

Co-Reasoning and Epistemic Inequality in AI Supported Medical Decision-Making

Holm, Søren; Ploug, Thomas

Published in:
American Journal of Bioethics

DOI (link to publication from Publisher):
[10.1080/15265161.2024.2377115](https://doi.org/10.1080/15265161.2024.2377115)

Creative Commons License
CC BY-NC-ND 4.0

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Holm, S., & Ploug, T. (2024). Co-Reasoning and Epistemic Inequality in AI Supported Medical Decision-Making. *American Journal of Bioethics*, 24(9), 79-80. <https://doi.org/10.1080/15265161.2024.2377115>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.



Co-Reasoning and Epistemic Inequality in AI Supported Medical Decision-Making

Søren Holm & Thomas Ploug

To cite this article: Søren Holm & Thomas Ploug (2024) Co-Reasoning and Epistemic Inequality in AI Supported Medical Decision-Making, The American Journal of Bioethics, 24:9, 79-80, DOI: [10.1080/15265161.2024.2377115](https://doi.org/10.1080/15265161.2024.2377115)

To link to this article: <https://doi.org/10.1080/15265161.2024.2377115>



© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 03 Sep 2024.



Submit your article to this journal [↗](#)



Article views: 265



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Co-Reasoning and Epistemic Inequality in AI Supported Medical Decision-Making

Søren Holm^{a,b}  and Thomas Ploug^c

^aUniversity of Manchester; ^bUniversity of Oslo; ^cAalborg University

Most of us do not doubt that our car mechanic knows more about the inner workings of the internal combustion engine or the synchronized gearbox than we do, and that they also know more about interpreting the outputs from the electronic diagnostic systems that continually monitor the performance of a large range of care systems and which are present in all modern cars. That is, we do not doubt that our car mechanic has epistemic superiority in certain areas of knowledge that are essential to diagnose car problems and propose ways to repair them. In Denmark part of the reason for not doubting this epistemic superiority is that it takes 4½ years to become a fully trained and certified car mechanic. This epistemic inequality between us and our car mechanic is not in itself a case of epistemic injustice. No injustice is done to us if we are given a very simplified explanation of why our car won't go into reverse gear, or why the car mechanic has come to the view that the root of the problem is with the 1st/2nd synchromesh ring. There may be cases where our interaction with the car mechanic involves testimonial epistemic injustice, e.g. cases where our description of how the car behaves is dismissed and not taken seriously, but these are not typical, and are not *a priori* more likely to occur after the advent of modern electronic car diagnostic systems. If we doubt the prescription for repair provided it is most often not because we doubt the depth of knowledge of the mechanic and their field specific epistemic superiority, but because we believe that they have a conflict of interest related to the fact that they make money out of repairing cars.

In our relation with our car mechanic an ethical issue would arise if they refused to explain their own diagnostic reasoning or the function and output of the electronic diagnostic system when asked, or did not respond appropriately to a polite contestation of their

conclusion. But does this mean that they have to involve us as co-reasoners and that we can no longer just book in the car for service in the morning and pick it up in the evening and simply rely on the knowledge and skills of the mechanic to sort things out?

The reason for describing this analogy between car repair and medicine in detail is that it decenters some possibly unrealistic, implicit assumptions about the typical health care professional-patient relationship in modern health care systems. Especially assumptions about the time available for health care professionals to interact with patients, and about the degree to which the health care professional-patient relationship is a deeply personal relationship. The analogy also illustrates that there are many areas of life where we accept that other people with whom we interact about important problems are experts and are, in the particular field of interaction, in a position of justified epistemic superiority. Let us now focus on the medical case of AI supported decision-making.

It is inherent in notions such as explainability and contestability that they are 'abilities', they are or should, as we argue elsewhere, be available options if an explanation is requested or a decision is contested (Ploug and Holm 2020, Ploug et al. 2021). They do not entail an obligation always to explain or always to ask for a contestation if no explanation is requested or no contestation made. There may, of course be cases where there are strong instrumental reasons for providing an explanation, e.g. if it makes it more likely that a patient will change behavior and avoid a recurrence of the problem, and where an explanation should therefore be given even when not asked for, but that does not undermine the fact that explainability is about providing options for explanation giving. This entails that neither a requirement for explainability nor a requirement for contestability creates an

CONTACT Søren Holm  soren.holm@manchester.ac.uk  Centre for Social Ethics and Policy, Department of Law, Williamson Building, Manchester University, Manchester M13 9PL, UK.

© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

obligation to engage in co-reasoning about AI advice for either the health care professional or the patient.

There may be other good reasons to conceptualize the encounter between a health care professional and a patient as an opportunity for co-reasoning. They are not actually, as O'Neill's 'fellow workers' referred to by the authors (Salloch and Eriksen 2024), engaged in developing or justifying maxims, but the outcome of the encounter is often a reasoned decision. There is no doubt that patients often have something to bring to the epistemic table and that health care professionals should be open to co-reasoning and facilitate it whenever a patient want to engage (Holm 2005, 2011). But, the arguments justifying that conclusion are completely independent of whether AI decision-support is involved or not. Yes, introducing co-reasoning "introduces a social and dialogical dimension that potentially exerts a different and more dynamic kind of pressure" (Salloch & Eriksen 2024, 72), but if introducing a social and dialogical dimension produces better reasoning in some circumstances, those circumstances are far wider than those that are AI related. It is also important to realize that even if health care professionals have an obligation to facilitate co-reasoning, patients (and car owners) have no obligation to initiate or participate in it.

There is also a potentially problematic account in Salloch and Eriksen's paper of what the aim of providing an explanation should be in the medical case. After having described some simplified explanations of AI outcomes the authors write:

Such rudimentary features can be enough to establish trust, which is the success condition of explanations in this context. (Salloch & Eriksen 2024, 74, our emphasis)

But that can't be true. Let us imagine that a health care professional has good knowledge about a particular AI system and knows that communicating some features of the system and its development is likely to engender trust in patients, whereas communicating other features is likely to undermine trust. If the success condition for an explanation in this context is to establish trust, then the health care professional should focus on the trust supporting features if asked for an explanation by the patient, and should suppress the trust undermining features. But that would be both deceptive and unethical, since it would render the health care professional untrustworthy.

It is, perhaps finally worth noting that there are many cases where it is primarily the health care professional who is in need of an explanation or an ability to engage in a co-reasoning process. These are the cases where a clinician relies on a report from a diagnostic procedure that has been provided with AI support, or perhaps in the future simply provided by an AI system. Clinicians routinely rely on pathology reports or scanning reports in many cases and may not have easy access to discuss the report with whoever provided it.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author(s).

FUNDING

This work was funded by the Independent Research Fund Denmark, Grant TP 2027-00140B.

ORCID

Søren Holm  <http://orcid.org/0000-0002-7200-5607>

REFERENCES

- Holm, S. 2005. Justifying patient self-management—evidence based medicine or the primacy of the first person perspective. *Medicine, Health Care, and Philosophy* 8 (2):159–64. doi:[10.1007/s11019-005-2280-x](https://doi.org/10.1007/s11019-005-2280-x).
- Holm, S. 2011. Final responsibility for treatment choice: The proper role of medical doctors? *Health Expectations: An International Journal of Public Participation in Health Care and Health Policy* 14 (2):201–9. doi:[10.1111/j.1369-7625.2011.00673.x](https://doi.org/10.1111/j.1369-7625.2011.00673.x).
- Ploug, T., and S. Holm. 2020. The four dimensions of contestable AI diagnostics—A patient-centric approach to explainable AI. *Artificial Intelligence in Medicine* 107:101901. doi:[10.1016/j.artmed.2020.101901](https://doi.org/10.1016/j.artmed.2020.101901).
- Ploug, T., A. Sundby, T. B. Moeslund, and S. Holm. 2021. Population preferences for performance and explainability of artificial intelligence in health care: Choice-based conjoint survey. *Journal of Medical Internet Research* 23 (12):e26611. doi:[10.2196/26611](https://doi.org/10.2196/26611).
- Salloch, S., and A. Eriksen. 2024. What are humans doing in the loop? Co-reasoning and practical judgment when using machine learning-driven decision aids. *The American Journal of Bioethics* 24 (9):67–78. doi:[10.1080/15265161.2024.2353800](https://doi.org/10.1080/15265161.2024.2353800).