders). In [13] the authors discussed how people of colour in the U.S. face disparities in access to healthcare, the quality of the care received, which consequently impacts the health outcome of these patients. According to [14], people of colour face more barriers in accessing care, which means less access to preventive services and treatments, and even when access to care is secured, patients of colour tend to have unsatisfactory interactions with healthcare providers. Racial bias has extended even to patient recommendations for bypass surgery [15]. Namely, some physicians were more likely to recommend White patients over Black patients, because they believed that Black patients would not adhere to the necessary physical activity needed after surgery.

There have also been reported cases of sex and gender bias. A study [16] showed that medical professionals were more likely to dismiss chronic pain in women than men, expressing it as difference of “brave men” compared with “emotional women”. These biases can lead to the silencing of patients in addressing important health problems. Such is also the case with transgender people, who feel reluctant to receive proper healthcare due to expected unfair treatment [17].

Other types of bias are also present in behaviour, attitudes, and opinions of medical professionals. Namely, medical professionals find working with older patients and their families as challenging, and have described these patients as demanding and offensive, and wanting to manage their own treatment [18]. Furthermore, disabled patients have limited access to certain areas in healthcare centres. The authors of [19] discussed how over 80% of medical professionals would rather work with people without disabilities. People with obesity are also likely to receive poor treatment by their health provider, as well as have their symptoms attributed to their weight [20]. People with lower socioeconomic status are more likely to experience delays in testing and treatment, which creates issues with regular and quality preventive treatments [21].

Altogether, the presented types of bias result in reduction of healthcare access and quality for certain patient groups based on prejudiced opinions and beliefs deeply rooted in the behaviour of medical personnel. Therefore, these groups are exposed to different serious risks which directly affect their health, due to delayed or non-existent treatment, incorrect diagnosis, which might overlook serious conditions and complications.

3.2 Bias in Medical Technology

Reporting of bias has also been extended to medical technology actively used in practice to evaluate patients. Inherent in the medical technology itself, due to the fact that the device's testing did not include diverse population, this type of bias has been long overlooked, whilst medical technologies incorporating it were being constantly used and were directly influencing patient statistics.

The first reported bias in medical technology dates back to 1968 [22]. Namely, guidelines given by manufacturers indicated the need for higher radiation exposure for Black patients, which was absorbed into the medical practice and resulted with X-ray technicians routinely exposing Black patients to higher doses of radiation compared to those received by White patients. Naturally, the guidelines were based on the false belief that Black patients have denser bones and thicker skin.

A study [23] found that forehead thermometers, which measure temperature through the skin using infrared technology, were 26% less likely to detect fever for Black patients compared to oral thermometers. Another study [24], which gained attention during the COVID-19 pandemic, found that pulse oximeters were three times less likely to detect low oxygen levels in Black patients, which delays necessary treatment and puts patients at risk.

In actuality, bias in medical equipment, and one that is being used on daily bases, can have unforgivable consequences for underserved patients. The problems with the equipment lead to penalisation of certain population groups, which can contribute to delay in diagnosis or treatment administration, that might lead to fatalities.

3.3 Bias in Medical Datasets

If left unchecked, systemic human biases, stigmatic opinions and bias in medical technology can be incorporated in medical datasets and AI-based medical algorithms, and heighten the presence of bias in a wide spectre of applications developed with the aim of assisting and bettering the healthcare process and experience for sick patients.

Biased datasets are either full of biased markers or have underrepresentation of certain patient groups, which can stem from one or more in a list of reasons:

- systematic discrimination arising from unequitable treatment of patients from medical personnel due to racial, socioeconomic, or additional aspects;
- bias imbedded during the data collection process;
- lack of diversity and interdisciplinarity in accessing medical equipment and technology quality;
- or simply, lack of detailed quality investigations in obtained technological, clinical, and scientific research.

Collecting data from medical institutions without mitigating biased opinions, practices, and treatments can effectively lead to bias in medical datasets. Considering the already presented issues in behaviour of medical professionals, it is understandable why there have been many reported cases of bias in medical datasets. Biased datasets can come from transference from implicit bias of medical professionals, or implicit bias during the data selection process and underrepresentation of diverse patient groups.

Data has to be representative of population variety, otherwise it can reinforce lack of generalisation and different forms of bias [10] [25]. In [26] the authors investigated the history and physical notes from 18,259 patients that were collected in an urban academic medical centre. The study analysed the presence of negative descriptors, like noncompliant or resistant, regarding patients and their behaviour. The results showed presence of racial bias in the analysed electronic health records; namely, Black patients were 2.5 times more likely to be given negative descriptors compared with White patients.

Medical imaging datasets have also been scrutinised for being ladened with different forms of bias. The authors in [27] address the importance of gender balance in medical imaging datasets by showing consistent decrease of performance for underrepresented genders. The authors also investigated the influence which imbalanced datasets had on model performance. They conclude that when working with 25%/75% imbalance ratio between classes model performance across the minority class is significantly lower compared to the majority class. On the other hand, that difference was not observed in balanced datasets.

Medical imaging datasets have been investigated for presence of racial and ethnic imbalance. The National Lung Screening Trial collected data from 53,000 smokers to investigate lung cancer diagnosis [28], however from the selected patients only 4% were Black. Another dataset targeted for its biased data is the International Skin Imaging Collaboration. The dataset is one of the most used open-access dataset on skin lesions in the diagnostic process of melanoma which the most serious form of skin cancer; however, the data was collected from mostly fair-skinned patients [29]. Another form of bias is geographic bias, as pointed by the researchers in [30]. The study was conducted in 2020 and it showed that 71% of studies in U.S. where geographic location was present were using data only from three states: California, Massachusetts, and New York. Additionally, they found that conducted studies used data from only 16 countries, whereas there were no datasets available from the remaining 34 countries.

Altogether, the presented cases in which bias has been noted in medical datasets show that open-access data available for researchers can be significantly limited on patients from a certain geographical area, and belonging to one dominant race or gender. This impacts uniform patient representation; thus, datasets carry limited knowledge which does not allow for thorough understanding of medical conditions and subtle changes which might occur across different patient populations.

3.4 Bias in AI-based Medical Applications

When AI algorithms use biased datasets during the training process, the algorithms have a limited view into the problem and have better understanding of the problem from the perspective of the dominantly present group. Therefore, the models can learn the bias which the data incorporates, and have lower performance accuracy over certain patient groups. Consequently, the trained algorithms reinforce inequities in healthcare in everything from cancer-detection algorithms which are less effective for Black patients [31] to cardiac risk-scores that underestimate the amount of care needed by Black patients. There have been reports of bias in algorithms used for maternal health.

Namely, a widely-used Vaginal Birth after Caesarean (VBAC) algorithm contributed to higher rates of c-section among women of colour because it was predicting lower successful VBAC rates for pregnant women of colour [32].

A case study [33] showed evidence of racial bias in an actively used algorithm, which carried decisions for more than 200 million people in the U.S. The origin of the bias came from using health costs as a proxy for health needs. Since less money was being spent on Black patients who have the same level of need as White patients, the algorithm learned this discrepancy, and therefore, assigned the same level of risk scores to White and Black patients, even though the Black patients were in worse medical condition compared to the White patients. According to an estimation made by the authors of the study, the number of Black patients identified for extra care was reduced 2.5 times compared with what it should have been.

Another study [34] evaluated the performance of state-of-the-art algorithms in detecting abnormalities (e.g., pneumonia, lung nodules, lesions, fractures, etc.) in chest X-rays. The results showed that young females had the highest rate of underdiagnosis, followed by Black patients, then by patients with public health insurance due to low income. This was even further pronounced with patients with fusion of more than one of the listed criteria, i.e., a Black woman with public insurance and low-income background had the highest rate of underdiagnosis.

An investigation into an AI-based tool for early detection of sepsis, actively used by more than 170 hospitals, showed the model's inability in predicting this life-threatening illness in 67% of patients who developed it [35]. Furthermore, the model also generated false sepsis alerts on thousands of patients who did not develop the illness.

Another algorithm was criticised for [36] suggesting extreme cuts to in-home care of disabled patients, which caused extreme disruption of patients' lives and resulted with increased hospitalisation. Another study which reported a form of socioeconomic bias aimed to assess the degree to which data quality of electronic health records related to socioeconomic status [37]. The machine learning models investigated in the study aimed to predict asthma exacerbation in children. The results of the study showed worse predictive model performance in patients with lower socioeconomic status. An AI-based model for Alzheimer's diagnosis from audio data, built in Canada, underperformed for patients with certain accents because the training process included speech samples from one accent, therefore making the application unusable for everyone else in the country [38].

Altogether, AI-based medical applications with biased performance across different patient groups have been widely reported only after actively being used in medical setting and severely impacting the quality of care offered to patients. Many reported cases have endangered patients' lives by missing disease diagnosis in life-threatening situations. Other cases show undue stress inflicted to patients by inaccurate diagnosis of illnesses which are later proven as non-existent. Furthermore, the algorithms' flaws are more expressed in patients which have a diminished access to healthcare therefore creating severe difficulties for them, i.e., patients with low income and limited access to medical care cannot afford a second opinion; this makes erroneous diagnosis in these cases a heavy-handed and punitive action towards the patient.

4 Summary of Bias and Lessons Learned

The allotted sections and the papers discussed in them are summarised in Table 1. For each of the papers referenced, the table lists the type of bias, a brief description of the paper, the source of bias (medical practice, datasets, AI applications, with a more detailed and concrete scope), and the implications which can be drawn from the observed.

From the ample examples referenced, it is evident bias is present in medical practice and technology. In addition, that bias can easily transfer from medical practice to medical datasets, and eventually, to AI-based applications unless adequate actions are taken. Briefly, data presumed to reflect different population groups equitably has on many occasions failed to do just that, AI-based algorithms presumed to have equal performance across different population groups have proven biased in vast scope, and understandably so, when they are impacted by human influence that embeds societal prejudices against patients of different race, gender, appearance, socioeconomic status, etc. Therefore, from these lessons learned, the question of mitigating bias emerges.

The primary problem from which all others derive is the human aspect in medical practice. For that reason, the most important step to take is countering biased practices by imposing mandatory training of all medical personnel with the purpose of imprinting fundamental understanding of bias in health and consequences. Non-functional medical equipment in the 21st century is another huge problem. Diversity in testing new medical equipment before releasing it for mass production and ensuring that equipment cannot do any inadvertent harm to some populations is another must.

Creating datasets which have blind spots across race, gender, socioeconomic status, and so on, should not be available to everyone around the world. Enabling diverse and multi-disciplinary teams, when creating medical datasets, can be beneficial in reducing, and even maybe eradicating, those cultural or academic blind spots, and thus allowing for fair and equitable datasets.

Every stage after the data collection can also succumb to bias. In order to provide responsible algorithm development, steps must be taken in the pre-processing, in-processing, and post-processing stages. Namely, when working on a model the development teams should require maximising the accuracy of the model while at the same time minimising the influence of biased markers. Post-processing mitigation also helps, in that, through thorough analysis of the model performance across different population groups issues with AI-based applications can be detected in the early stages of production. That would allow models to be retrained or retuned to work equitably for everyone.

Table 1. Summary of bias in Healthcare.

Ref.	Bias Type	Description	Source	Implications
		Human Aspect of Bias		
[12]	socio-economic, location, racial and ethnic	worsened care for poorer populations; subpar care for some racial and ethnic groups	medical practice	bias mitigation strategies not entirely effective
[13]	racial and ethnic	people of colour facing disparities in healthcare access and quality	medical practice	reduced preventive care, impact on outcomes for patients of colour
[14]	racial and ethnic	people of colour facing barriers in accessing care	medical practice	reduced preventive care, impact on outcomes for patients of colour
[15]	racial and ethnic	patient recommendations for bypass surgery based on skin colour	medical practice	Black patients were denied bypass surgery
[16]	sex and gender	dismissal of chronic pain in women	medical practice	can lead to silencing patients on important health problems
[17]	sex and gender	transgender people feel reluctant to receive proper healthcare	medical practice	transgender people might avoid visiting a healthcare professional unless urgent
[18]	age	medical professionals find working with older patients demanding	medical practice	unfair treatment of older patients
[19]	ableism	over 80% of medical professionals would rather work with people without disabilities	medical practice	favouring working with able-bodied people can create an issue for disabled patients and access to quality, objective healthcare
[20]	obesity	people with obesity have their symptoms attributed to weight	medical practice	medical conditions can be overlooked
[21]	socio-economic	poorer people are more likely to experience delays in testing and treatment	medical practice	chances of worsening medical conditions and complications due to wait time
		Bias in Medical Technology		
[22]	racial and ethnic	higher radiation exposure for Black patients	equipment	absorbed in medical practice and routinely applied
[23]	racial and ethnic	forehead thermometers were 26% less likely to detect fever in Black patients compared with oral thermometers	equipment	missed fevers could lead to delays in diagnosis and treatment, and possibly cause an increased death rate in Black patients
[24]	racial and ethnic	pulse oximeters were three times less likely to detect low oxygen levels in Black patients	equipment	could lead to delays in diagnosis treatment, and possibly cause an increased death rate in Black patients
		Bias in Medical Datasets		
[26]	racial and ethnic	physical notes from 18,259 patients showed negative	medical practice transferred	2.5 times more negative descriptors for Black patients compared with White patients can

		descriptors for certain racial and ethnic groups	to medical datasets	lead to AI applications learning that discrepancy and operating on that bias
[27]	gender	decrease of model performance in case of underrepresented classes in datasets	medical datasets transferred to AI applications	working with imbalanced ratio significantly overlooks the minority of the population and can result with AI applications performing worse for certain population groups
[28]	racial and ethnic	only 4% of selected patients for cancer diagnosis dataset were Black	medical datasets	underrepresentation can lead to AI applications performing well only for certain ethnic and racial groups
[29]	racial and ethnic	dataset for skin cancer collected from mostly fair-skinned people	medical datasets	underrepresented populations that will very likely lead to AI applications unfamiliar with different skin colours and understanding cancer only for White patients
[30]	location	71% of studies with geographic location included came from only three states	medical datasets	underrepresentation of patients coming from certain areas can lead to AI models operating accurately only for the areas in the data
		Bias in AI-based Medical Applications		
[31]	racial and ethnic	cancer detection algorithm less effective for Black patients	medical applications	loss of lives (in Black patients) which could have been prevented
[32]	racial and ethnic	VBAC algorithm predicts higher rates of c-section among women of colour	medical applications	higher rates of potentially unnecessary procedures for women of colour
[33]	racial and ethnic	same risk scores were assigned to White and Black patients, even though Black patients were in worse medical condition	medical application	number of Black patients identified for extra care was reduced 2.5 times compared with what it should have been
[34]	socio-economic, sex and gender, racial and ethnic	chest X-ray showed young females had the highest rate of underdiagnosis, followed by Black patients, then by patients with public health insurance	medical application	disregard for serious illnesses
[35]	ableism	extreme cuts to in-home care of disabled patients	medical application	extreme disruption of patients' lives which resulted with increased hospitalisation
[36]	socio-economic	researched the degree to which data quality of electronic health records related to socioeconomic status	medical dataset transferred to medical application	worse predictive model performance in patients with lower socioeconomic status
[37]	linguistic	Alzheimer's diagnostic tool underperformed for patients with certain accents	medical dataset transferred to medical application	the application was unusable for a large population of people

In summary, ensuring accurate medical equipment and adequate data gathering with wide representation and accurate labelling is extremely important, since with faulty data little can be done to prevent bias transfer to the AI application. Furthermore, regulations must be followed when creating the application. Teams must be equipped to handle different aspects of a problem, which is why a vast array of diversity, knowledge, and understanding is a must. In the end, even after all precautions are taken, the model must be rigorously analysed for bias before being put into practice. Once all requirements for fairness are met, sharing details on how the model was developed is essential, for several reasons: one, it allows the research community to better understand the steps which should be taken in order to develop unbiased applications; two, it would account for how the model should be used; three, additional bias assessments can be conducted by impartial teams; four, transparency would help patients with trusting the process; and more.

The brevity of it all is, there are fundamental issues to be considered and corrected, and with urgency, as they impact lives all over the world. And, that change should come from us all.

5 Conclusion

Certain patient groups are marginalised due to different aspects pertaining to their gender and sexuality, the colour of their skin, their socioeconomic status, etc., which affects the quantity and quality of care which they are offered. Biased practices impact patients' healthcare, and patients are subjected to opinions and behaviour which negatively influence their quality of life. Therefore, detecting the presence of different types of bias in the healthcare system, namely biases in medical technology, behaviour of medical professionals, datasets collected from patients, and AI-based medical applications, as well as understanding the sources of existing bias are important and needed steps for improving healthcare access and the quality of care offered to different patient groups.

Therefore, in this work we illustrated different types of bias present in the healthcare systems, focusing on surveying papers which illustrate four different types of bias: in healthcare professionals, in the technology used for medical procedures, in the datasets collected in medical setting, and in AI-based medical applications for wide use. Our survey showed cases of different forms of bias which have had significant impact on patient lives around the world.

With this reflective analysis on AI technologies, we wish to raise awareness for the need of creating clinically robust and safe medical applications, built on widely-representative datasets, which successfully address ethical complaints and are transparent all throughout the development process, and therefore can be successfully and safely integrated in healthcare practices.

Acknowledgements

Part of the study was supported by WideHealth project – Horizon 2020, under grant agreement No 95227.

References

1. Park, C.W., Seo, S.W.: Artificial Intelligence in Health Care: Current Applications and Issues. J Korean Med Sci 35(42), 379 (2020).
2. Kumar, Y., Koul, A., Singla, R., Ijaz, M.F.: Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. J Ambient Intell Humaniz Comput. 1-28 (2022).
3. Oren, O., et al.: Artificial intelligence in medical imaging: switching from radiographic pathological data to clinically meaningful endpoints. The Lancet Digital Health 2(9), 486–488 (2020).
4. Johnson, K.B., Wei, W.Q., Weeraratne, D., Frisse, M.E., Misulis, K., Rhee, K., Zhao, J., Snowdon, J.L.: Precision Medicine, AI, and the Future of Personalized Health Care. Clin Transl Sci 14(1), 86–93 (2021).
5. Rayan, Z., Alfonse, M., Salem, AB.M.: Machine Learning Approaches in Smart Health. Procedia Computer Science 154, 361–368 (2019).
6. Weissler, E.H., Naumann, T., Andersson, T. et al.: The role of machine learning in clinical research: transforming the future of evidence generation. Trials 22, 537 (2021).
7. Paul, D,. Sanap, G., Shenoy, S., Kalyane, D., Kalia, K., Tekade, R.K.: Artificial intelligence in drug discovery and development. Drug Discov Today 26(1), 80–93 (2021).
8. Bernal, J., Mazo, C.: Transparency of Artificial Intelligence in Healthcare: Insights from Professionals in Computing and Healthcare Worldwide. Appl. Sci.12(20), 10228 (2022).
9. Ghassemi, M.: Exploring Healthy Models in ML for Health. In: AI for Healthcare Equity Conference, AI & Health. MIT (2021).
10. Norori, N., Hu, Q., Aellen, F.M., Faraci, F.D., Tzovara, A.: Addressing bias in big data and AI for health care: A call for open science. Patterns 2(10), 100347 (2021).
11. Brooks, K.C.: A piece of my mind. A silent curriculum. JAMA 313(19), 1909–1910 (2015).
12. 2019 National Healthcare Quality and Disparities Report. Rockville, MD. Agency for Healthcare Research and Quality. AHRQ Pub. No. 20(21)-0045-EF (2020).
13. Fiscella, K., Franks, P., Gold, M.R., Clancy, C.M.: Inequality in quality: addressing socioeconomic, racial, and ethnic disparities in health care. JAMA 283(19), 2579–2584 (2000).
14. 2013 National Healthcare Disparities Report. Rockville, MD. Agency for Healthcare Research and Quality, US Dept of Health and Human Services. AGRQ Pub. No. 14-0006 (2014).
15. Dovidio, J.F., Eggly, S., Albrecht, T.L., Hagiwara, N., Penner, L.: Racial biases in medicine and healthcare disparities. TPM-Testing, Psychometrics, Methodology in Applied Psychology 23(4), 489–510 (2016).
16. Samulowitz, A., Gremyr, I., Eriksson, E., Hensing, G.: "Brave Men" and "Emotional Women": A Theory-Guided Literature Review on Gender Bias in Health Care and Gendered Norms towards Patients with Chronic Pain. Pain Res Manag, 6358624 (2018).
17. Casey, L.S., Reisner, S.L., Findling, M.G., Blendon, R.J., Benson, J.M., Sayde, J.M., Miller, C.: Discrimination in the United States: Experiences of lesbian, gay, bisexual, transgender, and queer Americans. Health Serv Res 54 (Suppl 2), 1454–1466 (2019).

18. Ben-Harush, A., Shiovitz-Ezra, S., Doron, I., Alon, S., Leibovitz, A., Golander, H., Haron, Y., Ayalon, L.: Ageism among physicians, nurses, and social workers: findings from a qualitative study. Eur J Ageing 14(1), 39–48 (2016).
19. VanPuymbrouck, L., Friedman, C., Feldner, H.: Explicit and Implicit Disability Attitudes of Healthcare Providers. Rehabilitation Psychology 65 (2), 101–112 (2020).
20. Phelan, S.M., Burgess, D.J., Yeazel, M.W., Hellerstedt, W.L., Griffin, J.M., van Ryn M.: Impact of weight bias and stigma on quality of care and outcomes for patients with obesity. Obes Rev 16(4), 319–326 (2015).
21. Arpey, N.C., Gaglioti, A.H., Rosenbaum, M.E.: How Socioeconomic Status Affects Patient Perceptions of Health Care: A Qualitative Study. Journal of Primary Care & Community Health 8(3), 169–175 (2017).
22. Bavli, I., Jones, D.S.: Race Correction and the X-Ray Machine — The Controversy over Increased Radiation Doses for Black Americans in 1968. New England Journal of Medicine 387(10), 947–952 (2022).
23. Bhavani S.V., Wiley Z., Verhoef P.A., Coopersmith C.M., Ofotokun I.: Racial Differences in Detection of Fever Using Temporal vs Oral Temperature Measurements in Hospitalized Patients. JAMA 328(9), 885–886 (2022).
24. Sjoding, M.W., Dickson, R.P., Iwashyna, T.J., Gay, S.E., Valley, T.S.: Racial Bias in Pulse Oximetry Measurement. New England Journal of Medicine 383(25), 2477–2478 (2020).
25. Tasci, E., Zhuge, Y., Camphausen, K., Krauze, A.V.: Bias and Class Imbalance in Oncologic Data-Towards Inclusive and Transferrable AI in Large Scale Oncology Data Sets. Cancers (Basel) 14(12), 2897 (2022).
26. Sun, M., Oliwa, T., Peek, M.E., Tung, E.L.: Negative Patient Descriptors: Documenting Racial Bias In The Electronic Health Record. Health Affairs 41(2), 203–211 (2022).
27. Larrazabal, A.J., Nieto, N., Peterson, V., Milone, D.H., Ferrante, E.: Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. Proceedings of the National Academy of Sciences 117(23), 12592–12594 (2020).
28. Ferryman, K., Pitcan, M.: Fairness in Precission Medicine. (2018).
29. Adamson, A.S., Smith, A.: Machine Learning and Health Care Disparities in Dermatology. JAMA Dermatol 154(11), 1247–1248 (2018).
30. Kaushal, A., Altman, R., Langlotz, C.: Geographic Distribution of US Cohorts Used to Train Deep Learning Algorithms. JAMA 324(12), 1212–1213 (2020).
31. Sourlos, N., Wang, J., Nagaraj, Y., van Ooijen, P., Vliegenthart, R.: Possible Bias in Supervised Deep Learning Algorithms for CT Lung Nodule Detection and Classification. Cancers (Basel) 14(16), 3867 (2022).
32. Taylor, L.M.: Race-Based Prediction in Pregnancy Algorithm Is Damaging to Maternal Health. (2021).
33. Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S.: Dissecting racial bias in an algorithm used to manage the health of populations. Science 366(6464), 447–453 (2019).
34. Seyyed-Kalantari, L., Liu, G., McDermott, M., Chen, IY., Ghassemi, M.: CheXclusion: Fairness gaps in deep chest X-ray classifiers. Pac Symp Biocomput 26, 232–243 (2021).
35. Wong, A., et. al.: External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. JAMA Intern Med 181(8), 1065–1070 (2021).
36. Lutz, R.: Incident Number 110. in McGregor, S. (ed.) Artificial Intelligence Incident Database. Responsible AI Collaborative. Retrieved on February 16, 2023 from incidentdatabase.ai/cite/110.
37. Juhn, Y.J., Ryu, E., Wi, C.I., King, K.S., Malik, M., Romero-Brufau, S., Weng, C., Sohn, S., Sharp, R.R., Halamka, J.D.: Assessing socioeconomic bias in machine learning

algorithms in health care: a case study of the HOUSES index. J Am Med Inform Assoc 29(7), 1142–1151 (2022).

38. Fraser, K.C., Meltzer, J.A., Rudzicz, F.: Linguistic Features Identify Alzheimer's Disease in Narrative Speech. J Alzheimers Dis 49(2), 407–422 (2016).