HOW AI BIAS IS EXACERBATING HEALTH INEQUALITY IN TIMES OF COVID-19

Mónica Fernández Peñalver

Artificial Intelligence Masters, Radboud University

13th January 2022

Abstract

In the face of the pandemic, Artificial Intelligence (AI) has become a key player in the fight against COVID-19. However, despite its efforts in predicting virus spread, and in medical decision making, researchers argue that there is still bias in the design and use of these systems that enhance existing health inequality [1,2]. Differences in health status and unequal access to healthcare is causing vulnerable and minority populations to be at a higher risk of infection and death from COVID-19 than any other demographic group. We believe that health inequality, in combination with COVID-19, and the deployment of AI, could make the situation worse. An analysis of the literature suggests that AI bias is exacerbating inequality in the fight against COVID-19. Our findings suggest that the rush to find technical solutions has been the main source of bias, giving rise to two other sources: poor methodological conduct, and the repurposing of AI systems. In this review, we evaluate how these situations can exacerbate health inequality in different ways, as well as possible measures that can be taken in other to mitigate AI bias through research and policy.

1. Introduction

After two years of fighting the pandemic, a lot of praise has been given to Artificial Intelligence (AI) for its crucial role in combating COVID-19. It has quickly become a key player in the operation of the healthcare system and the deployment of COVID-19 measures. The predictive and automated power of AI has allowed the prediction of the spread of the virus, and its application in treatment, diagnostics, vaccine discovery and more [3,4,5]. Yet, despite its recognition, researchers argue that there is still bias in the design and use of these systems that can enhance existing health inequality [1,2]. If models are not developed and deployed properly to tackle the pandemic, its consequences will place minority populations at a higher risk of infection and a higher risk of death.

This paper aims to evaluate how AI bias is exacerbating inequality in times of COVID-19. The topic is highly relevant given the present times of the pandemic that is causing millions of deaths globally. It should be of great concern and a top priority, to ensure that the systems we use to tackle the virus does not turn out to cause harm itself. If the issue is not addressed, biased data that stems from inequality and is fed into our AI systems will inevitably create outcomes that will exacerbate inequality itself. Especially in times of a pandemic, these outcomes can become deadly for those negatively affected.

The paper reviews studies and reports addressing health inequality, COVID-19, and AI bias, in combination, or alone, released over the past three years. It covers the topics of health inequality, disparities in the risks and deaths caused by COVID-19, and different sources of AI bias arising due to the pandemic. Finally, it covers approaches that could be taken in research and policy to mitigate AI bias and promote the development of responsible AI in times of crisis. The review does not go into technical detail about the different applications mentioned.

The organisation of the review will be split in three parts. First, it will explain how a greater disease burden driven by inequality puts vulnerable and minority populations at higher risk of infection and death from COVID-19. The second part will describe how this inequality, and the health disparities caused by it, is exacerbated by AI models containing different kinds of biases. Lastly, it will describe what can be done to develop and deploy responsible AI systems and mitigate AI bias.

2. How Health Inequality is giving rise to COVID-19 Inequality

Health inequality can be defined as the systemic differences in health status between different socio-economic groups [6]. Differences in health status and unequal access to healthcare has been shown to be driven by a variety of individual, community, and national factors, and is predominantly studied based on social class, gender, and ethnicity [7]. Research has evaluated the interaction between COVID-19 and inequality, concluding that "COVID-19 does not discriminate" is a complete myth [8, 9, 10]. They address the fact that the medical model of disease risk; having multiple comorbidities, is prioritized over social factors that also lead to higher risks of infection and death from COVID-19 [8].

Regarding differences in health status, low socio-economic status, including having a low/unstable income, poor living conditions, and specific occupations, can lead to higher risk of contracting COVID-19, placing an entire demographic group at a higher risk of death [8, 9, 10]. Financial uncertainty caused by government measures such as lockdowns can easily affect the mental health of individuals in the form of stress [8]. High stress can lower the immune system and put an individual at a higher risk of contracting diseases [8]. Poor living conditions also contribute to the expansion of epidemic diseases (due to overcrowding) and puts people at a higher risk of contracting other illnesses such as obesity, diabetes, cardiovascular disease, and high blood pressure [6, 9]. Since studies have defined multiple comorbidities as a major risk-factor of illness and death from COVID-19 [11, 12], social factors should be considered a high risk factor as it is to have multiple comorbidities alone.

Differences in health status have been reflected in COVID-19 statistical reports as shown by Public Health England, who released a report showing the disparities in risks and outcomes of COVID-19 across England. The review shows how risk of dying from COVID-19 was higher in those living in the more deprived areas than those living in the least deprived [11]. Risk was also shown to be higher in those in Black, Asian and Minority Ethnic (BAME) groups than in White ethnic groups [11]. Given the variability of reported cases, the study analysed confirmed diagnoses representing the population of people with severe disease, rather than all of those who get infected. The numbers, therefore, reflect differences in the risk of getting the infection, in presenting to hospital with a medical need and in the likelihood of being tested [11]. The analysis also supported that comorbidities such as diabetes and hypertensive disease was higher in all BAME groups when compared to White ethnic groups [11, 12].

Regarding the unequal access to healthcare, socioeconomically disadvantage groups, minoritized ethnicities and immigrants tend to show up to the doctors at more advance stages of the disease due to various reasons. For instance, confidence in that they will be treated with respect can be hindered by language barriers or biased attitudes and discrimination in the healthcare system [8]. Even if they do show up, health records collected about COVID-19 patients disproportionately represent the upper class who have access to "digitally mature" hospitals (because datasets require a degree of quality and integrity that only well-off hospitals can provide) [1]. Likewise, if COVID-related data is collected through contact tracing apps, it does not represent the people who do not have access to smartphones.

These different scenarios, reflecting the unequal to healthcare, causes data to become unrepresentative of minoritized populations. This supposes a problem because, as explained above, the prevalence of COVID-19 is higher in these population groups. If data from minority

groups aren't used to train AI systems that help with automated medical decision-making for instance, these groups will become negatively affected by such systems.

Overall, studies and reports suggest that social inequality is a major contributor to health inequality, that causes minority populations (such as socioeconomically disadvantage groups, minoritized ethnicities and immigrants) to have a lower health status and lower access to healthcare, both leading to a higher risk of infection and death from COVID-19. The situation gets worse when the data obtained from a discriminatory social system gets combined with a AI. Feeding this kind of data to AI systems will inevitably make AI to be biased and exacerbate existing health inequality present in our society.

3. AI Bias in times of COVID-19

Bias in AI can come in the form of statistical bias, social bias, or both [13]. Statistical bias, for example, refers to an algorithm that produces a result that differs from the true underlying estimate [13]. On the other hand, social bias, in the context of health care, refers to inequity in care delivery that systematically leads to worse outcomes for a particular demographic group [13]. Statistical bias can be caused by technical failures such as improper sampling or measurement error in predictor variable, whereas social bias can be caused by a statistically biased algorithm (as it will be explained in the examples below) or by human bias [13].

With the use of the above definitions, we will review three sources of AI bias that the literature has suggested to arise due to the presence of COVID-19: 1) the rush to find technical solutions, 2) poor methodological conduct, and 3) the repurposing of AI systems. Our findings also suggest that point 1 is the main contributor to point 2 and point 3 when it comes to tackling COVID-19 during time of crisis.

3.1. The Rush to find Technical Solutions

Developing an AI system is a laborious process, but the demand for a rapid response to tackle the virus sped up the task. Yet, the rush and the urgency to find solutions has led to a power imbalance in agenda setting [1]. Diversity statistics in the AI industry show that the demographic composition of those who set research agendas normally fall outside those that are negatively affected by a biased AI [1,14]. This suggests that agenda setting is likely being influenced by social bias.

Furthermore, when it comes to developing a system within time constraints, developers and deployers overlook the measures needed to ensure that the product is responsibly used and unbiased. The ALLAI group, an organisation working on promoting and fostering responsible AI, has shown that many systems developed and deployed during the pandemic have a negative impact on society and are being irresponsibly used [15]. Several examples include AI-driven proctoring, algorithmic grading, or the deployment of thermal cameras for COVID-risk detection [15]. For the latter example, the European Centre of Disease Prevention and Control has defined the use of thermal cameras as a "high-cost, low-efficient measure" [16], because these systems can only estimate elevated surface body temperature under specific circumstances and does not detect COVID-19 directly. This system has shown to have a high degree of invasiveness and be biased towards people who have elevated higher temperature for reasons other than COVID-19 [15]. Overlooking the low efficiency and invasiveness of an AI application is an example of how the pandemic is creating a world where speed to create technological "solutions" is prioritized over ethical and responsible AI deployment.

The rush and urgency to develop technical solutions has become a gateway for the other two sources of AI bias discussed below: poor methodological conduct, and the repurposing of AI systems.

3.2. Poor Methodological Conduct

The rush to find technical solutions can lead to poor methodological conduct that gives rise to statistical bias. The COVID-PRECISE (Precise Risk Estimation to optimise covid-19 Care for Infected or Suspected patients in diverse sEttings) group have published a living systematic review assessing the validity and usefulness of prediction models for diagnosing, prognosis, and detection of people with increased risk of COVID-19 infection [17]. By using PROBAST, a prediction model risk of bias assessment tool, their analysis indicated that 230 out of 232 prediction models are poorly reported and present a high risk of bias [17]. Poor methodological conduct was seen due to the use of data from a single country, thus lacking generalisability and representativeness of the population if the model were to be used in the rest of the world [17]. Moreover, models were likely to be overfitted due to the use of limited sample sizes and number of outputs of interest [17]. Lastly, many of the models lacked external validation and calibration was rarely assessed [17].

Notably, 50 models (22%), were found to be subjected to possible bias due to the use of subjective or proxy outcomes (i.e., non-covid-19 severe respiratory infections) [17]. Proxies can be an important source of algorithmic bias in many contexts. For example, an AI system was discovered to be biased because the algorithm used healthcare costs as a proxy for health status [18]. Using such proxy, however, would carry a large racial bias because even though Black patients are considerable sicker than White patients (a difference in health status), unequal access to care causes less money to be spent caring for Black patients than for White patients [18]. This inevitably made the algorithm predict that Black patients needed less care than White patients. In the context of COVID-19, proxies should be used and picked carefully, because certain comorbidities, despite being a high risk of contracting the virus, is mostly present in minority and vulnerable populations.

3.3. Repurposing AI Systems.

The rush to find technical solutions can also lead to the repurposing of AI systems. On one hand, repurposing AI systems has been argued to be promising, especially in environments like the pandemic, as simply modifying an existing application is way quicker than developing an entirely new system that might take years to finish [19].

On the other hand, repurposed AI systems can be susceptible to producing unpredictable, unexpected, or biased results, because the model are initially trained for the original purpose [19]. Worse, it can lead to fatal consequences when they are repurposed to fight a pandemic. Partnership of AI recently released an issue brief on the risks of repurposing PATTERN [20]. Pattern is an AI tool that predicts recidivism in the United States prison system, but during the pandemic it was used to determine which inmates should be released to home confinement [20]. They reported that due to the high racial bias of PATTERN, Black inmates were more likely to not be released and left exposed to a higher risk of COVID-19 infection, and consequently, death [20]. This example clearly shows how misusing AI systems to manage a deadly virus can lead to fatal consequences, especially when they are embedded with social and statistical biases.

4. Approaches to mitigate AI Bias

Given the examples above, the exacerbation of health inequality demands an urgent call for research and regulation against AI bias.

At research level, the PRECISE living review clearly shows that many of the predictor models suffer from unclear reporting and high bias [17]. The research group recommends authors to adhere to the TRIPOD (transparent reporting of a multivariable prediction model for individual prognosis or diagnosis; www.tripod-statement.org) reporting guideline to improve reporting and guide their modelling choices [17]. Similarly, a different reporting guideline MINIMAR (MINimum Information for Medical AI Reporting) has been proposed [21]. MINIMAR provides guidelines for describing the minimum information necessary to understand intended predictions, target populations, hidden biases, and the ability to generalize models [21]. To further tackle bias, existing standards such as PROBAST, as used by the PRECISE group, should be used to assess the validity of variables and data used, or to guide developers in choosing these during AI development [13, 17].

At policy level, policy makers, public health officials, and AI developers should come together with affected stakeholders to figure out how to include vulnerable populations in policies, initiatives, and innovations [21]. In addition, they should address the problem of diversity in AI research and industry. Despite the urgency to find technical solutions, the crisis should not give rise to irresponsible AI development and its irresponsible deployment. Instead, it should be an opportunity to critically assess how the newly developed technologies would affect and exacerbate the already existing inequalities present in our society. To do this, concepts of fairness and health equity should be taken as the starting point of the debate of how to approach problems of health inequality in the current and future pandemics. Lastly, the lessons learnt throughout the pandemic should stimulate efforts to implement meaningful and ethically guided regulation in legal frameworks like the Artificial Intelligence Act (AIA) that can stand strong even during times of crisis [22].

5. Conclusion

The studies and reports reviewed in this paper has brought us insight on the topics of health inequality, COVID-19, AI bias, and approaches to mitigate these. We can conclude that the use of AI to fight COVID-19 is exacerbating health inequality through different sources of AI bias arising due to the pandemic: 1) the rush to find technical solutions, 2) poor methodological conduct, and 3) the repurposing of AI systems.

The review brings up several points. Different aspects of health inequality point out that social factors should be considered a high-risk factor of contracting the virus, and one should pay attention to the unrepresentative data of vulnerable populations with little to no access to healthcare. Furthermore, the pandemic has created an environment where quickly finding "solutions" is prioritised over responsible AI development. The rush has also led to poor methodological conduct and to the repurposing AI systems, both which bring high risks of affecting vulnerable populations through statistical and social bias. Moreover, the literature suggests the use of reporting guidelines and bias assessment tools to minimise bias in research. Lastly, at policy level, efforts should be made for the inclusion of vulnerable groups among stakeholders for setting policies, initiatives and innovations that would stand strong even during times of crisis.

The current state of the literature has been addressing the issue of AI bias from different angles, but most recently through the angle of its interaction with COVID-19. Yet, there seems to be little research on how AI could be used to tackle health inequality in itself. Furthermore, the current regulations have created loopholes that have allowed the development and deployment of biased and irresponsible AI during the pandemic. Therefore, interdisciplinary researchers and policymakers should focus on developing a legal framework that ensures that AI is responsibly used during such times, as well as a regulative measure to re-assess the models that have been developed and deployed since the start of 2020 to fight the pandemic.

References

[1] D. 13, A. Mazumder, A. Peppin, M. Wolters and A. Hagerty, "Does “AI” stand for augmenting inequality in the era of covid-19 healthcare?", BMJ, p. n304, 2021. Available: 10.1136/bmj.n304 [Accessed 12 January 2022].

[2] E. Röösli, B. Rice and T. Hernandez-Boussard, "Bias at warp speed: how AI may contribute to the disparities gap in the time of COVID-19", Journal of the American Medical Informatics Association, vol. 28, no. 1, pp. 190-192, 2020. Available: 10.1093/jamia/ocaa210 [Accessed 11 January 2022].

[3] Y. Malik et al., "How artificial intelligence may help the Covid-19 pandemic: Pitfalls and lessons for the future", Reviews in Medical Virology, vol. 31, no. 5, pp. 1-11, 2020. Available: 10.1002/rmv.2205 [Accessed 10 January 2022].

[4] M. Belfiore et al., "Artificial intelligence to codify lung CT in Covid-19 patients", La radiologia medica, vol. 125, no. 5, pp. 500-504, 2020. Available: 10.1007/s11547-020-01195-x [Accessed 10 January 2022].

[5] A. Keshavarzi Arshadi et al., "Artificial Intelligence for COVID-19 Drug Discovery and Vaccine Development", Frontiers in Artificial Intelligence, vol. 3, 2020. Available: 10.3389/frai.2020.00065 [Accessed 10 January 2022].

[6] M. Whitehead, "The concepts and principles of equity and health", Health Promotion International, vol. 6, no. 3, pp. 217-228, 1991. Available: 10.1093/heapro/6.3.217 [Accessed 10 January 2022].

[7] "Inequalities in health (e.g. by region, ethnicity, soci-economic position or gender) and in access to health care, including their causes", Health Knowledge. [Online]. Available: https://www.healthknowledge.org.uk/public-health-textbook/medical-sociology-policy-economics/4c-equality-equity-policy/inequalities-distribution. [Accessed: 10- Jan- 2022].

[8] J. Patel et al., "Poverty, inequality and COVID-19: the forgotten vulnerable", Public Health, vol. 183, pp. 110-111, 2020. Available: 10.1016/j.puhe.2020.05.006 [Accessed 10 January 2022].

[9] V. Abedi et al., "Racial, Economic, and Health Inequality and COVID-19 Infection in the United States", Journal of Racial and Ethnic Health Disparities, vol. 8, no. 3, pp. 732-742, 2020. Available: 10.1007/s40615-020-00833-4 [Accessed 10 January 2022].

[10] S. Ali, M. Asaria and S. Stranges, "COVID-19 and inequality: are we all in this together?", Canadian Journal of Public Health, vol. 111, no. 3, pp. 415-416, 2020. Available: 10.17269/s41997-020-00351-0 [Accessed 10 January 2022].

[11] Public Health England. “COVID-19: review of disparities in risks and outcomes”, COVID-19: impacts on health – reports and tools, 2020. [Online] Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/908434/Disparities_in_the_risk_and_outcomes_of_COVID_August_2020_update.pdf [Accessed 10 January 2022].

[12] N. Sattar et al, “BMI and future risk for COVID-19 infection and death across sex, age and ethnicity: Preliminary findings from UK biobank”, Diabetes & Metabolic Syndrome: Clinical Research & Reviews, vol. 14, pp. 1149e1151, 2020. Available: https://doi.org/10.1016/j.dsx.2020.06.060

[13] R. Parikh, S. Teeple and A. Navathe, "Addressing Bias in Artificial Intelligence in Health Care", JAMA, vol. 322, no. 24, p. 2377, 2019. Available: 10.1001/jama.2019.18058 [Accessed 11 January 2022].

[14] S.M. Myers, M. Whittaker, K. Crawford, “DISCRIMINATING SYSTEMS Gender, Race, and Power in AI”, AI Now Institute, 2019. [Online]. Available: https://ainowinstitute.org/ discriminatingsystems.html [Accessed: 11- Jan- 2022].

[15] ALLAI, "AI and Corona Observatory", AI and Corona Project, 2021. [Online]. Available: https://allai.nl/corona-observatory-2/ [Accessed: 11- Jan- 2022].

[16] ECDC, “Considerations for Travel-related Measures to Reduce Spread of COVID-19 in the EU/EEA”, European Centre for Disease Prevention and Control, 2020. [Online]. Available: https://www.ecdc.europa.eu/sites/default/files/documents/Considerations-related-to-measures-for-travellers-reduce-spread-COVID-19-in-EUEEA.pdf [Accessed: 11- Jan- 2022].

[17] L. Wynants et al., "Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal", 2021.

[18] Z. Obermeyer, B. Powers, C. Vogeli and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations", Science, vol. 366, no. 6464, pp. 447-453, 2019. Available: 10.1126/science.aax2342 [Accessed 11 January 2022].

[19] J. Sipior, "Considerations for development and use of AI in response to COVID-19", International Journal of Information Management, vol. 55, p. 102170, 2020. Available: 10.1016/j.ijinfomgt.2020.102170 [Accessed 11 January 2022].

[20] A Xiang, R. Fogliato, A. Chouldechova, “Algorithmic Risk Assessment and COVID-19: Why PATTERN Should Not Be Used.”, Partnership on AI, 2020 [Online] Available: http://partnershiponai.org/wp-content/uploads/2021/07/Why-PATTERN-Should-Not-Be-Used.pdf [Accessed 11 January 2022].

[21] D. Leslie, "Tackling COVID-19 through responsible AI innovation: Five steps in the right direction", SSRN Electronic Journal, 2020. Available: 10.2139/ssrn.3652970 [Accessed 12 January 2022].

[22] “Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL: LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS”, European Commission, 2021 [Online] Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206 [Accessed 11 January 2022].