

A healthy debate

Exploring the views of medical doctors on the ethics of artificial intelligence

Martinho, Andreia; Kroesen, Maarten; Chorus, Caspar

DOI

[10.1016/j.artmed.2021.102190](https://doi.org/10.1016/j.artmed.2021.102190)

Publication date

2021

Document Version

Final published version

Published in

Artificial Intelligence in Medicine

Citation (APA)

Martinho, A., Kroesen, M., & Chorus, C. (2021). A healthy debate: Exploring the views of medical doctors on the ethics of artificial intelligence. *Artificial Intelligence in Medicine*, 121, Article 102190. <https://doi.org/10.1016/j.artmed.2021.102190>

Important note

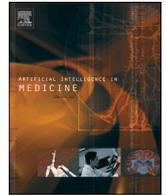
To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



A healthy debate: Exploring the views of medical doctors on the ethics of artificial intelligence

Andreia Martinho^{*}, Maarten Kroesen, Caspar Chorus

Delft University of Technology, Delft, the Netherlands

ARTICLE INFO

Keywords:

Artificial intelligence
Healthcare
Medicine
Ethics
Q-methodology

ABSTRACT

Artificial Intelligence (AI) is moving towards the health space. It is generally acknowledged that, while there is great promise in the implementation of AI technologies in healthcare, it also raises important ethical issues. In this study we surveyed medical doctors based in The Netherlands, Portugal, and the U.S. from a diverse mix of medical specializations about the ethics surrounding Health AI. Four main perspectives have emerged from the data representing different views about this matter. The first perspective (*AI is a helpful tool: Let physicians do what they were trained for*) highlights the efficiency associated with automation, which will allow doctors to have the time to focus on expanding their medical knowledge and skills. The second perspective (*Rules & Regulations are crucial: Private companies only think about money*) shows strong distrust in private tech companies and emphasizes the need for regulatory oversight. The third perspective (*Ethics is enough: Private companies can be trusted*) puts more trust in private tech companies and maintains that ethics is sufficient to ground these corporations. And finally the fourth perspective (*Explainable AI tools: Learning is necessary and inevitable*) emphasizes the importance of explainability of AI tools in order to ensure that doctors are engaged in the technological progress. Each perspective provides valuable and often contrasting insights about ethical issues that should be operationalized and accounted for in the design and development of AI Health.

1. Introduction

Artificial Intelligence (AI) is moving towards the health space. Given the abundance of data generated by health systems as a result of digitization efforts made over the last decade, a new data-driven approach to implement AI in healthcare has emerged. In contrast with previous and somewhat failed rule-based approaches to implement AI in healthcare [1,2], this new approach relies heavily on algorithms that detect patterns in data from clinical practice (e.g. medical imaging and electronic health records), clinical trials, genomics studies, and insurance, pharmaceutical, and pharmacy benefits management operations [3]. There is an expectation that these state-of-the-art-data-driven AI methods and algorithms will be able to use such data to address the complex problems of health systems [4,3].

The implementation of AI in healthcare holds great promise for expanding the medical knowledge and providing optimal yet cost-effective healthcare solutions [5,6]. In the clinical domain, expected results include identification of individuals at high risk for a disease, improved diagnosis and matching of effective personalized treatment,

and out-of-hospital monitoring of therapy response [4,7]. Despite the projected benefits associated with Health AI, it also raises important ethical issues [8,9].

It is well known that AI has the potential to threaten values such as Autonomy, Privacy, and Safety [10], which are core values in Medicine [11,12]. Therefore, in order for AI to promote quality of care and minimize potentially disruptive effects [13], its deployment must take ethics into account. An important step towards ethical deployment of disruptive AI technologies is to learn the views of practitioners about such technologies. This information allows a better operationalization of the ethical issues associated with AI in a particular domain, which eventually is expected to lead to more meaningful debates and robust policies.

The current academic literature provides interesting and valuable information on the perspectives of practitioners about the impact of AI technologies in the medical profession [14,15,16,17]. Most of these studies are particularly suited to medical fields with a strong image processing component, which is adequate for automated analysis, such as radiology [18,19,20,21,22,23,24], pathology [25], and dermatology

^{*} Corresponding author.

E-mail address: a.m.martinho@tudelft.nl (A. Martinho).

<https://doi.org/10.1016/j.artmed.2021.102190>

Received 12 February 2021; Received in revised form 22 September 2021; Accepted 29 September 2021

Available online 12 October 2021

0933-3657/© 2021 Delft University of Technology.

Published by Elsevier B.V. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

[26,27]. However, there is little knowledge on the views of medical doctors about the ethical issues associated with the implementation of AI in healthcare.

The aim of this study is to gain insight into the reasoning patterns and moral opinions about Health AI from those involved in the medical practice. By surveying medical doctors in The Netherlands, Portugal, and U.S. on the ethical issues associated with the implementation of AI in healthcare, we expect to enrich existing literature on the impact of AI technologies in medicine and provide valuable knowledge for the operationalization of Health AI Ethics.

We first provide a brief commentary about the ethics of AI in healthcare. Subsequently we explain the methods used in this research by outlining the basic steps of q-methodology and explaining how we established these steps in this study. Later we present the results of the study by describing the four different perspectives that have emerged from the data. These results are further analyzed and discussed. Finally we draw conclusions and present directions for further research.

2. The ethics of health AI

The empirical work about AI in healthcare that has been reported in the literature focuses mainly on issues directly related to the medical practice and career, such as Future of Employment, Education about AI, and Accountability.

It has been reported that medical students and practitioners understand the increasing importance of AI in healthcare and have positive attitudes towards the clinical use of AI [20,26,17], but mainly as a supportive system for diagnosis [18,19,26,27,25,24].

Despite the positive attitudes towards AI, it has also been reported that students and medical doctors are poorly trained on these technologies [20,28,29,30]. One study indicated that, although a small cohort of UK medical students who received AI teaching felt more confident in working with AI in the future compared to students that did not receive teaching, a significant number of taught students still felt inadequately prepared [20]. In order to take full advantage of these technologies, scholars seem to agree that medical school training on AI should be expanded and improved [18,20,21,26,25].

Regarding the impact of AI on career choice and reputation, it was reported that AI has an impact in the career intentions of students with respect to radiology [20], but radiologists would still choose this specialty if given that choice [21]. These specialists have, however, revealed concerns that AI might diminish their professional reputation [24].

Contrary to the perceptions of the general public that AI will completely or partially replace human doctors [31], medical students and doctors in general are not concerned about job replacement [18,26,17,32,24].

Another important issue related to medical practice and career is liability. In a study in which pathologists were surveyed, it was reported that, with respect to medico-legal responsibility for diagnostic errors made by a human/AI combination, opinions were split between those who believed that the platform vendor and pathologist should be held equally liable, and others who believed responsibility remains primarily that of the human, with only a minority reporting that the platform vendor should primarily be liable [25].

Clearly, the ethics surrounding implementation of AI in healthcare goes beyond issues related to medical practice and career. Health AI gives rise to higher level ethical issues such as Autonomy, Fairness, or Privacy [33,10] but, with the exception of fairness, these issues have received less attention in the scientific literature. Fairness concerns related to racial and gender bias in AI-powered medical applications have to do with the fact that AI algorithms are trained on predominantly male white patient data. Concerns have been raised both in popular and scientific literature about these algorithms perpetuating and amplifying existing bias and inequalities in healthcare [34,35,36,37,38]. It has been cautioned that medical data needs to be critically appraised in order to

avoid such bias [34].

In this empirical study we surveyed medical doctors on a wider scope of ethical issues about AI in Healthcare. We addressed Privacy, Fairness, Accountability, Transparency, Safety, Human Oversight, Explainability, Future of Employment, Responsible Research Funding, Education about AI, Human Autonomy, Certification of AI products, Ethical Design. The diverse array of Health AI ethics surveyed in this empirical study allows us to discern the views and moral opinions of medical doctors about the implementation of AI in healthcare.

3. Methods

3.1. Overview

In this research we used q-methodology, a systematic empirical approach to identify possibly conflicting perspectives of (stakeholder) individuals about a particular topic [39,40,41,42,43]. The core premise in q-methodology is that subjectivity is always self-referent, i.e. only the individual can measure his or her subjectivity, relational, i.e. the meaning of a statement is derived from its relation to other statements, and it can be demonstrated to have structure and form [41]. This method is therefore considered adequate for our purpose of systematically discerning and studying the subjective views of medical doctors about Health AI.

Q-methodology requires participants to sort a pre-defined set of items according to a subjective notion of agreement/disagreement. In this study medical doctors were invited to sort a set of statements retrieved from popular and scientific literature capturing key ethical issues about Health AI in a bell shaped distribution ranging from -5 to $+5$ and to provide additional comments about the statements they ranked highest ($+5$) and lowest (-5).¹ Using statistical techniques, coherent clusters are formed which present particular perspectives into the ethics of Health AI. We interpret these perspectives and discuss in what ways they relate to and differ from one another.

There are great advantages in using q-methodology when compared to other exploratory research methods, such as interviews, focus groups, and surveys. Unlike interviews, q-studies provide numerical results to support subjective perspectives about a particular topic thus combining quantitative and qualitative approaches [44]. Moreover, because participants in q-studies sort items individually, these studies are less affected by dominance effects, which are observed in other research methods administered in groups, such as focus groups [44]. And unlike standard surveys, in which the opinions of participants about each topic are extracted separately, q-studies require participants to consider such topics simultaneously thus uncovering latent connections and allowing for more nuanced and sophisticated opinions [44,45].

For the purposes of our study, which we recall is to reveal the diverse views about the ethics of Health AI, we also considered q-methodology to be a more suitable research method when compared to the Delphi method [46]. The latter is typically used for expert consultation and in that sense it is similar to our q-methodological study, in which we survey medical doctors about Health AI. However, the focus in the Delphi method is on reaching convergence (reducing heterogeneity) among experts about certain uncertain outcomes, whereas the q-method focuses on revealing the heterogeneity among stakeholders or experts.

This study followed the typical four phase sequence in q-methodological studies comprising (i) *definition of the concourse of communication*; (ii) *development of the set of statements (Q-set)*; (iii) *selection of participants (P-set)*; and (iv) *analysis and interpretation*. Below we provide further details about each one of these phases in this particular study.

¹ This study received ethics approval from the Human Research Committee of Delft University of Technology (letter of approval 1156).

3.2. Concourse of communication

The concourse of communication is a corpus of opinions related to a particular topic [47]. Such opinions can be gathered through direct sources, such as interviews and nominal group technique, or indirect sources, such as articles, discussion boards, and blogs. In this study we used quite varied indirect sources, including scientific publications as well as publications issued by popular science outlets, professional associations, consulting companies, and also blogs.

We reviewed scientific and gray literature on Health AI using combinations of keywords “Artificial Intelligence”, “Machine Learning”, and “Augmented Intelligence” along with connector “AND” and keywords “Healthcare”, “Medicine”, “physicians”, and “medical doctors” in Google, Google Scholar, and Web of Science.

Initially we selected 353 statements for our concourse of communication and subsequently we assigned these statements to fifteen clusters, using a list of ethical issues compiled from 22 major guidelines of AI ethics as a guidance tool [10]. Each cluster of statements was associated with a particular AI ethical issue from such list, namely Privacy; Fairness; Accountability; Transparency; Safety and Cybersecurity; Human Oversight; Explainability; Future of Employment; Responsible Research Funding; Education about AI; Human Autonomy; Certification of AI products; Ethical Design; Health specific deliberations; and one additional cluster was added concerning AI in the Covid-19 Pandemic.

Organizing the concourse of communication in clusters that map onto overarching AI ethical issues facilitated the definition of the q-set, since the statements in this set should reflect the entire space of ethical issues identified in the concourse. It should be remarked, however, that the list of ethical issues used in this research as a reference tool reflects a particular deontological-based approach to Ethics. Other potential relevant ethical approaches and principles related, for instance, to informed consent and risk acceptance are therefore not included in such list [48]. Future research may identify and further explore additional ethical issues and values about Health AI.

3.3. Set of statements (q-set)

The q-set is a comprehensive yet manageable subset of the concourse of communication. We analyzed each statement in the clusters defined within the concourse in order to select the relevant items for a structured, comprehensive, and balanced set of statements. This selection was guided by three main considerations, namely, (i) accounting for a broad scope of positions put forward in the AI Health popular and scientific literature; (ii) favoring clarity; and (iii) avoiding redundancy. Using this method of obtaining the concourse, we have aimed for a maximum of objectivity and neutrality.

The final q-set features 40 statements. Minor edits were made to these statements in order to ensure neutrality and also to meet the number of characters allowed by FlashQ [49], the software tool that was used in this study for administering the survey. The size of the set is at par with current q-methodology practices [41].

The landscape of statements in the q-set with respect to the pre-defined ethics clusters is composed of Privacy (statements 1–4); Fairness (5–8); Accountability (9–10,40); Transparency (11); Safety and Cybersecurity (12–13,39); Human Oversight (18); Explainability (15–17); Future of Employment (19–20, 22); Responsible Research Funding (23–24); Education About AI (25,34); Human Autonomy (18); Certification of AI products (29–30); Ethical Design (31–33); Health specific deliberations (14,21,26-27,36–38); and AI in the Covid-19 pandemic (28,35). The final set is listed below.

1. Privacy should not be the highest priority in AI-based Healthcare.
2. Confidentiality should not constrain the implementation of AI in Healthcare.
3. Without clear rules about data usage, storage, and anonymization, AI should never be used in Healthcare.
4. Confidentiality, as defined today, has little use in a future where Healthcare relies heavily in AI.
5. AI is more likely to resolve rather than amplify inequalities in healthcare.
6. Improving equity and inclusion should be the top priority when developing and deploying AI in healthcare.
7. AI will increase discrimination based on predicted future medical problems.
8. We should be conservative in promoting AI in healthcare because of the unresolved ethical issues.
9. AI developers must be bound by medical ethics.
10. For the sake of technology advancement AI companies should not be liable for medical errors.
11. AI medical tools should only be used if clinicians understand how AI decisions are made.
12. There is high risk for monopolistic behavior by private AI companies in the domain of Healthcare.
13. It is undesirable that big companies enter the health care space because they know little about Medicine.
14. The patient-physician relationship will change dramatically once AI is fully deployed in health systems.
15. Health professionals do not need to know how AI medical tools work but rather if they are reliable.
16. Health professionals have always trusted black boxes (e.g. MRI) and it will not be different with AI.
17. Appropriate informed consent is not possible if the medical doctor cannot explain to the patient how the AI medical device works.
18. AI will decrease the autonomy and authority of medical doctors.
19. AI will not replace doctors, but doctors who use AI will replace doctors who do not.
20. If AI tools work well, Hospitals should save money by hiring less highly skilled practitioners.
21. AI will worsen problems in healthcare such as overtesting, overdiagnosis, and overtreatment.
22. Automation may work well in factories, but not in Hospitals.
23. AI-based medical products won't be able to match the hype.
24. All the funding allocated for AI is worthwhile if it can take over bureaucratic shores, such as note-taking, coding, and pattern-finding.
25. Doctors are not interested in learning about AI and Computer Science.
26. In the medical field it is problematic that machines lack contextual knowledge and ability to read social clues.
27. It would be unethical not to use AI tools if they provide better decisions than medical doctors.
28. AI has already played a vital role in the COVID-19 pandemic.
29. The mantra of the tech industry “fail fast and fix it later” is putting patients at risk and regulators are not doing enough to keep consumers safe.
30. AI healthcare products must be tested in randomized clinical trials, which is the strongest source of medical evidence.
31. Because AI systems are designed mainly to increase profit, in the future health systems will have more resources and provide better care.
32. Healthcare AI technology must be aligned with bioethical principles.
33. Medical doctors must participate in the design process of AI for Healthcare.
34. Clinicians lack the time to learn how to use complex AI-based medical devices.
35. AI enhances medical decision making in situations of care rationing.
36. AI will allow providers, clinicians, and staff, to focus on more top-of-license skill sets and activities.
37. Most areas of healthcare can benefit from AI.

Table 1
Participants.

Specialization	Portugal	Netherlands	U.S.
Surgery	1	3	4
Anesthesiology	1	2	6
OBGYN	3	2	0
Ophthalmology	3	1	2
Rehabilitation Medicine	1	1	1
Intensive Medicine	1	2	0
Neurology	1	2	1
Family Medicine	2	1	2
Radiology/Nuclear Medicine/Neuroradiology	4	3	0
Pathology	1	6	0
Rheumatology	8	3	1
Oncology	1	1	2
Dermatology	1	3	0

38. It is not very difficult to operationalize clinical practice for a machine.
39. Medicine should never rely on AI because such computer systems are vulnerable to cybersecurity threats.
40. If a medical doctor makes a mistake as a result of the advice from an AI tool, he/she should be considered liable.

3.4. Participants (p-set)

In the recruitment of participants for this study three different approaches were used, which entailed reaching out to: (i) Hospital departments (through phone and subsequently by e-mail); (ii) medical doctors who are personal acquaintances; and (iii) medical doctors who are not personal acquaintances (through email addresses made available in publications found in Google Scholar related to various medical fields). Provided that approach (iii) proved to be much more successful, eventually the other strategies were dropped and we focused mainly in reaching out medical doctors through publications they had recently authored.

Rather than focusing on a particular medical field, in which the practitioners may share similar thoughts about AI, we aimed at including a diverse mix of specializations which would allow us to have a wider breadth of viewpoints in the data.

In order to select publications from different medical fields, we used keywords “Surgery”, “Anesthesiology”, “OBGYN”, “Gynecology”, “Ophthalmology”, “Intensive Medicine”, “Neurology”, “Family Medicine”, “Primary Care”, “Radiology”, “Nuclear Medicine”, “Neuroradiology”, “Pathology”, “Rheumatology”, “Oncology”, “Dermatology” along with connector “AND” and keywords “Netherlands”, “Portugal”, and “United States” in Google Scholar. Subsequently, through snowballing techniques, additional relevant articles and scholars were identified.

Provided that this study aimed at surveying medical doctors, when the corresponding author of a scientific article was not identified as a MD in the publication, we did additional searches in Google to confirm if the scholar was indeed a medical doctor. Each participant was therefore contacted through the e-mails made publicly available in the scientific publications, in the capacity of being an author or co-author of a particular publication as well as a medical doctor.

The final set of participants in this study comprised medical doctors (residents and specialists) from thirteen different specialties including medical specialties (Family Medicine, Rheumatology, Dermatology, Intensive Medicine, Oncology, Neurology), surgical specialties (Surgery, Ophthalmology, OBGYN, Anesthesiology, Rehabilitation Medicine, Neurology), and diagnosis specialties (Pathology, Radiology/ Nuclear Medicine/ Neuroradiology) based in The Netherlands, Portugal, and U.S. Further details are found in Table 1 below.

A total of 77 participants successfully completed the survey, which is an adequate number for a q-methodological study featuring a q-set of 40 items [41,39]. Indeed because q-methodology aims just at establishing

the existence of particular viewpoints, large numbers of participants are not required. Moreover, q-studies do not require a rigorously representative sample but rather a population sample that contains participants with relevant viewpoints on the matter.

We are confident that the p-set in this study includes scholars with relevant viewpoints on AI Health. However, we acknowledge that, by targeting medical doctors who had recently published scientific articles, the set of participants is mainly composed of practitioners who are involved in research and/or academic activities. We may therefore have failed to represent other perspectives from practitioners that are less involved in research.

In this context, it should also be clarified that a q-study typically makes no claim that the relative sizes of the perspectives (in terms of the number of respondents that adhere to them) reflect the population distribution. In keeping with the notion that q-methodology is an exploratory rather than a confirmatory technique, and acknowledging the way in which the sample was obtained, we will refrain from drawing any quantitative conclusions about sizes of perspectives and differences between countries and specializations. Follow up confirmatory research (e.g. aiming at establishing minority versus majority views) should be based on representative samples.

3.5. Survey collection tool

The data was collected through the html version of FlashQ (Figs. 1–3),² a software that allows online q-sorting. The distribution was coded as a 11-point distribution resembling a normal distribution $[-5, +5]$ with two cells placed under each tail (-5 and $+5$), three cells under both -4 and 4 , three cells under both -3 and 3 , four cells under both -2 and 2 , five cells under both -1 and 1 , and six cells under 0 . Participants were asked to arrange the 40 statements according to a subjective notion of disagreement/agreement and subsequently were asked to provide further comments on the statements they ranked -5 and $+5$. Each particular arrangement of the statements in the forced bell-shaped distribution is called a q-sort so in this study the collected data consisted of 77 q-sorts.

3.6. Analysis

As a derivation of factor analysis, q-methodology is a data reduction technique which aims to reduce a larger number of variables into fewer factors. Therefore, the analytic process of q-methodology relies on multivariate data-reduction techniques. In q-studies, data analysis entails three main steps: (i) factor extraction; (ii) factor rotation; and (iii) factor interpretation. In the analytic process (steps (i) and (ii)) we used PQMethod, a statistical program that accommodates the requirements of q-studies³ [50].

The first step consists of extracting factors from previously collected q-sorts thus summarizing all individual responses into a few representative responses [44]. In this study, the factors were extracted through Principal Component Analysis (PCA), a linear reduction technique that reduces the dimensionality of the data while retaining most of the variation in the dataset, which is often used in exploratory data analysis [51]. The identification of orthogonal vectors (principal components) along which variation is maximal allows the reduction of data into a few components that represent the dominant patterns in the data [52,51].

The extracted factors were subsequently rotated in order to position each factor so that its viewpoint closely approximates the viewpoint of a

² Q-methodology software packages and resources are available <https://qmeth.org/resources/software/>. The html version of FlashQ used in this research is no longer available at <http://www.hackert.biz/flashq>. A html version of FlashQ can be currently found in <https://shawnbanasick.github.io/ken-q-analysis/> [49].

³ <http://schmolck.org/qmethod/pqmanual.htm>

Table 2
Overview of factors.

	Factor 1	Factor 2	Factor 3	Factor 4
Defining sorts	15	17	6	9
Eigenvalues	17.77	11.55	8.47	10.78
Variance	17%	15%	11%	14%

three defining sorts.

Each factor is characterized by a factor array featuring 40 scores (one score per statement), which is a single q-sort configured to represent the viewpoint of the factor. Given that factors have different numbers of defining sorts, each score in the factor array is a standardized (z) score to allow cross-factor comparison [54]. The factor arrays of each factor are shown in Table 3.

Finally, the last step entailed analyzing and interpreting the factor arrays of the four perspectives in order to understand the key features of each perspective. For this purpose we used the crib sheet method [41]. By looking at the factor arrays, for each perspective we composed four basic categories: (i) items with highest ranking in the factor array; (ii) items with lowest ranking in the factor array; (iii) items ranked higher in factor *i* than in any of the other factors (by 2 or more units); (iv) items ranked lower in factor *i* than in any of the other factors (by 2 or more units). It is noted that the interpretation of the factors made by the authors is (inherently) subjective. It is possible that different people may arrive at different interpretations of the four factors based on the same factor scores. Yet, given that these interpretations are constrained by the factor scores, we would expect that other researchers would arrive at similar interpretations.

4. Results: perspectives & interpretations

Four different perspectives about Health AI were identified in this study (Table 4). The core characteristics of each perspective are derived from the statements ranked -5 and + 5 [(N : | 5 |)] where *N* is the number of the statement and (|5|) may be either -5 or + 5] as well as the statements ranked highest or lowest compared to the arrays of the other perspectives s [(N : | P_i |)] where *N* is the number of the statement, *P_i* is the perspective with *i* ∈ [1, 4], and |P_i| may either be -P_i or +P_i depending if the statement is ranked lowest or highest than in the arrays of other perspectives]. For the purpose of further illustrating each perspective, we also included statements written by participants associated with the defining sorts of each perspective, about the statements they ranked highest and lowest.

4.1. Perspective 1: AI is a helpful tool: let physicians do what they were trained for

In this perspective there is an overall positive outlook about the implementation of AI technology in healthcare. AI is regarded as a helpful tool that will allow doctors to have the time to focus on top-of-license skill sets and activities (36:+5).

Underlining this position, one participant wrote *That is the main aim! To let physicians do what they were trained for - medicine - and alleviate many of the potentially automatic and time-consuming processes they have to daily face.* Another participant noted that *AI means less time needed for boring work means more time for challenging work.* And yet another participant reflected on his early days in the medical field to make a point about the positive aspects of automation *Much like automation for lab tests, AI will free up the providers' hands and mind to focus on higher order issues. As an intern, I had to spin my own hematocrits at night. I do not miss that at all!*

Traditional arguments raised against Health AI are understated in this perspective. It is not problematic that AI is a black box technology since health professionals have been using other black box technologies,

Table 3

Four factor arrays where each array features the normalized scores [-5, +5] assigned to the statements in the q-set by participants who loaded significantly on the factor array.

Statement	Factor 1	Factor 2	Factor 3	Factor 4
1	0	-4	-5	-3
2	-1	-5	-3	-1
3	2	4	-3	2
4	-1	-4	-2	-4
5	1	-3	2	0
6	0	2	-1	1
7	0	0	-3	1
8	-1	1	1	0
9	4	5	4	4
10	-2	-5	-4	-2
11	-2	2	3	5
12	1	4	-1	2
13	-2	1	-5	-1
14	-2	-2	1	0
15	2	-1	1	-4
16	1	-2	-1	-2
17	-3	1	2	-2
18	-5	-1	0	-1
19	4	0	0	1
20	-5	-4	-1	1
21	-4	2	-2	-2
22	-4	-1	-4	-3
23	-1	1	-2	0
24	3	2	2	2
25	-4	-2	0	-5
26	0	3	4	4
27	3	0	1	2
28	1	-2	-1	0
29	0	3	1	0
30	3	5	5	-1
31	-1	-3	0	-1
32	4	3	5	4
33	5	4	3	5
34	0	0	0	-4
35	2	-1	3	1
36	5	0	2	3
37	2	1	4	3
38	-3	-3	-2	-5
39	-3	-1	-4	-3
40	1	0	0	3

such as MRI (16:1). And there is also a neutrality about AI's lack of contextual knowledge and ability to read social clues (26:0). Along these lines, one participant noted that *The role of a skilled physician is to take into consideration what a machine / AI tells him and make the correct connection with clinical reality.* Moreover, this perspective does not subscribe to the thought that AI will worsen problems in healthcare such as overtesting, overdiagnosis, and overtreatment (21:-4). About the role of AI in Hospitals, one participant wrote that *AI is going to play a pivotal role in stratification, thereby assigning patients into low-risk and high-risk groups or patients responding to a certain treatment or patients not responding to it. This will prevent testing or treating patients in whom it is deemed not efficient.*

Despite the positive outlook about AI, this perspective emphasizes that medical doctors must remain in charge not only in the medical decision process (18:-5) but also by participating in the technology design process (33:+5) *(AI will only help the physicians resolving their clinical doubts, but the last decision should never be given by AI; I see AI as an additional tool, not as something that will replace MDs or decrease autonomy*

Table 4

Four perspectives about Health AI.

P1	AI is a helpful tool: Let physicians do what they were trained for
P2	Rules & regulations are crucial: Private companies only think about money
P3	Ethics is enough: Private companies can be trusted
P4	Explainable AI tools: Learning is necessary and inevitable

and authority. A MD will always have the final verdict; AI designers know the technology, but need MDs to design relevant products; The goal of a certain AI tool has to be defined together with the medical doctors to ensure clinical relevance; As doctors lack informatics skills, engineers lack medical knowledge and hospital needs, therefore medical Doctors are key in the design of AI.)

Looking forward to the future of medical employment, according to this perspective AI-based hospitals should not save money by employing less skilled doctors (20:-5) (*With AI tools working well, Medicine will advance to a more precise act, with decision based on multidisciplinary team opinion, so highly skilled practioners will be most needed*) but it is considered that even though AI will not replace doctors, doctors who use AI will replace doctors who do not (19-4) *As in any area of technical progress, AI is a tool that will be embraced by those at the cutting edge. Those who don't – like surgeons who never mastered laparoscopy – will find their scope of practice diminishing.*

4.2. Perspective 2: Rules & regulations are crucial: private companies only think about money

In the second perspective identified in this study, there is a clear negative outlook about AI technology (21:2; 31-3) and a clear distrust in private health companies. There is a sentiment that the tech industry is not well aligned with core healthcare values (*Private companies only think about MONEY*), has little knowledge about Medicine (13:1), and poses a risk for monopolistic behavior (12:4).

About the risk for monopolistic behavior, there is a great concern about the implications of these big companies owning medical data. Privacy is heralded as a core ethical value in the medical field also in a future where healthcare relies heavily on AI (2:-5; 4:-4). In particular, there are concerns about the implications on the patient physician relationship. A participant noted that *when a patient is worried that confidentiality is breached towards a tech company or insurance company, they may not provide full information or avoid treatment leading to a more profound disease.* There are also concerns about the power that would come from owning such data. A participant cautioned that *Healthcare AI companies would have too much power if this information was not anonymous. They would sell information about specific people to drug companies, to hiring companies, to insurance companies*

Because in this perspective private companies are not to be trusted, there is a strong emphasis in rules and regulations to keep these corporations in check. It is considered that the tech mantra “fail fast and fix later” is putting patients at risk and that regulators are not doing enough to keep consumers safe (29:3). Therefore, even though Health AI technologies and its developers should be bound by core medical ethics (9:5), clear rules about liability, data, and product certification must also be in place.

Moreover, it is perceived that technology companies must be liable for medical errors even if such liability hampers technological advancement (10:-5). A participant elaborated that *profit demands risk and the companies must bear that risk. Developers must be accountable if they wish to enter the demanding arena of care.* These technologies should only be used in healthcare once clear rules and regulations about data usage, storage, and anonymization, are in place in AI (3:4).

Along the same lines, it is also emphasized that AI health products must be tested in randomized clinical trials, which is the strongest source of medical evidence (30: +5). One participant illustrated this point quite clearly by stating that *since the diagnostic and treatment AI tools affect directly the patients health, they should be held against the highest standards as usual in medicine for new diagnostic and treatment strategies. I do not see why this should be different for AI then for new” conventional diagnostic tests” or drugs.*

4.3. Perspective 3: Ethics is enough: private companies can be trusted

The most striking feature about this perspective is the overall

positive outlook about AI companies. According to this view, it is not undesirable that these companies start operating in the health space (13:-5). Moreover, there are no major concerns about the risk for monopolistic behavior (12:-1).

A potential explanation about this positive account on tech is the perception that current health systems already rely heavily on technology and tech companies. As one participant noted, about the potential of automation in Hospitals, *this is ridiculous. Automation already works in Hospitals.*

Rather than dwelling on rules and regulations (3:-3), ethics in itself is enough to ground the private sector. AI technology must be aligned with bioethical principles (32:+5) such as Privacy (1:-3; 2:-3), which should remain a core medical value (4:-2). Despite the trust in tech companies, also in this perspective the need for testing of AI health products is emphasized (30:5).

According to this perspective, AI will not increase discrimination based on predicted future medical problems (7:-3) and therefore improving equity and inclusion is not mandated to be the top priority when developing and deploying AI in healthcare (6:-1).

4.4. Perspective 4: Explainable AI tools: learning is necessary and inevitable

Explainability is a key value in this perspective. In order to reap the benefits of AI, medical doctors must understand and lead the AI technological progress. A participant wrote *AI should never be a “black box”.* *Doctors should be able to explain the results from AI tools with reasoning.* Not only health professionals need to know how AI medical tools work (15:-5), but in fact such tools should only be used if clinicians understand how AI decisions are made (11:5). Along these lines, one participant remarked that *the adoption of AI will be improved if doctors do understand the ‘black box’.*

According to this perspective, doctors are interested in learning about AI and Computer Science (25:+5) and have the time to learn how to use complex AI-based medical devices (3:-4). One participant noted that *overall, the intellect of doctors is underestimated and under evaluated by technicians;* and another highlighted that *doctors cannot work without computers and use them daily for registration. Learning is necessary and inevitable.*

It is considered problematic that machines lack contextual knowledge and ability to read social clues (26:4) and it is difficult to operationalize clinical practice for a machine (38:-5). Therefore, medical doctors must participate in the design process of AI for Healthcare (33:5). Accordingly, a participant wrote that *AI is here to stay (I think), and medical doctors are the most suited to adjust and improve the various algorithms etc. that are currently being designed.*

5. Discussion

The perspectives identified in this study reveal diverse and often contradictory viewpoints about Health AI. Understanding these underlying values and tensions is important for operationalizing the ethical issues associated with the implementation of AI technologies in healthcare. Ultimately, such operationalization is expected to lead to more meaningful debates and policies towards an ethically aligned deployment of Health AI.

Our study offers a systematic analysis of the perspectives of medical doctors about Health AI. It is possible to observe elements of the four reported perspectives in the current literature. Several articles have reported findings that trace back to P1 (AI is a helpful tool: Let physicians do what they were trained for) with respect to the positive attitudes about the use of AI as a supportive technology [18,19,26,27,25,24]. The need for AI medical school training to be expanded and improved is also well addressed in the literature [18,20,21,26,25] and relates to P4 (Explainable AI tools: Learning is necessary and inevitable), which contends that doctors must understand AI. Moreover, none of the

perspectives identified in this study reveal concerns about job replacement, which is a finding aligned with studies reported in the literature [18,26,17,32,24]. The strength of our work lies on the fact that, because participants were surveyed on a wide range of ethical issues related to Health AI, the perspectives that have emerged provide a more comprehensive picture of the moral views of practitioners.

Each perspective provides insights about the AI Health ethics outlook and also about particular ethical issues, such as Fairness, Explainability, and Ethical Design, which need to be accounted for on the implementation processes of AI technologies in healthcare.

About the outlook on AI Health Ethics, P1 and P2 (Rules & Regulations are crucial: Private companies only think about money) represent somewhat conventional views about AI technology, which contrast with P3 (Ethics is enough: Private companies can be trusted).

The perception that AI-based medical tools will improve efficiency in the clinical setting builds on decades of successful development of sophisticated medical technologies. In general, this perspective (P1) is aligned with the narratives presented by tech companies, which tend to focus on the benefits of technology and automation to take over repetitive tasks. Similar accounts are presented by developers of other AI-based technologies, such as the autonomous vehicle [55].

By projecting AI-based medical technologies as yet another type of medical tool, higher order conversations about ethical issues associated with this technology, such as Fairness or Human Autonomy, are to some degree avoided. However, because of the disruptive potential of these technologies, further thoughts about ethics are required. Tech companies and developers should indeed acknowledge the singularity of AI and ensure that the design process is to be guided by ethical considerations.

When the emphasis is not just on the benefits of AI technology, regulation is often seen as the solution to ensure the safety of consumers. Medical doctors who are hesitant about AI-based medical tools, consider that rules and regulations are a crucial element in the transition for AI-based healthcare. This perspective is well aligned with the tradition in the healthcare and medicine fields, which are notoriously heavily regulated [56]. The regulation of AI-based medical devices (also known as Artificial Intelligence/Machine Learning-Based Software as a Medical Device) has unique challenges and is known to be lagging behind the technology [57].

In contrast, a less conventional perspective with respect to the ethics of AI Health reported in this study (P3: *Ethics is enough: Private companies can be trusted*) considers that heavy regulation of tech companies is not needed. By refusing to demonize AI tech companies, this perspective provides a somewhat unusual positive outlook about these health stakeholders. It builds on the idea that ethical awareness is enough for corporations to be trusted. Recently it has been argued that AI corporations should indeed promote virtue ethics, rather than the traditional deontology-infused guidelines, as an effective form of ensuring ethical behavior in corporations [10].

While completely forgoing regulations may be unrealistic, private tech companies should indeed internalize that to be accepted as an ethical player in the health space while also striving for profits, must promote ethical environments and practices.

About particular ethical issues, our study unraveled tensions and contradictory viewpoints that should be accounted for in future debates, namely with respect to Fairness, Explainability and Ethical Design.

Fairness, non-discrimination, and justice relate to reasonableness and impartiality of actions. AI-based technologies are designed and produced by humans and rely extensively on data thus being exposed to errors, ill judgments, and prejudices which can enter into the innovation lifecycle and create biases [58,59]. There are several concerns about biased or discriminatory outcomes in the context of Health AI. A biased medical device operates in such a manner that produces disadvantages to certain demographic groups and influences health inequality [60]. Different types of bias are associated with medical AI-powered devices, namely physical bias (design of the medical device disadvantages

certain demographic groups based on physical traits such as skin color), computational bias (training datasets that serve as inputs in medical device are not representative of population), and interpretation bias (medical device is subject to biased inference of readings) [60].

Our study indicates that there are more concerns about these matters in P2 and less in P3, which is not surprising given the remaining features of these perspectives. However, in general, all perspectives are quite neutral when it comes to fairness and discrimination in Health AI. This neutrality could well be a short-coming of this study, namely related the selected statements about Fairness, but could also signal that medical doctors are ill-informed about these issues or they just do not consider them as relevant or pressing. Recently some concrete cases about unfair medical devices have been reported in the literature (for instance it has been reported that pulse oximeters are not as accurate in measuring blood oxygenation in Black patients [60]), but it could be the case that these matters remain largely abstract for the majority of practitioners. Future empirical research should further explore the views of practitioners on Fairness and bias issues.

Another important issue associated with the ethics of Health AI is Explainability. An explainable model provides *interpretable* (description of a system in a way that can be understood by humans) and *complete* (accurate description of the operation of a system) information about the system [61]. The challenge of explainability is therefore to reach both interpretability and completeness, given that accurate explanations are not easily interpretable and the latter often lack predictive power [61]. AI-powered medical technologies rely on complex algorithms which are not easily interpretable, thus known as black-boxes.

Our study revealed contrasting viewpoints with respect to explainable Health AI technologies. According to P1, the lack of explainability in AI-powered devices is not problematic, since health professionals have been using other complex technologies, such as MRI, which also resemble black-boxes. In contrast, P4 considers that Explainability is a key value and that in order to reap the benefits of AI, medical doctors must understand the intricacies of AI powered medical devices.

The comparison with MRI is often called for in this literature but it is not widely accepted. Indeed the MRI is a complex medical technology and practitioners are not expected to know the underlying physics and math of this technology. However, the algorithms that operate these systems are indeed explainable and understood by developers. Differently, the explainability challenge associated with AI is not contingent to medical practitioners but also to developers in general. Lack of knowledge about the decision rules that sustain a certain outcome is especially problematic in the healthcare setting. As one participant remarked, *A good health care professional will never blindly rely on any single measure without the story of the patient.*

The contrasting views reported in our study support the need for further empirical research in order to determine whether practitioners who share different perspectives with respect to Explainability would interact differently with the same algorithm [62].

Both values of Fairness and Explainability explored above should be accounted for in the ethical design of Health AI. Our study shows quite clearly that, regardless of the positive or negative industry outlook, all perspectives consider that medical doctors must participate in the design process of AI health technologies.

By further exploring the comments of participants, it seems that medical doctors even consider that the success and clinical relevance of AI Health depends on the involvement of practitioners in design and development of the technology. The reasons advanced by participants go beyond medical knowledge (*As doctors lack informatics skills, engineers lack medical knowledge and hospital needs, therefore Medical Doctors are key in the design of AI.*), and include also clinical reasoning (*Without knowing how a clinical thinks, than AI would not be a useful tool*), and the societal role of medical practitioners (*MDS are trained and dedicated in ethical and societal decision making. They are natural bridge builders between a complex medical/technical reality and the personal space of an individual patient. Crossing this bridge is fundamental for any novel*

development to have a reason of existence. Without MDs involved, the major stakeholders of the development are neglected).

The development of AI Health technology should be a multi-disciplinary effort. Whether practitioners will act as advisors or, as cautioned in P4, will take a more prominent role in the tech development is yet to be seen. Further research should explore models of technology development that are able to integrate the elements mentioned above in the design process.

Our exploratory study revealed four perspectives about Health AI, which we expect may help to shape future debates as well as ethical design processes. There are contrasting views about the Ethics of Health AI in general but also about particular ethical issues such as Explainability. It is observed that medical doctors are more concerned about the role of large companies in healthcare and less aware or concerned about higher level and often abstract issues such as Fairness, bias, and health inequalities.

There are important limitations in this study. The first limitation is related to the filtering process, in which the set of 353 statements retrieved from scientific and popular literature was reduced to 40 statements. This process was carried out by the authors, who do not have medical training. We acknowledge that having the input of a medical doctor in such filtering process would add value to this study by allowing us to have a better understanding of the relevance and knowledge of medical practitioners about the matters captured in the statements. Another limitation is related to the recruitment of participants. As mentioned earlier, the vast majority of participants in this study was recruited through recent scientific publications. We may therefore have failed to capture perspectives of practitioners that are less involved in academic and research activities. By recruiting medical doctors from three Western countries we also failed to include in our p-set medical doctors from developing countries which may have contributed with additional perspectives about Health AI. Finally, while in this study we have focused on perspectives regarding AI-applications in Healthcare in general, there is a need to expand the literature and look into particular domains and tasks of Health AI.

6. Conclusion

For AI to meet its potential in the complex Healthcare space, ethics needs to be taken into account. In this empirical study we surveyed medical doctors based in The Netherlands, Portugal, and U.S. on a wide scope of ethical issues about Health AI. This survey allowed us to discern different perspectives about the ethics surrounding the deployment of Health AI.

We identified four main perspectives: P1: *AI is a helpful tool: Let physicians do what they were trained for*; P2: *Rules & Regulations are crucial: Private companies only think about money!*; P3: *Ethics is enough: Private companies can be trusted*; and P4: *Explainable AI tools: Learning is necessary and inevitable!*

Each perspective provides valuable insights about ethical issues that should be operationalized and accounted for in the design and development of these technologies. Our study reveals contrasting viewpoints about the ethics associated with Health AI. It is also observed that medical doctors are mostly concerned about the role of large companies in healthcare and less aware or concerned about higher level issues such as Fairness, bias, and health inequalities. Regardless of the positive and negative industry outlook, our study revealed that medical doctors consider that they must participate in the design process. These findings are useful starting points for a fruitful discussion between medical professionals, industry stakeholders, and policy-makers.

Given the exploratory nature of this research, there is ample opportunity for confirmatory research directions and to explore how to translate these perspectives into actionable insights and design models for the different health stakeholders.

Acknowledgements

The authors acknowledge the anonymous reviewers for providing helpful comments on earlier versions of the manuscript, and the European Research Council for financial support of this research (ERC Consolidator grant BEHAVE – 724431).

References

- [1] Yu K-H, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng* 2018;2(10):719–31.
- [2] Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. *Gastrointest Endosc* 2020;92(4):807–12. <https://doi.org/10.1016/j.gie.2020.06.040>. <http://www.sciencedirect.com/science/article/pii/S001651720344667>.
- [3] Wallis C. How artificial intelligence will change medicine. *Nature* 2019;576(7787):S48.
- [4] van Hartskamp M, Consoli S, Verhaegh W, Petkovic M, van de Stolpe A. Artificial intelligence in clinical health care applications. *Interact J Med Res* 2019;8(2):e12100.
- [5] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017;2(4):230–43.
- [6] Wolff J, Pauling J, Keck A, Baumbach J. The economic impact of artificial intelligence in health care: systematic review. *J Med Internet Res* 2020;22(2):e16866. <https://doi.org/10.2196/16866>. <http://www.jmir.org/2020/2/e16866/>.
- [7] Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2014;2(1):3.
- [8] He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019;25(1):30–6.
- [9] Morley J, Machado CC, Burr C, Cows J, Joshi I, Taddeo M, et al. The ethics of ai in health care: a mapping review. *Soc Sci Med* 2020;260:113172. <https://doi.org/10.1016/j.socscimed.2020.113172>. <http://www.sciencedirect.com/science/article/pii/S0277953620303919>.
- [10] Hagendorff T. The ethics of ai ethics: an evaluation of guidelines. *Mind Mach* 2020:1–22.
- [11] Rigby MJ. Ethical dimensions of using artificial intelligence in health care. *AMA J Ethics* 2019;21(2):121–4.
- [12] Price WN, Cohen IG. Privacy in the age of medical big data. *Nat Med* 2019;25(1):37–43.
- [13] Grigger E, Khoury C. Making policy on augmented intelligence in health care. *AMA J Ethics* 2019;21(2):188–91.
- [14] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25(1):44.
- [15] Lai M-C, Brian M, Mamzer M-F. Perceptions of artificial intelligence in healthcare: findings from a qualitative survey study among actors in france. *J Transl Med* 2020;18(1):14. <https://doi.org/10.1186/s12967-019-02204-y>.
- [16] Blease C, Kaptchuk TJ, Bernstein MH, Mandl KD, Halamka JD, DesRoches CM. Artificial intelligence and the future of primary care: exploratory qualitative study of UK general practitioners' views. *J Med Internet Res* 2019;21(3):e12802.
- [17] Oh S, Kim JH, Choi S-W, Lee HJ, Hong J, Kwon SH. Physician confidence in artificial intelligence: an online mobile survey. *J Med Internet Res* 2019;21(3):e12422.
- [18] Santos DP Dos, Giese D, Brodehl S, Chon S, Staab W, Kleinert R, et al. Medical students' attitude towards artificial intelligence: a multicentre survey. *Eur Radiol* 2019;29(4):1640–6.
- [19] van Hoek J, Huber A, Leichte A, Härmä K, Hilt D, von Tengg-Kobligk H, et al. A survey on the future of radiology among radiologists, medical students and surgeons: students and surgeons tend to be more skeptical about artificial intelligence and radiologists may fear that other disciplines take over. *Eur J Radiol* 2019;121:108742.
- [20] Sit C, Srinivasan R, Amlani A, Muthuswamy K, Azam A, Monzon L, et al. Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: a multicentre survey. *Insights Imaging* 2020;11(1):14.
- [21] Ooi SKG, Makmur A, Soon AYQ, Fook-Chong S, Liew C, Sia SY, et al. Attitudes toward artificial intelligence in radiology with learner needs assessment within radiology residency programmes: a national multi-programme survey. *Singapore Med J* 2019;1:22.
- [22] Park CJ, Paul HY, Siegel EL. Medical student perspectives on the impact of artificial intelligence on the practice of medicine. In: *Current problems in diagnostic radiology*; 2020.
- [23] Jungmann F, Jorg T, Hahn F, dos Santos DP, Jungmann SM, Düber C, et al. Attitudes toward artificial intelligence among radiologists, it specialists, and industry. In: *Academic radiology*; 2020.
- [24] Coppola F, Faggioni L, Regge D, Giovagnoni A, Golfieri R, Bibbolini C, et al. Artificial intelligence: radiologists' expectations and opinions gleaned from a nationwide online survey. In: *La Radiologia medica*; 2020.
- [25] Sarwar S, Dent A, Faust K, Richer M, Djuric U, Van Ommeren R, et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. *NPJ Digit Med* 2019;2(1):1–7.
- [26] Polesie S, Gilstedt M, Kittler H, Lallas A, Tschandl P, Zalaudek I, et al. Attitudes towards artificial intelligence within dermatology: an international online survey. *Br J Dermatol* 2020.

- [27] Shen C, Li C, Xu F, Wang Z, Shen X, Gao J, et al. Web-based study on chinese dermatologists' attitudes towards artificial intelligence. *Ann Transl Med* 2020;8(11).
- [28] Chan KS, Zary N. Applications and challenges of implementing artificial intelligence in medical education: integrative review. *JMIR Med Educ* 2019;5(1):e13930. <https://doi.org/10.2196/13930>. <http://mededu.jmir.org/2019/1/e13930/>.
- [29] van der Niet AG, Bleakley A. Where medical education meets artificial intelligence: 'does technology care?'. *Med Educ* 2021;55(1):30–6 (arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/medu.14131>, doi:<https://doi.org/10.1111/medu.14131>). URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/medu.14131>.
- [30] Kolachalama VB, Garg PS. Machine learning and medical education. *NPJ Digit Med* 2018;1(1):54. <https://doi.org/10.1038/s41746-018-0061-1>.
- [31] Gao S, He L, Chen Y, Li D, Lai K. Public perception of artificial intelligence in medical care: content analysis of social media. *J Med Internet Res* 2020;22(7):e16649.
- [32] Doraiswamy PM, Blease C, Bodner K. Artificial intelligence and the future of psychiatry: insights from a global physician survey. *Artif Intell Med* 2020;102:101753.
- [33] Jobin A, Ienca M, Vayena E. The global landscape of ai ethics guidelines. *Nat Mach Intell* 2019;1(9):389–99.
- [34] Brault N, Saxena M. For a critical appraisal of artificial intelligence in healthcare: The problem of bias in mhealth. *J Eval Clin Pract* 2020. n/a (n/a). [arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/jep.13528>, doi:<https://doi.org/10.1111/jep.13528>]. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/jep.13528>.
- [35] Tat E, Bhatt DL, Rabbat MG. Addressing bias: artificial intelligence in cardiovascular medicine. *Lancet Digit Health* 2020;2(12):e635–6.
- [36] Cirillo D, Catuara-Solarz S, Morey C, Guney E, Subirats L, Mellino S, et al. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. *NPJ Digit Med* 2020;3(1):81. <https://doi.org/10.1038/s41746-020-0288-5>.
- [37] Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019;17(1):195.
- [38] Vellido A. Societal issues concerning the application of artificial intelligence in medicine. 2019. <https://doi.org/10.1159/000492428>.
- [39] McKeown B, Thomas DB. *Q methodology*. Sage publications; 2013.
- [40] Watts S, Stenner P. Doing q methodology: theory, method and interpretation. *Qual Res Psychol* 2005;2(1):67–91.
- [41] Watts S, Stenner P. Doing Q methodological research: theory, method & interpretation. Sage; 2012.
- [42] Stephenson W. The study of behavior; q-technique and its methodology. University of Chicago Press; 1953.
- [43] Stephenson W. Technique of factor analysis. *Nature* 1935.
- [44] Zabala A, Sandbrook C, Mukherjee N. When and how to use q methodology to understand perspectives in conservation research. *Conserv Biol* 2018;32(5):1185–94.
- [45] Kamal S, Kocór M, Grodzińska-Jurczak M. Quantifying human subjectivity using q method: when quality meets quantity. *Qual Sociol Rev* 2014;10(3).
- [46] Linstone HA, Turoff M, et al. The delphi method. Reading, MA: Addison-Wesley; 1975.
- [47] Stephenson W. Concourse theory of communication. *Communication* 1978;3:21–40.
- [48] Menon C, Alexander R. A safety-case approach to the ethics of autonomous vehicles. In: *Safety and reliability*. vol. 39. Taylor & Francis; 2020. p. 33–58.
- [49] S. Banasick, Ken-q analysis (version 1.0.6) (2019).
- [50] Schmolck P. Pq-method, version 2.11 manual, Neibiderg. Germany: University; 2002.
- [51] Ringné M. What is principal component analysis? *Nat Biotechnol* 2008;26(3):303–4. <https://doi.org/10.1038/nbt0308-303>.
- [52] Wold S, Esbensen K, Geladi P. Principal component analysis, chemometrics and intelligent laboratory systems. In: *Proceedings of the multivariate statistical workshop for geologists and geochemists*. 2 (1); 1987. p. 37–52. [https://doi.org/10.1016/0169-7439\(87\)80084-9](https://doi.org/10.1016/0169-7439(87)80084-9). <http://www.sciencedirect.com/science/article/pii/0169743987800849>.
- [53] Akhtar-Danesh N, et al. A comparison between major factor extraction and factor rotation techniques in q-methodology. *Open J Appl Sci* 2017;7(04):147.
- [54] Zabala A, Pascual U. Bootstrapping q methodology to improve the understanding of human perspectives. *PloS one* 2016;11(26845694):e0148087. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4742059/>.
- [55] Martinho A, Herber N, Kroesen M, Chorus C. Ethical issues in focus by the autonomous vehicles industry. *Transp Rev* 2021;0(0):1–22. <https://doi.org/10.1080/01441647.2020.1862355>.
- [56] Field RI. Why is health care regulation so complex? P & T: a peer-reviewed journal for formulary management 2008;33(19750043):607–8. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2730786/>.
- [57] Minssen T, Gerke S, Aboy M, Price N, Cohen G. Regulatory responses to medical machine learning. *J Law Biosci* 2020.
- [58] Leslie D. Understanding artificial intelligence ethics and safety. arXiv preprint [arxiv:190605684](https://arxiv.org/abs/190605684) 2019.
- [59] Bolukbasi T, Chang K-W, Zou JY, Saligrama V, Kalai AT. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In: *Advances in neural information processing systems*; 2016. p. 4349–57.
- [60] Kadambi A. Achieving fairness in medical devices. *Science* 2021;372(6537):30–1.
- [61] Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. Explaining explanations: an overview of interpretability of machine learning. In: *2018 IEEE 5th international conference on data science and advanced analytics (DSAA)*. IEEE; 2018. p. 80–9.
- [62] Gerke S, Babic B, Evgeniou T, Cohen IG. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. *NPJ Digit Med* 2020;3(1):53. <https://doi.org/10.1038/s41746-020-0262-2>.