Insights into Imaging

OPINION

Open Access

Letter to the editor: "Not all biases are bad: equitable and inequitable biases in machine learning and radiology"

Antoine Iannessi[1], Hubert Beaumont[2*] and Anne Sophie Bertrand[3]

Abstract

Artificial intelligence algorithms are booming in medicine, and the question of biases induced or perpetuated by these tools is a very important topic. There is a greater risk of these biases in radiology, which is now the primary diagnostic tool in modern treatment. Some authors have recently proposed an analysis framework for social inequalities and the biases at risk of being introduced into future algorithms. In our paper, we comment on the different strategies for resolving these biases. We warn that there is an even greater risk in mixing the notion of equity, the definition of which is socio-political, into the design stages of these algorithms. We believe that rather than being beneficial, this could in fact harm the main purpose of these artificial intelligence tools, which is the care of the patient.

Keywords: Artificial intelligence, Bias, Inequity

Keypoints

- 'Health equity' terminology is socio-politically based, committed to eliminating disparities in health.
- The patient's medical interest should prevail over social equity in debiasing strategy.
- Transparency in artificial intelligence can debias by providing contextualized information to radiologists.

Background

Dear Editor in Chief,

We read with interest the article entitled 'Not all biases are bad: equitable and inequitable biases in machine learning and radiology' by Pot et al. [1] recently published in *Insights into Imaging*.

The authors propose a framework to analyze how social inequities in health transition into artificial intelligence (AI) algorithms in radiology. To illustrate their topic, they use race, gender and wealth inequities as examples, drawing a parallel between the root of such unfair inequalities and potential biases existing in machine learning (ML) radiology. They state that distributive and relational inequities are at risk of being translated into dataset bias quantitatively and qualitatively, respectively. Moreover, they are concerned about specific 'socially related' cognitive biases transiting into ML algorithms.

Main text

To better understand what is as stake, 'health equity' must be understood as a political terminology and the principle underlying a commitment to eliminate disparities in health and its determinants, including social determinants [2].

Indeed, there is unequivocal strong evidence to link economic/social disadvantage with lack of healthcare opportunities, illness and disability. These inequalities are also unfair, because they could be reduced by the right mix of government policies, according to the World Health Organization [3].

Additionally, the authors are concerned with cultural bias (i.e., the interpretation of situations, actions or data

*Correspondence: Hubert.beaumont@mediantechnologies.com
[2] Median Technologies, 1800 route des crêtes, 06560 Valbonne, France
Full list of author information is available at the end of the article

Patient Wealth Status	AI powered medical device	Debiasing Strategies		
		Principle of Justice	Principle of Truth	Principle of Transparency
	Real Weight	BIAS Weight -5 kg		
	BIAS Minus 5 Kg		Real Weight	BIAS Minus 5 Kg Warning

Fig. 1 An artificial intelligence powered diagnostic tool in radiology presented as a weighing scale for understanding. The tool reduces the weight of a rich person by 5 kg but uses the actual weight of a poor person. It is unfairly biased in favor of rich people leading to these patients being diagnosed as overweight less frequently. To resolve the bias, if one prioritizes the principle of equity, and reduces the weight of a poor person by 5 kg as well, the outcome of the scale is wrong for both groups of patients. If one focuses on obtaining a correct weight for all the patients, one can cancel the error on the biased scale or inform the user that the result may be biased. In conclusion, in terms of debiasing strategy, the patient's medical interest prevails over the principle of social equity. *AI* Artificial intelligence

based on the standards of one's own culture). This bias is discriminative because it is associated with partiality to a sub-group value. Cognitive biases regroup under different names and have been previously described in radiology as attribution biases, as mentioned by the authors.

We agree with the authors that AI systems for radiology are not free of bias, which can be deleterious or useful, and that engineers, radiologists and politicians should be aware of bias when developing/using AI algorithms.

However, our opinions differ on how to consider and manage bias. We are aware of the emergence of equally discriminatory strategies in the fight against cultural bias, which we believe are not the best approach to tackling the issue.

Indeed, debias strategies discussed by the authors include introducing another bias to compensate for the one identified as being at risk of contaminating the AI algorithm in radiology. They suggest that a 'better' ratio of ethnicity, level of wealth or patient gender must be enforced in the dataset considered, to balance a 'socially inequitable' distribution. To handle qualitative cognitive biases, they suggest that one could positively discriminate algorithm developers or involved radiologists in order to promote a diversity of opinion.

This approach can be criticized for three major reasons:

- First, voluntarily injecting a correction inside the ML algorithm in radiology is driven by political motives.

- Second, if these corrections are not transparent to the end user, an additional bias would be introduced, which is not consistent with the initial objective. Indeed, unlike the systematic error controlled by the developer when training the algorithm, these biases are not perceived by the radiologist and are therefore very difficult to avoid without awareness.

- Third, enforcing homogeneity across identified sub-populations in the training data can lead to risky and uncontrolled situations. Unless evidence can be collected to the contrary, this runs the risk of jeopardizing the performance of the ML algorithm for other/unidentified subpopulations, when applied to the general population. This seems to be at odds with the primary objective of offering the most medically efficient algorithm possible to a patient as a non-political individual. Indeed, in a perspective to market an algorithm over a large territory, we believe that the population of validation should be representative of the population of utilization, thus limiting the 'equitable generalization process'.

Ultimately, collecting 'equitable' data for an ML algorithm is a political concern and, in our opinion, should not be considered without evidence of increased patient benefit. It is also essential that the end user radiologist is fully aware of any such corrections, if applied.

As an analogy, we can take the example of a weighing scale which systematically reduces weight by 5 kg for a rich person but not for a poor person (Fig. 1). If your decision is based on weight, you would reduce the weight by 5 kg for a poor person as well if following a principle of justice, or measure the actual weight of a rich person, if following a principle of truth. If you cannot correct the scale, the principle of transparency applies, and you should advertise the risk of an erroneous result for rich people [4].

Medicine is a branch of human science and is based on an ideal of neutrality. Regarding the social-related cognitive bias applying to radiologists facing inequities, we would like to refer to the original Hippocratic oath, i.e., 'Into whatever homes I go, I will enter them for the benefit of the sick, avoiding any voluntary act of impropriety or corruption' [5]. Swearing this oath does not eliminate bias in social individuals (radiologists included), but we believe that physicians fundamentally respect this oath by treating all patients equally.

Conclusion

In conclusion, we believe the strategies for debias suggested by the authors will not help solve the problem, and that radiology should be kept away from political interference. Social and cultural biases are deeply political, and we agree with the authors that there is a risk of such bias creeping into newly built algorithms in radiology. Such bias is supported by the concept of inequities, leading us to think that a good solution for debias would be to restore the equity. As we have explained, the concept of 'debiasing' does not mean 'to compensate for' but 'to remove' the bias. Therefore, we do not believe that there should be an aim to compensate for social or cultural bias in the conception of AI in radiology. The only possible relevant exception would be if the outcome of the algorithm results in increased medical benefit to the patient. Even then, the user should be fully aware that the algorithm will propose a diagnostic based on a 'corrected' population so that they can make an informed decision.

Abbreviations
AI: Artificial intelligence; ML: Machine learning.

Authors' contributions
We certify that all co-authors contributed equally and significantly to the study and to the design of the Letter to the Editor.

Declarations

Ethics approval and consent to participate
Our Letter to the Editor did not include interaction or intervention with human subjects or include any access to identifiable private information, no IRB approval was required. Written informed consent was not required for this study because not impacting patient management.

Consent for publication
We certify that this Letter to the Editor is not under consideration for publication elsewhere. All co-authors take public responsibility for the content of the present manuscript. The final version of the manuscript has been reviewed and approved by all co-authors.

Competing interest
The authors of this manuscript Hubert Beaumont and Antoine Iannessi declare relationships with the following companies: Median Technologies. The other author of this manuscript declares no relationships with any companies, whose products or services may be related to the subject matter of the article.

Author details
[1]Centre Antoine Lacassagne, 33 Avenue de Valombrose, 06100 Nice, France. [2]Median Technologies, 1800 route des crêtes, 06560 Valbonne, France. [3]Centre Hospitalier Princess Grace, 1 Avenue Pasteur, 98000 Monaco, Monaco.

References
1. Pot M, Kieusseyan N, Prainsack B (2021) Not all biases are bad: equitable and inequitable biases in machine learning and radiology. Insights Imaging 12(1):13
2. Braveman P (2014) What are health disparities and health equity? We need to be clear. Public Health Rep 129(Suppl 2):5–8
3. Whitehead M, Dahlgren GR, World Health Organization E (2006) Regional office for, Levelling up (part 1): a discussion paper on concepts and principles for tackling social inequities in health/by Margaret Whitehead and Göran Dahlgren. WHO Regional Office for Europe, Copenhagen.
4. Felzmann H, Fosch-Villaronga E, Lutz C, Tamò-Larrieux A (2020) Towards transparency by design for artificial intelligence. Sci Eng Ethics 26(6):3333–3361
5. Davenport ML, Lahl J, Rosa EC (2012) Right of conscience for health-care providers. Linacre Q 79(2):169–191

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.