

KEYWORDS:

SCIENTIFIC BIAS | HEALTH TECH | AI ERRORS | MEDIA DISCOURSE | PUBLIC AWARENESS

DOI:

<https://doi.org/10.5771/2747-5174-2022-1-34>



Veronica Barassi is a Professor of Media and Communication Studies in the School of Humanities and Social Sciences at the University of St. Gallen, Switzerland, as well as Chair of Media and Culture in the Institute of Media and Communications Management. She is an anthropologist who researches and writes about the impact of data technologies and artificial intelligence on human rights and democracy.



Rahi Patra is a PhD candidate and researcher in the School of Humanities and Social Sciences at the University of St. Gallen, Switzerland. Her research focuses on healthcare surveillance technologies, algorithmic bias and their implications on human rights.

Acknowledgements

This work emerged in the context of a wider research project titled, The Human Error Project: AI, Human Nature and the Conflict over Algorithmic Profiling, funded by the University of St. Gallen. The authors would like to thank their project team members, Ms. Antje Scharenberg and Ms. Marie Poux-Berthe for their contribution to the data collection. The authors also wish to thank the Institute of Media and Communications Management (MCM) and the University of St. Gallen for their support.

AI Errors in Health?

THE PROBLEM OF SCIENTIFIC BIAS AND THE LIMITS OF MEDIA DEBATE IN EUROPE

AUTHORS: Veronica Barassi & Rahi Patra

ABSTRACT:

The ever-greater use of AI-driven technologies in the health sector begs moral questions regarding what it means for algorithms to mis-understand and mis-measure human health and how as a society we are understanding AI errors in health. This article argues that AI errors in health are putting us in front of the problem that our AI technologies do not grasp the full pluriverse of human experience, and rely on data and measures that have a long history of scientific bias. However, as we shall see in this paper, contemporary public debate on the issue is very limited. Drawing on a discourse analysis of 520 European news media articles reporting on AI-errors the article will argue that the ‘media frame’ on AI errors in health is often defined by a techno-solutionist perspective, and only rarely it sheds light on the relationship between AI technologies and scientific bias. Yet public awareness on the issue is of central importance because it shows us that rather than ‘fixing’ or ‘finding solutions’ for AI errors we need to learn how to coexist with the fact that technologies – because they are human made, are always going to be inevitably biased.

1. INTRODUCTION

The collection, analysis and processing of data has been historically crucial to healthcare. Translation of ancient Egyptian hieroglyphic inscriptions and papyri indicate that the use of medical records has been a common practice since ancient civilization. Even if historically health data has been central to the functioning of healthcare, in the last two decades the rise of big data and AI technologies have radically transformed the ways in which health data is produced, processed, and analyzed. This transformation occurred at different levels in our societies. At macro level, national governments and public health institutions have invested largely in the datafication of the health sector and the use of AI technologies for both the monitoring of individual patients and the tracking of public health. The rise of the Covid-19 pandemic, for instance, has radically transformed our reliance on these technologies for public and health surveillance (Calvo et al., 2020). AI-driven technologies were not only used to trace contacts and even Covid-19 patients themselves, through the use of face recognition technologies, as in South Korea (Parry, 2021), but they were also used to predict Covid-19 infections by the sound of a cough (Knowles, 2021) or decide how Covid-19 vaccinations would be allocated (Singer, 2021).

Over the last years we have also seen the rise of AI technologies in medical practice. According to Hamet and Tremblay (2017), broadly speaking the AI impact in medicine happened on two different – yet interconnected dimensions, the virtual dimension (which includes informatics approaches from deep learning information management to control of systems' guidance of physicians in their treatment decisions) and the physical dimension (which includes robots and technologies used to assist patients or doctors). The benefits of this transformation can be many. This is particularly true if we think about how AI systems are used for diagnosis, disease monitoring and the creation of personalized treatment plans – see for example, the key developments in treating oncology patients (Jaryd et al., 2021), clinical drug discovery (Fleming, 2018) or genomics research (Dias and Torkamani, 2019). The transformation did not only affect our society at macro level but also at micro level, with the rapid expansion in use of health apps, wearables, and other technologies that gather, process and analyze large amounts of users' data. These types of technologies fall broadly into the market definition of mHealth (mobile health). A Grandview Research survey – published in February 2021 – estimates the total value of the mHealth market at \$45.7 billion in 2020 and forecasts an annual growth rate of 17.6% between 2021 and 2028 (Grandview Research, 2021).

The ever-greater use of AI in health technologies begs moral questions regarding what it means for algorithms to mis-understand and mis-measure human health and the current discourses that are emerging around AI errors in health. This article argues that if we want to understand contemporary AI errors in health, we need to investigate the history of scientific bias and appreciate that the technologies we are creating today are often based on scientific data and measures that carry with them a long history of racial and gender bias. Yet we will show that public debate of this issue – although of central importance – is still limited. Relating the findings of a critical discourse analysis of 520 European media articles reporting on AI-errors in 3 different European countries (Germany, France, and the UK) and 2 countries of

interest (US and Switzerland), the article will show that reporting on AI errors in health is often defined by a techno-solutionist perspective and rarely addresses the problem of scientific bias. Rather than falling into the trap of techno-solutionism, this paper will conclude that we need to acknowledge the fact that our technologies do not grasp the pluriverse of human experience and because they are human made, they are always going to be inevitably biased (Barassi, 2021).

2. LITERATURE REVIEW

2.1. THE HEALTH SECTOR AND SCIENTIFIC BIAS: A BRIEF OVERVIEW

The need to better understand the morals and ethics of AI in health technologies has perhaps never been as urgent as during and since the Covid-19 pandemic. According to Madianou (2020), since the pandemic we may be witnessing a “second-order disaster” here, in that Covid-19 related data practices amplify already existing inequalities (Madianou, 2020). Besides the normalisation of surveillance through technologies like contact tracing apps there is also the risk that data collected in this context may be used in different situations, such as digitised welfare provision, which could further harm already marginalised communities. Similarly, Milan and Treré (2020), raise the question of the “data poor”, that is cases in which processes of quantification and a severe lack of Covid-19-related data in low-income countries and nations in the Global South, perpetuates already existing geopolitical inequalities. The questions that we are being confronted with today are technically not new. If we look at the history of Western scientific thought and especially the social critique of the social sciences, we realize that the monitoring, measuring and evaluation of human health has never been ‘fair or objective’. This is particularly true if we consider the issue of *scientific racism*, or the ways in which body measures have been used to make claims not only about health, but also about racial superiority. Originally published in 1981, Gould, for example, wrote a book called *The Mismeasure of Man*, where he demonstrates how Western scientific thought has often relied on scientific measures

modeled on the white man (Gould, 2006). Gould was particularly fascinated by IQ tests and the idea that intelligence could be measured biologically (e.g. skull measurements) and shows how the benchmarks of these tests use the white man as a reference point. In the 19th century, before Darwin, for instance Morton gathered human skulls measures to argue that: “*ranking of races could be established objectively by physical characteristics of the brain, particularly by its size*” (p. 83). According to his study, the Caucasians were on the top of the racial hierarchy, Black people were at the bottom while the Indians were in the middle of the hierarchy (Gould, 2006). Part of the reason of this so-called superiority was connected to the belief that a simple form of data such as the size of a skull signified “intelligence”.

Of course, these positions have long been discredited. Already in 1917, in his review of Furguson’s work, Robert Park (1917) indicated that the intelligence tests that were being done in schools lacked other important data. and pointed out how Ferguson’s (1916) conclusions were ignorant of the social and economic aspects such as, poverty, lack of quality education, poor living conditions, lack of financial support to schools as contributing factors to their performance levels in tests. Similarly, Jensen (1969) regarded intelligence as highly hereditary and ignored the influence of cultural factors or even the impact of educational inequalities on the academic achievements of Black people.

Racial bias has defined the history of public health initiatives and healthcare provision in Western countries, and studies have shown significant role of racism in creating inequalities towards racial minorities (Gee et al., 2012; Williams et al., 2019). The Tuskegee Syphilis Study is one of the infamous cases of scientific racism in the 20th century. This study highlighted a severe violation of human rights and trust towards the Government. For four decades, the U.S. Public Health Service conducted a study of untreated syphilis on Black people in Alabama. The study began in 1932, was based on the findings of medical literature published in the early 20th century suggesting “nervous resistance to syphilis” as a “natural inferiority of blacks”. Furthermore, this resistance to neurosyphilis was considered as an indication that Black people have rudimentary brain and nervous system development. This unregulated medical research aimed to identify how syphilis impacts Black men who

did not receive proper treatment and then monitor their progress in untreated conditions (Crenner, 2012). The participants were even deceived with the assurance that U.S. Government is providing them with free healthcare treatment (Paludi and Haley, 2014).

Scientific racism has not only defined the history of public health initiatives or research in Western countries but also clinical practice. This is true still today in Europe, where studies have shown that clinical uncertainty or assumptions about patients from other cultures or linguistic backgrounds could be yet another potential factor in differential treatment of racial minor groups (Nelson, 2002; Bowler, 1993). In Netherlands for instance, underestimation of patient pain and lack of communication skills misconstrued as exaggeration, lead to ignorance of an Afro-Caribbean woman’s early symptoms of cervical cancer (Suurmond et al., 2010). Similar cases have been identified in a study conducted in Sweden by Hamed et al. (2020), where ignorance of patients’ pain during labor led to the death of their babies. A woman with severe migraine attack was not attended to until she fainted with pain because Swedish patients with less severity were deemed more important and thus prioritized by the health system. Often patients from racial minor backgrounds have been annotated with stereotypical notions like, they are “bad patients” with “low pain threshold” who “make fuss about nothing” (Bowler, 1993, pp. 172–173) without any scientific evidence backing such claims.

All these examples show clearly that scientific bias still today operates at the intersection of different inequalities (Crenshaw, 1989), and often scientific racism intersects with *scientific sexism*. As decades of feminist scholarship – not least since Simone de Beauvoir’s ground-breaking *The Second Sex* from 1949 – have shown, Western thought has a track record of understanding the female body as *secondary*, that is as a deviation from what is perceived as the norm: the male body (see also Stephen, 2002). More recent publications by feminist activists like Criado-Perez’s *Invisible Women: exposing data bias in a world designed for men* (2019), and scholars like D’Ignazio and Klein’s *Data Feminism* (2020) have shown that we still live in a world designed with a standard (male) human body in mind.

The health-sector in the West has thus been dominated by mis-readings and mis-measures of ‘other bodies’. As critical theorist Ahmed explains in her book *Racialised Body* (2002), often the bodies of colonised “others” were appropriated as “objects of knowledge” (p. 48) production to enlist the differences against that of the “standard” white body to establish the “superiority of the white race”. Here, scientific knowledge and experimentation played an important role in establishing this superiority as,

“...scientific discourse sought to constitute what was normal and ideal. Rather than finding evidence of racial difference, science was actually constructing or even inventing the very idea of race itself as bodily difference and bodily hierarchy.”

Ahmed 2002, p. 50

Given the fact that the health-sector has been shaped by a long history of scientific bias it is not surprising that these inequalities are being translated in our technological designs.

2.2. THE BIAS IN OUR HEALTH TECH

What seems to be emerging clearly from the most recent research is the fact that the technologies we are creating today are based on scientific data and measures that very often carry with them a long history of human reductionism and implicit bias. A particularly important example right now that demonstrates the human reductionism of our systems is found in COVID-19 contact tracking apps. In her fascinating work, Milan has shown that most of these apps rely on a “standard” experimental subject that hardly allows for the exploration of other variables such as gender, ethnicity, race, or low income (Milan, 2020). Milan shows how the roots of this reductionism, stem from the very practice of design. It is for this reason that she draws on the anthropologist Arturo Escobar who advanced a new vision of design theory that takes into account the complex and intersectional pluriverse in which we live (Escobar, 2018).

Rather than taking into account the pluriverse of human experience, our technologies seem to be replicating the social inequalities of the past in contemporary data-driven technologies (Milner and Traub, 2021, p. 4). In 2020, for instance, an article by Boston Review indicated that the pulse oximeters being used to monitor Covid-19 patients are racially biased against non-white people (Moran-Thomas, 2020). Although previous studies (Bickler et al., 2005; Feiner et al., 2007) have suggested that skin colour has an impact on the pulse oximeter accuracy and is a source of racial bias, but most of these technologies are “calibrated using light-skinned individuals, with the assumption that skin pigmentation does not matter” (Bickler et al., 2005; see also Moran-Thomas, 2020). Similarly, there have been cases where Fitbits used for measuring heart-beat tend to show better results for lighter skin tones as compared to the darker ones (Hailu, 2019). Later in the same year, a study by Obermeyer et al. (2019), provided critical insights on how algorithms can create conditions of racial bias against black patients in predicting healthcare emergencies. The risk-prediction algorithm for identifying patients for the “high-risk management” programs, was used extensively by big health systems as well as insurers and applied to around 200 million people in the US per year. The study highlighted that the algorithm relied on the data about the healthcare costs involved in treating a particular patient as an indication for higher health risk (higher the costs involved, greater the risk), rather than considering the level of illness of a patient. The algorithm failed to acknowledge the discrepancies due to unequal access to healthcare services by black patients as compared to white patients, resulting in less money being spent on Black patients (Obermeyer et al., 2019). Therefore, the algorithm was not only responsible for generating racial bias but also for targeting marginalized sections of the society. Scientific bias is not only affecting the technologies that we use in the health sector but also our consumer technologies. In Strings’ book, *Fearing the Black Body* (2019) for instance, the author demonstrates how body mass index calculations (BMI), on which a lot of our medical apps and wellness programs are based, were not established following years of study on healthy weight-form in different ethnicities and cultural contexts, but were modeled exclusively on white bodies.

The understanding that our technologies rely on a long history of scientific bias and human reductionism, begs for critical questions about how as a society we are negotiating with the fact that algorithms often

mis-understand and mis-measure human health and how much awareness there is in public debate about the relationship between AI in health technologies and scientific racism and sexism.

3. THE HUMAN ERROR PROJECT: MAPPING THE DEBATE ON AI ERROR IN EUROPE

The findings discussed in this article have emerged in the context of a research project entitled *The Human Error Project: AI, Human Nature and the Conflict Over Algorithmic Profiling*. The project is based on the belief that one of the most fundamental questions of our times is to map, study and analyze the emerging debates and conflicts over AI errors and algorithmic profiling. Over the last years, we have seen the emergence of academic researchers, civil society organizations, policy makers, and critical tech entrepreneurs who have argued that AI systems are often shaped by systemic inequalities (Eubanks, 2018; Crawford, 2021), by racial biases (Noble, 2018; Richardson et al. 2019) and an inaccurate analysis of human practices and intentions (Barassi, 2020). The combination of bias, inaccuracy and lack of transparency implies that AI systems will always be somehow fallacious in reading humans (Barassi, 2021). The ‘human error of AI’ can have profound impact individual lives and human rights. However, as Aradau and Blanke have shown (2021), little attention has been placed on the study of error and its political life. The aim of the Human Error Project is to investigate the struggles, practices and discourses emerging on the issue of AI error and highlight how – in Europe – we are negotiating with these problems and looking for solutions.

Whilst we use different methods to understand this debate, including organisational mapping and interviews with civil society organisations and ethical tech actors whose work is aimed at developing alternatives to the status quo of algorithmic prediction, this article draws on findings that emerged throughout a critical discourse analysis of how AI errors are discussed in news media. As the sociologist Bourdieu (1991), has shown, news media can be understood as the centre of symbolic power in our societies, and they can have a fundamental role in “constructing our sense of reality”. Of course,

as media researchers and analysts, have shown media power is a subject of great complexity and controversy (Curran and Seaton, 1997, p. 270), and rather than top-down forms of meaning construction, the media often become the places where dominant meanings in society are transmitted, negotiated and contested. We thus believe that news media are a key field of interest when it comes to ‘map the debate’ on AI errors.

Our discourse analysis critically investigated articles published in the European news media between February 2020 and February 2022, with a particular focus on three core countries: Germany, France, and the UK as core countries. Our rationale to focus on the three European media was twofold: firstly, we were interested in how the debate is shaping in Europe in particular in order to decentre and develop an alternative view to the often US-centric discourse of AI; and secondly, we believed that understanding the debate in European media matters in particular because of Europe’s legal and political particularities, which – notably due to EU legislations like the GDPR, Digital Services Act, the Digital Markets Act and the AI Act¹ – presents a legislative environment that could allow for legal and political possibilities to emerge.

For each of our key focus countries, we selected five key media that met these criteria and ensured that they were balanced across the political spectrum (*left leaning, centre left, centre, centre-right, right leaning*) we have also selected articles based on a predefined number of key words (*AI ethics; Algorithmic bias/ inaccuracy/ error; Algorithmic profiling; Algorithmic error; Data rights/ human rights; Data (in)justice/ harms/ (in) equality; Algorithmic (un)accountability*), which varied slightly according to the three languages used (German, French and English). In this way we selected and analyzed articles from other media, including both other mainstream media, such as public broadcasters like *Deutschlandfunk Kultur* in Germany, the *SRF* in Switzerland or the *BBC* in the UK, from each of the selected countries, as well as other media,

including tech-specific publications like *Wired* in the English-speaking world and alternative media like *Netzpolitik* in Germany. We did this in order to understand where the mainstream media discourse fell short of crucial topics. At times we would complement our analysis on a specific topic (e.g. health app) with additional articles taken from the US, Switzerland and other countries, as well as specialist media.

Over the period of two years, we have analysed 520 articles in total, which have been coded alongside nine thematic categories (see forthcoming research report for further details on data gathering and analysis). While this analytical process has informed our overall findings on respective issues, for this article, we will draw on 77 analysed articles, which covered issues relating to the use of algorithm and AI technologies in the health sector. Our research led us to the collection of different case studies of AI errors in health, and to the conclusion that even if the news media are becoming an important space for the discussion of how AI technologies can mis-understand and mis-measure human health, the ‘media frame’ on AI errors in health is often defined by a techno-solutionist perspective, and only rarely it sheds light on the relationship between AI technologies and scientific bias.

4. REPORTING AI ERRORS IN HEALTH: FROM THE FEAR OF MISDIAGNOSIS TO PUBLIC AWARENESS OF SCIENTIFIC BIAS

Of the many different cases collected during our research the case of recently launched dermatology apps which was discussed by selected media in France, Germany, Switzerland and the UK has been particularly interesting. While it is not the only provider of a respective application launched in recent years (other examples European media reported about were *Skinvision*, *Aysa* or *Miiskin*), the perhaps most prominent case, here, is Google’s *Derm Assist*. The application was launched in Europe in 2021 and asks users to upload images of their skin conditions which are then analysed by an AI for possibly matching conditions. The *Financial Times* optimistically called the app a “[b]reakthrough development [that] will assist users in self-diagnosing issues ranging from acne to melanoma” (Murgia, 2021, emphasis added). In the article, the reporter focuses primarily on the views, and opinions of the researchers, and developers of the app or experts that focus on the benefits for early cancer detection and demonstrates a certain level of uncritical awareness and techno-optimism. The approach of the *Financial Times* was not isolated. One of the first striking findings in our analysis of European media coverage on AI-driven dermatology applications was how respective applications by Google and others were met with a sense of techno-optimism and discussed entirely uncritically by various media. Of the thirteen articles reporting about AI-driven dermatology apps which we analysed, only half of them discussed the applications somewhat critically, with some of them almost reading like Google press releases (such as the *Daily Mail*’s headline which read “Hey Google, what’s this rash? Google unveils AI tool that can identify 288 skin conditions including warts, cysts and lipoma from just THREE photos”, Chadwick, 2021), while others, like the aforementioned *Financial Times* article, struck a particularly optimistic tone.

¹ The Artificial Intelligence Act was first proposed on 21 April 2021 by the European Commission as a regulatory framework for AI and is currently under discussion. The AI Act seeks to regulate the use of AI technologies based on their level of risks, classified as: (1) minimal risk, (2) limited risk, (3) high risk and, (4) unacceptable risk. More information can be found here: <https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/>

When articles used a more critical and biased approach they focused on the issue of misdiagnosis. As reported by the *FT*, Google specified that its tool could not replace the diagnosis from a ‘real’ dermatologist, yet the enthusiasm for the app of the *FT* was counterbalanced by different newspaper articles that highlighted the anxiety and panic in which the patients would be left after the app’s verdict following its potential for under- or over-diagnoses of skin diseases. An article by *The Guardian*, for instance reported the opinion of Dr Ray Moynihan, assistant professor and NHMRC early career fellow with Bond University’s Institute for Evidence-Based Healthcare explains this (Davey, 2021):

“What we have at the moment are puffed up press releases and promotional media stories that make no mention of the potential downsides of these experimental AI tools – and one of the biggest downsides is unnecessary diagnosis, and the harm, anxiety and waste that can cause.” Davey, 2021, para. 19

Similarly, an article that appeared on the *Deutsche Welle* flagged the concern that “Google’s app may have a negative knock on effect due to a possible influx of requests for dermatological examination” (Bleiker, 2021). The article explains how an app like Google’s is always more likely to “make **overcautious mistakes**” because, “when in doubt, would rather classify something as a cause for concern than overlook a disease”. Citing Prof. Eoin McKinney from the Center for Artificial Intelligence in Medicine at Cambridge University in Great Britain, the article concludes that: “panic among patients and high costs for the health system”.

Most of the articles that we encountered during our research seemed to be influenced by a techno-solutionist perspective in reporting AI errors in health, one which focused on broader societal impacts of AI errors (overload of health systems) or implications of mis-diagnosis that could be fixed with better technology or better health systems. Only few articles engaged in a critical understanding of the relationship between AI errors in health and scientific bias. In an article that appeared on *Reuters* (Dave, 2021) for instance, the reporter shows how most of the dermatology apps in the market rely on the Fitzpatrick skin types scale, but the scale misses important varieties in human skin especially as far as darker skin tones are concerned, and explains:

“The concern over FST is that its limited scale for darker skin could lead to technology that, for instance, works for golden brown skin but fails for espresso red tones... “Many people would assume I am skin type V, which rarely to never burns, but I burn,” said Dr. Susan Taylor, a University of Pennsylvania dermatologist who founded Skin of Color Society in 2004 to promote research on marginalized communities. “To look at my skin hue and say I am type V does me disservice.” Dave, 2021, para. 7-15

A similar point is made by The Guardian (Madhusoodanan, 2021) which highlights the scientific bias of the databases that are used to train derm apps:

“Another fundamental flaw lies in databases that algorithms study – particularly for skin conditions. Common databases of skin images rarely capture the myriad variations in skin tones and textures from around the world. That’s in part because compared with white patients, only half as many Black or Hispanic people see dermatologists. Patients with less education or lower socioeconomic status are also far less likely to be represented in these image libraries.” para. 12

Whilst the derm apps were widely reported in the media as significant case studies, reflections of scientific bias and AI in health covered other different health applications. An article on *Wired US* (Knight, 2020), which focused on eye-conditions diagnosis for instance argued that these technologies are often biased. The article refers to the analysis of 94 datasets by a team of UK researchers, which indicated that almost all the data were from North America, China, and Europe indicating bias against “*racial groups from under-represented countries*”. The lack of diversity within the datasets used for training the algorithm, lack of key demographic data and lack of transparency of data were some of the key factors identified as a means of racial bias. Amit Kaushal, an assistant professor of medicine at Stanford University, which is cited in the article explains the consequences of such exclusionary data (Knight, 2020):

“When subgroups of the population are systematically excluded from AI training data, AI algorithms will tend to perform worse for those excluded groups... Problems facing underrepresented populations may not even be studied by AI researchers due to lack of available data.” para. 11

During our research we also found reflections on scientific sexism. Between 2020 and February 2021, we found different articles mentioning the gender bias in AI and the female actors who are seeking to address it, including in the French *Le Temps* (Atmani, 2020) and in Germany's *Die Zeit* (Heuser, 2021) and *Süddeutsche Zeitung* (Brühl, 2021).

Our research led us to the conclusion that whilst within mainstream media discourse we are seeing an increased awareness of the relationship between AI errors and health and scientific bias, the debate was still poorly developed in mainstream media reporting. Our analysis also put us in front of another fundamental problem which was evident also in the reporting of other AI errors: the lack of the views of civil society actors who actively engage with questions about algorithmic injustice and scientific bias. This finding is in line with a previous discourse analysis conducted by the Bertelsmann Foundation in Germany, which looked at the reporting on algorithms and AI technologies between 2005 and 2020. Here, too, researchers found that “An economic perspective has dominated the discourse with a generally very positive framing, while political and civil society actors are rarely mentioned” (Fischer and Puschmann, 2021, p.10). Consequently, especially more radical visions like Indigenous AI or Data Feminism, which we have come across in other datasets, such as interviews or organisational mapping, have hardly been mentioned in the mainstream media (besides Atmani, 2020). An exception, in our analysis, are the two organizations *Big Brother Watch* in the UK and *Algorithm Watch*, which is repeatedly quoted by – particularly left-leaning – German, French and Swiss media. Indeed, one article in *Die Zeit* even sees a new “movement” of civil society actors for a more progressive data politics emerging – a movement that is noticeably led by women, in contrast to the male dominated world of coding and AI, according to the article. Greater inclusions of these voices could be crucial in expanding the debate withing mainstream media and raising public awareness of this important topic.

5. CONCLUSION

AI in health technologies are bringing many important benefits, yet critical questions need to be raised about how they often mis-understand and mis-measure humans. These mismeasurements and misunderstandings can lead to significant errors. These errors are not ‘glitches’ that can be fixed, they need to be problematised by looking at the epistemic and ideological foundations of Western health data and health practices. As this article has shown, the health-sector in the West has been dominated by mis-readings and mis-measures of ‘other bodies’ which were rooted in a long history of scientific racism and sexism. Given the fact that the health sector has been shaped by a long history of scientific bias it is not surprising that these inequalities are being translated, and amplified in our health tech. In this sense it would be naïve to believe that we can ‘fix’ AI errors in health. Rather than ‘fixing’ or ‘finding solutions’, these errors teach us that we need to learn how to ‘coexist’ with the fact that Technologies – because they are human made, are always going to be inevitably biased. (Barassi, 2021).

One important step in this direction is to improve public debate on the issue, which is still limited. As we have shown, through a discourse analysis of European media, even if the news media are becoming an important space for the discussion of how AI technologies can mis-understand and mis-measure human health, the ‘media frame’ on AI errors in health is often defined by a techno-solutionist perspective, and only rarely it sheds light on the relationship between AI technologies and scientific bias. Yet public awareness, we conclude, is of central importance because we need to realize that AI technologies will always be entangled in human culture and its meaning-making processes with all its flaws and partial readings of what it means to be human. Such a more in-depth understanding of these problematic histories, the article concludes, will allow us to recognise that we cannot really *free ourselves* from AI errors in health and that we cannot fall into the trap of techno-solutionism.

